



ESCUELA SUPERIOR POLITECNICA DEL LITORAL

Instituto de Ciencias Matemáticas

Ingeniería en Estadística Informática

“Estimadores Robustos para el Vector de Medias y la Matriz de Varianzas y Covarianzas de Vectores Aleatorios Multivariados”

TESIS DE GRADO

Previa a la obtención del Título de:

INGENIERO EN ESTADÍSTICA INFORMÁTICA

Presentado por:

Néstor Rafael Montaña Pulzara

Guayaquil – Ecuador

Año

2009

AGRADECIMIENTO


A la vida, entendiéndola como el conjunto de condiciones iniciales al nacer, eventos ocurridos y personas con las que he interactuado. Cada elemento ha influido en mi formación para ser lo que hoy soy y llegar hasta aquí.

DEDICATORIA

A mi familia, a mis amigos, a mis camaradas y compañeros de Integración Estudiantil.

A todo aquel que ingresa a la Universidad no para mejorar su “nivel” de vida sino para servir, construir, crear y compartir.

TRIBUNAL DE GRADUACIÓN



ING. ROBERT TOLEDO

PRESIDENTE



M.SC. GAUDENCIO ZURITA

DIRECTOR DE TESIS



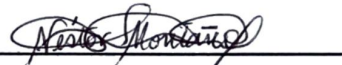
ING. MARGARITA MARTÍNEZ

VOCAL

DECLARACIÓN EXPRESA

"La responsabilidad del contenido de esta tesis de grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la Escuela Superior Politécnica del Litoral"

(Reglamento de graduación de la ESPOL)



Néstor Rafael Montaña Pulzara



CIB E. 1. 1

RESUMEN

En las últimas décadas se ha desarrollado la teoría de la Estimación Robusta, cuya idea es obtener estimadores que cumplan que “un pequeño cambio en la muestra genere sólo un pequeño cambio en el valor de las estimaciones”, siendo este un campo en constante desarrollo se han propuesto diversos métodos de Estimación, en el presente trabajo se realiza un estudio comparativo de varios Métodos de Estimación Robusta para el Vector de Medias y Matriz de Varianzas y Covarianzas además del Método Clásico de Máxima Verosimilitud; el propósito es evaluar el comportamiento de los estimadores bajo diversas condiciones; para ello se realiza una Simulación Matemática en la misma que se generan muestras que luego son contaminadas de diversas formas con el objetivo de simular los valores aberrantes que pueden ocurrir en un estudio real.

En el primer capítulo se expone el problema que motiva el presente trabajo, además se presenta, una serie de definiciones que darán al lector una base teórica que le permitirá comprender los procesos y razonamientos que se emplean al desarrollar el trabajo

En el segundo capítulo se trata de manera somera la teoría básica de robustez, se profundiza en el estudio del método de los estimadores-M univariados, para finalmente se presentan algunos aspectos importantes sobre la estimación Robusta Multivariada y los estimadores a evaluar en el presente estudio

El tercer capítulo corresponde a la teoría detrás de la Simulación matemática mientras que en el capítulo cuatro se detallan los parámetros considerados en la simulación además de los escenarios considerados en la misma, explicando las razones de dichas elecciones; luego de lo cual se presenta un resumen de los resultados de la simulación para cada uno de los estimadores que conforman el presente estudio; efectuando cuadros comparativos y eligiendo el mejor de los estimadores en cada escenario estudiado.

Al final se presentan varias Conclusiones y se plantean varias Recomendaciones en base a los resultados y limitaciones del presente trabajo.

TABLA DE CONTENIDOS

1. BÁSICO

1.1	Introducción	1		
1.2	Planteamiento del Problema – Bosquejo inicial	2		
1.3	Conceptos Básicos.- Probabilidad	7		
	Definición 1.1	Álgebra y σ -álgebra		
	Definición 1.2	σ -álgebra generada por E		
	Definición 1.3	σ -Álgebra de Borel (Boreliano)		
	Definición 1.4	Espacio de Probabilidad		
	Definición 1.5	Variable		
	Definición 1.6	Independencia de Variables Aleatorias conjunta		
	Definición 1.7	Vector Aleatorio y Función de distribución		
	Definición 1.8	Distribución Normal Multivariada		
1.4	Conceptos Básicos.- Estimadores Puntuales	15		
	1.4.1	Estimador	15	
	Definición 1.9	Estadístico y Estimador		
	1.4.2	Métodos de estimación	17	
		1.4.2.1	Método de los momentos	18
	Definición 1.10	Momentos de orden k y Función generadora de momentos		
	Definición 1.11	Función Característica		
	Teorema 1.1	Existencia de la Función Característica		
	Teorema 1.2	Unicidad de la Función Característica		

1.4.2.2	Método de Máxima Verosimilitud	26
Definición 1.12	Función de verosimilitud	
1.4.3	Convergencias de variables aleatorias	30
Definición 1.13	Convergencia puntual de funciones	
Definición 1.14	Convergencia uniforme de funciones	
Definición 1.15	Convergencia en distribución	
Teorema 1.3	Límite Central (Lévy-Lindeberg)	
Definición 1.16	Convergencia en probabilidad	
Teorema 1.4	Ley Débil de los Grandes Números de Tchebychev	
Definición 1.17	Convergencia segura de variables aleatorias	
Definición 1.18	Convergencia casi segura de variables aleatorias	
Definición 1.19	Convergencia en media cuadrática	
1.4.3.1	Algunas relaciones entre los tipos de convergencia	40
Teorema 1.5	De la Convergencia en Probabilidad a Convergencia en Distribución	
1.4.4	Propiedades de los estimadores	43
Definición 1.20	Estimador Insesgado	
Definición 1.21	Estimador asintóticamente insesgado	
Definición 1.22	Eficiencia relativa	
Definición 1.23	Estimador consistente	
Definición 1.24	Estadístico suficiente	
Teorema 1.6	Teorema de Factorización	
Definición 1.25	Estimador insesgado de mínima varianza	
Definición 1.26	Error cuadrático medio	

2 ESTIMACIÓN ROBUSTA

2.1	Introducción	53
2.2	Preliminares	54
2.3	Teoría de la Robustez	57
2.3.1	Generalidades	58
	Definición 2.1 Familia de Contaminación	
	Definición 2.2 Eficiencia	
	Definición 2.3 Estabilidad	
2.3.2	Robustez Cualitativa	60
	Definición 2.4 Robustez según Hampel	
2.3.3	Robustez Cuantitativa	62
	Definición 2.5 Sesgo Asintótico Máximo, Varianza Asintótica Máxima y el Punto de Ruptura Asintótico	
2.3.4	Robustez Infinitesimal	65
	Definición 2.6 Curva de Influencia y sensibilidad a errores groseros	
2.4	Tres métodos para construir estimadores robustos	68
2.4.1	Estimadores L	69
	Definición 2.7 Media Truncada y Mediana	
	Definición 2.8 Estadísticos de orden y Bisagras	
	Definición 2.9 Trimedia	
	Definición 2.10 Media α -Winsorizada	
	Definición 2.11 Estimador de Salto Múltiple	
2.4.2	Estimadores R	72
	Definición 2.12 Estimador de Lehman & Hodges	

2.5 Estimadores M	74
2.5.1 Ideas generales y definición	74
Definición 2.13 Estimadores de tipo Máxima Verosimilitud (Estimadores M)	
2.5.2 Propiedades de los estimadores M	77
Consistencia de los estimadores M	
Teorema 2.1 Consistencia de los estimadores M	
Teorema 2.2 Normalidad asintótica de los estimadores M	
2.5.3 Estimadores M de Huber	82
Definición 2.14 Estimador M Bicuadrático.	
2.5.4 Dos algoritmos para calcular Estimadores M	86
Algoritmo basado en Medias Ponderadas Iteradas (MPI)	
Algoritmo de Newton Raphson (NR)	
2.6 Estimadores de Escala	92
Definición 2.15 Mediana Absoluta de las Desviaciones	
2.6.1 Estimadores M de Escala	93
Definición 2.16 Estimador M Bicuadrático de escala	
2.6.2 Estimadores M de localización y dispersión.	96
2.7 Caso Multivariado	97
2.7.1 Eficiencia de los Estimadores Robustos Multivariados	99
Definición 2.17 Punto de ruptura Asintótico	
2.7.2 Estimadores M para muestras multivariadas	102
2.7.3 Estimadores basados en una escala robusta	106
Definición 2.18 Estimadores S	
Definición 2.19 Estimador S Bicuadrático Multivariado	
Definición 2.20 Estimador S T-Bicuadrático	
Definición 2.21 Estimador Covarianza de Mínimo Determinante MCD	

Definición 2.22	Estimador Stahel-Donoho (DS)	
2.7.4	Algoritmos para calcular Estimadores Robustos Multivariados	111
	Estimadores M Monótonos	
	Soluciones locales para los Estimadores S	

3 MÉTODO MONTE CARLO

3.1	Introducción	114
3.2	¿Porqué Simular?	115
3.3	Método Monte Carlo	117
3.3.1	Monte Carlo bajo hipótesis de Independencia	118
3.3.2	Generación de Variables Aleatorias No Uniformes	122
	Método de Transformación Inversa	
	Generación de muestras aleatorias a partir de una Distribución Normal	
	Generación de una Muestra Normal p-variada	
3.3.3	Generación de números Pseudoaleatorios	126
	Generador Congruencial	

4 SIMULACIÓN

4.1	Introducción	128
4.2	Generalidades	129
4.3	Detalles de la Simulación	132
4.3.1	Parámetros de la Simulación	132

4.4 Resultados de la Simulación	136
4.4.1 Análisis por Escenario	136
4.4.2 Análisis en función del Tamaño de la Muestra, Sesgo y Curtosis Muestral	258

INDICE DE CUADROS

1. BÁSICO

Cuadro 1.1	Cómo influyen los valores extremos	4
------------	------------------------------------	---

2. ESTIMACIÓN ROBUSTA

Cuadro 2. 1	Eficiencia asintótica relativa (ARE) de d_n en relación a S_n	56
Cuadro 2. 2	Sensibilidad de la Media Muestral Vs. Mediana Muestral	66
Cuadro 2. 3	Funciones $\rho(x)$ y $\psi(x)$ para la Media Muestral, Mediana Muestral y el Estimador M; para $c = 1.4$	83
Cuadro 2. 4	Varianza asintótica del estimador M propuesto por Huber	84
Cuadro 2. 5	Densidad de la Distribución Exponencial, Normal estándar y t_{10}	86
Cuadro 2. 6	Procedimiento de Newton Raphson	91
Cuadro 2. 7	Influencia de valores extremos en una distribución Bivariada	98

3. MÉTODO MONTE CARLO

Cuadro 3. 1	Convergencia del Método Monte Carlo en función de n	122
-------------	---	-----

4. SIMULACIÓN

Cuadro 4. 1	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, sin Contaminar	139
Cuadro 4. 2	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ sin Contaminar	142
Cuadro 4. 3	Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ sin Contaminar	146
Cuadro 4. 4	Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ sin Contaminar	149
Cuadro 4. 5	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_1)$	152
Cuadro 4. 6	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_2)$	155
Cuadro 4. 7	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = Uesf(d = 5)$	158
Cuadro 4. 8	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = Uesf(d = 5)$	161
Cuadro 4. 9	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$	164
Cuadro 4. 10	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.05$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$	167

Cuadro 4. 11	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = N(0,9)$	170
Cuadro 4. 12	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = N(0,9)$	173
Cuadro 4. 13	Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y $H = U(-5,5)$	176
Cuadro 4. 14	Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.05$ y $H = U(-5,5)$	179
Cuadro 4. 15	Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_1)$	182
Cuadro 4. 16	Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_2)$	185
Cuadro 4. 17	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.10$ y $H = Uesf(d = 5)$	186
Cuadro 4. 18	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = Uesf(d = 5)$	191
Cuadro 4. 19	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(0.537\mathbf{\alpha}_p, 9\Sigma_1)$	194
Cuadro 4. 20	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(0.537\mathbf{\alpha}_p, 9\Sigma_2)$	197
Cuadro 4. 21	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(0,9)$	200

Cuadro 4. 22	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(0,9)$	203
Cuadro 4. 23	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = U(-5,5)$	206
Cuadro 4. 24	Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = U(-5,5)$	209
Cuadro 4. 25	Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(\mathbf{0}, 9\Sigma_1)$	212
Cuadro 4. 26	Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(\mathbf{0}, 9\Sigma_2)$	216
Cuadro 4. 27	Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, contaminada $\varepsilon = 0.05$ y 0.10 $H = Uesf(d = 5)$	218
Cuadro 4. 28	Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, contaminada $\varepsilon = 0.05$ y 0.10 $H = Uesf(d = 5)$	220
Cuadro 4. 29	Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.05$ y 0.10 $H = N(0.537\alpha_p, 9\Sigma_1)$	223
Cuadro 4. 30	Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.05$ y 0.10 $H = N(0.537\alpha_p, 9\Sigma_2)$	225
Cuadro 4. 31	Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(0,9)$	227
Cuadro 4. 32	Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(0,9)$	229

Cuadro 4. 33	Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = U(-5, 5)$	231
Cuadro 4. 34	Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = U(-5, 5)$	234
Cuadro 4. 35	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.30$ y $H = N(\mathbf{0}, 9\Sigma_1)$	237
Cuadro 4. 36	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.30$ y $H = Uesf(d = 5)$	239
Cuadro 4. 37	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.30$ y $H = Uesf(d = 5)$	241
Cuadro 4. 38	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.30$ y $H = N(0.537\alpha_p, 9\Sigma_1)$	243
Cuadro 4. 39	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.30$ y $H = N(0.537\alpha_p, 9\Sigma_2)$	245
Cuadro 4. 40	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.30$ y $H = N(0, 9)$	247
Cuadro 4. 41	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.30$ y $H = N(0, 9)$	249
Cuadro 4. 42	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.30$ y $H = U(-5, 5)$	251
Cuadro 4. 43	Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.30$ y $H = U(-5, 5)$	253

Cuadro 4. 44	Sesgo vs Error Relativo de los Valores Propios: Caso Población $N(\mathbf{0}, \Sigma)$, sin Contaminar	255
Cuadro 4. 45	Sesgo vs Error Relativo de los Valores Propios: Caso Población $CAU(\mathbf{0}, \Sigma)$, sin Contaminar	257
Cuadro 4. 46	Estimadores Robustos en función del Sesgo y la Curtosis para $k=5$	260
Cuadro 4. 47	Estimadores Robustos en función del Sesgo y la Curtosis para $k=10$	262
Cuadro 4. 48	Estimadores Robustos en función del Sesgo y la Curtosis para $k=20$	264

CAPÍTULO 1

1 BÁSICO

1.1 Introducción

En este primer capítulo se expone el problema que motiva el presente trabajo explicando la metodología a utilizar.

Se encontrará además, algunas definiciones y demostraciones que darán al lector una base para comprender los procesos y razonamientos que se emplean al desarrollar el trabajo. Se ha tratado de definir lo más genéricamente posible cada elemento, prestando atención a los enfoques como la teoría de la medida, de conjuntos, mostrando su aplicación a la teoría de la probabilidad.

1.2 Planteamiento del Problema – Bosquejo inicial

La Estadística como ciencia tiene entre sus objetivos la descripción de fenómenos no explicables determinísticamente; esto es, se ocupa de las mediciones o procesos cuyos resultados presentan variabilidad o, usando un término estadístico, dispersión. Para dicha descripción se ha ideado una serie de características teóricas llamadas parámetros inherentes a la población, las mismas que son clasificadas como Medidas de Tendencia Central o Medidas de Dispersión.

Estos parámetros, en general, no son conocidos (ni fácilmente determinables) y por tanto deben ser estimados utilizando la información disponible, o que se pueda obtener por medio de subconjuntos de la población llamados Muestras.

Para la tarea de estimar se han desarrollado diversas teorías, métodos y técnicas, las mismas que parten de supuestos esenciales para su aplicación, supuestos que no necesariamente se cumplen cuando se desea emplear las técnicas en problemas reales, creándose así la duda de saber si el estimador que se obtiene de la aplicación de la técnica sirve, a pesar de no cumplirse algún supuesto, y el problema de desarrollar técnicas resistentes a

desviaciones de los supuestos iniciales; por ejemplo, un resultado teórico indica que si se tiene $\{X_i\}$, $i = 1, 2, \dots, n$; siendo $\{X_i\}$ una sucesión de variables aleatorias con Distribución Normal con media μ y Varianza σ^2 entonces $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$, pero se desconoce el comportamiento de $\frac{(n-1)S^2}{\sigma^2}$ si no todas las variables aleatorias X_i siguen una Distribución Normal.

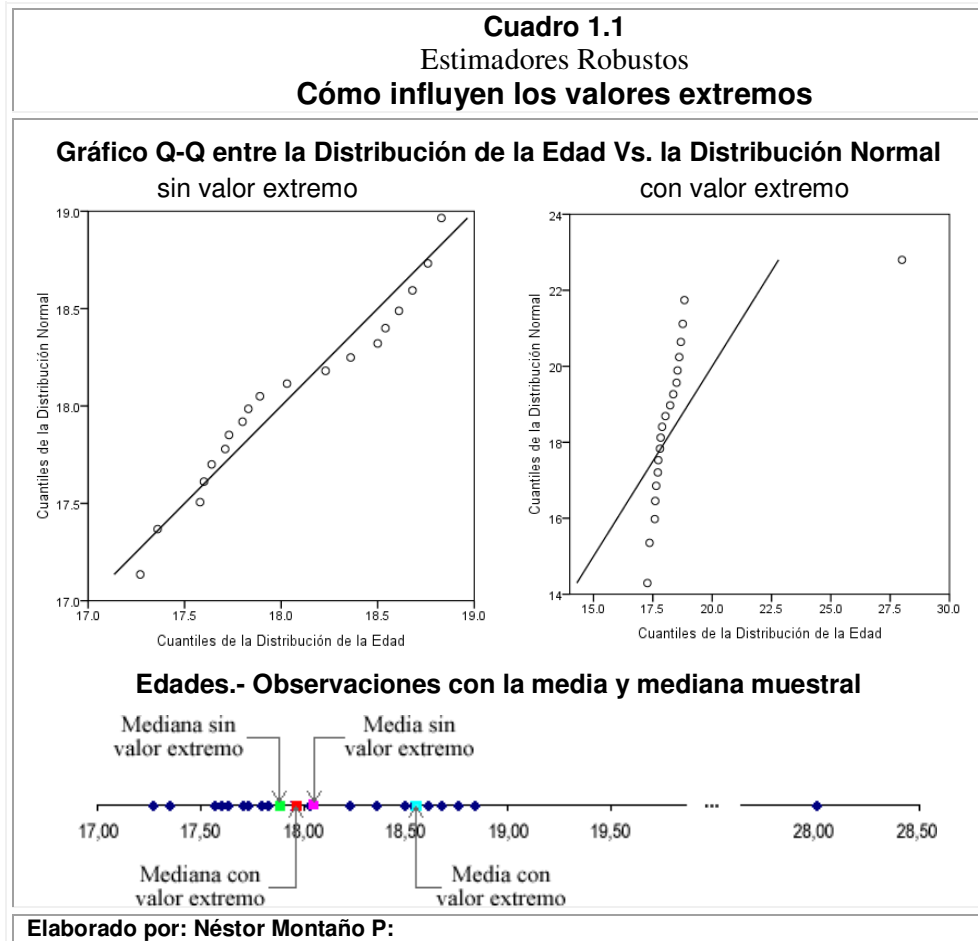
Se mencionó que, para estimar los parámetros poblacionales se utiliza información obtenida a partir de los datos que nos proporciona una Muestra; situémonos ahora en el caso en el que la Muestra presenta observaciones atípicas y/o erróneas, llamadas “valores aberrantes” o “extremos” porque se alejan acentuadamente del comportamiento general de los demás datos, ¿el estimador seguirá siendo una “buena” aproximación, o se verá afectado por este particular?

Se debe recalcar que la idea de la presencia de “valores aberrantes” no es absurda, pues se verifica que, en la práctica, un alto porcentaje de las investigaciones que se efectúan contienen errores ya sea al efectuar la medición, o al transcribir los datos, etc. Se ilustra esta situación para un caso univariado:

Considere las edades, en años, de un grupo de 20 estudiantes, mostradas a continuación en orden ascendente

17,27	17,36	17,58	17,60	17,64
17,71	17,73	17,80	17,83	17,89
18,03	18,23	18,36	18,50	18,54
18,61	18,68	18,76	18,83	28,00

Como se observa en el Cuadro 1.1 las 19 primeras observaciones se ajustan adecuadamente a una Distribución Normal, sin embargo basta un solo valor extremo para que dicho ajuste se deteriore.



Además, se ilustra cómo al estimar la media aritmética y desviación estándar utilizando sólo las primeras diecinueve observaciones se tiene $\bar{x}_{19} = 18.05$ y $s_{19} = 0.49$ mientras que considerando los veinte datos se tiene $\bar{x}_{20} = 18.55$ y $s_{20} = 2.28$, se puede observar entonces cómo el valor extremo, influye en \bar{x} y s ; por otro lado al obtener la mediana para ambos casos tenemos $\tilde{x}_{19} = 17.89$ y $\tilde{x}_{20} = 17.96$ respectivamente, aquí se nota en cambio la poca influencia del valor extremo en la mediana.

En general, sea \bar{x}_n y \bar{x}_{n+1} la media aritmética basada en n y $n+1$ observaciones respectivamente, se puede escribir que

$$\bar{x}_{n+1} = \frac{n}{n+1} \bar{x}_n + \frac{1}{n+1} x_{n+1} = \bar{x}_n + \frac{1}{n+1} (x_{n+1} - \bar{x}_n),$$

por lo tanto \bar{x}_{n+1} puede tomar valores tan altos (o tan bajos) como se quiera, basta con tomar x_{n+1} suficientemente alejado de \bar{x}_n .

Haciendo un análisis parecido con la mediana se puede ver que para que la mediana tome un valor tan alto como se quiera no es suficiente un valor extremo, sino se requiere por lo menos $\frac{n}{2}$ valores extremos; se concluye entonces que la mediana es mucho más resistente a valores extremos que la media aritmética y es un ejemplo de un estimador robusto.

Lo desarrollado hasta aquí es vagamente una muestra de la importancia y necesidad de estimadores que se comporten “bien” aún cuando existan variaciones en los supuestos iniciales o cuando es posible que el modelo esté “contaminado” por valores aberrantes que producen influencias en los resultados que conducen a estimaciones erróneas; a estos estimadores se los llama Estimadores Robustos.

El objetivo del presente trabajo es evaluar el comportamiento de algunos Estimadores Robustos para el vector de medias y la matriz de varianzas y covarianzas de un vector p variado \mathbf{X} , particularmente se estudiarán los “Estimadores M” propuestos por Huber en [14] y las variaciones propuestas a los mismos. La metodología que se empleará será: un entorno controlado de simulación, generando poblaciones conocidas que serán contaminadas en diversos niveles variando así el peso de las colas, todo ello al objeto de asimilar la presencia de valores atípicos para finalmente comparar los resultados y finalmente concluir sobre la calidad de los estimadores estudiados.

1.3 Conceptos Básicos.- Probabilidad

Chow y Teicher en [3] indican que una Clase es un conjunto cuyos elementos son en sí mismos conjuntos y un Espacio es un conjunto no vacío cualquiera que es usualmente utilizado como una referencia o punto de partida. Con estos antecedentes se presenta a continuación definiciones y propiedades referidas a este tema

Definición 1.1 Álgebra y σ -álgebra

Sea Ω un espacio, una clase no vacía \mathcal{A} de subconjuntos de un espacio Ω es un álgebra si

- i. $A^c \in \mathcal{A}$ cuando $A \in \mathcal{A}$,
- ii. $A_1 \cup A_2 \in \mathcal{A}$ cuando $A_j \in \mathcal{A}$, $j=1,2$.

Además, \mathcal{A} es llamado σ -álgebra si

- iii. $\left(\bigcup_{n=1}^{\infty} A_n\right) \in \mathcal{A}$ cuando $A_n \in \mathcal{A}$, $n \geq 1$

Como \mathcal{A} es no vacío de (i) y (ii) se concluye que $\emptyset \in \mathcal{A}$ y $\Omega \in \mathcal{A}$

Además (ii) implica que $\left(\bigcup_{i=1}^n A_i\right) \in \mathcal{A}$ cuando $A_i \in \mathcal{A}$.

Como $\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c$ utilizando (iii) y (i) se concluye que

$\left(\bigcap_{n=1}^{\infty} A_n\right) \in \mathcal{A}$; se puede probar también que $\left(\bigcap_{i=1}^n A_i\right) \in \mathcal{A}$. Por

último se puede demostrar que la σ -álgebra más “pequeña” posible es $\mathcal{A}_0 = \{\emptyset, \Omega\}$ y la más “grande” es $\mathcal{A}_{max} = \mathcal{P}(\Omega)$, donde $\mathcal{P}(\Omega)$ denota el conjunto potencia de Ω , se puede deducir entonces que para cualquier σ -álgebra \mathcal{A} , $\mathcal{A}_0 \subseteq \mathcal{A} \subseteq \mathcal{A}_{max}$

Por otra parte se puede mencionar que en el contexto de la Teoría de la Medida, si \mathcal{A} es una σ -álgebra relativa al espacio Ω , entonces el par (\mathcal{A}, Ω) es llamado Espacio Medible, y \mathcal{A} es llamado Conjunto Medible.

Por último, se puede demostrar que cualquier intersección de σ -álgebras es también una σ -álgebra. Esto nos permite definir.

Definición 1.2 σ -álgebra generada por \mathcal{E}

La Mínima σ -álgebra \mathcal{E}' que contiene a la clase no vacía \mathcal{E} de subconjuntos de Ω , es la σ -álgebra que cumple

- i. $\mathcal{E} \subset \mathcal{E}'$,
- ii. $\mathcal{E}' \subset \mathcal{E}''$, donde $\mathcal{E} \subset \mathcal{E}''$ donde \mathcal{E}'' es un σ -álgebra.

La mínima σ -álgebra que contiene a \mathcal{E} es también llamada la σ -álgebra generada por \mathcal{E} y es denotada por $\sigma(\mathcal{E})$.

Se puede demostrar que

$$\sigma(\mathcal{E}) \equiv \bigcap \left\{ \mathcal{E}_\alpha : \mathcal{E}_\alpha \text{ es una } \sigma\text{-álgebra de } \Omega \text{ para la cual } \mathcal{E} \subset \mathcal{E}_\alpha \right\}$$

Definición 1.3 σ -Álgebra de Borel (Boreliano)

Según Shorack en [29], si $\Omega = \mathbb{R}$ y \mathcal{E} es la unión de los subconjuntos de la forma $(-\infty, a]$ donde $a \in \mathbb{R}$, entonces $\mathcal{B} = \sigma(\mathcal{E})$ es llamada la σ -álgebra de Borel o Boreliano.

Nótese que se cumple lo siguiente: que si $a < b$, entonces $(-\infty, a] \in \mathcal{B}$; $(-\infty, b] \in \mathcal{B}$, luego $(-\infty, a]^c = (a, \infty) \in \mathcal{B}$ y $(a, \infty) \cap (-\infty, b] = (a, b] \in \mathcal{B}$, es decir, cualquier intervalo real está en \mathcal{B}

Además, si $\Omega = \mathbb{R}^n$ y $\mathcal{E} = \{(-\infty, a_1] \times (-\infty, a_2] \times \dots \times (-\infty, a_n] \mid a_1, a_2, \dots, a_n \in \mathbb{R}\}$ entonces $\mathcal{B}^n = \sigma(\mathcal{E})$ es la σ -álgebra de Borel en \mathbb{R}^n

Definición 1.4 Espacio de Probabilidad

Yohai en [30] define Espacio de Probabilidad como una terna (Ω, \mathcal{S}, P) donde Ω es un conjunto, \mathcal{S} es una σ -álgebra sobre Ω y $P: \mathcal{S} \rightarrow [0, 1]$ es una función que satisface:

1. $P(\Omega) = 1$

2. (σ -aditiva) Si $\{S_n\}_{n \geq 1}$ es una sucesión de elementos en \mathcal{S}

disjuntos dos a dos ($S_i \cap S_j = \emptyset$ si $i \neq j$), entonces $P\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} P(S_i)$

El conjunto Ω se denomina Espacio Muestral y se interpreta como el conjunto de resultados posibles de un experimento, los elementos de \mathcal{S} se denominan eventos, y corresponden a los subconjuntos de Ω para los cuales la probabilidad está definida. Finalmente P se denomina Función de Probabilidad, y dado $S \in \mathcal{S}$, $P(S)$ se interpreta como la probabilidad de que el resultado S esté en \mathcal{S}

En el contexto de la Teoría de la Medida, la terna (Ω, \mathcal{S}, P) corresponde a un Espacio de Medida donde la medida P asigna el valor uno al espacio Ω , esto es $P(\Omega) = 1$.

Como señala Zurita en [33], los axiomas planteados no indican como escoger a la función o medida P , son simplemente requisitos que debe cumplir P para poder ser llamada Función de Probabilidad; esto deriva en que a un Espacio Medible (Ω, \mathcal{S}) se le puede definir diferentes medidas P , como se ilustra a continuación

Considere un experimento que consiste en lanzar una moneda y observar si sale cara o sello, entonces $\Omega = \{\text{cara, sello}\} = \{c, s\}$.

Tomaremos \mathcal{S} como $\mathcal{P}(\Omega)$ entonces $\mathcal{S}=\{\emptyset, \{c\}, \{s\}, \Omega\}$; se plantea en seguida dos medidas P_1 y P_2 sobre el Espacio Medible (Ω, \mathcal{S})

Suponga que la moneda es “legal”, esto es que existe la misma probabilidad de que salga cara o sello, entonces se puede definir la Función de Probabilidad P_1 talque $P_1(\emptyset)=0$, $P_1(\{c\})=0.5$, $P_1(\{s\})=0.5$ y $P_1(\Omega)=1$.

Ahora supóngase que la moneda esta “arreglada” de tal manera que el peso del lado de la cara sea el doble que el peso del lado del sello, planteado así, es razonable pensar que la probabilidad que ocurra cara ahora será la mitad de la probabilidad que salga sello, entonces se puede definir otra Función de Probabilidad P_2 talque $P_2(\emptyset)=0$, $P_2(\{c\})=\frac{1}{3}$, $P_2(\{s\})=\frac{2}{3}$ y $P_2(\Omega)=1$. \square

Definición 1.5 Variable Aleatoria

Según Giraldo en [10] dado un espacio de probabilidad (Ω, \mathcal{S}, P) y el espacio medible (R, \mathcal{B}) donde \mathcal{B} es σ -álgebra de Borel en R , se dice que una aplicación $X: \Omega \rightarrow R$ es una Variable Aleatoria (v.a.) si es medible, es decir, si

$$X^{-1}(B) \in \mathcal{S}, \quad \forall B \in \mathcal{B} \quad \text{donde } X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$$

En particular, si \mathbb{R} es el conjunto de los Reales y $B=(-\infty, a]$ entonces X es una variable aleatoria real y $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \leq a\}$ y se denota por $(X \leq a)$. Además, si $\mathbb{R} = \mathbb{R}^p$ entonces \mathbf{X} representa a un Vector Aleatorio p-variado.

El concepto de variable aleatoria es esencialmente el mismo que el de función medible en la teoría de la medida, pues, Si (Ω, \mathcal{S}, P) es un espacio de medida, entonces $f: \mathcal{S} \rightarrow \mathbb{R}$ se dice “medible” si y solo si para todo x se cumple que $f^{-1}((-\infty, x]) \in \mathcal{S}$

Definición 1.6 Independencia de Variables Aleatorias

Shorack en [29], expresa que dado un Espacio de Probabilidad (Ω, \mathcal{S}, P)

a) Sea $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n$ varios sub σ -álgebra de \mathcal{S} . Se dice que $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n$ son σ -álgebra independientes si satisfacen

$$P(S_1 \cap \dots \cap S_n) = \prod_{i=1}^n P(S_i) \text{ cuando } A_i \in \mathcal{S}_i$$

b) Variables aleatorias X_1, X_2, \dots, X_n son llamadas v.a. independientes si las σ -álgebra $\mathcal{F}(X_i) \equiv X_i^{-1}(\mathcal{B})$, $i=1,2,\dots,n$ son independientes.

La idea de la independencia es que el hecho que ocurra o no un evento no influye en que ocurra o no otro, o que el comportamiento de la variable aleatoria X_1 no influye en el comportamiento de la variable aleatoria X_2 .

Nótese que al hablar de independencia, se trabaja con más de una variable, y en la práctica, en los estudios se analizan varias variables, así que, efectuándose análisis multivariados además de los univariados; por lo que a continuación se presenta la definición de Vector Aleatorio

Definición 1.7 Vector Aleatorio y Función de Distribución Conjunta

Según Yohai en [30] Sea (Ω, \mathcal{S}, P) un espacio de probabilidad. Se dice que $\mathbf{X}^T = (X_1, X_2, \dots, X_p)$ es un vector aleatorio p-variado si para cada $j = 1, 2, \dots, p$ se tiene que $X_j: \Omega \rightarrow \mathbb{R}$ es una variable aleatoria, además se define la Función de Distribución Conjunta del vector \mathbf{X} como la función $F_{\mathbf{X}}: \mathbb{R}^p \rightarrow [0, 1]$ dada por

$$F_{\mathbf{X}}(x_1, x_2, \dots, x_p) = P_{\mathbf{X}}((-\infty, x_1] \times (-\infty, x_2] \times \dots \times (-\infty, x_p])$$

$$F_{\mathbf{X}}(x_1, x_2, \dots, x_p) = P\left(\bigcap_{i=1}^p \{\omega / X_i(\omega) \leq x_i\}\right)$$

Utilizando la Función de Distribución Conjunta se puede inferir sobre la independencia de Variables Aleatorias, utilizando el siguiente criterio.

Criterio para Independencia

Las Variables Aleatorias X_1, X_2, \dots, X_n son independientes sí y sólo sí

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n) \quad \text{para todo } x_1, \dots, x_n$$

Definición 1.8 Distribución Normal Multivariada

Según Flury en [7] el vector aleatorio p-variado $\mathbf{X}^T = (X_1, X_2, \dots, X_p)$ sigue una Distribución Normal Multivariada si se cumple que para todo $\mathbf{c} \in \mathbb{R}^p$, la combinación lineal de la forma $U = \mathbf{c}^T \mathbf{X}$ es una Variable Aleatoria Normal univariada o una constante.

Si $\boldsymbol{\mu} = E(\mathbf{X})$ es el vector de medias de \mathbf{X} y $\boldsymbol{\Sigma} = E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T]$ es su matriz de varianzas y covarianzas (que debe ser semidefinida positiva); entonces denotaremos lo anterior como

$$\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \equiv (y = \mathbf{c}^T \mathbf{X} \sim N(\mu = \mathbf{c}^T \boldsymbol{\mu}, \sigma^2 = \mathbf{c}^T \boldsymbol{\Sigma} \mathbf{c}), \forall \mathbf{c} \in \mathbb{R}^p)$$

La función de densidad para este tipo de variables esta dada por:

$$f(\mathbf{X}) = f(x_1, x_2, \dots, x_p) = (2\pi)^{-p/2} (\det \boldsymbol{\Sigma})^{-1/2} \exp\left(-\frac{1}{2} (\mathbf{X} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})\right) \text{ para}$$

$$\mathbf{X} \in \mathbb{R}^p$$

1.4 Conceptos Básicos.- Estimadores Puntuales

Según Zamar en [31] una práctica muy común en Estadística, en Estadística Paramétrica para ser preciso, es suponer que los datos han sido generados por un mecanismo aleatorio (modelo matemático) totalmente conocido, a excepción de un conjunto finito θ de valores a los cuales se les llama Parámetros Poblacionales. Se crea así el problema de estimar (aproximar) dichos Parámetros Poblacionales θ .

1.4.1 Estimador

Dada la Variable Aleatoria X sobre el Espacio de Probabilidad (Ω, \mathcal{S}, P) ; sea F_X la Función de Distribución asociada a X , la misma que es conocida excepto por un número finito de parámetros (todos ellos desconocidos) que forman un vector que se llamará Vector de Parámetros y que se denota por θ ; así, se dirá que F_θ es la Función de Distribución de X , y que P_θ es la distribución de probabilidad de X ; además, se denota al conjunto de los valores posibles para θ como Θ que se denomina Espacio de Parámetros.

García en [8] proporciona la siguiente definición para estadístico y estimador

Definición 1.9 Estadístico y Estimador

Sea X una Variable Aleatoria con distribución de probabilidad P_θ donde $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ el vector de parámetros desconocidos, sea $\{X_i\}$, donde $i = 1, 2, \dots, n$ una sucesión de Variables Aleatorias Independientes Idénticamente distribuidas. Una aplicación $T(X_1, X_2, \dots, X_n)$ en Θ (medible) se define como un Estadístico. Se dice además que el Estadístico $T(X_1, X_2, \dots, X_n)$ es un Estimador del vector de parámetros θ .

A continuación se ilustrará lo definido hasta el momento

Sea $\{X_i\}$, donde $i = 1, 2, \dots, 19$ una muestra aleatoria que sigue una distribución normal con media μ y varianza σ^2 , ambos desconocidos, entonces:

$$P_\theta = N(\mu, \sigma^2)$$

$$\theta = (\mu, \sigma^2)$$

$$\Theta = \{(x, y) / x \in \mathbb{R}, y > 0\}$$

Definimos $T: \mathbb{R}^{19} \rightarrow \Theta$ como

$$T(X_1, X_2, \dots, X_{19}) = \{(\bar{x}, s^2)\}; \text{ donde}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{19} x_i$$

$$s_n^2 = \frac{1}{n} \sum_{i=1}^{19} (x_i - \bar{x})^2$$

□

En el ejemplo planteado anteriormente, se ha escogido a \bar{x} y s_n^2 como estimadores para μ y σ^2 respectivamente, pero se podría escoger otras funciones como estimadores, por lo que se advierte el problema de decidir qué función usar; en otras palabras:

- La dificultad de encontrar el estadístico T que proporcione la "mejor aproximación" del vector de parámetros θ , y
- Los criterios en base a los cuales se mide la bondad de la aproximación;

Estos cuestionamientos son tratados en lo que resta del presente estudio.

1.4.2 Métodos de estimación

Los principales métodos de estimación de parámetros de distribuciones de probabilidad son los siguientes:

- Método de los Momentos
- Método de Máxima Verosimilitud

En los estudios de regresión se utiliza principalmente el

- Método de Mínimos Cuadrados

Mendenhall, Wackerly y Scheaffer en [21] explican el Método de los Momentos y de Máxima Verosimilitud que serán los que se desarrollarán en el presente capítulo.

1.4.2.1 Método de los Momentos

El Método de los Momentos es uno de los métodos más sencillos de estimación, fue formalizado por Pearson a finales del siglo XIX.

Definición 1.10 Momentos de orden k y Función Generadora de Momentos

Sea X una variable aleatoria, definimos el momento de orden $k > 0$ de X como:

$$\mu'_k = E[X^k] \text{ cuando este valor existe}$$

y el momento absoluto de orden $k > 0$ de X como

$$\mu'_k = E[|X^k|]$$

Además, se define la función generadora de momentos como

$$\mathcal{M}_X(t) = E[e^{xt}] \text{ siempre que tal valor exista} \quad (1.1)$$

La función generadora de momento cumple con

$$\mu'_k = \left. \frac{\partial^k}{\partial t^k} \mathcal{M}_X(t) \right|_{t=0}$$

Además de permitir calcular momentos de una variable aleatoria, la Función Generadora de Momentos permite identificar la función de densidad o de probabilidad de una variable aleatoria, pues ésta cumple con la propiedad de unicidad, la misma que establece que

hay una correspondencia uno a uno entre Funciones de Densidad o Probabilidad y Funciones Generadoras de Momentos.

Por tanto si dos variables aleatorias tienen los mismos momentos, entonces dichas variables tienen o siguen la misma Función de Densidad, razonamiento que se utiliza para estimar sus respectivos parámetros.

Además, se debe indicar que el k-ésimo momento m_k de una Muestra Aleatoria X_1, X_2, \dots, X_n es la media aritmética de sus k-ésimas potencias, esto es:

$$m_k = \frac{1}{n} \sum_{i=1}^n x_i^k$$

Para ilustrar la diferencia entre el Momento Muestral y el Momento Poblacional, tomemos una variable aleatoria $X \sim N(\mu, \sigma^2)$, entonces el primer momento poblacional es $\mu_1' = E[X] = \mu$. Por otro lado si se tiene la siguiente muestra de tamaño $n = 5$; $\{2, 3, 5, 7, 8\}$

para lo cual $m_1 = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{5}(2+3+5+7+8) = 5$. □

Se procede ahora a desarrollar el Método de los Momentos

El Método de los Momentos consiste en igualar los primeros momentos de una población a los momentos correspondientes de una muestra, entonces si una distribución tiene k parámetros desconocidos, para su estimación se tendrá lo siguiente:

$$\mu'_1 = m_1$$

$$\mu'_2 = m_2$$

$$\vdots$$

$$\mu'_k = m_k$$

Por ejemplo, sea X_1, X_2, \dots, X_n una muestra aleatoria de una distribución $N(\mu, \sigma^2)$ y se desea estimar μ, σ^2 . Entonces, utilizando el Método de los Momentos, se debe igualar los dos primeros momentos poblacionales a los dos primeros momentos muestrales.

El primer momento poblacional corresponde a $E[X] = \mu$, mientras que el primer momento muestral es $\frac{1}{n} \sum_{i=1}^n X_i = \bar{x}$; por tanto el estimador de $\hat{\mu}$ es igual a \bar{x} .

Por otro lado, el segundo momento poblacional corresponde a $E[X^2] = \sigma^2 + E[X]^2$, de donde se obtiene que $\sigma^2 = E[X^2] - E[X]^2$;

el segundo momento muestral es igual a $\frac{1}{n} \sum_{i=1}^n X_i^2$, entonces el

estimador $\hat{\sigma}^2$ que notaremos por s^2 será igual a

$$\frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 \quad \square$$

Antes de presentar el denominado Método de Máxima Verosimilitud, se estudia la Función Característica, función que tiene cierta similitud con la Función Generadora de Momentos y que se la utilizará principalmente en la demostración del Teorema del Límite Central que se realizará más adelante. Chow en [3], Shorack en [29], Perelló en [23] definen y desarrollan varias propiedades de esta función.

Definición 1.11 Función Característica

Sea X una variable aleatoria en el Espacio de Probabilidad (Ω, \mathcal{S}, P) y sea F_X su Función de Distribución. Se define la Función Característica de X como la función $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$ tal que

$$\begin{aligned} \varphi_X(t) &= E[e^{itX}] & (1.2) \\ &= E[\cos(tX)] + iE[\text{sen}(tX)] \\ &= \int \cos(tx)dF(x) + i \int \text{sen}(tx)dF(x) \end{aligned}$$

Se puede observar, vea (1.1), que $\varphi_X(t) = M_X(it)$, sin embargo la Función Generadora de Momentos puede no existir, mientras que la

Función Característica sí, es más, es una función que existe aunque $E[X]$ no exista.

Teorema 1.1 Existencia de la Función Característica

Toda variable aleatoria X en el Espacio de Probabilidad (Ω, \mathcal{S}, P) posee Función Característica $\varphi_X(t)$.

Demostración: Del desarrollo de (1.2), $\varphi_X(t)$ existirá si $E[\cos(tX)]$ y $E[\sin(tX)]$ son finitos, y se puede hacer

$$|E[\cos(tX)]| \leq E[|\cos(tX)|] \leq E[1] = 1$$

De donde se concluye que $E[\cos(tX)]$ es finito, lo mismo ocurre con $E[\sin(tX)]$. Por tanto $\varphi_X(t)$ siempre existe \square

La Función Característica φ_X de una variable aleatoria X (discreta o absolutamente continua) satisface las siguientes propiedades:

1. $\varphi_X(0) = 1$ y $|\varphi_X(t)| \leq 1$, para todo $t \in \mathbb{R}$
2. $\varphi_X(-t) = \overline{\varphi_X(t)}$, en donde la barra denota el complejo conjugado de $\varphi_X(t)$.
3. Si $Y = aX + b$, entonces $\varphi_Y(t) = e^{itb} \varphi_X(at)$, para todo $a, b \in \mathbb{R}$.
4. Si X, Y son variables aleatorias independientes, entonces

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) \quad (1.3)$$

Además, se puede demostrar que, al derivar sucesivamente el argumento del valor esperado se obtiene como resultado.

$$\varphi_X^{(1)}(t) = E(iXe^{itX}) = iE(Xe^{itX})$$

$$\varphi_X^{(2)}(t) = E(i^2 X^2 e^{itX}) = i^2 E(X^2 e^{itX})$$

...

$$\varphi_X^{(n)}(t) = E(i^n X^n e^{itX}) = i^n E(X^n e^{itX})$$

De donde se obtiene que $E[X^n] = i^{-n} \varphi_X^{(n)}(0)$, es decir, también es posible calcular los momentos utilizando la Función Característica.

Teorema 1.2 Unicidad de la Función Característica

Existe una correspondencia biunívoca entre Función de Distribución y la Función Característica.

Demostración.- Se utiliza la Fórmula de Inversión de Lévy la misma que se enuncia de la siguiente manera

Fórmula de Inversión de Lévy. Si X es una variable aleatoria con función característica φ , entonces para $-\infty < a < b < \infty$

$$\lim_{c \rightarrow \infty} \frac{1}{2\pi} \int_{-c}^c \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt = P\{a < X < b\} + \frac{P\{X = a\} + P\{X = b\}}{2}$$

Si F es continua, la parte derecha de la igualdad se reduce a $F(b) - F(a)$.

Sea

$$I(C) = \frac{1}{2\pi} \int_{-C}^C \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt = \frac{1}{2\pi} \int_{-C}^C \frac{e^{-ita} - e^{-itb}}{it} \mathbb{E} \left[e^{itX} \right] dt$$

Donde $\left[(e^{-ita} - e^{-itb}) / it \right] e^{itX}$ está definido para todo t

$$\begin{aligned} I(C) &= \frac{1}{2\pi} \mathbb{E} \left[\int_{-C}^C \frac{e^{it(X-a)} - e^{it(X-b)}}{it} dt \right] \\ &= \frac{1}{\pi} \mathbb{E} \left[\int_0^C \frac{\text{sen} t(X-a) - \text{sen} t(X-b)}{t} dt \right] \\ &= \frac{1}{\pi} \mathbb{E} \left[\int_0^{C(X-a)} \frac{\text{sen} t}{t} dt - \int_0^{C(X-b)} \frac{\text{sen} t}{t} dt \right] = \mathbb{E} J_C(X) \end{aligned}$$

donde

$$J_C(u) = \frac{1}{\pi} \int_{C(u-b)}^{C(u-a)} \frac{\text{sen} t}{t} dt \xrightarrow{C \rightarrow \infty} \begin{cases} 1, & a < u < b \\ \frac{1}{2}, & u = \{a, b\} \\ 0, & u < a, u > b \end{cases}$$

Entonces $|J_C(u)| \leq 2$ para todo $-\infty < u, C < \infty$ por convergencia.

$$\lim_{C \rightarrow \infty} I(C) = \lim_{C \rightarrow \infty} \mathbb{E} [J_C(X)] = \mathbb{E} \lim_{C \rightarrow \infty} J_C(X)$$

$$\mathbb{E} \left[\frac{1}{2} I_{[X=a,b]} + I_{[a < X < b]} \right] = \frac{P\{X=a\} + P\{X=b\}}{2} + P\{a < X < b\}$$

..

Con esto queda demostrado la Fórmula de Inversión de Lévy, recordando que $\varphi_X(t) = \int \cos(tx)dF(x) + i \int \sin(tx)dF(x)$ entonces se dirá que idénticas Funciones de Distribución tienen idénticas Funciones Características mientras la conversión cumpla

$$F(b) = \lim_{a \rightarrow -\infty} \lim_{C \rightarrow \infty} \frac{1}{2\pi} \int_{-C}^C \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt, \quad b \in C(F) \quad \square$$

A continuación se obtendrá la función característica para una variable aleatoria X con distribución Normal de parámetros $\mu = 0$ y $\sigma = 1$.

Si $X \sim N(0, 1)$ entonces

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Se obtiene $\varphi_X(t) = E[e^{itX}]$

$$\begin{aligned} \varphi(t) &= \int_{-\infty}^{+\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \int_{-\infty}^{+\infty} e^{-(x-it)^2/2} dx \\ &= e^{-t^2/2} \quad \square \end{aligned}$$

1.4.2.2 Método de Máxima Verosimilitud

García en [8] expone cómo Fisher en su obra “Theory of Statistical Estimation” (1925), establece las bases para la estimación puntual, por la que debe considerársele descubridor del Método de Máxima Verosimilitud para la estimación de parámetros, método cuya aplicación ofrece en general estimadores con mejores propiedades que el método de los momentos.

Definición 1.12 Función de Verosimilitud

Sea $\mathbf{X}^T=(X_1, X_2, \dots, X_n)$, donde X_1, \dots, X_n es una sucesión de variables aleatorias independientes idénticamente distribuidas con distribución de probabilidad P_θ donde $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Sea $f_\theta(X_1, X_2, \dots, X_n)$ la Función de Distribución asociada a \mathbf{X} , donde $\theta \in \Theta$

La función,

$$L(X_1, X_2, \dots, X_n, \theta) = f_\theta(X_1, X_2, \dots, X_n)$$

Se denomina Función de Verosimilitud de \mathbf{X} , θ

Recordando una de las propiedades de la Función Generadora de Momentos, se dice que dos variables aleatorias X y Y son Idénticamente Distribuidas si se cumple que $M_X(t) = M_Y(t)$, o utilizando la Función Característica la condición es $\varphi_X = \varphi_Y$.

Si X_1, X_2, \dots, X_n son Independientes e Idénticamente Distribuidas con función de distribución $f_\theta(X)$ la función de verosimilitud queda:

$$L(X_1, X_2, \dots, X_n, \theta) = \prod_{i=1}^n f_\theta(x_i)$$

Los Estimadores de Máxima Verosimilitud (EMV) de $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ son los valores $T = (T_1, T_2, \dots, T_k)$ que maximizan la Función de Verosimilitud, o sea los valores tales que

$$L(X_1, X_2, \dots, X_n, T) = \sup_{\theta \in \Theta} L(X_1, X_2, \dots, X_n, \theta) \quad (1.4)$$

Como generalmente la función de verosimilitud es difícil de evaluar, y dado que existe una relación biunívoca entre ésta y su logaritmo, entonces se prefiere trabajar con el logaritmo de la función de verosimilitud en lugar de con la propia Función de Verosimilitud y nuestro objetivo será ahora encontrar T tal que

$$\log L(X_1, X_2, \dots, X_n, T) = \sup_{\theta \in \Theta} \log L(X_1, X_2, \dots, X_n, \theta) \quad (1.5)$$

Considérese la situación en que Θ es un subconjunto abierto de \mathbb{R}^k y $f_\theta(X)$ es una función diferenciable en θ (esto es, que existen las derivadas parciales de primer orden en las componentes de θ). Si el supremo en (1.4) existe, se lo puede hallar derivando parcialmente la función de verosimilitud con respecto a cada componente θ_i del

vector de parámetros θ para luego igualar cada ecuación obtenida a cero, es decir:

$$\frac{\partial}{\partial T_i} L(X_1, X_2, \dots, X_n, \theta) = 0 \quad \forall i=1,2,\dots,p$$

Así por ejemplo, sean X_1, \dots, X_n v.a. iid $\sim N(\mu, \sigma^2)$ donde tanto μ como σ^2 son desconocidos. Se pretende determinar los Estimadores de Máxima Verosimilitud de μ y σ^2

$$\begin{aligned} L(\mu, \sigma^2) &= f(x_1, x_2, \dots, x_n | \mu, \sigma^2) \\ &= f(x_1 | \mu, \sigma^2) f(x_2 | \mu, \sigma^2) f(x_3 | \mu, \sigma^2) \cdots f(x_n | \mu) \\ &= \left\{ \frac{1}{(\sigma\sqrt{2\pi})} \exp\left(-\frac{(x_1 - \mu)^2}{2\sigma^2}\right) \right\} \cdots \left\{ \frac{1}{(\sigma\sqrt{2\pi})} \exp\left(-\frac{(x_n - \mu)^2}{2\sigma^2}\right) \right\} \\ &= \frac{1}{(\sigma^2 2\pi)^{\frac{n}{2}}} \exp\left(-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}\right) \end{aligned}$$

Aplicando logaritmo natural, se tiene,

$$\ln[L(\mu, \sigma^2)] = -\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

Los Estimadores de Máxima Verosimilitud de μ y σ^2 son los valores que maximizan $\ln[L(\mu, \sigma^2)]$. Al obtener las derivadas respecto a μ y σ^2 , resulta

$$\frac{\partial \{ \ln [L(\mu, \sigma^2)] \}}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)$$

y

$$\frac{\partial \{\ln[L(\mu, \sigma^2)]\}}{\partial \sigma^2} = -\left(\frac{n}{2}\right)\left(\frac{1}{\sigma^2}\right) + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2$$

Se iguala estas derivadas a cero para entonces resolver; así de la primera ecuación obtenemos:

$$\frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{\mu}) = 0 \text{ que resolviendo queda } \sum_{i=1}^n x_i - n\hat{\mu} = 0 \text{ de donde}$$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

Se sustituye \bar{x} por $\hat{\mu}$ en la segunda ecuación para luego obtener de manera explícita $\hat{\sigma}^2$, así

$$-\left(\frac{n}{\hat{\sigma}^2}\right) + \frac{1}{\hat{\sigma}^4} \sum_{i=1}^n (x_i - \bar{x})^2 = 0 \text{ entonces } s^2 = \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Por consiguiente, \bar{x} y s^2 son respectivamente los Estimadores de Máxima Verosimilitud de μ y σ^2 . \square

Obsérvese que dada una muestra X_1, X_2, \dots, X_n , un Estimador Puntual $\hat{\theta}$ del Vector de Parámetros θ (obtenido a partir de la muestra dada) es un Vector Aleatorio, y como tal verifica ciertas propiedades que revisaremos luego de considerar algunos tipos de convergencia que son necesarios para el estudio de ciertas propiedades de los estimadores.

1.4.3 Convergencias de variables aleatorias

Definición 1.13 Convergencia puntual de funciones

Sea $\{f_n\}_{n \in \mathbb{N}} : \Omega \rightarrow \mathbb{R}$ una sucesión de funciones. Se dice que f_n *converge puntualmente* a otra función f si para todo $\omega \in \Omega$ y para todo $\varepsilon > 0$; existe $n_0 \in \mathbb{N}$ dependiente de ε y de ω tal que si $n \geq n_0$ entonces $|f_n(\omega) - f(\omega)| < \varepsilon$.

Cuando la elección de n_0 puede hacerse con independencia de ω , se tiene la siguiente noción de convergencia.

Definición 1.14 Convergencia uniforme de funciones

Sea $\{f_n\}_{n \in \mathbb{N}}$ una sucesión de funciones definidas sobre un conjunto Ω y que toma valores reales. Se dice que la sucesión $\{f_n\}$ *converge uniformemente* en Ω a otra función f si para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que: Si $n \geq n_0$ entonces $|f_n(\omega) - f(\omega)| < \varepsilon$ para todo $\omega \in \Omega$.

Se puede probar que si $\{f_n\}_{n \in \mathbb{N}}$ converge uniformemente en Ω , entonces $\{f_n\}_{n \in \mathbb{N}}$ *converge puntualmente*. La recíproca es falsa.

Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias definidas sobre un mismo Espacio de Probabilidades (Ω, \mathcal{S}, P) . Para cada $\omega \in \Omega$, la

sucesión $\{X_n(\omega)\}$ es una sucesión de números reales que puede ser convergente o divergente según sea ω . En otras palabras, la existencia o no del límite de la sucesión $\lim_{n \rightarrow \infty} x_n$ dependerá de los eventos $\omega \in \Omega$. Esta dependencia origina tres tipos de convergencia de Variables Aleatorias: la Convergencia Casi Segura, la Convergencia en Probabilidad y la Convergencia en Media Cuadrática.

Por otra parte, si $\{F_i\}$ es una sucesión de funciones de distribución asociadas a la sucesión de variables aleatorias $\{X_i\}$, se estudiará también la convergencia de $\{F_i\}$, convergencia denominada Convergencia en Distribución.

Definición 1.15 Convergencia en Distribución

Considérese un espacio de probabilidades (Ω, \mathcal{S}, P) . Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias definidas sobre este espacio, con funciones de distribución $\{F_n\}_{n \in \mathbb{N}}$. Sea X otra Variable Aleatoria definida también sobre (Ω, \mathcal{S}, P) con Función de Distribución F . Se dice que la sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ “*Converge en Distribución o en Ley*” a la variable aleatoria X si

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

en todos los puntos x donde F es continua, y se lo denota como

$$X_n \xrightarrow{D} X; \text{ ó, } X_n \xrightarrow{L} X.$$

A continuación se presenta el Teorema de Límite Central el mismo que es una aplicación de este resultado.

Teorema 1.3 Límite Central (Lévy-Lindeberg)

Sea $\{X_n\}$ una sucesión de v.a. i.i.d. con media μ y varianza σ^2 finita.

Entonces

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Z \sim N(0,1)$$

Demostración: Sea $Y_n = X_n - \mu$; (n=1, 2, 3,...). Evidentemente la sucesión Y_n está constituida por Variables Aleatorias Independientes con

$$E(Y_n) = 0 \quad \text{y} \quad E(Y_n^2) = \sigma^2 \quad (n=1, 2, 3, \dots)$$

Consideremos ahora las variables aleatorias

$$Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n Y_i = \frac{1}{\sigma\sqrt{n}} \left(\sum_{i=1}^n X_i - n\mu \right)$$

Primero se va a demostrar que la Función Característica de Z_n tiende a la función característica de la distribución normal estándar

$$\varphi(t) = e^{-t^2/2}$$

En efecto, si φ_n es la Función Característica de Z_n , entonces se tiene que

$$\begin{aligned}\varphi_n(t) &= E[e^{itZ_n}] = E[\exp(e^{itZ_n})] \\ &= E\left[\exp\left(\frac{it}{\sigma\sqrt{n}} \sum_{i=1}^n Y_i\right)\right]\end{aligned}$$

siendo Y_n variables independientes, se puede aplicar (1.3) obteniendo

$$\begin{aligned}\varphi_n(t) &= \prod_{i=1}^n E\left[\exp\left(\frac{itY_i}{\sigma\sqrt{n}}\right)\right] \\ &= \prod_{i=1}^n \varphi_{Y_i}\left(\frac{t}{\sigma\sqrt{n}}\right)\end{aligned}$$

Además, las variables Y_n son idénticamente distribuidas y, por tanto, tienen la misma función característica φ . Así, tenemos

$$\varphi_n(t) = \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \quad (1.6)$$

Desarrollando ahora φ como una Serie de Taylor en $t=0$, se obtiene

$$\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) = \varphi(0) + \frac{t}{\sigma\sqrt{n}} \varphi'(0) + \frac{t^2}{2\sigma^2 n} \varphi''(0) + o\left(\frac{1}{n^{3/2}}\right)$$

en donde el símbolo

$$o\left(\frac{1}{n^{3/2}}\right)$$

indica que el resto de los términos de la serie son infinitésimos de orden igual o superior a $1/n^{3/2}$ cuando n tiende a infinito.

Recordando las propiedades de la función características se obtiene

$$\varphi(0) = 1; \quad \varphi'(0) = iE(Y_n) = 0; \quad \varphi''(0) = i^2 E(Y_n^2) = -\sigma^2$$

Así se tiene

$$\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n^{3/2}}\right)$$

Por tanto de (1.6), se tiene

$$\begin{aligned} \lim_{n \rightarrow \infty} \varphi_n(t) &= \lim_{n \rightarrow \infty} \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) \right]^n \\ &= \lim_{n \rightarrow \infty} \left[1 - \frac{t^2}{2n} + o\left(\frac{1}{n^{3/2}}\right) \right]^n \\ &= \lim_{n \rightarrow \infty} \left[1 - \frac{t^2}{2n} \right]^n \\ &= e^{-t^2/2} \end{aligned}$$

Es decir, la función característica de

$$Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n Y_i = \frac{1}{\sigma\sqrt{n}} \left(\sum_{i=1}^n X_i - n\mu \right)$$

tiende a la función característica de la distribución normal estándar, y recordando la Unicidad de la Función Característica entonces se dice que la función de distribución de Z_n tiende a la función de

distribución normal estándar y en consecuencia Z_n converge en Distribución a $Z \sim N(0,1)$ \square

Definición 1.16 Convergencia en Probabilidad

Considérese un espacio de probabilidades (Ω, \mathcal{S}, P) . Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias definidas sobre este espacio, sea además X otra variable aleatoria definida también sobre (Ω, \mathcal{S}, P) . La sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ se dice que “converge en probabilidad” a la variable aleatoria X si para todo $\varepsilon > 0$ se cumple

$$\lim_{n \rightarrow \infty} P(\omega \in \Omega / |X_n(\omega) - X(\omega)| \geq \varepsilon) = 0$$

o de forma equivalente

$$\lim_{n \rightarrow \infty} P(\omega \in \Omega / |X_n(\omega) - X(\omega)| < \varepsilon) = 1$$

Se denota por $X_n \xrightarrow{P} X$

El concepto de convergencia en probabilidad afirma que dado $\varepsilon > 0$, la probabilidad de que X_n difiera de X en más de ε es menor que una cantidad tan pequeña como se quiera, a partir de un n suficientemente grande. En particular, $\{X_n\}$ “Converge en Probabilidad” a la constante a si

$$\lim_{n \rightarrow \infty} P(\omega \in \Omega / |X_n(\omega) - a| \geq \varepsilon) = 0$$

Se utiliza la Convergencia en Probabilidad para demostrar el teorema conocido como Ley Débil de los Grandes Números, que se presenta a continuación.

Teorema 1.4 Ley Débil de los Grandes Números de Tchebychev

.Sea $\{X_n\}$ una sucesión de variables aleatorias mutuamente independientes que tienen varianzas finitas y acotadas por una misma constante α

$$\sigma_n^2 = E[(X_n - E(X_n))^2] \leq \alpha$$

Para todo $n \in \mathbb{N}$, entonces

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

Demostración: Aplicando la desigualdad de Tchebyshev a

$\frac{S_n - E(S_n)}{n}$ se obtiene:

$$P\left(\left|\frac{S_n - E(S_n)}{n}\right| \geq \varepsilon\right) \leq \frac{E(|S_n - E(S_n)|^2)}{n^2 \varepsilon^2}$$

Para todo $\varepsilon > 0$. Ahora bien, como las variables aleatorias X_n ($n=1,2, 3,\dots$) son mutuamente independientes, se cumple

$$E((S_n - E(S_n))^2) = Var(S_n)$$

$$\begin{aligned}
&= \text{Var} \left(\sum_{i=1}^n X_i \right) \\
&= \sum_{i=1}^n \text{Var}(X_i) \\
&\leq n\alpha
\end{aligned}$$

Por tanto,

$$P \left(\left| \frac{S_n - E(S_n)}{n} \right| \geq \varepsilon \right) \leq \frac{E \left(|S_n - E(S_n)|^2 \right)}{n^2 \varepsilon^2} \leq \frac{\alpha}{n^2 \varepsilon^2}$$

y, en consecuencia,

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{S_n - E(S_n)}{n} \right| \geq \varepsilon \right) = 0 \quad ; \quad \text{para todo } \varepsilon > 0 \quad \square$$

Se menciona ahora tres tipos más de convergencias de variables aleatorias.

Definición 1.17 Convergencia segura de variables aleatorias

Considérese un espacio de probabilidades (Ω, \mathcal{S}, P) . Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias definidas sobre este espacio, sea además X otra variable aleatoria definida también sobre (Ω, \mathcal{S}, P) . Se dice que la sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ “*Converge Seguro*” a X si para todo $\omega \in \Omega$ la sucesión $\{X_n(\omega)\}$ converge a $X(\omega)$, es decir que para todo $\omega \in \Omega$ se cumple que

$$\lim_{n \rightarrow \infty} \{X_n(\omega)\} = X(\omega) \quad (1.7)$$

Este tipo de convergencia se da en muy pocos casos, en general lo que ocurre es que (1.7) se cumple sólo para un cierto subconjunto C de Ω , es decir, un suceso \mathcal{S} , en otras palabras, en el caso de *Convergencia Segura* $C=\Omega$.

El siguiente es un criterio de convergencia un poco menos fuerte que el anterior.

Definición 1.18 Convergencia casi segura de variables aleatorias

Considérese un espacio de probabilidades (Ω, \mathcal{S}, P) . Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias definidas sobre este espacio, sea además X otra variable aleatoria definida también sobre (Ω, \mathcal{S}, P) . Se dice que la sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ “*Converge Casi Seguro*” a la variable aleatoria X si y sólo si (1.7) se cumple sólo para un suceso C talque $P(C) = 1$, es decir, si

$$C = \left\{ \omega \in \Omega / \lim_{n \rightarrow \infty} \{X_n(\omega)\} = X(\omega) \right\}, \text{ entonces } P(C) = 1$$

También se puede decir que la sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ “*Converge Casi Seguro*” a la variable aleatoria X si y sólo si

$$P\left(\limsup\left\{\omega \in \Omega / |X_n(\omega) - X(\omega)| \geq \frac{1}{k}\right\}\right) = 0$$

La Convergencia Casi Segura se denota por $X_n \xrightarrow{c.s.} X$ y se la puede interpretar expresando que $\{X_n\}$ converge a X como una sucesión de números, excepto en un conjunto C^c con $P(C^c)=0$.

Definición 1.19 Convergencia en Media Cuadrática

Considérese un espacio de probabilidades (Ω, \mathcal{S}, P) . Sea $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias definidas sobre este espacio, sea además X otra variable aleatoria definida también sobre (Ω, \mathcal{S}, P) . Una sucesión de variables aleatorias $\{X_n\}_{n \in \mathbb{N}}$ se dice que “*Converge en Media Cuadrática*” a la variable aleatoria X si $E[X_n^2]$ y $E[X^2]$ son finitos y

$$\lim_{n \rightarrow \infty} E[(X_n - X)^2] = 0$$

Se lo denotará por $X_n \xrightarrow{m.c.} X$

En general se dice que X_n converge en media de orden r ($r > 0$) a X si

$$E[|X_n|^r] \text{ y } E[|X|^r] \text{ son finitos y } \lim_{n \rightarrow \infty} E[(X_n - X)^r] = 0$$

1.4.3.1 Algunas relaciones entre los tipos de convergencia

Algunos tipos de convergencia son más “fuertes” que otros, esto es, unos tipos incluyen a otro y en consecuencia un tipo de convergencia puede implicar otra. Se presenta primero la relación entre Convergencia en Probabilidad y Convergencia en Distribución

Teorema 1.5 De la Convergencia en Probabilidad a Convergencia en Distribución

Sea X una variable aleatoria y $\{X_n\}_{n \in \mathbb{N}}$ una sucesión de variables aleatorias todas definidas sobre el Espacio de Probabilidades (Ω, \mathcal{S}, P) . Si $\{X_n\}_{n \in \mathbb{N}}$ Converge en Probabilidad a X , entonces $\{X_n\}_{n \in \mathbb{N}}$ Converge en Distribución a X

Demostración: Sea F_X la función de distribución de X y x un punto de continuidad.

Probemos primero que

$$\{X_n \leq x\} \subset \{X \leq x + \varepsilon\} \cup \{|X_n - X| \geq \varepsilon\} \quad (1.8)$$

Para esto basta demostrar que si ω no está en ninguno de los dos conjuntos que forman la unión en el miembro derecho, entonces no está en $\{X_n \leq x\}$. Sea ω tal que $X(\omega) > x + \varepsilon$ y $|X_n(\omega) - X(\omega)| < \varepsilon$.

Luego

$$X_n(\omega) = X(\omega) + (X_n(\omega) - X(\omega))$$

$$\begin{aligned}
&\geq X(\omega) - |X_n(\omega) - X(\omega)| \\
&> x + \varepsilon - \varepsilon \\
&= x
\end{aligned}$$

Probado (1.8)

Tomando probabilidades en ambos miembros se obtiene

$$F_{X_n}(x) \leq F_X(x + \varepsilon) + P(|X_n - X| \geq \varepsilon)$$

Tomando límite superior en ambos miembros y teniendo en cuenta que

$$\lim_{n \rightarrow \infty} P(|X_n - X| \geq \varepsilon) = 0 \quad (1.9)$$

Se obtiene

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n}(x) \leq F_X(x + \varepsilon)$$

Y haciendo que $\varepsilon \rightarrow 0$, en virtud de la continuidad de F en x se tiene

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n}(x) \leq F_X(x) \quad (1.10)$$

Ahora hacemos un razonamiento similar a izquierda de x . Se considera la inclusión

$$\{X \leq x - \varepsilon\} \subset \{X_n \leq x\} \cup \{|X_n - X| \geq \varepsilon\}$$

Tomando probabilidades en ambos miembros se obtiene

$$F_X(x - \varepsilon) \leq F_{X_n}(x) + P(|X_n - X| \geq \varepsilon)$$

Tomando límite inferior en ambos miembros y usando (1.9) se obtiene

$$F_X(x) \leq \liminf_{n \rightarrow \infty} F_{X_n}(x) \quad (1.11)$$

De (1.10) y (1.11) resulta

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n}(x) \leq F_X(x) \leq \liminf_{n \rightarrow \infty} F_{X_n}(x)$$

Y como

$$\lim_{n \rightarrow \infty} F_{X_n}(x) \leq \overline{\lim}_{n \rightarrow \infty} F_{X_n}(x)$$

Debe ser

$$\overline{\lim}_{n \rightarrow \infty} F_{X_n}(x) = \liminf_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

Luego existe el límite de $\{F_{X_n}(x)\}$ en el punto x y además

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F(x) \quad \square$$

Ahora se presenta un resumen de otras relaciones existentes entre los tipos de convergencia

La Convergencia en Media Cuadrática (si existe) implica la convergencia en probabilidad

$$X_n \xrightarrow{m.c.} X \implies X_n \xrightarrow{P} X$$

La Convergencia en Probabilidad implica la Convergencia en Distribución

$$X_n \xrightarrow{P} X \implies X_n \xrightarrow{D} X$$

Si $\{X_n\}$ Converge en Distribución a un escalar c , entonces $\{X_n\}$ también Converge en Probabilidad a c

Como se mencionó, un estimador puntual $\hat{\theta}$ del Vector de Parámetros θ es un Vector Aleatorio, es deseable entonces que cumpla con ciertas características que estudiaremos a continuación.

Según Boente y Yohai en [2] y Mendenhall en [21] los estimadores presentan las siguientes propiedades.

1.4.4 Propiedades de los estimadores

La diferencia entre el parámetro poblacional y el estimador

$$|\theta - T| \tag{1.12}$$

es el “error de estimación”, obviamente una estimación será más precisa cuanto menor sea el error y por tanto es deseable que el error sea siempre cero, más eso es irreal; por ello, se esperarí que en promedio el error sea cero, esto es que la esperanza del error sea cero o que la esperanza del estimador sea igual al parámetro.

Definición 1.20 Estimador Insesgado

Sea X_1, \dots, X_n una sucesión de Variables Aleatorias independientes idénticamente distribuidas con distribución de probabilidad P_θ donde $\theta = (\theta_1, \theta_2, \dots, \theta_k)$. Sea T un Estimador del vector de parámetros θ . Se dice que T es un Estimador insesgado de θ si y sólo si

$$E[T] = \theta$$

Se denomina “Sesgo” del Estimador T a la diferencia entre el valor esperado de T y θ , denotado por

$$B(T) = E[T] - \theta \quad (1.13)$$

Así, un Estimador es Insesgado si y sólo si $B(T) = 0$; caso contrario el estimador es sesgado.

Para ilustrar la diferencia de un estimador sesgado y uno insesgado se presenta el siguiente ejemplo:

Sea X_1, X_2, \dots, X_n muestra aleatoria, con $E[X_i] = \mu$ y $Var(X_i) = \sigma^2$

Se define $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ y $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ dos

estimadores para σ^2 . Se procede a calcular el sesgo para ambos estimadores

Se desarrolla primero la sumatoria, se puede ver que

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - \frac{1}{n} \left(\sum_{i=1}^n X_i \right)^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$$

Por consiguiente,

$$E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] = E \left[\sum_{i=1}^n X_i^2 \right] - nE[\bar{X}^2] = \sum_{i=1}^n E[X_i^2] - nE[\bar{X}^2]$$

Obsérvese que $E(X_i^2)$ es igual para $i = 1, 2, \dots, n$. y además la varianza de una variable aleatoria está determinada por $Var(X) = E(X^2) - \mu^2$, entonces

$$E[X_i^2] = Var(X_i) + (E[X_i])^2 = \sigma^2 + \mu^2,$$

$$E[\bar{X}^2] = Var(\bar{X}) + (E[\bar{X}])^2 = \sigma^2 / n + \mu^2$$

Luego

$$\begin{aligned} E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] &= \sum_{i=1}^n (\sigma^2 + \mu^2) - n \left(\frac{\sigma^2}{n} + \mu^2 \right) \\ &= n(\sigma^2 + \mu^2) - n \left(\frac{\sigma^2}{n} + \mu^2 \right) \\ &= n\sigma^2 - \sigma^2 = (n-1)\sigma^2 \end{aligned}$$

Así, el sesgo de S_n^2 es

$$E(S_n^2) = \frac{1}{n} E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{1}{n} (n-1)\sigma^2 = \left(\frac{n-1}{n} \right) \sigma^2$$

Por lo tanto S_n^2 es sesgado porque $E[S_n^2] \neq \sigma^2$.

Mientras que el sesgo de S^2 es

$$E[S^2] = \frac{1}{n-1} E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{1}{n-1} (n-1) \sigma^2 = \sigma^2$$

Por lo tanto, se concluye que S^2 es un estimador insesgado para σ^2

□

Definición 1.21 Estimador asintóticamente insesgado

Sea $\{P_\theta, \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^k$ una familia no vacía de distribuciones de probabilidad y T un estimador de θ . Sean X_1, X_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas con distribución de probabilidad P_θ . Sean $T(X_1, X_2, \dots, X_n) = T_n$, $n \geq 1$ una sucesión de Estimadores puntuales de θ .

Un Estimador T_n del vector de parámetros θ se dice Asintóticamente Insesgado si y sólo si

$$E[T_n] \xrightarrow{n \rightarrow \infty} \theta$$

Que equivale a

$$\lim_{n \rightarrow \infty} B(T_n) = 0$$

Un Estimador insesgado acierta en promedio al parámetro estudiado, pero, ¿cuánto las estimaciones se dispersan del parámetro?. Cuando se tiene más de un estimador insesgado para un mismo parámetro θ , ¿cómo decidimos cuál conviene usar?

Por ejemplo, sea X_1, X_2, \dots, X_n una muestra aleatoria (sucesión de variable aleatoria independientes e idénticamente distribuidas) de una distribución $N(\mu, \sigma^2)$. Se puede verificar que los siguientes estimadores de μ son todos insesgados:

$$\hat{\mu}_1 = \bar{x}$$

$$\hat{\mu}_2 = \frac{X_1 + X_2}{2}$$

$$\hat{\mu}_3 = X_1$$

Sin embargo, las varianzas de estos estimadores son

$$V(\hat{\mu}_1) = \frac{\sigma^2}{n}$$

$$V(\hat{\mu}_2) = \frac{\sigma^2}{2}$$

$$V(\hat{\mu}_3) = \sigma^2$$

Es lógico elegir el estimador más preciso, es decir el de menor varianza, en este caso $\hat{\mu}_1$ \square

Definición 1.22 Eficiencia relativa

Los estimadores insesgados suelen compararse en términos de sus respectivas varianzas. Si T_1 y T_2 son dos estimadores insesgados del parámetro θ con $E[T_1^2] < \infty$ y $E[T_2^2] < \infty$. Si la varianza de T_1 es menor que la varianza de T_2 , se dice que T_1 es más eficiente que T_2 . También se suele usar la siguiente relación

$$ef_{\theta}(T_1|T_2) = \frac{V(T_2)}{V(T_1)}$$

para medir la eficiencia relativa de T_2 con respecto a T_1 .

Es razonable esperar que un estimador se aproxime más al parámetro poblacional a medida que se aumenta el tamaño de la muestra. Cuando el tamaño de la muestra es muy grande los estimadores tomarán, por lo general, valores muy próximos a los parámetros respectivos. Este concepto de proximidad se generalizará mediante la siguiente definición de consistencia.

Definición 1.23 Estimador consistente

Sea $\{P_{\theta}, \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^p$ una familia no vacía de Distribuciones de Probabilidad y T un Estimador de θ . Sean X_1, X_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas con

Distribución de Probabilidad P_θ . Sean $T(X_1, X_2, \dots, X_n) = T_n$, $n \geq 1$ una sucesión de Estimadores puntuales de θ .

Un Estimador $\hat{\theta}$ del vector de parámetros θ se dice consistente si y sólo si

$$T_n \xrightarrow{P} \theta$$

Es decir,

$$\lim_{n \rightarrow \infty} P(|T_n - \theta| > \varepsilon) = 0$$

La Ley de los Grandes Números nos permite comprobar que la Media Muestral $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ Converge en Probabilidad a la media poblacional; por tanto se dice que \bar{X}_n es un Estimador Consistente para μ .

Definición 1.24 Estadístico suficiente

Sea T un estadístico que va desde el espacio (Ω, \mathcal{A}, P) en el espacio probabilizable $(\mathbb{R}, \mathcal{B})$. Diremos que T es Suficiente para la familia P_θ de funciones de probabilidad si, para cada par de conjuntos de Borel, $A \in \mathcal{A}$ y $B \in \mathcal{B}$,

$$P[A \cap T^{-1}(B)] = \int_B P(A / T = t) P^T(dt),$$

donde $P(A / T = t)$ es independiente de la familia de medidas de probabilidad.

En otras palabras, se dice que un Estimador T es Suficiente si utiliza toda la información relevante de la muestra para estimar el parámetro θ de la población. Es decir, un estimador T es suficiente si todo el conocimiento que se obtiene acerca del parámetro θ es mediante la especificación real de todos los valores de la muestra.

Teorema 1.6 Teorema de Factorización

Sea X un vector aleatorio con función de densidad o función de probabilidad puntual $p(x, \theta)$, $\theta \in \Theta$. Entonces, el estadístico $T=r(X)$ es suficiente para θ si y sólo si existen dos funciones g y h tales que

$$p(x, \theta) = g(r(x), \theta)h(x)$$

Demostración: Se presenta la demostración sólo para el caso discreto. Se supone que existen dos funciones g y h tales que $p(x, \theta)$ se factoriza según $p(x, \theta) = g(r(x), \theta)h(x)$.

Entonces la función de densidad conjunta vale

$$p_{XT}(x, t, \theta) = \begin{cases} g(t, \theta)h(x) & \text{si } r(x) = t \\ 0 & \text{si } r(x) \neq t \end{cases}$$

Y la densidad marginal $p_T(t, \theta) = \sum_{r(x)=t} p_{XT}(x, t, \theta) = \sum_{r(x)=t} g(r(x), \theta)h(x)$

$$= g(t, \theta) \sum_{r(x)=t} h(x) = g(t, \theta) h^*(t)$$

Donde las sumatorias se realizan sobre todos los $x=(x_1, x_2, \dots, x_n)$

tales que $r(x) = t$. Así resulta la función de densidad condicional.

$$p_{xT}(x, \theta/t) = \begin{cases} h(x)/h^*(t) & \text{si } r(x) = t \\ 0 & \text{si } r(x) \neq t \end{cases}$$

Y por tanto la distribución de X dado $T = t$ es independiente de θ para todo t .

Recíprocamente, si suponemos que $T = r(x)$ es suficiente para θ se tiene

$$\begin{aligned} P_{\theta}(X=x) &= P_{\theta}(X=x, T=r(x)) = p_{xT}(x, r(x), \theta) \\ &= p_{X|T}(x, \theta | r(x)) p_T(r(x), \theta) \end{aligned}$$

El primero de los factores del último miembro es por hipótesis independiente de θ y por eso podemos llamarlo $h(x)$; mientras que el segundo miembro (que depende de x a través de t) puede denominarse $g(r(x), \theta)$. Por tanto el teorema queda demostrado. \square

Definición 1.25 Estimador insesgado de mínima varianza

Sea $\{P_{\theta}, \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}^p$ una familia no vacía de distribuciones de probabilidad y T un estimador de θ . Sean X_1, X_2, \dots una sucesión de variables aleatorias independientes idénticamente distribuidas con distribución de probabilidad P_{θ} . Sean $T(X_1, X_2, \dots, X_n) = T_n$, $n \geq 1$ una

sucesión de estimadores puntuales de la función escalar $c(\theta_1, \theta_2, \dots, \theta_n)$.

Un estimador T_n se dice estimador uniformemente insesgado de mínima varianza si y sólo si

$$E[T_n] = c(\theta)$$

$$V(T_n) \leq V(t(X)), \quad \text{para todo } t(X) \text{ estimador insesgado de } c(\theta)$$

Finalmente, se define el Error Cuadrático Medio como sigue

Definición 1.26 Error Cuadrático Medio

Sea T_n cualquier estimador de un parámetro θ . Se define el error cuadrático medio como el valor esperado del cuadrado de la diferencia entre el estimador T y el parámetro θ que trata de estimar.

$$ECM(T) = E[(T - \theta)^2]$$

El Error Cuadrático Medio puede escribirse como sigue:

$$ECM(T_n) = E[T_n - E[T_n]]^2 + (\theta - E[T_n])^2 = V(T_n) + B(T_n)^2$$

Esto es, el Error Cuadrático Medio de T_n es igual a la Varianza del Estimador más el Sesgo al cuadrado, así, si el estimador es insesgado el Error Cuadrático Medio es igual a la Varianza del Estimador.

CAPÍTULO 2

2 ESTIMACIÓN ROBUSTA

2.1 Introducción

En este segundo capítulo se amplía la idea de cuan importante es la estimación robusta, luego de la cual se trata de manera somera la teoría básica de robustez.

Se explican tres de los métodos para construir estimadores robustos y se proporcionan algunas ilustraciones de estimadores obtenidos a través de dichos procedimientos, profundizando en el estudio del método de los estimadores-M, para los que se desarrollan algunas propiedades definidas en la parte de la teoría básica de robustez. Finalmente se presentan algunos aspectos importantes sobre la estimación Robusta Multivariada.

2.2 Preliminares

En el Capítulo I se expone cómo en Estadística es común suponer que los datos han sido generados de una función de distribución dependiente del vector de parámetros poblacionales θ , y se explican además varios aspectos sobre la estimación de θ (enfoque que es la base de la llamada Estadística Paramétrica); enfatizando que en la práctica es frecuente que el modelo aleatorio real no se ajuste fielmente al modelo paramétrico supuesto siendo éste último sólo una moderada aproximación. Ésta es una de las razones que hacen necesaria a la teoría de la Robustez.

Para entender de mejor manera lo expuesto, se presenta a continuación un ejemplo, el mismo que fue planteado por Tukey en [27]

Se supone que se trabaja con un grupo aleatoriamente mezclado entre “buenas” observaciones normales $N(\mu, \sigma^2)$ y “malas” observaciones que son normales $N(\mu, 9\sigma^2)$; es decir, todas las observaciones tienen la misma media, pero las desviaciones de algunas son incrementadas por un factor 3; así cada observación X_i es buena con probabilidad $(1-\varepsilon)$ y mala observación con probabilidad ε , donde $\varepsilon(100)\%$ representa un porcentaje de “contaminación”. Dos medidas tradicionales de dispersión son:

El promedio de la desviación absoluta

$$d_n = \frac{1}{n} \sum |x_i - \bar{x}|$$

y la raíz del promedio de las desviaciones cuadráticas

$$S_n = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$$

Alrededor de 1920 se desarrolló una disputa entre Eddington y Fisher, acerca de las ventajas relativas de d_n y S_n , Fisher entonces hizo notar que para observaciones normales, S_n es 12% más eficiente que d_n , y esto pareció resolver el asunto.

Por supuesto, ambos estadísticos miden diferentes características para la distribución del error; por ejemplo, si los errores son normales, S_n converge a σ , mientras que d_n converge a

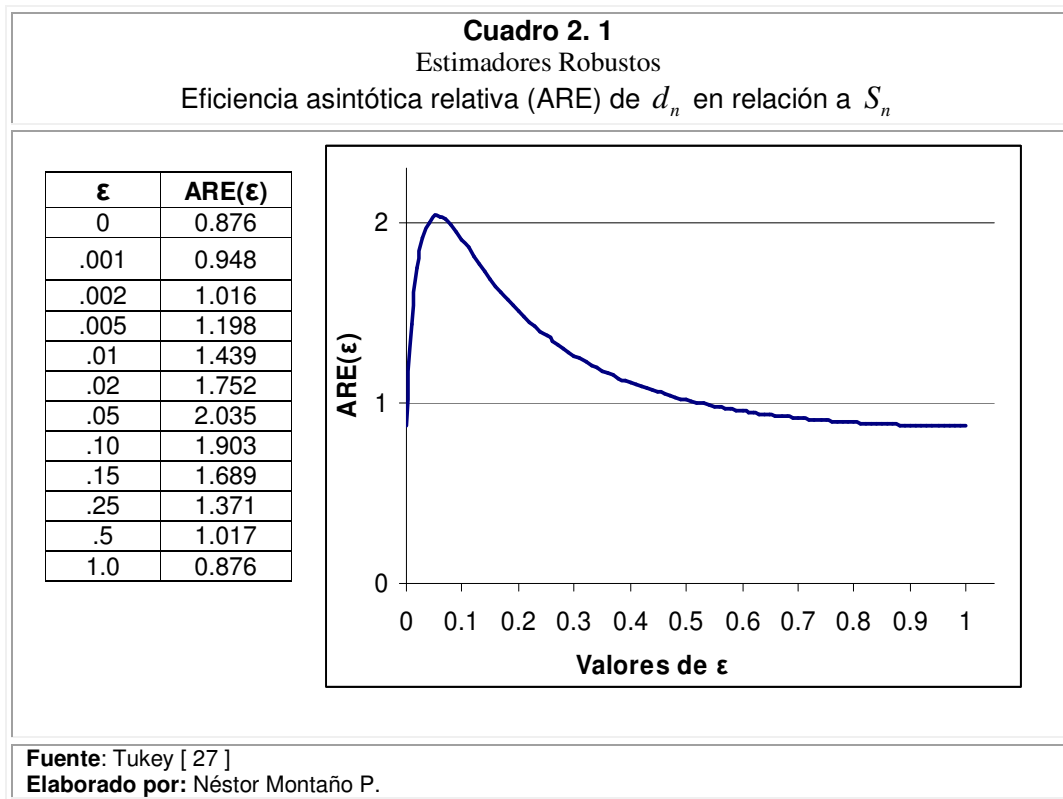
$$\sqrt{\frac{2}{\pi}} \sigma \approx 0.80 \sigma.$$

Tukey compara el rendimiento de estas dos medidas en base a la eficiencia asintótica relativa (ARE) de d_n en relación a S_n , de la siguiente manera:

$$ARE(\varepsilon) = \lim_{n \rightarrow \infty} \frac{\text{var}(S_n) / [E(S_n)]^2}{\text{var}(d_n) / [E(d_n)]^2}$$

$$= \frac{(3(1+80\varepsilon)/(1+8\varepsilon)^2 - 1)/4}{(\pi/2)(1+8\varepsilon)/(1+2\varepsilon)^2 - 1}$$

Se presenta a continuación la tabla de valores de $ARE(\varepsilon)$ para diferentes valores de ε y además el gráfico correspondiente



El resultado es inquietante: Bastan dos observaciones “malas” en 1000 para contrarrestar el 12% de ventaja de la media cuadrática del error y el ARE alcanza valores mayores a 2 alrededor de $\varepsilon=0.05$

Por otro lado, cuando se tienen valores extremos, para obtener una mejor estimación de los parámetros deseados, se pueden utilizar criterios de rechazo para dichos valores, sin embargo Rencher en [25] menciona que, Relles y Rogers (1977) realizaron un estudio en el cual compararon, usando una simulación tipo Monte Carlo, el rendimiento de algunos criterios de rechazo para valores extremos versus algunos de los procedimientos robustos. Ellos concluyeron que los criterios de rechazo son al menos un 20% menos eficiente que el mejor estimador robusto en todos los casos estudiados y que en el contexto multivariado el rendimiento relativo de los criterios podría ser aún peor; además los criterios de rechazo deberían “podar” una o dos observaciones más de las que en estos se consideran.

Se pasa ahora a una revisión de la teoría de los estimadores Robustos.

2.3 Teoría de la Robustez

Huber en [14], [15] y [16] Hampel en [11] y [12], Tukey en [27] y Zamar en [31], desarrollan y proponen varios aspectos a ser considerados, los mismos que se presentan a continuación.

2.3.1 Generalidades

El problema derivado de la naturaleza aproximada del modelo paramétrico F_θ es modelado por Tukey, quien define la “Familia de Contaminación” para tratar el problema de la presencia de valores aberrantes.

Definición 2.1 Familia de Contaminación

Se llama Familia de Contaminación a la familia de distribuciones \mathcal{F}_ε definida como:

$$\mathcal{F}_\varepsilon = \{(1 - \varepsilon)F_\theta + \varepsilon H; \theta \in \Theta\} \quad (2.1)$$

donde $0 < \varepsilon < 0.5$ es fijo y H es arbitraria y desconocida.

Es visible que $F_\theta \in \mathcal{F}_\varepsilon$.

El modelo (2.1) refleja la situación en que la mayoría de las observaciones provienen de F_θ , pero una pequeña fracción ε de las observaciones son valores atípicos generados por H , donde $\varepsilon(100)\%$ representa el porcentaje de contaminación; siendo evidente que una Distribución $F \in \mathcal{F}_\varepsilon$ no está completamente especificada, puesto que H es desconocida.

Por otro lado, además de clasificar a los estimadores como robustos o no, se debe comparar un estimador con otro, indicando por ejemplo que el estimador $\hat{\theta}_1$ de θ es más “Robusto” que $\hat{\theta}_2$; esto implica que se debe tener la capacidad de medir la robustez de un estimador.

De manera informal se puede decir, que el estimador $T(F_n)$ es “Robusto” si su comportamiento es relativamente “bueno” y “estable” cuando F varía sobre el entorno \mathcal{F}_ε del modelo paramétrico F_θ ; esto es, que el estimador $T(F_n)$ debe poseer las propiedades:

Eficiencia y

Estabilidad.

Definición 2.2 Eficiencia

Intuitivamente eficiencia se puede considerar como el hecho que $T(F_n)$ se comporte “bien” cuando el modelo paramétrico central $F = F_\theta$ se satisface.

Así, la eficiencia implica que $T(F_n)$ debe ser comparable con el estimador de máxima verosimilitud $EMV(F_n)$ cuando $F = F_\theta$ es decir

a) $T(F_n)$ debe ser “Consistente” en el sentido de Fisher, esto es

$$T_n \xrightarrow{p} T(F)$$

b) $T(F_n)$ debe ser asintóticamente normal, esto es

$$\mathcal{L}_F(\sqrt{n}(T_n - T(F))) \xrightarrow{L} N(0, A(F, T))$$

donde: F es una medida de probabilidad para (Ω, A)

y $\mathcal{L}_F(T_n)$ denota la distribución de T_n bajo F

Definición 2.3 Estabilidad

Un estimador $T(F_n)$ es estable cuando su “buen” comportamiento se preserva cuando F varía sobre \mathcal{F}_ε ; esta propiedad se puede explicar como: un pequeño cambio en la distribución debería causar sólo un pequeño cambio en el funcionamiento de un procedimiento estadístico.

Existen varios criterios para medir la Estabilidad de un Estimador; es decir para medir el buen comportamiento que se desea que presente cuando la muestra está contaminada, varios de los enfoques que se han desarrollado se tratan a continuación

2.3.2 Robustez Cualitativa

Definición 2.4 Robustez según Hampel.

En 1971, Hampel en [11] propuso una definición del concepto de robustez que responde a la noción intuitiva desarrollada hasta este

momento; su enfoque considera que el funcional $T(F_n)$ es estable (cualitativamente robusto) si para una adecuada métrica d_m en un espacio de probabilidad medible se cumple que

$$\forall \epsilon > 0 \exists \delta > 0 \forall G \text{ y } \forall n; \text{ tal que } n \geq n_0, d_m(F, G) < \delta \implies [\mathcal{L}_F(T_n), \mathcal{L}_G(T_n)] < \epsilon$$

donde: F es una medida de probabilidad para (Ω, \mathcal{A})

y $\mathcal{L}_F(T_n)$ denota la distribución de T_n bajo F

Huber en [16] indica que “esta situación es bastante análoga a la estabilidad de ecuaciones diferenciales ordinarias, donde la solución debería depender continuamente en los valores iniciales”; además simplifica la anterior formalización de robustez de la siguiente manera:

Supóngase que el estimador $T_n = T_n(X_1, X_2, \dots, X_n)$ donde X_1, X_2, \dots, X_n son variables aleatorias idénticamente distribuidas, puede ser escrito como un funcional T de la función de Distribución Empírica F_n o de la Medida Empírica

$$m_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$$

donde (x_1, x_2, \dots, x_n) es la Muestra y δ_{x_i} es el punto unitario de masa de x . Entonces, según Huber “un pequeño cambio en la muestra

resultaría en un pequeño cambio en el valor $T_n = T(m_n)$ de las estimaciones”.

Esto, sin embargo, no permite realizar comparaciones dentro de la clase de estimadores que son cualitativamente robustos, así las investigaciones en robustez se dirigen principalmente a cuantificar la misma.

2.3.3 Robustez Cuantitativa

En el Capítulo 1 se menciona la importancia de los estimadores insesgados y asintóticamente insesgados, midiendo el sesgo como (1.13); para los Estimadores Robustos el Sesgo Asintótico es un aspecto importante, puesto que si se supone que la Distribución F_θ es un elemento de \mathcal{F}_ε con $\varepsilon > 0$, entonces como resultado de la contaminación generada por H, un Estimador Consistente T_n de θ en general será Asintóticamente Sesgado.

El Sesgo Asintótico del estimador $T(F)$ cuando F varía sobre \mathcal{F}_ε , puede definirse como

$$b_\theta(F, \theta) = d [T((1-\varepsilon)F_\theta + \varepsilon H), \theta]$$

donde d es una distancia que respeta las invarianzas o equivarianzas naturales del espacio Θ .

Para el caso del modelo simple de posición y dispersión $\theta = (\mu, \sigma^2)$, las medidas de Sesgo Asintótico del estimador $\hat{\mu}(F_n)$ de μ y del estimador $\hat{\sigma}(F_n)$ de σ deben ser invariantes ante cambios de posición y escala. Por ejemplo, las medidas de sesgo

$$\frac{|\hat{\mu}(F) - \mu|}{\sigma}$$

y

$$\left| \frac{\hat{\sigma}(F)}{\sigma} - 1 \right|$$

poseen la propiedad de invarianza deseada, en virtud de la cual, se puede establecer un vector canónico θ_0 de θ , así, para $\theta_0^T = (0 \ 1)$ las medidas de sesgo planteadas son:

$$|\hat{\mu}(F)|$$

y

$$|\hat{\sigma}(F) - 1|$$

Con esa medida de sesgo, se define Sesgo Máximo como

$$b_1(\mathcal{E}) = \sup_{F \in \mathcal{F}_{\mathcal{E}}} |T(F) - T(F_0)|$$

y la Varianza Máxima como

$$v_1(\mathcal{E}) = \sup_{F \in \mathcal{F}_{\mathcal{E}}} A(F, T)$$

donde $A(F, T)$ proviene del requerimiento de normalidad asintótica

$$\mathcal{L}_F(\sqrt{n}(T_n - T(F))) \xrightarrow{L} N(0, A(F, T))$$

La idea de el Sesgo Máximo y de la Varianza Máxima es medir el “peor” comportamiento del estimador para cada nivel de contaminación ε ; sin embargo estas medidas no permiten establecer que para un n “grande” el estimador T_n se comporta bien para todo $F \in \mathcal{F}_\varepsilon$, para ello Huber propone lo siguiente

Definición 2.5 Sesgo Asintótico Máximo, Varianza Asintótica Máxima y el Punto de Ruptura Asintótico.

Se define, respectivamente, el Sesgo Asintótico Máximo y la Varianza Asintótica Máxima como:

$$b(\varepsilon) = \limsup_n \sup_{F \in \mathcal{F}_\varepsilon} |M(F, T_n)| \geq b_1(\varepsilon) \tag{2.2}$$

$$v(\varepsilon) = \limsup_n \sup_{F \in \mathcal{F}_\varepsilon} Q_t(F, T_n)^2 \geq v_1(\varepsilon) \tag{2.3}$$

donde $M(F, T_n)$ es la mediana de $\mathcal{L}_F(\sqrt{n}(T_n - T(F_0)))$ y $Q_t(F, T_n)$ es el rango t-cuantil normalizado de $\mathcal{L}_F(\sqrt{n} T_n)$. Para la Distribución Normal, Q_t coincide con la Desviación Estándar, así Q_t^2 será llamado pseudo-varianza.

Las desigualdades (2.2) y (2.3) son complicadas de manejar, por lo que se trabaja con b_I y v_I , gracias a que generalmente $b=b_I$ y $v=v_I$

La Varianza Asintótica Máxima generalizada (tomando el caso multivariado) se define como:

$$\overline{AV}_T(\varepsilon) = \sup_{F \in \mathcal{F}_\varepsilon} \text{traza}(A(F, T))$$

Ambas medidas miden el peor comportamiento del estimador, para n “grande”.

Huber define además el Punto de Ruptura Asintótico de T sobre F_0 , como

$$\varepsilon^* = \varepsilon^*(F_0, T) = \sup\{\varepsilon / b(\varepsilon) < b(1)\}$$

que intuitivamente representa la mayor fracción de contaminación que el estimador puede tolerar.

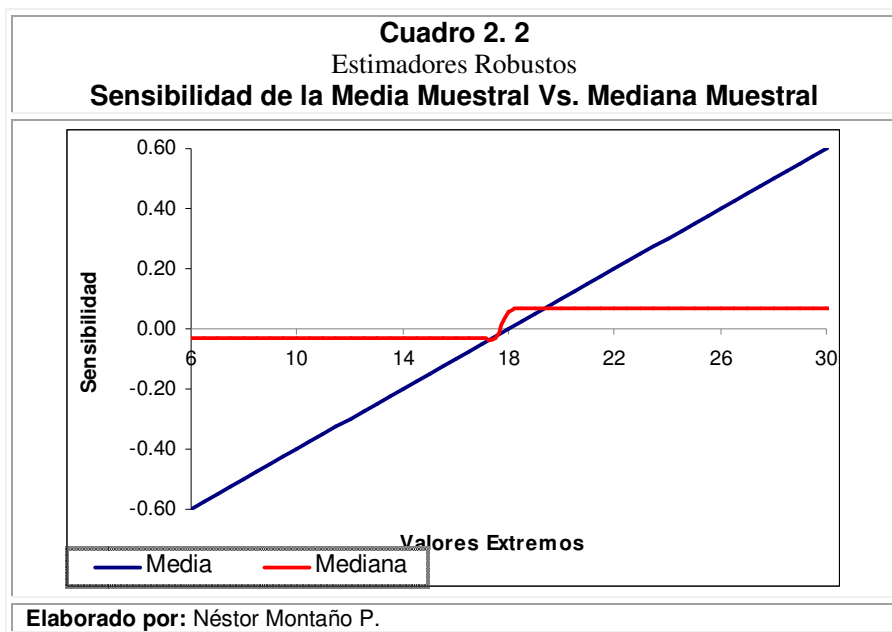
2.3.4 Robustez Infinitesimal

Recuerde el ejemplo dado en el Bosquejo inicial del Capítulo I, donde, sin considerar el valor atípico se tiene que $\bar{x}_{19} = 18.05$ y $\tilde{x}_{19} = 17.89$; mientras que incluyendo la observación $x_{20} = 28$ se tiene, $\bar{x}_{20} = 18.55$ y $\tilde{x}_{20} = 17.96$. Para medir el efecto de x_{20} se define la Curva de Sensibilidad del Estimador de Localización como

$$T(x_1, x_2, \dots, x_{19}, x_{20}) - T(x_1, x_2, \dots, x_{19}) .$$

que es la diferencia entre el parámetro estimado usando los 20 datos y 19 datos (excluyendo el valor extremo).

El cuadro 2.2 muestra la sensibilidad de la Media Muestral y la Mediana Muestral para el ejemplo dado, es claro que x_{20} tiene poca incidencia en la Mediana mientras que en Media ocurre lo contrario.



Ahora suponga que se desea medir el efecto de m observaciones iguales a un x_0 dentro de la muestra de tamaño n , es decir

$$T(x_0, x_0, \dots, x_0, x_{m+1}, \dots, x_n) - T(x_1, x_2, \dots, x_n)$$

Nótese que en este caso se reemplazan las primeras m observaciones por un valor x_0 .

De esta manera, si a las 19 primeras observaciones del ejemplo anterior, se le calcula la sensibilidad del estimador de localización reemplazando m observaciones con un valor de, por ejemplo, 100, veremos que con $m=1$ la Media se ve seriamente afectada, mientras que la Mediana se mantiene estable hasta $m=10$.

Hampel en [12] presenta una versión asintótica de la curva de sensibilidad.

Definición 2.6 Curva de Influencia y sensibilidad a errores groseros

Se define la llamada Curva de Influencia o Función de Influencia de la siguiente manera:

$$CI(x, F, T) = \lim_{\varepsilon \rightarrow 0} \frac{T((1-\varepsilon)F + \varepsilon\delta_x) - T(F)}{\varepsilon} \quad (2.4)$$

donde δ_x es la distribución tal que $P(x=x)=1$

La CI en (2.4) da la influencia diferencial adecuadamente escalada de una observación adicional con valor x cuando el tamaño de la muestra $n \rightarrow \infty$, es decir, es una aproximación del comportamiento de T cuando la muestra contiene una “pequeña” fracción ε de valores atípicos iguales a x .

Además Hampel define (suponiendo que existe el límite) la sensibilidad a errores groseros γ^* como

$$\gamma^*(T) = \sup_x |CI(x, F, T)|$$

En el presente trabajo se estudia el Modelo de Posición y Dispersión; esto es, cuando $\theta = (\mu, \sigma)$ donde μ es el estimador de posición, el mismo que proporciona una estimación del centro de los datos (una medida de tendencia central); y σ el estimador de Dispersión, el cual brinda una medida de la desviación de los datos con respecto al valor μ escogido. La Distribución Normal pertenece a la familia de modelos de posición y escala, de ahí la importancia de centrarnos en dicho modelo.

2.4 Tres métodos para construir Estimadores Robustos

Huber en [15] menciona tres métodos para la construcción de estimadores: Estimadores M, Estimadores L y Estimadores R; en lo que sigue, se menciona cada método y se proporciona ejemplos de estimadores obtenidos con los mismos.

Para una mejor comprensión se trabaja primero con un modelo puro de posición univariado, esto es

$$x_i = \mu + u_i \tag{2.5}$$

donde u_i son v.a. i.i.d y μ es el parámetro a estimar.

2.4.1 Estimadores L

A partir de una muestra ordenada $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, se define un estimador L como:

$$T_n = \sum_{i=1}^n a_i X_{(i)}$$

donde los ponderadores a_i son generados por $a_i = \int_{(i-1)/n}^{i/n} J(t) dt$ para

alguna función J que satisface $\int_0^1 J(t) dt = 1$.

La idea de los Estimadores L es construir estimadores a partir de combinaciones lineales de los estadísticos de orden.

Escribiendo el estimador L como un funcional de la función de distribución empírica $T_n = T(F_n)$ (exactamente o al menos aproximadamente) tenemos

$$T(F) = \int J(t) F^{-1}(t) dt$$

A continuación se presenta algunos Estimadores L, definidos según Gento en [9], Méndez en [22], Seber en [28].

Definición 2.7 Media Truncada y Mediana

Sea X_1, X_2, \dots, X_n una muestra aleatoria; $X_{(i)}$ el i -ésimo Estadístico de Orden y sea $\alpha \in [0, 0.50]$ el nivel de truncamiento. Se define entonces la media truncada como:

$$\alpha - trunc = \frac{1}{1 - 2\alpha} \int_{\alpha}^{1-\alpha} F^{-1}(t) dt \quad \circ$$

$$\alpha - trunc = \frac{1}{n - 2b} \sum_{i=b+1}^{n-2b} X_{(i)}, \text{ donde } b = [\alpha n]$$

Es decir, la Media Truncada a nivel α se construye como la media de las observaciones de la muestra eliminando el $\alpha\%$ de las observaciones en el extremo superior y el $\alpha\%$ en el extremo inferior.

Se puede notar que la Media y la Mediana podrían ser considerados como elementos extremos de la familia de las Medias Truncadas, pues un truncamiento con $\alpha=0$ coincide con la media muestral, y si $\alpha=0.50$ entonces se obtendría la mediana muestral.

Definición 2.8 Estadísticos de orden y Bisagras

Unos Estimadores parecidos a los Cuartiles fueron definidos por Tukey como Bisagras, que se denotan por H_1 y H_2 .

Sea $[\lambda]$ la parte entera de λ y $k = \frac{\left[\frac{n+1}{2} \right] + 1}{2}$, se define

$$H_1 = X_{(k)} \qquad H_2 = X_{(n+1-k)} \qquad , \text{si } k \in \mathbb{N}$$

$$H_1 = \frac{1}{2} (X_{([k])} + X_{([k]+1)}) \qquad H_2 = \frac{1}{2} (X_{([n+1-k])} + X_{([n+1-k]+1)}) , \text{si } k \notin \mathbb{N}$$

Definición 2.9 Trimedia

Sea X_1, X_2, \dots, X_n una Muestra Aleatoria, sea $X_{(i)}$ el i -ésimo estadístico de orden, \tilde{x} la Mediana Muestral y sea H_1 y H_2 las Bisagras de Tukey. Se define la Trimedia como:

$$\text{Trimedia} = \frac{1}{2} \tilde{x} + \frac{1}{4} H_1 + \frac{1}{4} H_2$$

Definición 2.10 Media α -Winsorizada

Sea X_1, X_2, \dots, X_n una muestra aleatoria, $X_{(i)}$ el i -ésimo estadístico de orden. Se define la Media α -Winsorizada como:

$$\bar{X}_{\alpha W} = \frac{1}{n} (gX_{(g+1)} + X_{(g+1)} + X_{(g+2)} + \dots + X_{(n-g)} + gX_{(n-g)}), \text{ donde } g = \alpha n$$

Esto significa realizar un simple reemplazo de cada uno de los g valores mayores ($X_{(n-g+1)}, X_{(n-g+2)}, \dots, X_{(n)}$) por $X_{(n-g)}$; y cada uno de los g valores menores ($X_{(1)}, X_{(2)}, \dots, X_{(g)}$) por $X_{(g+1)}$; así que, la muestra ordenada es $X_{(g+1)}, \dots, X_{(g+1)}, X_{(g+2)}, \dots, X_{(n-g)}, X_{(n-g)}, \dots, X_{(n-g)}$

Definición 2.11 Estimador de Salto Múltiple

Para este estimador se consideran los datos contenidos en las “patillas” de los gráficos Box-plot y posteriormente se elimina una proporción del número de datos excluidos. Sea el intervalo $[H_1 - 1,5(H_2 - H_1), H_2 + 1,5(H_2 - H_1)]$, sea k el número de datos fuera del intervalo y $L = \min\{\max\{1, 2k\}, (0,6n - k)\}$. Entonces se eliminan L valores extremos a cada lado del intervalo considerado, y se estima el parámetro como la media aritmética del resto de los valores. La principal particularidad es que en este estimador la proporción de datos a eliminarse depende de la muestra misma.

2.4.2 Estimadores R

Para estos Estimadores se considera una prueba de rangos para el modelo de desplazamiento de dos muestras. Sea Y_1, Y_2, \dots, Y_n y Z_1, Z_2, \dots, Z_n dos muestras independientes con distribuciones $F(x)$ y $F(x-\Delta)$ respectivamente. Se elabora la muestra combinada de tamaño $N=2n$, considere

$$H_0 : \Delta=0$$

$$H_a : \Delta>0$$

$$W(Y_1, Y_2, \dots, Y_n ; Z_1, Z_2, \dots, Z_n) = \sum_{i=1}^n J\left(\frac{i}{N+1}\right) V_i$$

donde $V_i = 1$ si la i -ésima entrada más pequeña en la muestra combinada es Y , y $V_i = 0$ en otro caso.

Se puede derivar estimadores de localización de dichas pruebas definiendo $T_n(X_1, X_2, \dots, X_n)$ tal que

$$W(X_1 - T_n, X_2 - T_n, \dots, X_n - T_n; -(X_1 - T_n), -(X_2 - T_n), \dots, -(X_n - T_n)) = 0$$

entonces, el estimador T_n que satisfaga la ecuación anterior es un Estimador R

Que al escribirlo como funcional (fórmula implícita en este caso) se tiene

$$\int J \left(\frac{F(x) + 1 - F(2T(F) - x)}{2} \right) F(dx) = 0$$

Un ejemplo de estos estimadores es el Estimador propuesto por Hodges Jr y Lehman en [13]

Definición 2.12 Estimador de Lehman & Hodges

Se define el Estimador \bar{X}_{HL} como:

$$\bar{X}_{HL} = \text{mediana}_{1 \leq i \leq j \leq n} \frac{1}{2}(Z_i - Z_j)$$

es decir, es la mediana del conjunto formado por los promedios de los pares de observaciones

2.5 Estimadores M

2.5.1 Ideas generales y definición

Huber en [14] propone (para la época) un nuevo método de estimación; él retoma la idea de estimación por medio de Mínimos Cuadrados propuesta por Gauss; esto es, ajustar el parámetro desconocido tal que la Suma de los Cuadrados de la diferencia entre los valores observados y los valores estimados sea minimizada, así, para el problema de estimar el parámetro de posición, se tiene que minimizar la expresión $\sum_i (x_i - T)^2$; lo que implica que T es igual a la media muestral.

Huber indica además que es natural preguntarse si uno puede obtener Estimadores Robustos al minimizar otra función de errores que no sea la suma de sus cuadrados, es entonces cuando define a los estimadores M de la siguiente forma.

$$T = T_n(x_1, \dots, x_n) \text{ que minimiza } \sum_i \rho(x_i - T); \quad (2.6)$$

donde ρ es una función no constante

Si se escoge $\rho(t) = \frac{t^2}{2}$ entonces el estimador que satisface (2.6) es la Media Muestral; de igual manera escogiendo $\rho(t) = |t|$, el estimador que resulta es la Mediana Muestral; así, la Media Muestral

y la Mediana Muestral son estimadores M y en general todos los estimadores de máxima verosimilitud son estimadores M, puesto que si la densidad de cada x_i es.

$$f(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right),$$

entonces la función de verosimilitud correspondiente a la muestra x_1, \dots, x_n será

$$L(\mu) = \frac{1}{\sigma^n} \prod_{i=1}^n f\left(\frac{x_i - \mu}{\sigma}\right)$$

Tomando logaritmos se tendrá que el estimador de máxima verosimilitud de μ estará dado por el valor que maximiza

$$\sum_{i=1}^n \log f\left(\frac{x_i - \mu}{\sigma}\right)$$

o equivalente a decir el valor que minimiza

$$\sum_{i=1}^n \rho_f\left(\frac{x_i - \mu}{\sigma}\right)$$

es decir, todos los Estimadores de Máxima Verosimilitud son Estimadores M con

$$\rho_f(u) = -\log f(u)$$

Formalmente se define a los estimadores M de la siguiente manera.

Definición 2.13 Estimadores de tipo Máxima Verosimilitud (Estimadores M)

Sea ρ una función de valor real de parámetro real con derivada $\psi = \rho'$. Se define un estadístico $T_n = T_n(X_1, X_2, \dots, X_n)$ por

$$\sum_{i=0}^n \rho(X_i - T_n) = \inf_t \sum_{i=0}^n \rho(X_i - t)$$

o

$$\sum_{i=0}^n \psi(X_i - T_n) = 0$$

Se reescribe la operación anterior de la siguiente manera:

$$\sum w_i \cdot (x_i - T_n) = 0$$

con

$$w_i = \frac{\psi(x_i - T_n)}{x_i - T_n}$$

de lo que se obtiene una representación de T_n como una media ponderada.

$$T_n = \frac{\sum w_i x_i}{\sum w_i}$$

Para incluir el problema de escala desconocida, se presenta una versión de escala invariante de este estimador si se reemplaza las ecuaciones definidas por

$$\sum_{i=1}^n \psi \left(\frac{X_i - T_n}{S_n} \right) = 0$$

donde $S_n = S(F_n)$ es algún estimador robusto de escala.

Para que los estimadores M sean robustos, deben poseer las propiedades de Eficiencia y Estabilidad mencionadas en la sección 2.3.

2.5.2 Propiedades de los estimadores M

Consistencia de los estimadores M

La condición de consistencia de Fisher, requerida para que el estimador M converja a μ está dada por

$$E_F \left(\psi \left(\frac{x - \mu}{\sigma} \right) \right) = 0 \quad (2.7)$$

donde $E_F(X)$ significa la esperanza de X con Distribución $F(x)$

La ecuación (2.7) es equivalente a:

$$E_F(\psi(u)) = 0 \quad (2.8)$$

Esta condición se cumple automáticamente si F tiene una densidad simétrica respecto a 0 ya que en ese caso se tiene:

$$E_F(\psi(u)) = \int_{-\infty}^{+\infty} u f(u) du = 0$$

ya que $uf(u)$ es una función impar.

Luego, se tiene el siguiente Teorema que muestra la consistencia de los estimadores M.

Teorema 2.1 Consistencia de los estimadores M

Sean x_1, \dots, x_n variables aleatorias independientes. Se dice que un estimador $\hat{\mu}_n$ converge en casi todo punto a μ en cualquiera de los siguientes casos:

- 1) La función ψ es estrictamente creciente.
- 2) La función ψ es no decreciente, $\psi(u) > \psi(0)$ y $F(u) > F(0)$ para todo $u > 0$.

Demostración: Solamente mostraremos para el caso (1).

Consideremos $\varepsilon > 0$. Luego, como ψ es estrictamente creciente tenemos que $\psi(u - \varepsilon) < \psi(u)$, y luego

$$E_F \psi(u - \varepsilon) < E_F \psi(u) = 0$$

Por tanto

$$E_F \psi\left(\frac{x - (\mu + \varepsilon)}{\sigma}\right) = E_F \psi(u - \varepsilon) < 0$$

De manera similar, se puede probar que

$$E_F \psi\left(\frac{x - (\mu - \varepsilon)}{\sigma}\right) = E_F \psi(u + \varepsilon) > 0$$

Sea ahora

$$G_n(\mu^*) = \frac{1}{n} \sum \psi\left(\frac{x_i - \mu^*}{\sigma}\right)$$

Luego el estimador M $\hat{\mu}_n$ satisface

$$G_n(\hat{\mu}_n) = 0$$

Por otro lado usando la Ley de los Grandes Números se tiene que, con probabilidad 1 existe un n_0 tal que, para todo $n > n_0$ se tiene que:

$$G_n(\mu + \varepsilon) < 0, \quad G_n(\mu - \varepsilon) > 0$$

Y por lo tanto como G_n es monótona decreciente, se tiene que el valor $\hat{\mu}_n$ tendrá que satisfacer que

$$\mu - \varepsilon < \hat{\mu}_n < \mu + \varepsilon$$

Esto prueba la consistencia de $\hat{\mu}_n$ □

También se debe probar la Normalidad Asintótica de los Estimadores M, por lo que se presenta el siguiente teorema

Teorema 2.2 Normalidad asintótica de los estimadores M

Sean x_1, \dots, x_n variables aleatorias independientes. Consideremos un estimador $\hat{\mu}_n$, donde ψ y F satisfacen (2.8). Supongamos

además que ψ tiene dos derivadas continuas y ψ^n es acotado. Luego se tiene que:

$$n^{1/2}(\hat{\mu}_n - \mu) \xrightarrow{D} N(0, \sigma^2 V(\psi, F))$$

donde

$$V(\psi, F) = \frac{E_F \psi^2(u)}{(E_F \psi^2(u))^2}$$

Demostración.- El estimador M $\hat{\mu}_n$ satisface

$$\sum_{i=1}^n \psi\left(\frac{x_i - \hat{\mu}_n}{\sigma}\right) = 0 \quad (2.9)$$

Haciendo un desarrollo de Taylor en el punto μ se tiene

$$0 = \sum_{i=1}^n \psi\left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^n \psi'\left(\frac{x_i - \mu}{\sigma}\right) \frac{\hat{\mu}_n - \mu}{\omega} + \frac{1}{2} \sum_{i=1}^n \psi''\left(\frac{x_i - \mu_n^*}{\sigma}\right) \frac{(\hat{\mu}_n - \mu)^2}{\sigma^2}$$

donde μ_n^* es un punto intermedio entre $\hat{\mu}_n$ y μ .

Luego, se tiene la siguiente ecuación implícita

$$(\hat{\mu}_n - \mu) = \frac{\sum_{i=1}^n \psi((x_i - \mu) / \sigma)}{\frac{1}{\sigma} \sum_{i=1}^n \psi'((x_i - \mu) / \sigma) - \frac{1}{2} \frac{(\hat{\mu}_n - \mu)}{\sigma^2} \sum_{i=1}^n \psi''((x_i - \mu_n^*) / \sigma)}$$

Por lo que:

$$n^{1/2}(\hat{\mu}_n - \mu) = \frac{\frac{1}{n^{1/2}} \sum_{i=1}^n \psi((x_i - \mu) / \sigma)}{\frac{1}{n\sigma} \sum_{i=1}^n \psi'((x_i - \mu) / \sigma) - \frac{1}{2\sigma^2}(\hat{\mu}_n - \mu) \frac{1}{n} \sum_{i=1}^n \psi''((x_i - \mu^*_n) / \sigma)}$$

Sea además,

$$A_n = \frac{1}{n^{1/2}} \sum_{i=1}^n \psi((x_i - \mu) / \sigma) = \frac{1}{n^{1/2}} \sum_{i=1}^n \psi(u_i)$$

$$B_n = \frac{1}{n^{1/2}} \sum_{i=1}^n \psi'((x_i - \mu) / \sigma) = \frac{1}{n^{1/2}} \sum_{i=1}^n \psi'(u_i)$$

$$C_n = \frac{1}{2}(\hat{\mu}_n - \mu) \frac{1}{n} \sum_{i=1}^n \psi''((x_i - \mu^*_n) / \sigma)$$

Luego

$$n^{1/2}(\hat{\mu}_n - \mu) = \frac{A_n}{\sigma^{-1} B_n + \sigma^{-2} C_n}$$

Aplicando el Teorema del Límite Central se tiene:

$$A_n \xrightarrow{D} N(0, E_F(\psi^2(u)))$$

Por la Ley Fuerte de los Grandes Números se concluye además que:

$$B_n \xrightarrow{p} E(\psi^2(u))$$

Finalmente, por hipótesis existe una constante K tal que $|\psi''(u)| < K$.

Luego $|C_n| < (K/2)(\hat{\mu}_n - \mu)$. Usando el hecho de que $\hat{\mu}_n \xrightarrow{p} \mu$, se

tiene que

$$C_n \xrightarrow{p} 0$$

□

2.5.3 Estimadores M de Huber

Huber propuso una familia de funciones intermedias entre las correspondientes a la distribución $N(0,1)$ y a la doble exponencial $e^{\frac{|x|}{2}}$, de tal manera que esta función es cuadrática para valores absolutos pequeños y lineal para valores absolutos grandes.

El estimador M de Huber es aquel que satisface (2.6) donde

$$\rho(x) = \begin{cases} x^2 / 2 & \text{para } |x| \leq c \\ c|x| - c^2 / 2 & \text{para } |x| > c \end{cases}$$

en consecuencia

$$\psi(x) = \begin{cases} -c & ; x < -c \\ x & ; -c \leq x \leq c \\ c & ; x > c \end{cases}$$

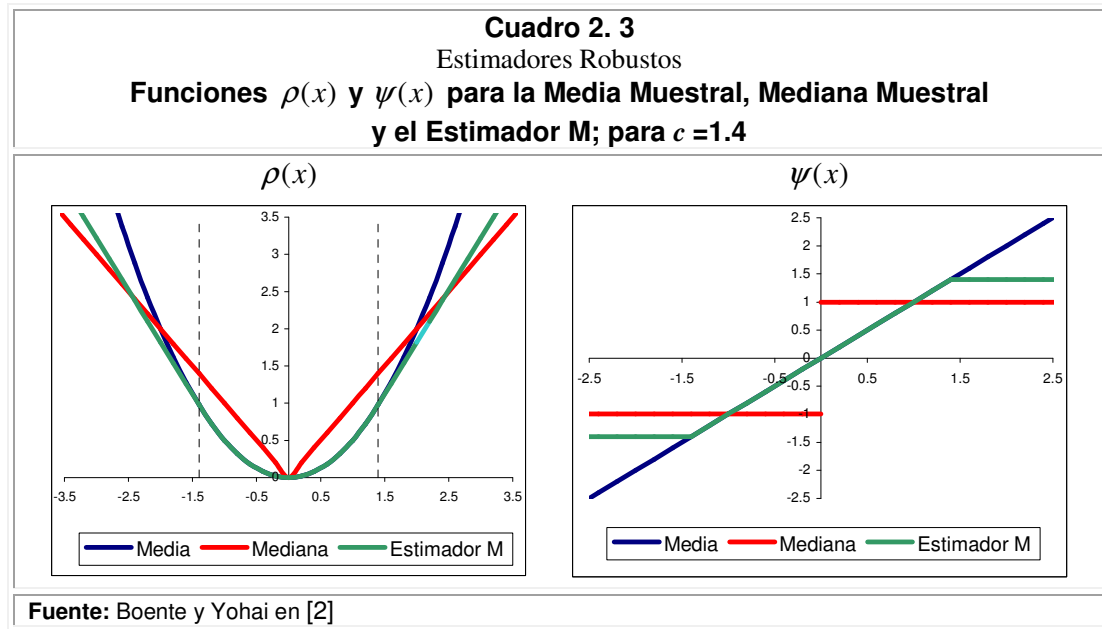
y, a manera de las ponderaciones, es:

$$w_i = \begin{cases} 1 & \text{para } |x_i - T_n| \leq c \\ \frac{c}{|x_i - T_n|} & \text{para } |x_i - T_n| > c \end{cases}$$

donde c depende de ε .

El siguiente gráfico muestra las funciones $\rho(x)$ y $\psi(x)$ para la Media Muestral, Mediana Muestral y el Estimador M de Huber para $c=1.4$; se puede ver que las funciones correspondientes al Estimador M de

Huber es idéntico a la Media Muestral cuando $|x| \leq c$ y es similar al de la Mediana Muestral para $|x| > c$



Se puede ver que el estimador M para un valor de $c=0$ corresponde a la Mediana muestral mientras que cuando c tiende a infinito, éste corresponde a la Media muestral

En general la Función de Ponderaciones w_i es par, monótona, decreciente en $|x_i - T_n| > c$, así el Estimador M de Huber se puede interpretar como Media Pondera que penaliza las observaciones para las cuales $|x_i - T_n|$ es grande.

Maronna, Martin y Yohai en [18] presentan una tabla con la Varianza Asintótica del estimador propuesto por Huber (bajo el entorno de contaminación \mathcal{F}_ϵ) con $F_{\theta=N(0,1)}$ y $H=N(0, 10)$ para diferentes valores de c .

Cuadro 2. 4			
Estimadores Robustos			
Varianza asintótica del estimador M propuesto por Huber			
c	$\epsilon = 0$	$\epsilon = 0.05$	$\epsilon = 0.10$
0	1.571	1.722	1.897
0.7	1.187	1.322	1.501
1.0	1.107	1.263	1.443
1.4	1.047	1.227	1.439
1.7	1.023	1.233	1.479
2.0	1.010	1.259	1.550
∞	1.000	5.950	10.900

Fuente: Maronna, Martin y Yohai en [18]
Elaborado por: Néstor Montaña P.

En la tabla se puede ver la relación entre robustez y eficiencia, así para $c=1.4$ y $\epsilon = 0$ (Normal Estándar sin Contaminación) la Varianza del estimador M es solo 4.7% mayor a la de la Media ($c \rightarrow \infty$) y mucho menor a la de la Mediana ($c=0$) mientras que para $\epsilon > 0$ (Normal contaminada) los estimadores M presentan menor Varianza Asintótica que la Media y la Mediana.

Además de los Estimadores M de Huber, Hampel y Tukey también han propuesto estimadores de tipo Máxima Verosimilitud, a continuación se presenta el estimador sugerido por Tukey.

Definición 2.14 Estimador M Bicuadrático.

Mientras que Huber propone una familia de funciones intermedias entre las correspondientes a la distribución $N(0,1)$ y a la doble exponencial $e^{-|x|}$, Tukey considera a la distribución T, la cual posee colas “más pesadas” que la normal, véase el Cuadro 2.5. Para este caso Tukey considera mejor usar ψ que sean “redescendientes”, esto es, que $\lim_{x \rightarrow \infty} \psi(x) = 0$, así el Estimador M Bicuadrático o Bponderado es aquel que satisface (2.6) donde

$$\rho(x) = \begin{cases} 1 - [1 - (x/k)^2]^3 & \text{si } |x| \leq k \\ 1 & \text{si } |x| > k \end{cases} \quad (2.10)$$

con derivada $\rho'(x) = 6 \frac{\psi(x)}{k^2}$ donde

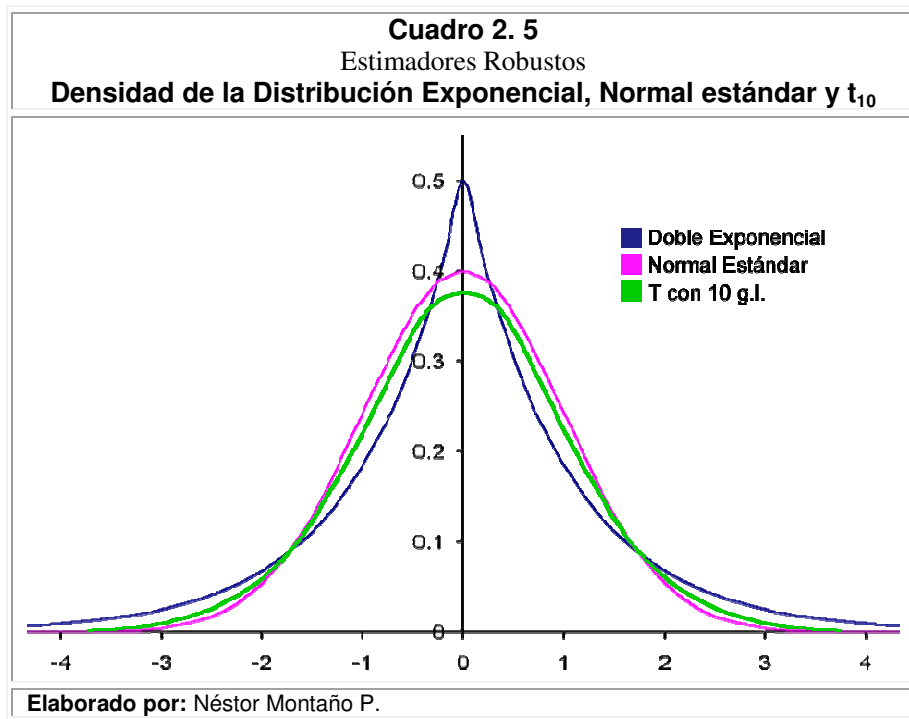
$$\psi(x) = x \left[1 - \left(\frac{x}{k} \right)^2 \right]^2 \mathbf{I}(|x| \leq k)$$

donde $\mathbf{I}(|x| \leq k) = \begin{cases} 1 & |x| \leq k \\ 0 & |x| > k \end{cases}$

así, a manera de ponderaciones es:

$$w_i = \left[1 - \left(\frac{x_i}{k} \right)^2 \right]^2 \mathbf{I}(|x_i| \leq k)$$

La idea de Tukey es que, las observaciones que estén muy alejadas de la concentración de datos, no solo sean ponderadas sino, en cierto punto rechazadas (con ponderación cero).



2.5.4 Dos algoritmos para calcular Estimadores M

A continuación se describen dos algoritmos para calcular el Estimador M de acuerdo a como indica Boente, G. & Yohai en [2].

Algoritmo basado en Medias Ponderadas Iteradas (MPI)

Sea $w(u) = \psi(u) / u$. La ecuación (2.9) se puede escribir como:

$$\sum_{i=1}^n (x_i - \hat{\mu}) w\left(\frac{x_i - \hat{\mu}_n}{\hat{\sigma}}\right) = 0$$

esto es

$$\sum_{i=1}^n x_i w\left(\frac{x_i - \hat{\mu}_n}{\hat{\sigma}}\right) = \hat{\mu} w\left(\frac{x_i - \hat{\mu}}{\hat{\sigma}}\right),$$

expresando $\hat{\mu}$ de manera implícita,

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i w((x_i - \hat{\mu}) / \hat{\sigma})}{\sum_{i=1}^n w((x_i - \hat{\mu}) / \hat{\sigma})}$$

Como se puede observar, esta expresión sugiere un algoritmo iterativo para calcular $\hat{\mu}$.

En efecto, considérese como un estimador inicial $\hat{\mu}_0$ de μ , a la Mediana Muestral, entontes la siguiente iteración es

$$\hat{\mu}_1 = \frac{\sum_{i=1}^n x_i w((x_i - \hat{\mu}_0) / \hat{\sigma})}{\sum_{i=1}^n w((x_i - \hat{\mu}_0) / \hat{\sigma})}$$

y en general $\hat{\mu}_{h+1}$ está dada por

$$\hat{\mu}_{h+1} = \frac{\sum_{i=1}^n x_i w((x_i - \hat{\mu}_h) / \hat{\sigma})}{\sum_{i=1}^n w((x_i - \hat{\mu}_h) / \hat{\sigma})} \quad (2.11)$$

Se puede mostrar que si ψ es continua, entonces cuando éste algoritmo iterativo converge, lo hace a una solución de

$$\sum_{i=1}^n \psi \left(\frac{x_i - \mu}{\hat{\sigma}} \right) = 0 \quad (2.12)$$

Se supone que $\lim_{h \rightarrow \infty} \hat{\mu}_h = \hat{\mu}$, luego tomando límite en ambos lados de (2.11) se tiene.

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i w((x_i - \hat{\mu}) / \hat{\sigma})}{\sum_{i=1}^n w((x_i - \hat{\mu}) / \hat{\sigma})}$$

Esta expresión define a $\hat{\mu}$ como promedio ponderado de las x_i con pesos proporcionales a $w((x_i - \hat{\mu}) / \hat{\sigma})$. Para la media aritmética se tiene $w(u) = 1$, y para el estimador M de Huber, la correspondiente función de peso está dada por

$$w_k^H(u) = \begin{cases} 1 & \text{si } |u| \leq k \\ \frac{k}{|u|} & \text{si } |u| > k \end{cases}$$

Algoritmo de Newton Raphson (NR)

Boente, G. & Yohai en [2] muestra que el algoritmo de Newton Raphson para calcular la raíz de (2.12) tiene la siguiente expresión recursiva

$$\hat{\mu}_{h+1} = \hat{\mu}_h + \hat{\sigma} \frac{\sum_{i=1}^n \psi((x_i - \hat{\mu}_h) / \hat{\sigma})}{\sum_{i=1}^n \psi'((x_i - \hat{\mu}_h) / \hat{\sigma})} \quad (2.13)$$

Para el caso del Estimador M de Huber, esta expresión toma una forma particularmente interesante, pues para calcular el valor de μ se puede dividir el conjunto de observaciones en tres subconjuntos

$$D_- = \{i : (x_i - \hat{\mu}_k) / \hat{\sigma} < -k\},$$

$$D_0 = \{i : |x_i - \hat{\mu}_k| / \hat{\sigma} \leq k\},$$

$$D_+ = \{i : (x_i - \hat{\mu}_k) / \hat{\sigma} > k\},$$

Así, se tiene

$$\psi((x_i - \hat{\mu}_h) / \hat{\sigma}) = \begin{cases} -k & \text{si } i \in D_- \\ (x_i - \hat{\mu}_h) / \hat{\sigma} & \text{si } i \in D_0 \\ k & \text{si } i \in D_+ \end{cases},$$

y con derivada

$$\psi'((x_i - \hat{\mu}_h) / \hat{\sigma}) = \begin{cases} 0 & \text{si } i \in D_-(\hat{\mu}_h) \\ 1 & \text{si } i \in D_0(\hat{\mu}_h) \\ 0 & \text{si } i \in D_+(\hat{\mu}_h) \end{cases} \frac{\Delta y}{\Delta x}$$

Sea n_- , n_0 y n_+ el número de elementos de D_- , D_0 y D_+

respectivamente, al reemplazar en (2.13) se tiene:

$$\hat{\mu}_{h+1} = \hat{\mu}_h + \hat{\sigma} \frac{k(n_+ - n_-) + \sum_{i \in D_0} (x_i - \hat{\mu}_h) / \hat{\sigma}}{n_0} = \frac{n_+ - n_-}{n_0} \hat{\sigma} k + \frac{1}{n_0} \sum_{i \in D_0} x_i$$

Obsérvese que el miembro derecho de esta última expresión solo depende de D_- , D_0 y D_+ . Estos tres conjuntos forman una partición del conjunto $\{1, 2, \dots, n\}$. Es claro que hay un número finito de éstas particiones, y por lo tanto si $\hat{\mu}_h$ converge, lo debe hacer en un número finito de pasos.

Ejemplo:

Se han generado 18 datos que siguen una Distribución Normal Estándar, los mismos que han sido contaminados con 2 datos; es decir, esta muestra corresponde a un entorno de contaminación descrito en (2.1) donde F_θ es una Normal Estándar y $\varepsilon = 0.1$

Los datos son:

2.800	-0.540	0.028	-0.381	0.017
2.289	-0.053	-1.711	1.327	-0.711
-0.879	-0.027	-0.998	-0.838	1.234
-0.229	-2.135	1.207	-0.040	-3.600

Donde la Media aritmética es igual -0.162 y la Mediana muestral igual a -0.141 están bastante alejados del valor μ que es igual a cero. Como se puede observar en el Cuadro 2.6, el Estimador M de Huber para el parámetro de posición, aplicando el procedimiento de Newton Raspón, resulta igual a -0.059; es decir, en este caso brinda una mejor estimación que la Media y la Mediana Muestral.

A continuación se presenta una tabla con los pesos asignados a cada observación para cada iteración generada con $c=1.5$

Cuadro 2. 6
Estimadores Robustos
Procedimiento de Newton Raphson

	Muestra inicial	Iteracion 1		Iteracion 2		Iteracion 3		Iteracion 4	
	X_i	peso	X_i *peso	peso	X_i *peso	peso	X_i *peso	peso	X_i *peso
1	2.800	1.500	0.000	0.156	0.000	0.150	0.000	0.149	0.000
2	2.289	1.500	0.000	0.156	0.000	0.150	0.000	0.149	0.000
3	-0.879	-0.738	-0.879	-0.724	-0.879	-0.729	-0.879	-0.730	-0.879
4	-0.229	-0.088	-0.229	-0.074	-0.229	-0.080	-0.229	-0.080	-0.229
5	-0.540	-0.399	-0.540	-0.385	-0.540	-0.391	-0.540	-0.391	-0.540
6	-0.053	0.088	-0.053	0.103	-0.053	0.097	-0.053	0.096	-0.053
7	-0.027	0.114	-0.027	0.129	-0.027	0.123	-0.027	0.122	-0.027
8	-2.135	-1.500	0.000	0.156	0.000	0.150	0.000	0.149	0.000
9	0.028	0.170	0.028	0.184	0.028	0.178	0.028	0.178	0.028
10	-1.711	-1.500	0.000	0.156	0.000	0.150	0.000	0.149	0.000
11	-0.998	-0.857	-0.998	-0.843	-0.998	-0.849	-0.998	-0.849	-0.998
12	1.207	1.349	1.207	1.363	1.207	1.357	1.207	1.357	1.207
13	-0.381	-0.239	-0.381	-0.225	-0.381	-0.231	-0.381	-0.231	-0.381
14	1.327	1.468	1.327	1.483	1.327	1.477	1.327	1.476	1.327
15	-0.838	-0.697	-0.838	-0.683	-0.838	-0.689	-0.838	-0.689	-0.838
16	-0.040	0.102	-0.040	0.116	-0.040	0.110	-0.040	0.110	-0.040
17	0.017	0.159	0.017	0.173	0.017	0.167	0.017	0.167	0.017
18	-0.711	-0.570	-0.711	-0.555	-0.711	-0.561	-0.711	-0.562	-0.711
19	1.234	1.376	1.234	1.390	1.234	1.384	1.234	1.384	1.234
20	-3.600	-1.500	0.000	0.156	0.000	0.150	0.000	0.149	0.000
μ-estimada	-0.141		-0.159		-0.059		-0.059		-0.059

Elaborado por: Néstor Montaña P.

2.6 Estimadores de Escala

Hasta el momento se ha tratado sólo el problema de la estimación del parámetro de posición, siendo necesario aún considerar la estimación del parámetro de dispersión. Considérese el modelo

$$x_i = \sigma u_i \quad (2.14)$$

donde las u_i son i.i.d con densidad f_0 y $\sigma > 0$ es el parámetro a estimar.

Se debe resaltar que en el modelo (2.14) σ es un parámetro de escala para u_i , sin embargo bajo el modelo $x_i = \mu + \sigma u_i$, σ representa un parámetro de dispersión para las x_i .

A continuación se presentan algunos ejemplos de estimadores robustos de escala.

Definición 2.15 Mediana Absoluta de las Desviaciones

Análogo a la Mediana se presenta la Mediana Absoluta de las Desviaciones, que se define como

$$MAD(x) = Med|x - Med(x)|$$

Pero este valor no converge a la Desviación Estándar en poblaciones Normales, por ello debe normalizarse de la siguiente manera

$$MADN(x) = \frac{MAD(x)}{0.675}$$

2.6.1 Estimadores M de escala

Siguiendo el modelo (2.14), la densidad correspondiente a x_i es

$$\frac{1}{\sigma} f_0\left(\frac{x}{\sigma}\right)$$

Conociendo que los estimadores M son una generalización de los Estimadores de Máxima Verosimilitud, se procede a obtener el Estimador de Máxima Verosimilitud bajo el modelo (2.14); así, la función de Verosimilitud es

$$\prod_{i=1}^n \frac{1}{\sigma^n} f_0\left(\frac{x_i}{\sigma}\right)$$

aplicando logaritmo

$$-n \ln \sigma + \ln f_0\left(\frac{x_1}{\sigma}\right) + \ln f_0\left(\frac{x_2}{\sigma}\right) + \dots + \ln f_0\left(\frac{x_n}{\sigma}\right) \quad (2.15)$$

Se maximiza la ecuación (2.15)

$$-\frac{n}{\sigma} + \frac{1}{f_0\left(\frac{x_1}{\sigma}\right)} f_0'\left(\frac{x_1}{\sigma}\right) (-x_1 \sigma^{-2}) + \dots + \frac{1}{f_0\left(\frac{x_n}{\sigma}\right)} f_0'\left(\frac{x_n}{\sigma}\right) (-x_n \sigma^{-2}) = 0$$

$$\frac{1}{\sigma^2} \sum_{i=1}^n \frac{x_i f_0' \left(\frac{x_i}{\sigma} \right)}{f_0 \left(\frac{x_i}{\sigma} \right)} = \frac{n}{\sigma}$$

$$\frac{1}{n} \sum_{i=1}^n \frac{\frac{x_i}{\sigma} f_0' \left(\frac{x_i}{\sigma} \right)}{f_0 \left(\frac{x_i}{\sigma} \right)} = 1 \quad (2.16)$$

Sea $\rho(t) = t\psi(t)$ con $\psi(t) = \frac{-f_0'(t)}{f_0(t)}$, entonces (2.16) equivale a:

$$\frac{1}{n} \sum_{i=1}^n \rho \left(\frac{x_i}{\hat{\sigma}} \right) = 1$$

En general, si un estimador satisface la ecuación

$$\frac{1}{n} \sum_{i=1}^n \rho \left(\frac{x_i}{\hat{\sigma}} \right) = \delta, \quad \delta > 0 \quad (2.17)$$

entonces se le llama estimador M de escala.

Nótese que, para que (2.17) tenga solución es necesario que $0 < \delta < \rho(\infty)$; entonces, si ρ es acotado, sin pérdida de generalidad se puede asumir que:

$$\rho(\infty) = 1; \quad \delta \in (0,1)$$

Mayormente se utiliza una función ρ cuadrática cerca del origen; esto es, $\rho'(0) = 0$ y $\rho''(0) > 0$, y en tales casos un Estimador M de escala puede ser representado como un promedio ponderado de la

Raíz de las Sumas Cuadráticas. Se define la Función de Ponderadores como:

$$w_i = \begin{cases} \rho(x_i) / x_i^2 & \text{si } x_i \neq 0 \\ \rho''(0) & \text{si } x_i = 0 \end{cases}$$

Entonces (2.17) es equivalente a:

$$\hat{\sigma}^2 = \frac{1}{n\delta} \sum_{i=1}^n w\left(\frac{x_i}{\hat{\sigma}}\right) x_i^2$$

Definición 2.16 Estimador M Bicuadrático de escala

Un estimador de escala que se utiliza frecuentemente es el estimador Bicuadrático o también llamado Biponderado, donde ρ está dado por (2.10) con $k=1$, es decir

$$\rho(x) = \min\{1 - (1 - x^2)^3, 1\} \text{ y } \delta = 0.5.$$

Se procede a considerar el modelo completo de posición y dispersión, con ambos parámetros desconocidos. Supóngase que se trabaja con el modelo

$$x_i = \mu + \sigma u_i \tag{2.18}$$

donde u_i tiene densidad f_0 . Donde μ y σ son los parámetros a estimar.

Entonces x_i tiene densidad:

$$f(x) = \frac{1}{\sigma} f_0\left(\frac{x - \mu}{\sigma}\right)$$

2.6.2 Estimadores M de localización y dispersión.

Bajo el modelo (2.18), $\hat{\mu}$ y $\hat{\sigma}$ se denominan estimadores M si satisfacen:

$$\sum_{i=0}^n \psi\left(\frac{x_i - \hat{\mu}}{\hat{\sigma}}\right) = 0$$

$$\frac{1}{n} \sum_{i=1}^n \rho_{dis}\left(\frac{x_i - \hat{\mu}}{\hat{\sigma}}\right) = \delta$$

donde $\psi(t) = \rho_0'(t)$, $\rho_{dis}(x) = x\psi(x)$, y $\delta = 1$

No es necesario buscar los estimadores de posición y dispersión al mismo tiempo, por ejemplo se podría calcular sólo el estimador de posición y tomar a $\hat{\sigma} = \text{MADN}(x)$

Hasta ahora se ha tratado de explicar, con cierto detalle, la teoría de la Estimación Robusta para el Caso Univariado; esto se ha hecho con la finalidad de facilitar al lector la comprensión de los aspectos básicos la Estimación Robusta, pues el propósito del presente trabajo es medir la calidad de algunos Estimadores Robustos Multivariados.

2.7 Caso Multivariado

Considere la estatura (cm) y el peso (kg) de veinte estudiantes, datos que se muestran a continuación

cm	kg	cm	kg	cm	kg	cm	kg
164	65	179	69	175	77	178	74
178	81	166	69	189	85	175	79
187	88	169	72	179	84	175	70
179	84	185	86	178	79	175	75
178	77	172	70	173	74	189	65

Un primer pensamiento puede llevar a verificar la existencia de valores extremos en cada variable de forma independiente, los gráficos Q-Q expuestos en el Cuadro 2.6 no presentan existencia de valores aberrantes, es más, los datos se ajustan a una Distribución Normal; sin embargo en el Gráfico de Dispersión para las dos variables estudiadas se observa un valor atípico, el mismo que corresponde al punto (189, 65).

Sea $\bar{\mathbf{X}}_{19} = (176.53, 76.73)$ el vector de medias correspondientes a los datos sin el valor atípico, y $\bar{\mathbf{X}}_{20} = (177.16, 76.14)$ el vector de medias correspondiente a los veinte datos, se puede ver que la diferencia entre ambos vectores es “pequeña”; sin embargo, considere las Matrices de Correlación $\mathbf{R}_{19} = \begin{pmatrix} 1 & \\ 0.833 & 1 \end{pmatrix}$ y $\mathbf{R}_{20} = \begin{pmatrix} 1 & \\ 0.549 & 1 \end{pmatrix}$ correspondientes a los datos sin valor atípico y a los datos

completos respectivamente, se puede notar que existe “gran” influencia de este valor atípico en la correlación entre las variables.

Cuadro 2. 7
Estimadores Robustos
Influencia de valores extremos en una distribución Bivariada

Gráfico Q-Q entre la distribución de la Estatura (cm) y el Peso (kg) Vs. La Distribución Normal

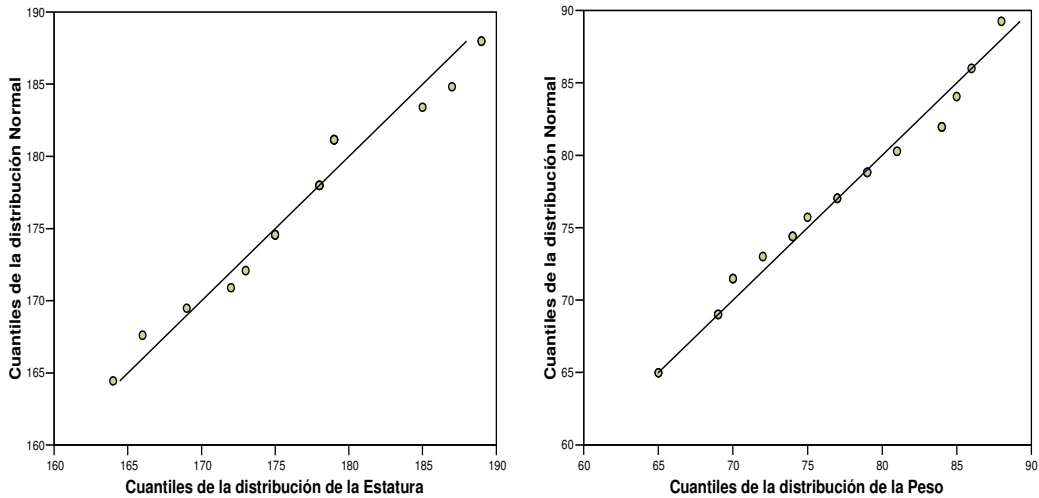
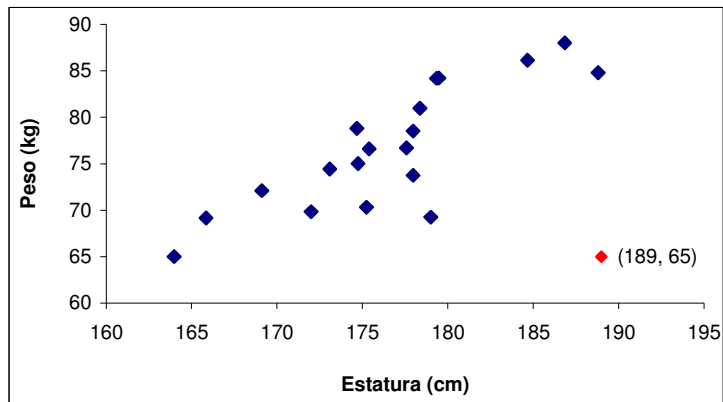


Gráfico de dispersión: Estatura Vs. Peso



Elaborado por: Néstor Montaña P.

La Matriz de Varianzas y Covarianzas y su correspondiente Matriz de Correlación son el punto inicial de muchos de los análisis

multivariados como Componentes Principales, Análisis de Correlación Canónica, etc. Así, este ejemplo permite entender la importancia de los estimadores robustos en muestras multivariadas

Análogo al caso univariado, el Estimador Robusto Multivariado debe cumplir las propiedades de Eficiencia y Estabilidad; además en procedimientos como Análisis de Discriminante, Correlación Canónica, etc. se trabaja con transformaciones lineales de las variables, entonces se desea que los estimadores cumplan propiedad de equivarianza, esto es:

Sea $\mathbf{y} = \mathbf{Ax} + \mathbf{b}$ entonces

$$\hat{\boldsymbol{\mu}}(Y) = \mathbf{A}\hat{\boldsymbol{\mu}}(X) + \mathbf{b}$$

$$\hat{\boldsymbol{\Sigma}}(Y) = \mathbf{A}\hat{\boldsymbol{\Sigma}}(X)\mathbf{A}^T$$

2.7.1 Eficiencia de los Estimadores Robustos Multivariados

El concepto y las condiciones son las mismas que en el caso univariado considerando como modelo paramétrico central a la Normal p-variada, es decir

Sea $(\hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n)$ los estimadores de localización y dispersión para una muestra de tamaño n y sea $(\hat{\boldsymbol{\mu}}_\infty, \hat{\boldsymbol{\Sigma}}_\infty)$ sus valores asintóticos, entonces $(\hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n)$ son consistentes si se cumple que

Si $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ entonces $\hat{\boldsymbol{\mu}}_\infty = \boldsymbol{\mu}$ y $\hat{\boldsymbol{\Sigma}}_\infty = c\boldsymbol{\Sigma}$ donde c es una constante

Además, los estimadores $\hat{\boldsymbol{\mu}}_n, \hat{\boldsymbol{\Sigma}}_n$ deben ser asintóticamente normales

$$\sqrt{n}(\hat{\boldsymbol{\mu}}_n - \hat{\boldsymbol{\mu}}_\infty) \xrightarrow{L} N_p(\mathbf{0}, \mathbf{V}_\mu)$$

$$\sqrt{n}(\text{vec}(\hat{\boldsymbol{\Sigma}}_n - \hat{\boldsymbol{\Sigma}}_\infty)) \xrightarrow{L} N_q(\mathbf{0}, \mathbf{V}_\Sigma)$$

donde $q = \frac{p(p+1)}{2}$ y $\text{vech}(\boldsymbol{\Sigma})$ es el vector que contiene los q

elementos de la triangular inferior de $\boldsymbol{\Sigma}$

Un aspecto a considerar es el siguiente:

Suponga que $\boldsymbol{\Sigma} = \text{Var}(\mathbf{x})$ y \mathbf{a} es un vector constante, entonces

$\text{Var}(\mathbf{a}^T \mathbf{x}) = \mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}$. Así, si $\boldsymbol{\Sigma}$ es cualquier matriz de dispersión

“robusta”, entonces $\sqrt{\mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}}$ puede ser considerada como una

Medida Robusta de Dispersión de la combinación lineal $\mathbf{a}^T \mathbf{x}$. Sea

$\lambda_1(\boldsymbol{\Sigma}) \geq \dots \geq \lambda_p(\boldsymbol{\Sigma})$ los Valores Propios de $\boldsymbol{\Sigma}$ en orden descendiente y

$\mathbf{e}_1 \dots \mathbf{e}_p$ los correspondientes Vectores Propios, se puede demostrar

que, para cualquier matriz simétrica $\boldsymbol{\Sigma}$, el mínimo y máximo de

$\mathbf{a}^T \boldsymbol{\Sigma} \mathbf{a}$ con $\|\mathbf{a}\|=1$ es equivalente a λ_p y λ_1 respectivamente y con

$\mathbf{a} = \mathbf{e}_p$ y $\mathbf{a} = \mathbf{e}_1$ correspondientemente. Al ser importante las

relaciones lineales entre variables, se debe considerar que no sólo es peligroso que el mayor Valor Propio se vuelva muy grande sino también el menor Valor Propio se vuelva más pequeño. El primer caso ocurre cuando existen valores extremos, muy alejados de la concentración de datos, y el segundo caso es provocado por observaciones concentradas en el mismo punto.

Definición 2.17 Punto de ruptura Asintótico

Con el entorno de contaminación \mathcal{F}_ε en (2.1), donde F_θ y H son Distribuciones Multivariadas, se define el Punto de Ruptura Asintótico de $(\hat{\mu}, \hat{\Sigma})$ como el mayor valor $\varepsilon^* \in (0,1)$ para el cual existen valores finitos positivos a, b, c tales que, para todo G :

$$\|\hat{\mu}_\infty((1-\varepsilon)F + \varepsilon G)\| \leq a$$

$$b \leq \lambda_p(\hat{\Sigma}((1-\varepsilon)F + \varepsilon G)) \leq \lambda_1(\hat{\Sigma}((1-\varepsilon)F + \varepsilon G)) \leq c$$

Es decir, el Punto de Ruptura Asintótico representa la mayor Contaminación; tal que, la norma del Estimador de Posición sea menor o igual a a y los valores propios correspondientes al Estimador de Dispersión estén en el intervalo $[b, c]$.

En general, las ideas y definiciones dadas en 2.3 se mantienen en el contexto multivariado por lo que no se continuará con definiciones

relativas a medir la robustez de un Estimador, se procede a definir los estimadores M para el caso multivariado.

2.7.2 Estimadores M para muestras multivariadas

En el caso univariado es posible definir o hallar por separado Estimadores robustos de Posición y Dispersión que tengan la propiedad de equivarianzas. Esto es más complicado de realizar en el caso multivariado por lo que es mejor estimar Posición y Dispersión simultáneamente. Considere la Función de Densidad Multivariada,

$$f(\mathbf{x}, \mu, \Sigma) = \frac{1}{\sqrt{|\Sigma|}} h(d(\mathbf{x}, \mu, \Sigma)) \quad (2.19)$$

Nótese que escogiendo $h(s) = c \exp(-s / 2)$ con $c = (2\pi)^{-p/2}$ y

$$d(\mathbf{x}, \mu, \Sigma) = (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \quad (2.20)$$

se tiene la Función de Densidad Normal Multivariada, donde $d(\mathbf{x}, \mu, \Sigma)$ es la Distancia de Mahalanobis. De manera similar al procedimiento utilizado en los Estimadores M de Escala, se encontrará el Estimador de Máxima Verosimilitud para la Función de Densidad (2.19) para luego generalizarla a un Estimador M.

Sea $\mathbf{x}_1, \dots, \mathbf{x}_n$ una muestra i.i.d de una función f de la forma (2.19) donde h se asume positiva. Se procede a calcular el Estimador de

Máxima Verosimilitud de μ y Σ , nótese que la Función de Verosimilitud es:

$$L(\mu, \Sigma) = \frac{1}{|\Sigma|^{n/2}} \prod_{i=1}^n h(d(\mathbf{x}_i, \mu, \Sigma))$$

Aplicando logaritmo se obtiene:

$$\log(L(\mu, \Sigma)) = -\frac{n}{2} \log|\Sigma| + \sum_{i=1}^n \log(h(d(\mathbf{x}_i, \mu, \Sigma)))$$

$$-2\log(L(\mu, \Sigma)) = n \log|\Sigma| + \sum_{i=1}^n \rho(d(\mathbf{x}_i, \mu, \Sigma))$$

Donde $\rho(s) = -2\log(h(s))$

$$-\frac{2}{n} \log(L(\mu, \Sigma)) = \log|\Sigma| + \frac{1}{n} \sum_{i=1}^n \rho(d(\mathbf{x}_i, \mu, \Sigma)) \quad (2.21)$$

Derivando la expresión (2.21) con respecto a μ , para maximizarla, se tiene:

$$\frac{\delta}{\delta\mu} \left(\log|\Sigma| + \frac{1}{n} \sum_{i=1}^n \rho(d(\mathbf{x}_i, \mu, \Sigma)) \right) = 0$$

$$\frac{1}{n} \sum_{i=1}^n \rho'(d(\mathbf{x}_i, \mu, \Sigma)) \frac{\delta}{\delta\mu} d(\mathbf{x}_i, \mu, \Sigma) = 0$$

Se puede probar que: $\frac{\delta(b^T A b)}{\delta b} = (A + A^T)b$

$$\text{donde } \frac{\delta}{\delta\mu} d(\mathbf{x}_i, \mu, \Sigma) = \frac{\delta}{\delta\mu} (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = 2\Sigma^{-1}(\mathbf{x} - \mu)$$

$$2\Sigma^{-1} \frac{1}{n} \sum_{i=1}^n \rho'(d(\mathbf{x}_i, \mu, \Sigma)) (\mathbf{x} - \mu) = 0 \quad (2.22)$$

Derivando la expresión (2.21) con respecto a Σ^{-1}

$$\frac{\delta}{\delta \Sigma^{-1}} \left(-\log |\Sigma^{-1}| + \frac{1}{n} \sum_{i=1}^n \rho(d(\mathbf{x}_i, \mu, \Sigma)) \right) = 0$$

$$-\frac{\delta}{\delta \Sigma^{-1}} (\log |\Sigma^{-1}|) + \frac{1}{n} \sum_{i=1}^n \rho'(d(\mathbf{x}_i, \mu, \Sigma)) \frac{\delta}{\delta \Sigma^{-1}} d(\mathbf{x}_i, \mu, \Sigma) = 0$$

Se puede demostrar que $\frac{\delta \log |A|}{\delta A} = (A^T)^{-1}$ y que $\frac{\delta}{\delta A} b^T A b = b b^T$

Se tiene además que:

$$\frac{\delta}{\delta \Sigma^{-1}} d(\mathbf{x}_i, \mu, \Sigma) = \frac{\delta}{\delta \Sigma^{-1}} (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) = (\mathbf{x} - \mu)(\mathbf{x} - \mu)^T$$

Entonces,

$$-\Sigma + \frac{1}{n} \sum_{i=1}^n \rho'(d(\mathbf{x}_i, \mu, \Sigma)) (\mathbf{x} - \mu)(\mathbf{x} - \mu)^T = 0 \quad (2.23)$$

Así, (2.22) y (2.23) son equivalentes a:

$$\sum_{i=1}^n W(d_i) (\mathbf{x}_i - \hat{\mu}) = 0$$

$$\frac{1}{n} \sum_{i=1}^n W(d_i) (\mathbf{x}_i - \hat{\mu})(\mathbf{x}_i - \hat{\mu})^T = \hat{\Sigma}$$

con $W = \rho'(d(\mathbf{x}_i, \mu, \Sigma))$

En general, se define a los estimadores M como soluciones de:

$$\sum_{i=1}^n W_1(d_i) (\mathbf{x}_i - \hat{\mu}) = 0 \quad (2.24)$$

$$\frac{1}{n} \sum_{i=1}^n W_2(d_i) (\mathbf{x}_i - \hat{\mu})(\mathbf{x}_i - \hat{\mu})^T = \hat{\Sigma} \quad (2.25)$$

donde W_1 y W_2 no son necesariamente iguales. Nótese que en (2.25) se puede interpretar a $\hat{\Sigma}$ como una Matriz de Covarianza Ponderada y en (2.24) se puede expresar a $\hat{\mu}$ como una Media Ponderada.

$$\hat{\mu} = \frac{\sum_{i=1}^n W_1(d_i) \mathbf{x}_i}{\sum_{i=1}^n W_1(d_i)}$$

donde las ponderaciones dependen de la Distancia de Mahalanobis (2.20).

La unicidad de las soluciones de (2.24) y (2.25) requieren que $W_2(d)$ sea una función no decreciente de d .

Se denomina un Estimador M de Localización y Dispersión Monótono si $W_2(d)$ es no decreciente y Redescendiente si no lo es. Los Estimadores M son definidos como soluciones para las ecuaciones (2.24) y (2.25) , mientras que los Estimadores M Redescendientes deben ser definidos para la minimización de alguna función objetivo como los Estimadores que se definen a continuación:

2.7.3 Estimadores basados en una escala robusta

El objetivo de los Estimadores robustos de Escala es definir Estimadores Multivariados de Posición y Dispersión que hagan “pequeñas” las distancias d_i . Para este fin se minimiza algunas distancias, lo que puede implicar que el menor valor propio de $\hat{\Sigma}$ tienda a cero; para prevenir que esto suceda, se define $|\hat{\Sigma}|=1$. Se denomina S_p al conjunto de matrices simétricas definidas positivas $p \times p$. Para un conjunto de datos X , sea $d(X, \hat{\mu}, \hat{\Sigma})$ el vector con elementos $d(x_i, \hat{\mu}, \hat{\Sigma}); i=1, \dots, n$, y sea $\hat{\sigma}$ un Estimador de Escala Robusto. Entonces se define los estimadores $\hat{\mu}$ y $\hat{\Sigma}$ como:

$$\hat{\sigma}(d(X, \hat{\mu}, \hat{\Sigma})) = \min ; \text{ donde } \hat{\mu} \in R^p, \hat{\Sigma} \in S_p, |\hat{\Sigma}|=1 \quad (2.26)$$

Se puede mostrar que los estimadores definidos por (2.26) son equivariantes.

Definición 2.18 Estimadores S

Un ejemplo de Estimadores Robustos de Escala son los denominados Estimadores S, los mismos que cumplen con (2.26) tomando $\hat{\sigma}$ como un Estimador M de Escala que satisface

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{d_i}{\hat{\sigma}}\right) = \delta \quad (2.27)$$

donde ρ es una función acotada.

Si ρ es diferenciable, puede mostrarse que la solución para (2.26) debe satisfacer las ecuaciones (2.24) y (2.25) de los Estimadores M; es decir,

$$\sum_{i=1}^n W\left(\frac{d_i}{\hat{\sigma}}\right)(\mathbf{x}_i - \hat{\mu}) = 0 \quad (2.28)$$

$$\frac{1}{n} \sum_{i=1}^n W\left(\frac{d_i}{\hat{\sigma}}\right)(\mathbf{x}_i - \hat{\mu})(\mathbf{x}_i - \hat{\mu})^T = c \hat{\Sigma} \quad (2.29)$$

donde

$$W = \rho' \text{ y } \hat{\sigma} = \hat{\sigma}(d_1, \dots, d_n)$$

y c es un escalar tal que $|\hat{\Sigma}| = 1$.

Esto claramente muestra a los Estimadores S como una particularización de los Estimadores M.

Nótese que si ρ es acotada (como es lo usual), $dW(d)$ no puede ser monótona; en realidad, para este tipo de estimadores, la Función de Ponderaciones $W(d)$ tiende a cero para distancia “grandes”. Así, un Estimador S no es un Estimador M monótono, y por lo tanto las ecuaciones (2.28) y (2.29) brindan sólo un mínimo local de $\hat{\sigma}$

A continuación se presentan dos ejemplos de Estimadores S a ser utilizados en el presente trabajo.

Definición 2.19 Estimador S Bicuadrático Multivariado

Se define al Estimador S Bicuadrático o Bponderado multivariado como un estimador con escala dada por (2.27) con

$$\rho(t) = \min\{1, 1 - (1-t)^3\}$$

Que tiene la siguiente Función de Ponderaciones

$$W(t) = 3(1-t)^2 I(t \leq 1)$$

Definición 2.20 Estimador S T-Bicuadrático

Mientras mayor sea el número de variables, los estimadores S con función de ponderaciones continua se aproximan al vector de medias muestral y la matriz de varianzas y covarianzas muestral, esto implica una pérdida de robustez; Rocke en [26] considera este problema y propone un tipo de estimador con función ρ dependiente del número de variables, en particular propuso una familia de estimadores cuya función ρ cumple que

$$\lim_{p \rightarrow \infty} \rho(d) = I(d > 1)$$

Se denomina $\chi_{p,\alpha}^2$ al α -cuantil de χ_p^2 . Para valores p grandes la distribución χ_p^2 es aproximadamente simétrica, con $\chi_{p,0.5}^2$ y

$$\chi_{p,(1-\alpha)-p}^2 \approx p - \chi_{p,\alpha}^2$$

Un estimador se denomina Estimador S de Roche si cumple (2.26) y (2.27) con

$$\rho(t) = \begin{cases} 0 & \text{para } 0 \leq t \leq 1 - \gamma \\ \left(\frac{t-1}{4\gamma}\right) \left[3 - \left(\frac{t-1}{\gamma}\right)^2 \right] + \frac{1}{2}; & \text{para } 1 - \gamma < t < 1 + \gamma \\ 1 & \text{para } t \geq 1 + \gamma \end{cases}$$

que tiene como derivada la Función de Ponderación

$$W(t) = \frac{3}{4\gamma} \left[1 - \left(\frac{t-1}{\gamma}\right)^2 \right] \mathbf{I}(1 - \gamma \leq t \leq 1 + \gamma)$$

$$\text{donde } \gamma = \min \left(\frac{\chi_{p, (1-\alpha)}^2}{p} - 1, 1 \right).$$

Definición 2.21 Estimador Covarianza de Mínimo Determinante MCD

Zuo en [32] indica que los estimadores MCD propuestos por Rousseeuw escogen h observaciones las cuales minimizan el determinante de la Matriz de Covarianzas Clásica, así, los estimador MCD de localización es el promedio de las h observaciones y el estimador MCD de escala es un múltiplo escalar de su Matriz de Covarianzas correspondiente a las h observaciones.

Este estimador es talvez el más popular debido (en parte) al “rápido” cálculo de los estimadores por parte de varios algoritmos desarrollados.

Definición 2.22 Estimador Stahel-Donoho (DS)

La idea del estimador Stahel-Donoho, propuesto por Stahel (1981) y Donoho (1982) es que un “valor aberrante” multivariado debe serlo también en alguna proyección univariada, de ahí que el estimador DS, es un vector de Media y Matriz de Covarianzas ambos ponderados de tal manera que la ponderación de \mathbf{x}_i es una función de la “lejanía” de \mathbf{x}_i , denotada por $t(\mathbf{x}_i)$.

Es decir, sean W_1 y W_2 dos funciones de ponderaciones y se define

$$\hat{\mu} = \frac{1}{\sum_{i=1}^n w_{i1}} \sum_{i=1}^n w_{i1} \mathbf{x}_i \tag{2.30}$$

$$\hat{\Sigma} = \frac{1}{\sum_{i=1}^n w_{i2}} \sum_{i=1}^n w_{i2} (\mathbf{x}_i - \hat{\mu})(\mathbf{x}_i - \hat{\mu})^T \tag{2.31}$$

con

$$w_{ij} = W_j(t(\mathbf{x}_i)), \quad j = 1, 2. \tag{2.32}$$

Donde $t(\mathbf{x}) = \max_{\mathbf{a}} t(\mathbf{x}, \mathbf{a})$, dada una dirección $\mathbf{a} \in R^p$ con $\|\mathbf{a}\| = 1$;

$\mathbf{a}^T \mathbf{X} = \{\mathbf{a}^T \mathbf{x}_1, \dots, \mathbf{a}^T \mathbf{x}_n\}$ representa la proyección del conjunto de datos

\mathbf{X} a lo largo de \mathbf{a} . Sea $\hat{\mu}$ y $\hat{\sigma}$ estadísticos univariados robustos de localización y dispersión; por ejemplo, la mediana y la Mediana Absoluta de las Desviaciones respectivamente. La “lejanía” con respecto a \mathbf{X} de un punto $\mathbf{x} \in R^p$ a lo largo de \mathbf{a} es definido por

$$t(\mathbf{x}, \mathbf{a}) = \frac{\mathbf{x}^T \mathbf{a} - \hat{\mu}(\mathbf{a}^T \mathbf{x})}{\hat{\sigma}(\mathbf{a}^T \mathbf{x})}$$

La lejanía de \mathbf{x} es definido por

$$t(\mathbf{x}) = \max_{\mathbf{a}} t(\mathbf{x}, \mathbf{a}) \quad (2.33)$$

2.7.4 Algoritmos para calcular Estimadores Robustos Multivariados

Estimadores M Monótonos

Se inicia con los estimadores iniciales $\hat{\mu}_0$ y $\hat{\Sigma}_0$; por decir, el vector de coordenadas de la mediana y la matriz diagonal con la Mediana Absoluta de las Desviaciones Cuadráticas Normalizadas de las variables en la diagonal. En la iteración k se define

$d_{ki} = d(\mathbf{x}_i, \hat{\mu}_k, \hat{\Sigma}_k)$ y se calcula

$$\hat{\mu}_{k+1} = \frac{\sum_{i=1}^n W_1(d_{ki}) \mathbf{x}_i}{\sum_{i=1}^n W_1(d_{ki})}, \quad \hat{\Sigma}_{k+1} = \frac{1}{n} \sum_{i=1}^n W_2(d_{ki}) (\mathbf{x}_i - \hat{\mu}_{k+1})(\mathbf{x}_i - \hat{\mu}_{k+1})^T \quad (2.34)$$

Como la solución es única para los Estimadores M Monótonos, el punto de inicio influye sólo en el número de iteraciones pero no en el resultado final.

Soluciones locales para los Estimadores S

Como el Mínimo Local de $\hat{\sigma}$ son soluciones de las ecuaciones (2.28) y (2.29), un procedimiento natural para minimizar $\hat{\sigma}$ es usar el procedimiento iterativo (2.34) para resolver ecuaciones, con $W_1 = W_2 = W$ donde $W = \rho'$. Además siendo $tW(t)$ redescendiente, este par de ecuaciones serían sólo un Mínimo Local de σ , por lo tanto los valores iniciales son esenciales. Se asume por el momento que se tiene los valores iniciales $\hat{\mu}_0$ y $\hat{\Sigma}_0$.

En la iteración k se calcula:

$$d_{ki} = d(\mathbf{x}_i, \hat{\mu}_k, \hat{\Sigma}_k), \hat{\sigma}_k = \hat{\sigma}(d_{k1}, \dots, d_{kn}), \omega_{ki} = W \left(\frac{d_{ki}}{\hat{\sigma}_k} \right) \quad (2.35)$$

Entonces se calcula

$$\hat{\mu}_{k+1} = \frac{\sum_{i=1}^n \omega_{ki} \mathbf{x}_i}{\sum_{i=1}^n \omega_{ki}}, \hat{C}_k = \sum_{i=1}^n \omega_{ki} (\mathbf{x}_i - \hat{\mu}_{k+1})(\mathbf{x}_i - \hat{\mu}_{k+1})^T, \hat{\Sigma}_{k+1} = \frac{\hat{C}_k}{|\hat{C}_k|^{1/p}} \quad (2.36)$$

Se puede detener la iteración cuando el cambio relativo $(\hat{\sigma}_k - \hat{\sigma}_{k+1}) / \hat{\sigma}_k$ es menor a una tolerancia dada. Experiencias

muestran que como el decrecimiento de $\hat{\sigma}_k$ es generalmente lento, no es necesario recalcularlo en cada paso.

Si W no es monótono, los pasos iterativos (2.35) y (2.36) no son garantizados para generar un decrecimiento de $\hat{\sigma}_k$ en cada paso, sin embargo, el algoritmo puede ser modificado para asegurar un decrecimiento en cada iteración, situación que no será tratada aquí.

CAPÍTULO 3

3 Método Monte Carlo

3.1 Introducción

En este Capítulo se exponen ciertos aspectos existentes detrás del Método Monte Carlo, particularmente cuando se asume independencia entre las variables aleatorias generadas.

Se presenta el Método de Transformación Inversa que permite generar muestras aleatorias a partir de Distribuciones específicas pero que no es aplicable para el caso de la Distribución Normal, por lo que se expone un algoritmo para la generación de Normales aleatorias univariadas y p-variadas. Por último, se trata de manera somera la generación de números pseudoaleatorios.

3.2 ¿Porqué Simular?

En el capítulo anterior se presenta varias formas de cuantificar la calidad del estimador robusto, sin embargo estas medidas no son fáciles de obtener analíticamente, por ello se explora otro campo "reciente" de la Estadística: la Simulación, particularmente el conjunto de técnicas denominadas Método Monte Carlo, cuyo algoritmo Metropolis fue considerado como el más destacado del siglo veinte, según [4].

Como se menciona en el primer capítulo, el objetivo del presente trabajo es evaluar el comportamiento de varios Estimadores Robustos para el vector de medias y la matriz de varianzas y covarianzas de un vector p variado \mathbf{X} , para ello se generan muestras aleatorias con Función de Distribución conocida a partir de las cuales se calculan los estimadores a analizar. Para observar el comportamiento de los estimadores bajo diferentes condiciones, se cambiarán varios de los parámetros de las Distribuciones utilizadas para generar las muestras. Para esto en el presente capítulo se desarrollan ciertos aspectos referentes al Método Monte Carlo.

Según Jimenez en [17], entre 1940 y 1950 aproximadamente en el Laboratorio de los Alamos se planteó un problema de difícil solución.

Se trataba de determinar el recorrido de los neutrones con los diferentes medios. Una solución que recurriese a los procedimientos “clásicos” resultaría laboriosa y complicada. Los técnicos conocían los datos fundamentales que se necesitaban para resolver el problema; sin embargo, la dificultad surgía al tratar de relacionar los datos en una sola fórmula. Ulam y Neumann idearon una solución que esencialmente consistía en que una ruleta resolviera el problema propuesto. Se fueron agrupando las probabilidades de los distintos sucesos, obteniéndose una solución que quedaba dentro de la aproximación exigida por los técnicos.

El Método se lo denominó Monte Carlo, que es utilizado para aproximar valores de integrales, sumatorias y minimización de funciones. Según Fishman en [6] el algoritmo de cada técnica tipo Monte Carlo se puede resumir en:

- i. Generar una muestra aleatoria
- ii. Convertir los datos muestrales a un estimador de la solución del problema estudiado
- iii. Evaluar el error de aproximación

3.3 Método Monte Carlo

Formalizando lo mencionado:

Sea $\mathbf{X}^T = (x_1, x_2, \dots, x_m)$ un vector perteneciente a \mathbf{R}^m , sea $g: \mathbf{R}^m \rightarrow \mathbf{R}$ una función y sea $\mathcal{X} \subseteq \mathbf{R}^m$ una región para la cual la integral

$$\lambda = \int_{\mathcal{X}} g(\mathbf{x}) d\mathbf{x} \quad (3.1)$$

existe; cuando \mathcal{X} es un conjunto de puntos discretos, entonces

$$\lambda = \sum_{\mathbf{x} \in \mathcal{X}} g(\mathbf{x})$$

El propósito del Método Monte Carlo es estimar λ , para ello selecciona un número finito de puntos pertenecientes a \mathcal{X} y los utiliza para hallar una aproximación de λ en (3.1).

Ahora, supóngase que

$$\int_{\mathcal{X}} g^2(\mathbf{x}) d\mathbf{x} < \infty$$

y sea $\omega = \left\{ \omega(\mathbf{x}) \geq 0: \mathbf{x} \in \mathcal{X}; \int_{\mathcal{X}} \omega(\mathbf{z}) d\mathbf{z} = 1 \right\}$ una Función de Probabilidad

con la siguiente propiedad

$$\int_{\mathcal{X}} \frac{g^2(\mathbf{x})}{\omega(\mathbf{x})} d\mathbf{x} < \infty \quad (3.2)$$

Para un vector aleatorio m -dimensional $\mathbf{X}^T = (x_1, x_2, \dots, x_m)$ con

Función de Probabilidad ω , el ratio $\frac{g(\mathbf{X})}{\omega(\mathbf{X})}$ tiene momentos

$$E\left[\left(\frac{g(\mathbf{X})}{\omega(\mathbf{X})}\right)^t\right] = \int_{\mathcal{X}} \left(\frac{g(\mathbf{x})}{\omega(\mathbf{x})}\right)^t \omega(\mathbf{x}) d\mathbf{x} ; \quad t = 1, 2, \dots$$

en particular, $\frac{g(\mathbf{X})}{\omega(\mathbf{X})}$ tiene media

$$E\left[\frac{g(\mathbf{X})}{\omega(\mathbf{X})}\right] = \lambda \tag{3.3}$$

y varianza

$$\text{var}_{\omega}\left[\frac{g(\mathbf{X})}{\omega(\mathbf{X})}\right] = \sigma^2 = \int_{\mathcal{X}} \left[\frac{g(\mathbf{x})}{\omega(\mathbf{x})}\right]^2 \omega(\mathbf{x}) d\mathbf{x} - \lambda^2 \tag{3.4}$$

3.3.1 Monte Carlo bajo hipótesis de Independencia

A continuación se presenta el algoritmo que describe el Método Monte Carlo en una forma muy simple que genera n vectores aleatorios independientes pertenecientes a \mathbb{R}^m , a partir de los cuales se calcula un promedio $\bar{\lambda}_n$ como una aproximación de λ .

Algoritmo IMC

Ingresar n (número de vectores aleatorios)

$i=1$

While $i \leq n$;

generar aleatoriamente $\mathbf{X}^{(i)} = (X_1^{(i)}, \dots, X_m^{(i)})$ a partir ω

evaluar $\frac{g(\mathbf{X}^{(i)})}{\omega(\mathbf{X}^{(i)})}$

$i = i + 1$

Calcular:

$$\bar{\lambda}_n = \frac{1}{n} \sum_{i=1}^n \frac{g(\mathbf{X}^{(i)})}{\omega(\mathbf{X}^{(i)})} \quad (3.5)$$

Calcular s_n^2 como una aproximación de σ^2

$$\begin{aligned} s_n^2 &= \frac{1}{n-1} \left\{ \sum_{i=1}^n \left(\frac{g(\mathbf{X}^{(i)})}{\omega(\mathbf{X}^{(i)})} - \bar{\lambda}_n \right)^2 \right\} \\ &= \frac{1}{n-1} \left\{ \sum_{i=1}^n \left(\frac{g(\mathbf{X}^{(i)})}{\omega(\mathbf{X}^{(i)})} \right)^2 - n\bar{\lambda}_n^2 \right\} \end{aligned} \quad (3.6)$$

□

Así planteado, $\bar{\lambda}_n$ en (3.5) cumple las siguientes propiedades.

i. Por la Ley de los Grandes Números se tiene que,

$$\bar{\lambda}_n \xrightarrow{P} \lambda$$

- ii. Como los datos son independientes e idénticamente distribuidos, entonces

$$\text{var}_{\omega}(\bar{\lambda}_n) = \frac{\sigma^2}{n}$$

En vista de que n se incrementa, esta propiedad muestra que la varianza del estimador decrece a razón de n^{-1} .

- iii. Por el Teorema del Limite Central se tiene que,

$$\frac{\bar{\lambda}_n - \lambda}{\sqrt{\sigma^2 / n}} \xrightarrow{D} N(0,1)$$

Sin embargo σ es desconocido, por lo que es necesario estimarlo a través de S_n^2 , así, si la condición

$$\int_{\mathcal{X}} \frac{g^4(\mathbf{x})}{\omega^3(\mathbf{x})} d\mathbf{x} < \infty$$

se cumple, entonces:

iv. $S_n^2 \xrightarrow{P} \sigma^2$

v. $\frac{\bar{\lambda}_n - \lambda}{\sqrt{\frac{S_n^2}{n}}} \xrightarrow{D} N(0,1)$

El estimador $\bar{\lambda}_n$ tiene Error Estándar (EE) igual a $\frac{\sigma}{n^{\frac{1}{2}}}$, el mismo que

es aproximado por

$$\text{EE} = \frac{S_n}{n^{\frac{1}{2}}},$$

y Error Relativo (ER) igual a $\frac{\sigma}{\lambda n^{\frac{1}{2}}}$, aproximado por

$$\text{ER} = \frac{s_n}{\bar{\lambda}_n n^{\frac{1}{2}}}$$

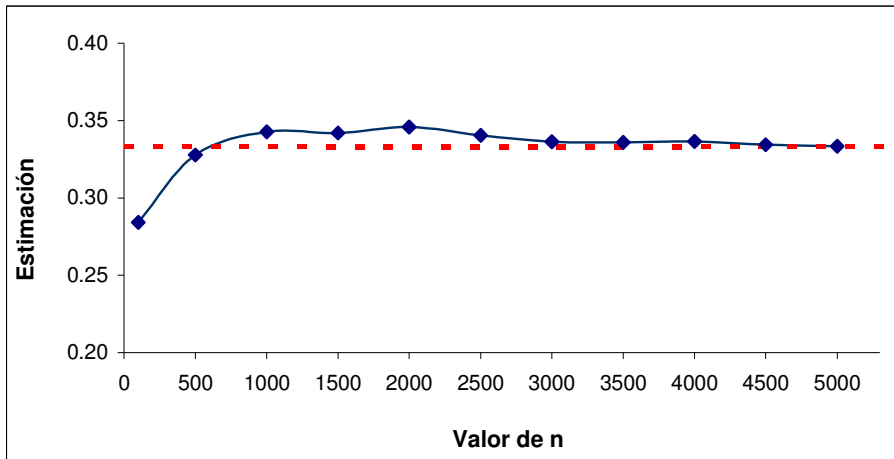
Para ilustrar el Método Monte Carlo se aproxima el valor de una integral, la cual es sencillamente evaluada analíticamente

Sea $\lambda = \int_0^1 g(x)dx$ donde $g(x) = x^2$. Se escoge ω como la Distribución Uniforme con parámetros 0 y 1, es decir

$$\omega(x) = \begin{cases} 1; & 0 \leq x \leq 1 \\ 0; & \text{resto de } x \end{cases}$$

Luego, se genera 100 números aleatorios con Función de Distribución $\omega(x)$, a partir de los cuales se evalúa (3.5) obteniendo $\bar{\lambda}_{100} = 0.284$. Para ilustrar la convergencia del Método Monte Carlo se repite el experimento para diferentes valores de n, el Cuadro 3.1 muestra la estimación resultante para valores de n=500,1000,...5000; se puede ver que n=5000 la estimación presenta un error menor a 0.001 pues $\bar{\lambda}_{5000} = 0.3334$ mientras que $\lambda = 0.3333$.

Cuadro 3. 1
 Estimadores Robustos
Convergencia del Método Monte Carlo en función de n



Elaborado por: Néstor Montaña P.

3.3.2 Generación de Variables Aleatorias No Uniformes

Nótese que en el ejemplo anterior se eligió como $\omega(x)$ una Distribución Uniforme pero podría ser cualquier otra Distribución; así se han desarrollado diversos métodos para generar muestras aleatorias de cualquier Distribución. A continuación se presenta el método de la Transformación Inversa, que es el método mas conocido.

Método de Transformación Inversa

Este método utiliza la Distribución Acumulada $F(x)$ de la Distribución que se utiliza para generar muestras aleatorias. Puesto

que $F(x)$ está definida en el intervalo $(0,1)$, se puede generar un número aleatorio uniforme R y tratar de determinar el valor de la variable aleatoria para la cual su Distribución Acumulada es igual a R , lo cual lleva a resolver la siguiente ecuación:

$$F(x) = R \qquad x = F^{-1}(R)$$

Así por ejemplo, la aplicación del Método de Transformada Inversa para la Distribución Exponencial se presenta de la siguiente manera:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$$

$$1 - e^{-\lambda x} = R$$

$$e^{-\lambda x} = 1 - R$$

$$e^{-\lambda x} = R$$

$$x = -\frac{1}{\lambda} \ln R$$

□

En el caso particular de la Distribución Normal, $F(z)$ no se puede hallar analíticamente, entonces, para generar muestras aleatorias a partir de una Distribución Normal no se podría utilizar el Método de Transformación Inversa, por lo que se desarrollan métodos alternativos como el siguiente.

Generación de muestras aleatorias a partir de una Distribución Normal

Del Teorema del Límite Central se puede decir que la suma de n variables aleatorias independientes se aproxima a una Distribución Normal a medida que n se aproxima al infinito.

Sea u_1, u_2, \dots, u_n es una sucesión de n variables aleatorias $U(0;1)$

entonces $E(u_i) = \frac{1}{2}$ y $\text{var}(u_i) = \frac{1}{12}$ entonces

$$Z = \frac{\sum_{i=1}^n u_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$$

Tiene una Distribución Normal Estándar a medida que n se aproxima al infinito.

Entonces, para generar muestras aleatorias se sigue el siguiente sería:

- i. Ingresar μ y σ
- ii. Generar n números aleatorios $U(0;1)$
- iii. Calcular

$$x = \mu + \sigma \left(\frac{\sum_{i=1}^n u_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \right) \quad (3.7)$$

Se ha comprobado que utilizando $n = 12$ la confiabilidad de los valores generados es aceptable. La ecuación (3.7) queda de la siguiente manera:

$$x = \mu + \sigma \left(\sum_{i=1}^n u_i - 6 \right)$$

En el presente trabajo se necesita simular muestras a partir de una Distribución Normal Multivariada, para ello utilizaremos el siguiente algoritmo que utiliza la descomposición de Cholesky

Generación de una Muestra Normal p-variada

Sea $\mathbf{z} \sim N_p(\mathbf{0}, \mathbf{I}_p)$, sea $\mathbf{y} = \mathbf{z}\mathbf{A} + \boldsymbol{\mu}$ entonces $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma} = \mathbf{A}^T \mathbf{A})$. Así, para generar vectores aleatorios que sigan una distribución Normal p-variada $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ se sigue el siguiente algoritmo

Ingresar $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$

Obtener la descomposición de Cholesky de $\boldsymbol{\Sigma}$, esto es

$$\boldsymbol{\Sigma} = \mathbf{A}^T \mathbf{A}$$

Generar $\mathbf{Z}_{n \times p}$ donde $z_{ij} \sim N(0,1)$

Transformar \mathbf{Z} a \mathbf{Y} usando la expresión $\mathbf{Y} = \mathbf{Z}\mathbf{A} + \mathbf{U}$ donde \mathbf{U} es la matriz resultante de repetir el vector $\boldsymbol{\mu}^T$ n veces.

□

El resultado final es una matriz $\mathbf{Y}_{n \times p}$ cuyas filas representan los vectores aleatorios generados.

Se puede observar que la generación de una muestra aleatoria a partir de cualquier Distribución depende trascendentalmente de la generación de una secuencia de números aleatorios independientes idénticamente distribuidos a partir de una $U(0;1)$; sin embargo ninguna técnica tipo Monte Carlo puede literalmente generar números aleatorios por lo que se utilizan los llamados Generadores de Números Pseudoaleatorios

3.3.3 Generación de números Pseudoaleatorios

Se presenta el algoritmo mas conocido para la Generación de números Pseudosaleatorios

Generador Congruencial

Se define

$$x_{n+1} = ax_n \text{ mod } m$$

donde:

$$x_0, a, c > 0 \text{ y } m > x_0, m > a \text{ y } m > c$$

x_0 : representa la semilla; $x_0 > 0$

a : es un número entero positivo llamado multiplicador

m : es el módulo y representa una cota superior entera para $x_i; i = 1, 2, \dots$

Así, la secuencia de fracciones $u_i = \frac{x_i}{m}$ es usada para aproximar una secuencia de variables aleatorias independientes e idénticamente distribuidas a partir de una $U(0;1)$.

Existen otros generadores de números pseudoaleatorios, según Fishman en [6] el generador que presenta mejor rendimiento es el Mersenne Twister desarrollado por Matsumoto y Nishimura en [20] el mismo que se utiliza en el desarrollo del presente trabajo.

CAPÍTULO 4

4 Simulación

4.1 Introducción

En el capítulo final del presente estudio se compara el comportamiento de algunos de los Estimadores Robustos para el Vector de Medias y Matriz de Varianzas y Covarianzas definidos en el segundo capítulo .

Se explican los escenarios considerados en la simulación, luego de lo cual se presentan un resumen de los resultados de la simulación para cada uno de los estimadores que conforman el presente estudio; efectuando cuadros comparativos y eligiendo el mejor de los estimadores en cada escenario.

4.2 Generalidades

En el segundo capítulo se definen algunos de los Estimadores Robustos para el Vector de Medias y Matriz de Varianzas y Covarianzas más conocidos, además se presentan varias maneras de medir la Robustez de un Estimador como: Sesgo Asintótico Máximo (2.2), Varianza Asintótica Máxima (2.3), etc; si bien, estas medidas analíticas nos permiten realizar comparaciones entre los Estimadores Robustos, es necesario recordar que las mismas son definidas Asintóticamente, quedando sin explicar el comportamiento de dichos Estimadores en Muestras Finitas, ello es lo que se explora en el presente capítulo utilizando Simulación Matemática.

En este tipo de estudio, se generan muestras aleatorias a partir de distribuciones en las que se manipula una serie de parámetros con el objetivo de simular varias condiciones y así estudiar el comportamiento de los Estimadores Robustos bajo estas condiciones; tal vez el primer estudio comparativo de estimadores robustos fue el realizado por Andrews et al. en [1], en dicho estudio se dio a lugar a la noción de las tres "esquinas" para representar las posibles situaciones que se puede encontrar en la práctica, se propuso probar los Estimadores Robustos bajo un enfoque muy optimista (Generando una Distribución Normal), muy pesimista (a

través de una *t* de Student con un grado de libertad, llamada Cauchy) y la última prueba consistía en 20 datos, 19 de los cuales eran generados a partir de una Normal Estándar y el último generado a partir de una $N(0,100)$. Los dos primeros casos son el punto de partida del presente estudio, es decir, se consideraran muestras generadas a partir de la Distribución Normal Multivariada $N(\mathbf{0}, \Sigma)$ y la Distribución Cauchy $CAU(\mathbf{0}, \Sigma)$, cuyas diferencia, recordando el segundo capítulo, es el “peso” de sus colas.

Las muestras generadas serán contaminadas, para ello se considera el Modelo de Contaminación (2.1) definido por $\mathcal{F}_\varepsilon = \{(1 - \varepsilon)F_\theta + \varepsilon H; \theta \in \Theta\}$, donde $\varepsilon = 0; 0.05; 0.1$ y 0.3 de tal manera que se observa el comportamiento de los estimadores analizados bajo condiciones de:

- No Contaminación $\varepsilon = 0$ caso que permitirá compararlos con los Estimadores de Máxima Verosimilitud,
- Contaminación Moderada $\varepsilon = 0.05$ y 0.1 que parece ser lo más cercano a la realidad y
- Contaminación Extrema $\varepsilon = 0.3$ cuyos resultados se pueden utilizar para verificar si el Punto de Ruptura de los Estimadores es menor a 0.3.

Además, para la Distribución H que genera la contaminación en el modelo (2.1) se ha escogido tres Distribuciones:

- Distribución Normal Multivariada $N(\mathbf{0}, 9\Sigma)$, que sigue la idea propuesta por Tukey en [27] y que es la Distribución frecuentemente utilizada en este tipo de estudios.
- Distribución Normal Multivariada $N(0.537\alpha_p, 9\Sigma)$, es decir una Contaminación Asimétrica donde α_p representa el vector propio correspondiente al menor valor propio asociado a la Matriz de Varianzas y Covarianzas; este tipo de contaminación es utilizada por Devlin en [5]; y
- Distribución Uniforme Esférica $U_{esf}(d)$, donde d representa la distancia hacia el origen; esta Distribución se utiliza bajo el supuesto que una observación errónea puede producirse en cualquier punto con igual probabilidad, por lo que se escoge esta Distribución para modelar lo planteado.

Por otro lado, como se menciona en el capítulo dos del presente trabajo, el Entorno de Contaminación (2.1) indica que una proporción ε de vectores no siguen la distribución original, suponiendo así, que todas las componentes de dichos vectores están contaminadas o No están contaminadas; esto sin embargo no es necesariamente lo que ocurre en la realidad, pues se puede pensar que los errores se dan

en una o varias componentes de la observación, difícilmente en todas; en base a esto se considera otro tipo de contaminación, la misma que contamina cada variable independientemente, es decir, mientras en el modelo (2.1) una observación tiene probabilidad ε de estar contaminada, el segundo modelo considerado indica que cada componente de la observación tiene probabilidad ε de estar contaminada. Las distribuciones utilizadas para generar la contaminación por variable son: Distribución Normal $N(0,9\sigma_i)$ y la Distribución Uniforme $U(-5,5)$.

4.3 Detalles de la Simulación

4.3.1 Parámetros de la Simulación

Número de Variables y Tamaño de la Muestra

Para el presente estudio se trabaja con $p = 6$, donde p representa el número de variables y con tamaño de muestra $n = kp$ donde $k = 5; 10$ y 20

Características de la Población sin Contaminar

Como se menciona previamente, las muestras son generadas a partir de la Distribución Normal Multivariada $N(\mathbf{0}, \Sigma)$ y la Distribución Cauchy $CAU(\mathbf{0}, \Sigma)$. Donde para Σ se utilizan dos Matrices de Varianzas y Covarianzas:

$$\Sigma_1 = \begin{pmatrix} 1 & & & & & & \\ .950 & 1 & & & & & \\ .300 & .100 & 1 & & & & \\ 0 & 0 & 0 & 1 & & & \\ 0 & 0 & 0 & -.499 & 1 & & \\ 0 & 0 & 0 & -.499 & -.499 & 1 & \end{pmatrix} \quad (4.1)$$

a la misma que le corresponde los valores propios $\lambda_1 = 2.029, \lambda_2 = \lambda_3 = 1.499, \lambda_4 = 0.943, \lambda_5 = 0.028, \lambda_6 = 0.002$. La segunda matriz considerada es

$$\Sigma_2 = \begin{pmatrix} 1 & & & & & & \\ .08 & 1 & & & & & \\ .10 & .12 & 1 & & & & \\ .12 & -.10 & -.08 & 1 & & & \\ -.10 & -.08 & .08 & .10 & 1 & & \\ -.08 & .10 & -.01 & .08 & .12 & 1 & \end{pmatrix} \quad (4.2)$$

cuyos valores propios son

$$\lambda_1 = 1.282, \lambda_2 = 1.253, \lambda_3 = 1.120, \lambda_4 = 1.056, \lambda_5 = 0.719, \lambda_6 = 0.697$$

La matriz definida por (4.1) fue utilizada por Devlin en [5], la misma tiene correlaciones que varían entre .10 y .95, de tal manera que se prueba la habilidad del estimador para detectar esta variedad de correlaciones, además, λ_6 es muy cercano a cero lo que indica que la matriz es “casi” singular, esto permite medir el comportamiento de los estimadores en condiciones de “casi” singularidad, finalmente, al

igual que las correlaciones, los valores propios asociados a Σ_1 también varían ampliamente.

La matriz definida por (4.2) en cambio, presenta correlaciones y valores propios asociados a Σ_2 con poca diferencia, así se mide la capacidad del estimador para detectar las pequeñas variaciones que se muestran en Σ_2 .

Detalles Técnicos

El software utilizado para la simulación es R versión 2.8.0 [24],

El Algoritmo para generar los números aleatorios uniformes es el Mersenne Twister, y el algoritmo utilizado para generar la muestras Normales p-variadas es el Método de la Descomposición de Cholesky, ambos mencionados en el capítulo anterior.

Estimadores Robustos a ser estudiados

Para el presente estudio se han considerado los siguientes estimadores:

- Estimador de Máxima Verosimilitud para el Vector de Medias y Matriz de Varianzas y Covarianzas, al mismo que se le llamará Clásico
- Estimador M de Huber
- Estimador Bicuadrático

- Estimador T-Bicuadrático
- Estimador MCD
- Estimador DS, Stahel y Donoho

Criterios de Evaluación

Con el propósito de evaluar la Eficiencia y Estabilidad de los Estimadores de Máxima Verosimilitud y Robustos estudiados; se

define $\bar{\bar{\mathbf{x}}}^{(e)} = \left\{ \frac{1}{N} \sum_1^N \bar{\mathbf{x}}^{(e)} \right\}$ el Vector Promedio de las estimaciones para

el Vectores de Medias correspondiente al estimador e en las N

repeticiones, $\bar{\mathbf{x}}_{\max}^{(e)} = \max \{ \bar{\mathbf{x}}^{(e)} \}$ el vector que representa el Vector “más

alejado”, en términos de la distancia Euclidiana, de los Vectores de

Medias correspondientes al estimador e en las N repeticiones,

entonces se obtiene el Sesgo Promedio Total y Sesgo Máximo Total

calculando la Distancia Euclidiana entre el Vector de Medias Real

con $\bar{\bar{\mathbf{x}}}^{(e)}$ y $\bar{\mathbf{x}}_{\max}^{(e)}$ respectivamente; se define también

$sd_{Total}(\bar{\mathbf{x}}^{(e)}) = \sum sd[\bar{\mathbf{x}}^{(e)}]_i$ la Suma de las Desviaciones Estándar de

cada componentes del Vector de Medias correspondiente al

estimador e , tal que $sd_{Total}(\bar{\mathbf{x}}^{(e)})$ representa la Desviación Total.

Por último, bajo el supuesto que el mejor estimador es el que

presente un mejor balance entre: su comportamiento promedio, su

“peor” comportamiento y la dispersión de sus estimaciones, se establece un índice que es la media aritmética entre el Sesgo Promedio Total, Sesgo Máximo Total y la Desviación Total, de tal manera que el estimador que presente menor índice será el más Eficiente y Estable en cada escenario estudiado, estableciendo igual importancia a las tres componentes.

Sin embargo, en situaciones reales, no se conoce cuan contaminada esta la muestra, ni que Distribución genera la contaminación, etc. solo se tiene la matriz de observaciones y el tamaño muestral, por lo que se explora el comportamiento de los estimadores en función del Sesgo y Curtosis Muestral, con el propósito de que dichas medidas junto con el tamaño de la muestra sirvan como criterios para decidir que estimador utilizar.

4.4 Resultados de la Simulación

4.4.1 Análisis por Escenario

Se empieza con dos casos donde la muestra es generada a partir de una Distribución Normal sin contaminar, estos casos son particularmente útiles para comparar los Estimadores Robustos con la Estimación de Máxima Verosimilitud, luego se presenta los resultados de la Simulación para los escenarios estudiados.

El caso de Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, sin Contaminar

En el Cuadro 4.1a se muestra la Tabla de Resultados de la Simulación para una población Normal No Contaminada, en las columnas se presenta el Sesgo Promedio Total, Desviación Total, Sesgo Máximo Total e Índice correspondiente a cada Estimador agrupados en función de los valores de k , donde $k=5,10$ y 20 , mientras que las filas corresponden a cada estimador agrupados por el parámetro a estimar, esto es Vector de Medias, Valores Propios de Σ y Matriz de Covarianzas Σ .

De la simulación se concluye que para $k=5$ el Estimador Clásico para el Vector de Medias presenta un Sesgo Promedio de 0.023 ± 0.693 y Sesgo Máximo de 1.486 ; es decir, el promedio del Vector de Medias resultante al utilizar el Método Clásico de Estimación se encuentra a una distancia 0.023 con respecto al Vector de Medias Poblacional, la estimación más alejada está a 1.489 unidades de distancia y la Desviación Total es 0.693 , valores que implican un Índice de 0.734 . Al considerar el Índice planteado se aprecia que para los valores de $k=10$ y 20 , los Estimadores Clásico y M de Huber son los que mejor se aproximan al valor del Vector de Medias Poblacional.

Utilizando el mismo procedimiento para los Valores Propios, se encuentra que los Estimadores M de Huber y Bicuadrático son los que presentan un mejor balance entre su Sesgo Promedio Total, Sesgo Máximo Total y Desviación Total.

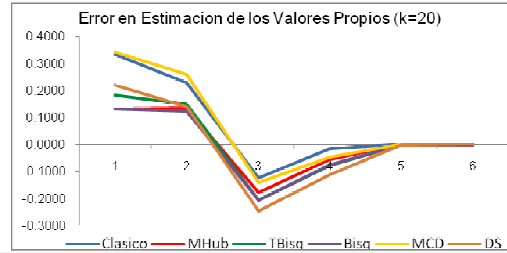
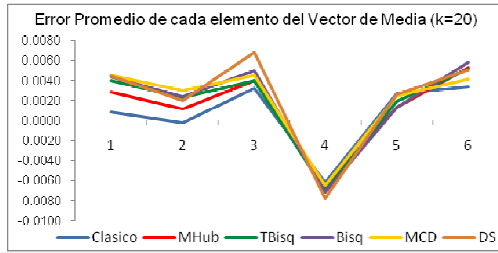
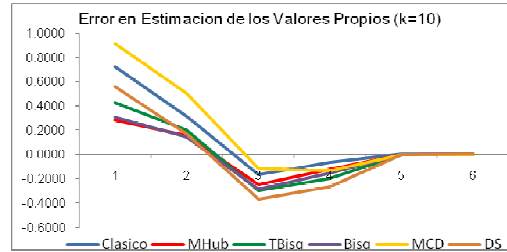
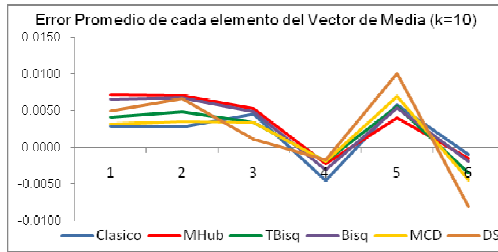
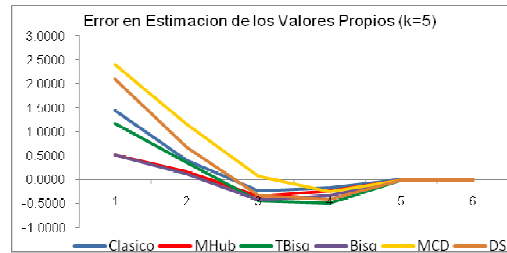
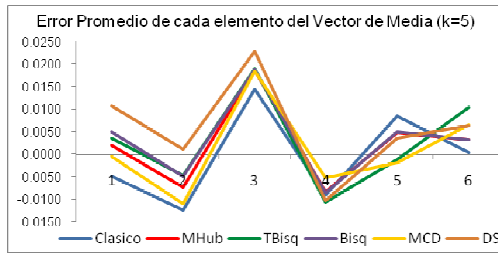
En la parte inferior del Cuadro 4.1a se observa el error promedio de la estimación de cada componente del Vector de Medias y de los Valores propios agrupados de acuerdo al tamaño muestral k , así, para $k=5$ se puede ver que en promedio el Estimador MCD subestima la primer y segunda componente del Vector de Medias, sobreestima la tercer componente, subestiman la cuarta y quinta componente y sobreestima la sexta componente; además, para todos los valores de k los estimadores presentan un comportamiento similar pues, ordenando los estimadores de mayor a menor, sobreestiman el primer y segundo valor propio, subestiman los 3 siguientes para finalmente estimar con “poco” error el menor valor propio.

Cada Matriz mostrada en el Cuadro 4.1b presenta en su triangular inferior el valor de la Matriz de Varianzas y Covarianzas Σ (en negro) y la estimación promedio del estimador correspondiente en la triangular inferior de cada matriz presentada en el Cuadro 4.1b se puede observar el valor promedio de la estimación de cada

Cuadro 4.1 a
Estimadores Robustos
Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, sin Contaminar

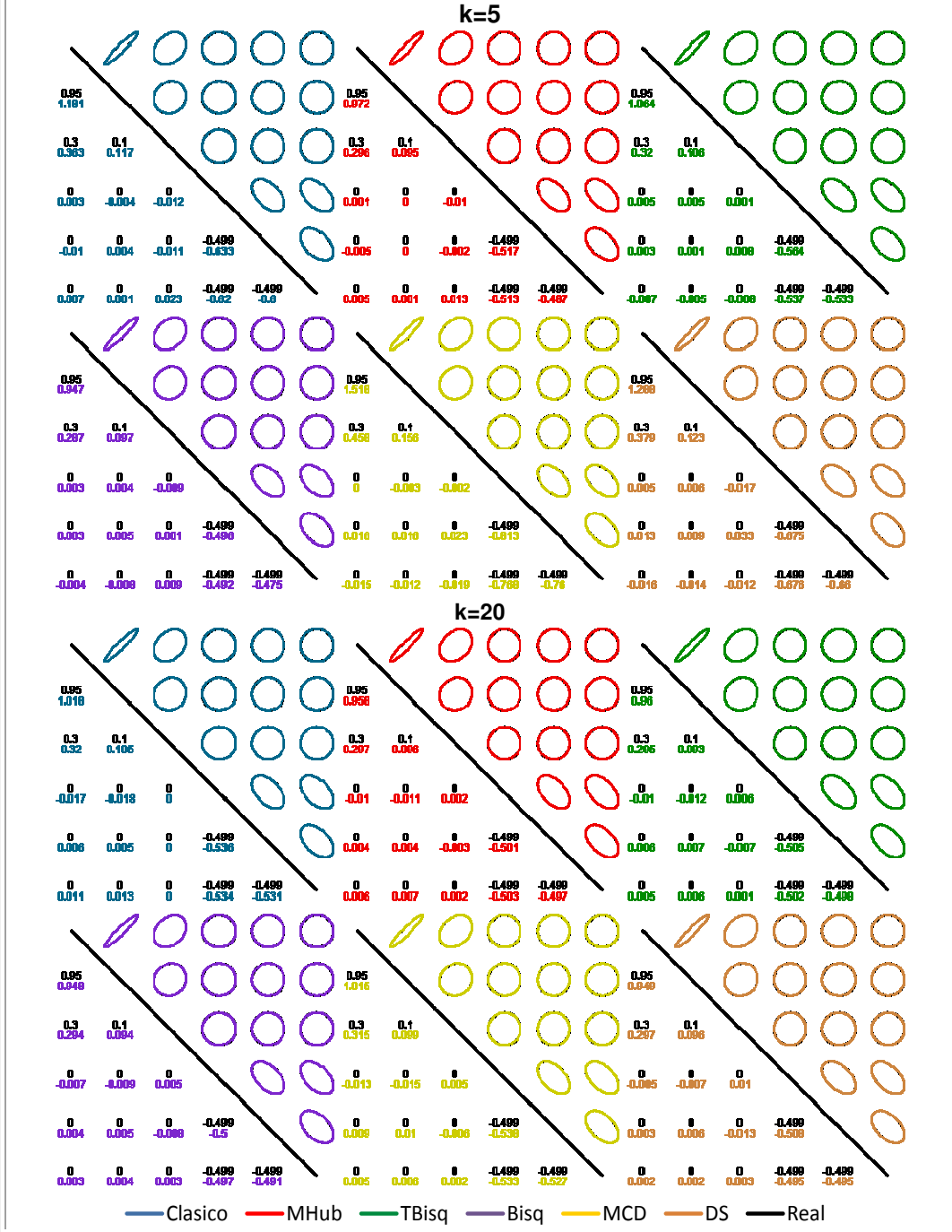
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

	k												
	5				10				20				
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	0.023 ±	0.693	1.486	0.734	0.009 ±	0.433	0.945	0.463	0.008 ±	0.307	0.697	0.338
	Mhuber	0.022 ±	0.636	1.459	0.706	0.012 ±	0.433	0.966	0.470	0.010 ±	0.322	0.761	0.364
	T-Bic	0.025 ±	0.811	1.747	0.861	0.010 ±	0.479	1.098	0.529	0.011 ±	0.337	0.805	0.384
	Bic	0.023 ±	0.650	1.495	0.722	0.012 ±	0.443	1.019	0.491	0.012 ±	0.328	0.815	0.385
	MCD	0.023 ±	0.779	1.801	0.868	0.010 ±	0.507	1.223	0.580	0.011 ±	0.356	0.828	0.398
	DS	0.028 ±	0.848	1.961	0.946	0.015 ±	0.523	1.219	0.586	0.013 ±	0.366	0.867	0.415
Valores Propios	Clásico	1.532 ±	1.425	7.920	3.626	0.808 ±	0.870	3.719	1.799	0.423 ±	0.542	2.200	1.055
	Mhuber	0.683 ±	0.857	2.552	1.364	0.423 ±	0.672	2.122	1.072	0.266 ±	0.440	1.264	0.656
	T-Bic	1.384 ±	1.449	4.843	2.559	0.590 ±	0.807	2.610	1.336	0.323 ±	0.525	1.636	0.828
	Bic	0.755 ±	0.896	3.922	1.858	0.469 ±	0.725	2.639	1.278	0.285 ±	0.505	1.416	0.736
	MCD	2.692 ±	1.744	7.660	4.032	1.060 ±	0.942	4.041	2.014	0.456 ±	0.546	2.164	1.055
	DS	2.266 ±	2.392	18.588	7.748	0.739 ±	0.860	3.171	1.590	0.376 ±	0.596	1.955	0.976
Matriz de Covarianzas	Clásico	0.666 ±	1.965	9.778	4.136	0.373 ±	1.229	4.539	2.047	0.197 ±	0.779	2.778	1.251
	Mhuber	0.060 ±	1.357	3.566	1.661	0.041 ±	0.958	2.422	1.141	0.026 ±	0.659	1.710	0.798
	T-Bic	0.288 ±	2.200	5.945	2.811	0.069 ±	1.246	3.164	1.493	0.033 ±	0.782	2.122	0.979
	Bic	0.065 ±	1.445	4.243	1.918	0.027 ±	1.080	2.872	1.327	0.025 ±	0.725	1.850	0.867
	MCD	1.593 ±	2.940	9.632	4.721	0.548 ±	1.532	4.731	2.270	0.193 ±	0.884	2.491	1.189
	DS	0.952 ±	2.761	17.308	7.007	0.062 ±	1.455	3.921	1.813	0.027 ±	0.884	2.305	1.072



Elaborado por: Néstor Montaña P.

Cuadro 4.1b
 Estimadores Robustos
 Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, sin Contaminar



Elaborado por: Néstor Montaña P.

componente de Σ , mientras que en su triangular superior se presenta, para $k=5$ y 20 , las elipses que representan la correlación existente en la nube de puntos correspondiente al cruce de cada par de variables, se representa en negro la elipse generada con los verdaderos valores y en color la elipse generada con el valor de la estimación promedio, así, la estimación de Σ correspondiente al Estimador M de Huber y $k=5$ converge (según la teoría del Método Monte Carlo) a los valores 0.972 , 0.296 , 0.001 , -0.005 y 0.005 en su primer columna, en la misma cuyos valores verdaderos son 0.95 , 0.30 , 0 , 0 , 0 ; por último las elipses correspondientes a este estimador se superponen casi perfectamente con la elipse generada por los valores poblacionales.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, sin Contaminar

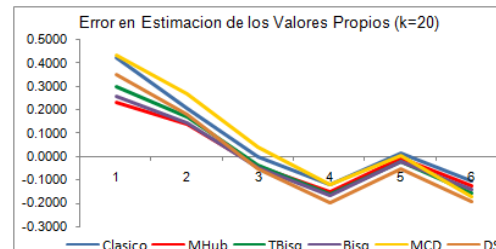
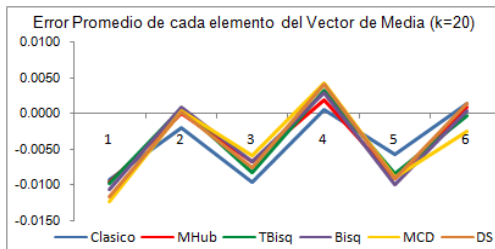
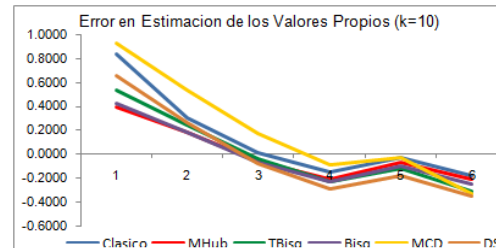
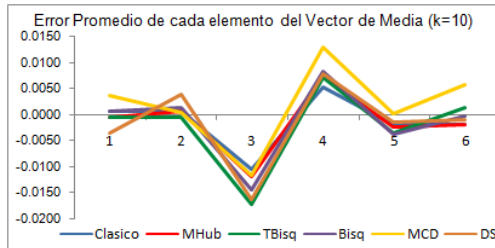
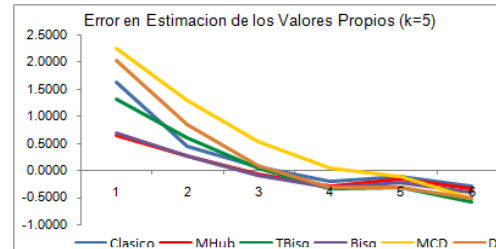
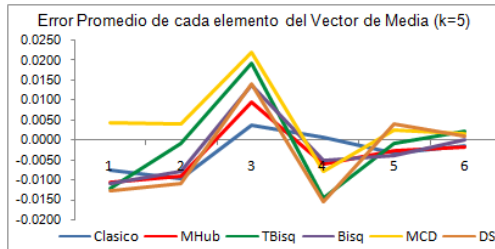
El Estimador M de Huber brinda, en todos los tamaños muestrales considerados la mejor estimación, en función del índice planteado, de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para $k=20$ obtiene un Índice de 0.555 resultado del Sesgo Promedio igual a 0.339 , Sesgo Máximo de 1.003 y Desviación Total de 0.323 (Véase Cuadro 4.2a)

Al analizar los resultados concernientes a la estimación del Vector de Medias, se encuentra que a pesar de que el Sesgo Promedio

Cuadro 4.2a
Estimadores Robustos
Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$ sin Contaminar

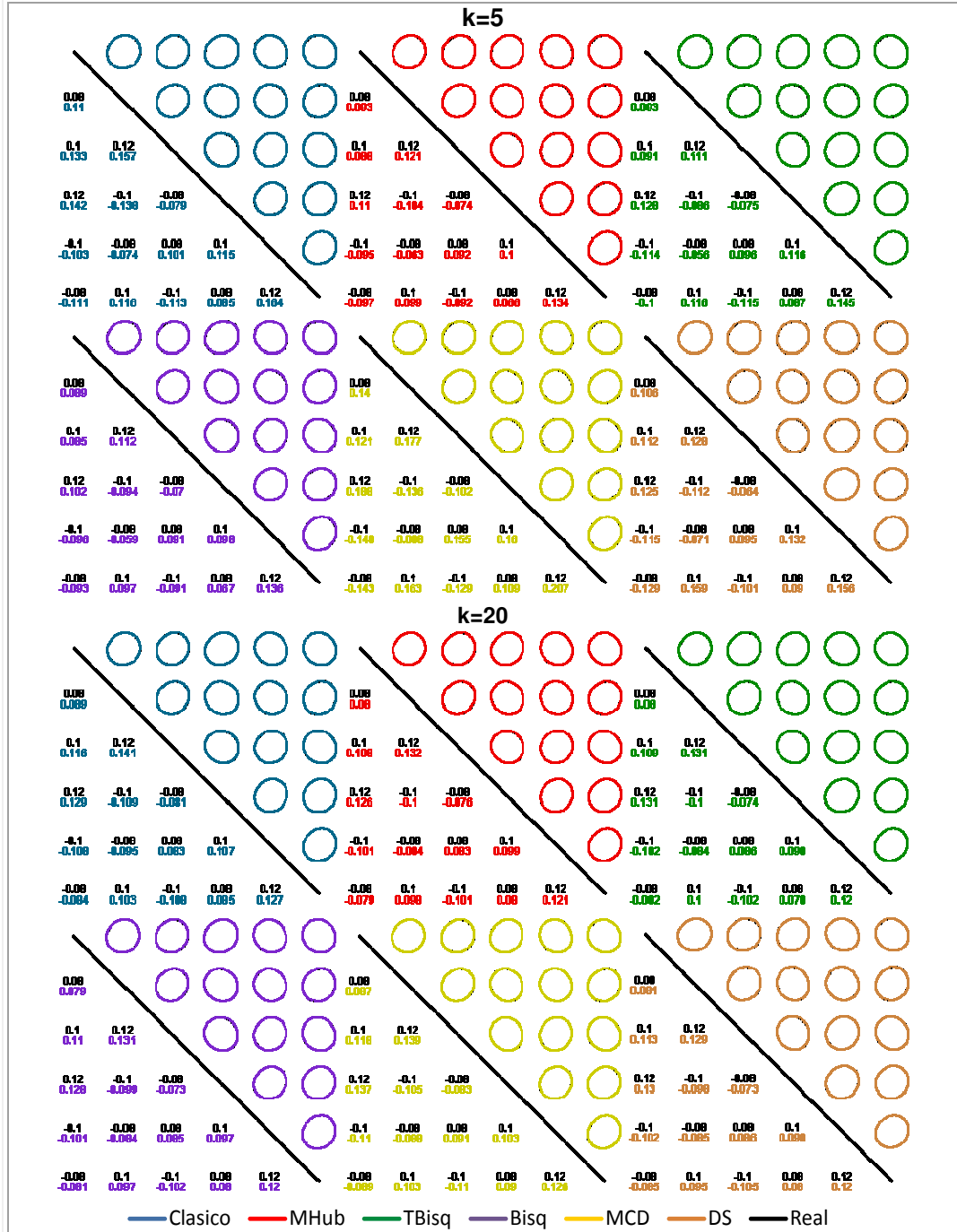
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.014	± 0.670	1.635	0.773	0.012	± 0.443	0.964	0.473	0.015	± 0.312	0.681	0.336
	Mhuber	0.018	± 0.637	1.435	0.696	0.014	± 0.442	0.921	0.459	0.015	± 0.316	0.728	0.353
	T-Bic	0.027	± 0.832	1.742	0.867	0.019	± 0.466	1.130	0.539	0.016	± 0.322	0.737	0.358
	Bic	0.020	± 0.680	1.492	0.731	0.017	± 0.451	0.970	0.479	0.016	± 0.322	0.760	0.366
	MCD	0.024	± 0.803	1.785	0.871	0.019	± 0.501	1.100	0.540	0.017	± 0.342	0.804	0.388
	DS	0.027	± 0.855	1.766	0.883	0.019	± 0.526	1.167	0.571	0.017	± 0.339	0.794	0.383
Valores Propios	Clásico	1.740	± 1.252	6.773	3.255	0.923	± 0.634	3.251	1.603	0.497	± 0.401	1.686	0.861
	Mhuber	0.833	± 0.617	2.448	1.299	0.532	± 0.422	1.503	0.819	0.339	± 0.323	1.003	0.555
	T-Bic	1.619	± 1.126	5.961	2.902	0.714	± 0.522	2.017	1.085	0.411	± 0.382	1.038	0.610
	Bic	0.934	± 0.716	2.706	1.452	0.584	± 0.486	1.408	0.826	0.373	± 0.360	1.099	0.611
	MCD	2.713	± 1.303	6.283	3.433	1.134	± 0.684	2.809	1.542	0.551	± 0.418	1.552	0.840
	DS	2.288	± 1.892	18.519	7.567	0.867	± 0.615	2.469	1.317	0.488	± 0.407	1.472	0.789
Matriz de Covarianzas	Clásico	0.633	± 1.890	11.196	4.573	0.326	± 1.195	5.358	2.293	0.180	± 0.761	2.696	1.212
	Mhuber	0.066	± 1.305	3.819	1.730	0.028	± 0.950	2.346	1.108	0.029	± 0.652	1.628	0.770
	T-Bic	0.288	± 2.196	7.722	3.402	0.051	± 1.172	3.072	1.432	0.043	± 0.765	1.911	0.906
	Bic	0.059	± 1.455	4.037	1.850	0.032	± 1.036	2.551	1.207	0.023	± 0.713	1.888	0.875
	MCD	1.454	± 3.024	10.150	4.876	0.481	± 1.534	4.354	2.123	0.191	± 0.891	2.387	1.156
	DS	0.759	± 2.663	20.144	7.855	0.042	± 1.362	3.980	1.795	0.028	± 0.869	2.225	1.041



Elaborado por: Néstor Montaña P.

Cuadro 4.2b
 Estimadores Robustos
 Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$ sin Contaminar



Elaborado por: Néstor Montaña P.

Total correspondiente al Estimación Clásico es siempre el menor, este estimador, para los tamaños muestrales iguales a 30 y 60, no es el de menor índice consecuencia de su Desviación total y Sesgo Máximo Total.

Según el Cuadro 4.2b la estimación de la Matriz de Covarianzas correspondiente al Estimador Bicuadrático y $k=20$ converge (según la teoría del Método Monte Carlo) a los valores 0.079, 0.11, 0.12, -0.101, -0.081 en su primer columna, obteniendo el menor Sesgo Promedio Total, y siendo junto al Estimador M de Huber los que ofrecen una mejor estimación.

El caso de una Población Cauchy $CAU(0, \Sigma)$, $\Sigma = \Sigma_1$, sin Contaminar

En el cuadro 4.3a parte inferior se observa que el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador MCD para $k=10$ subestima la primer componente en 0.0083 unidades, la segunda, tercera, cuarta y sexta por 0.0064, 0.0035, 0.0051 y 0.0004 respectivamente, además sobreestima la quinta componente por 0.0053 unidades, valores que proporcionan un Sesgo Promedio Total de 0.013.

Los valores representados por "---" son aquellos que indican que el estimador presenta un Sesgo "grande", así, se puede ver que el Estimador Clásico brinda estimaciones distantes del valor real en todos los casos, esto es también mostrado en el Gráfico 4.3b donde por medio de las elipses se puede observar como la estimación proporcionada por el Estimador Clásico se aparta del valor real.

Se puede apreciar además como el error "pequeño" en la estimación de la Matriz de Varianzas y Covarianzas causa una "buena" estimación de los valores propios, siendo el Estimador Bicuadrático y MCD los que menor Índice presentan para $k=5$ y 10 ; mientras que para $k=20$ los Estimadores MCD y DS son los que mejor balance entre Sesgo Promedio Total, Sesgo Máximo Total y Desviación Total muestran.

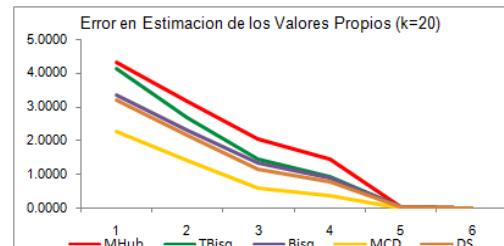
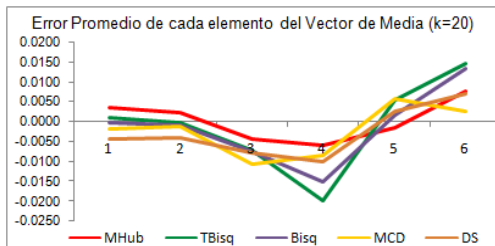
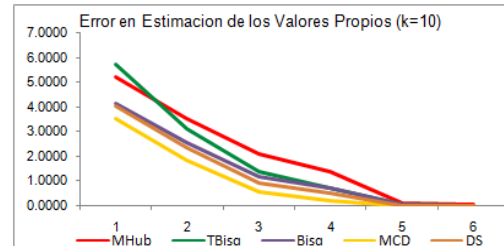
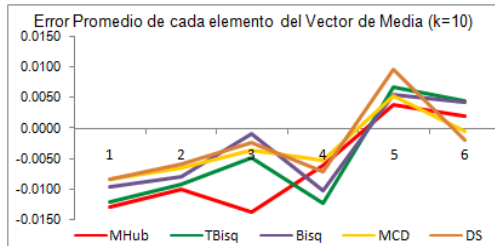
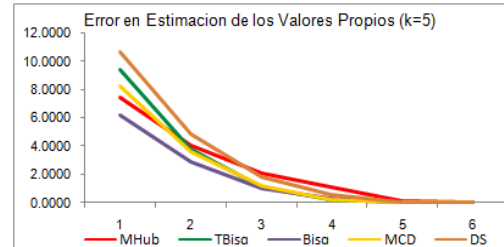
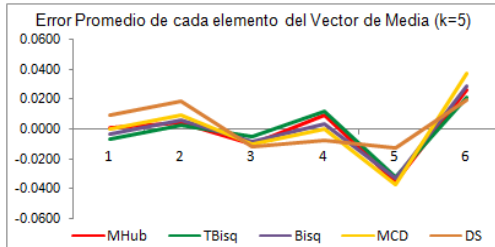
Nótese en el Cuadro 4.3b como las elipses generadas por el estimador Clásico se alejan marcadamente de las generadas por el verdadero valor de la Matriz de Varianzas y Covarianzas.

Por último, en las 500 repeticiones de este caso correspondientes a $k=5$ el Estimador M de Huber no convergió 12 veces, esto es que el algoritmo no encontró solución numérica en las iteraciones realizadas.

Cuadro 4.3a
Estimadores Robustos
Caso Población $CAU(0, \Sigma)$, $\Sigma = \Sigma_1$ sin Contaminar

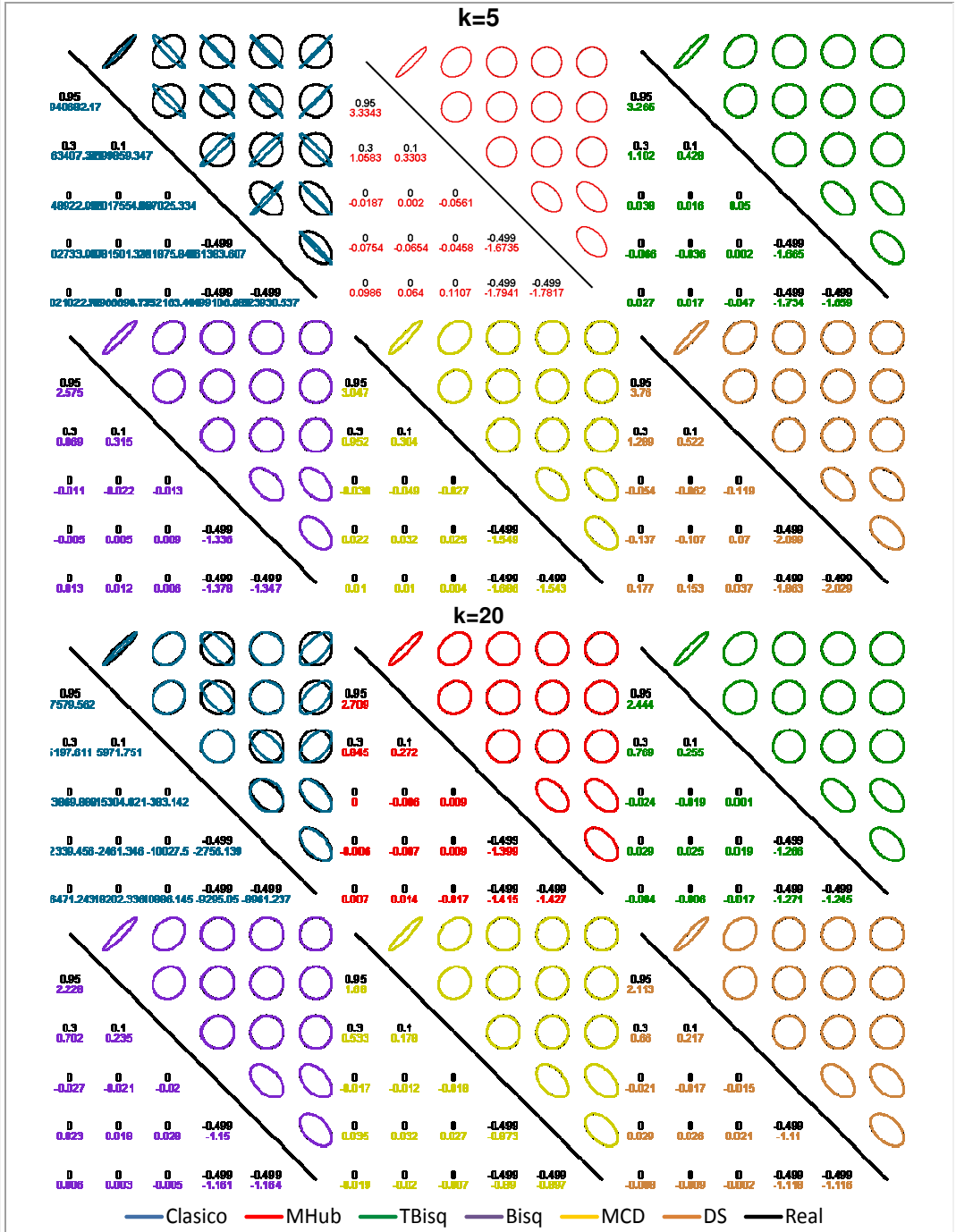
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

	5				k				20			
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	77.378 ± 4.978	---	---	1.913 ± 4.655	716.591	241.053	0.984 ± 4.843	391.010	132.279		
	Mhuber	0.047 ± 0.898	2.361	1.102	0.023 ± 0.620	1.428	0.690	0.011 ± 0.429	0.967	0.469		
	T-Bic	0.041 ± 0.962	3.014	1.339	0.022 ± 0.663	1.613	0.766	0.026 ± 0.479	1.048	0.518		
	Bic	0.045 ± 0.840	2.218	1.035	0.018 ± 0.590	1.367	0.658	0.022 ± 0.407	0.935	0.454		
	MCD	0.054 ± 0.939	2.241	1.078	0.013 ± 0.629	1.621	0.754	0.015 ± 0.439	1.080	0.512		
	DS	0.034 ± 0.835	1.940	0.937	0.016 ± 0.597	1.514	0.709	0.016 ± 0.425	0.950	0.464		
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.841 ± 6.787	51.060	22.229	6.753 ± 4.074	26.143	12.323	5.906 ± 2.740	19.202	9.283		
	T-Bic	10.176 ± 8.344	93.705	37.409	6.729 ± 5.045	25.051	12.275	5.225 ± 2.886	23.001	10.371		
	Bic	6.918 ± 6.212	40.263	17.798	5.017 ± 3.649	18.338	9.001	4.390 ± 2.411	14.457	7.086		
	MCD	9.065 ± 7.386	42.272	19.575	4.009 ± 3.068	13.959	7.012	2.746 ± 1.789	9.358	4.631		
	DS	11.871 ± 10.876	190.074	70.940	4.765 ± 3.569	36.838	15.057	4.115 ± 2.382	12.980	6.492		
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	6.998 ± 6.972	46.980	20.317	5.665 ± 4.496	23.372	11.178	5.121 ± 3.079	16.931	8.377		
	T-Bic	6.705 ± 8.220	79.492	31.473	5.107 ± 5.390	28.590	13.029	4.298 ± 3.556	19.172	9.009		
	Bic	4.797 ± 6.068	42.879	17.915	3.970 ± 3.992	19.123	9.029	3.698 ± 2.841	13.856	6.798		
	MCD	6.130 ± 7.542	45.319	19.663	2.858 ± 3.637	15.785	7.426	2.144 ± 2.295	9.708	4.716		
	DS	8.329 ± 9.559	170.344	62.744	3.630 ± 4.011	30.214	12.619	3.413 ± 2.799	13.110	6.441		



Elaborado por: Néstor Montaña P.

Cuadro 4.3b
Estimadores Robustos
 Caso Población $CAU(0, \Sigma)$, $\Sigma = \Sigma_1$ sin Contaminar



Elaborado por: Néstor Montaña P.

El Caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma = \Sigma_2)$, sin Contaminar

Nótese en el Cuadro 4.4a que para $k=10$ el promedio del Vector de Medias resultante al utilizar el Método Bicuadrático se encuentra a una distancia 0.015 con respecto al Vector de Medias Real; y la estimación más alejada está a 1.311 unidades de distancia; además este método presenta una Desviación Total de 0.556, por lo tanto el Índice correspondiente es 0.734. Al considerar el Índice planteado se aprecia que para $k=5$ y 10, los Estimadores DS y Bicuadrático son los que mejor se aproximan al valor real, mientras que para $k = 20$ el Estimador M de Huber y Bicuadrático son los que presentan menor Índice.

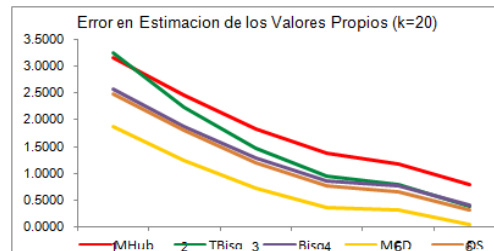
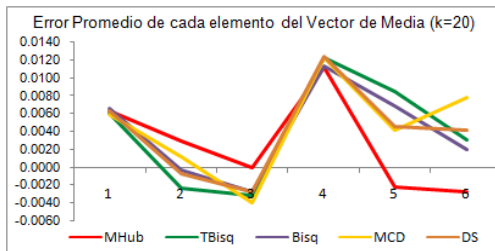
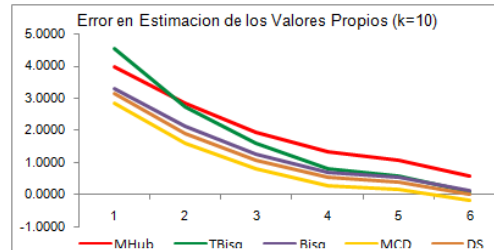
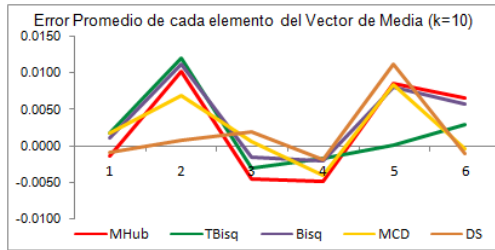
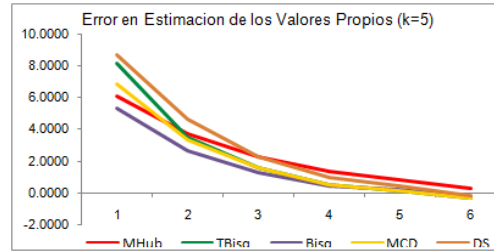
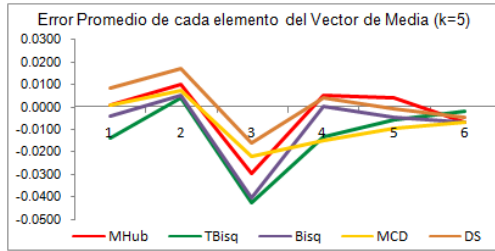
Además, con tamaño muestral 30 y 60 con respecto a los Valores Propios los Estimadores sobreestiman los cinco mayores Valores Propios y subestiman el menor Valor Propio, mientras que para tamaño muestral 120 todos los Valores Propios son sobreestimado.

Al igual que el caso en el cual la Muestra es generada a partir de una Población Cauchy $CAU(\mathbf{0}, \Sigma = \Sigma_1)$ No Contaminada, al Estimador Clásico le corresponden altos valores para su Sesgo Promedio Total, Sesgo Máximo Total y Desviación Total.

Cuadro 4.4a
Estimadores Robustos
Caso Población CAU(0, Σ), Σ = Σ₂ sin Contaminar

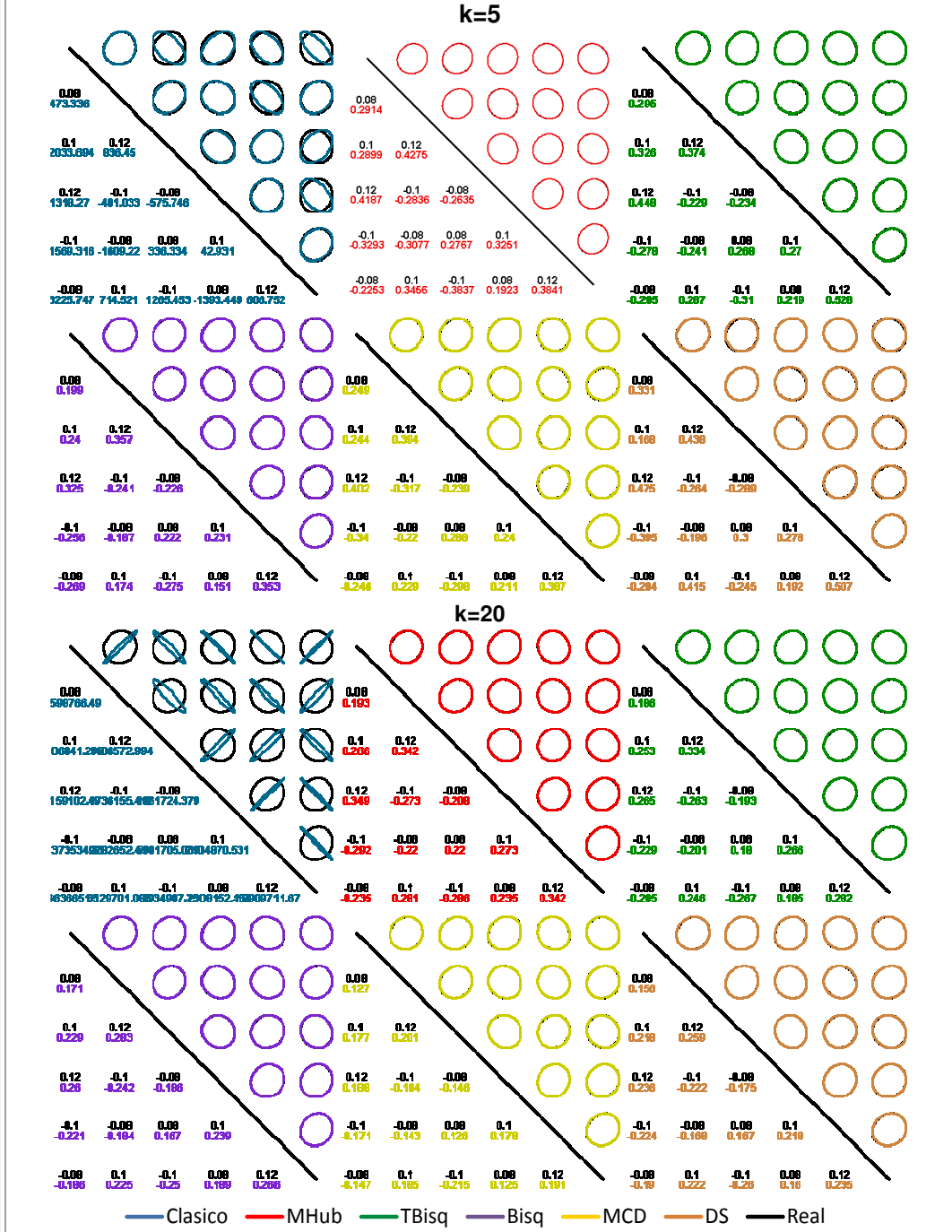
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

	5				10				20			
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	1.154 ± 5.069	284.199	96.808	1.732 ± 5.011	---	---	---	44.020 ± 4.492	---	---	---
	Mhuber	0.032 ± 0.896	2.152	1.027	0.016 ± 0.607	1.401	0.675	0.014 ± 0.433	0.971	0.473		
	T-Bic	0.047 ± 1.013	2.819	1.293	0.013 ± 0.638	1.459	0.703	0.017 ± 0.470	1.108	0.532		
	Bic	0.042 ± 0.887	2.131	1.020	0.015 ± 0.556	1.311	0.627	0.015 ± 0.418	0.988	0.474		
	MCD	0.030 ± 0.932	2.584	1.182	0.012 ± 0.618	1.492	0.708	0.017 ± 0.446	1.105	0.523		
	DS	0.026 ± 0.827	2.162	1.005	0.012 ± 0.575	1.327	0.638	0.015 ± 0.429	1.003	0.483		
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.652 ± 4.709	34.671	15.677	5.514 ± 3.004	30.932	13.150	4.840 ± 2.071	12.341	6.417		
	T-Bic	9.028 ± 6.508	59.509	25.015	5.598 ± 3.832	24.224	11.218	4.394 ± 2.298	13.349	6.680		
	Bic	6.099 ± 4.705	31.860	14.221	4.190 ± 2.841	23.960	10.330	3.640 ± 1.899	10.483	5.341		
	MCD	7.789 ± 5.546	33.671	15.669	3.405 ± 2.231	19.102	8.246	2.403 ± 1.318	6.132	3.285		
	DS	10.142 ± 10.870	78.345	33.119	3.908 ± 2.825	38.514	15.082	3.450 ± 1.881	9.878	5.070		
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	6.009 ± 5.858	39.610	17.159	4.797 ± 3.657	29.609	12.688	4.471 ± 2.766	13.397	6.878		
	T-Bic	5.583 ± 7.340	74.270	29.064	4.226 ± 4.682	26.573	11.827	3.742 ± 3.157	15.347	7.415		
	Bic	4.039 ± 5.661	39.616	16.439	3.306 ± 3.675	23.297	10.093	3.202 ± 2.581	12.091	5.958		
	MCD	5.043 ± 6.709	43.096	18.283	2.317 ± 3.221	18.953	8.164	1.884 ± 2.110	8.485	4.160		
	DS	6.959 ± 9.312	86.525	34.265	2.946 ± 3.602	36.217	14.255	2.971 ± 2.552	11.515	5.679		



Elaborado por: Néstor Montaña P.

Cuadro 4.4b
 Estimadores Robustos
 Caso Población $CAU(0, \Sigma)$, $\Sigma = \Sigma_2$ sin Contaminar



Elaborado por: Néstor Montaña P.

**El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada
bajo el modelo (2.1) con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_1)$**

De acuerdo al Cuadro 4.5a, el Estimador M de Huber brinda, en todos los tamaños muestrales considerados, la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para $k=10$ obtiene un Índice de 1.124 resultado del Sesgo Promedio igual a $0.525 \pm$, Sesgo Máximo de 2.188 y Desviación Total de 0.658.

La estimación de la Matriz de Covarianzas correspondiente al Estimador T-Bicuadrático y $k=5$ converge a los valores 1.052, 0.317, 0.002, 0.012, -0.013 en su primer columna, mientras que para $k=20$ esos valores son 0.989, 0.317, -0.005, -0.002, 0.007 (Véase el Cuadro 4.5b)

Se encuentra además que para $k=5$ el error de la Estimación promedio de cada Valor Propio correspondiente al Estimador Clásico es menor a los errores correspondientes al Estimador MCD, esto sin embargo no se mantiene para mayores tamaño muestrales donde el Estimador Clásico sobreestima el primer Valor Propio con error promedio mucho mayor a los Estimadores Robustos.

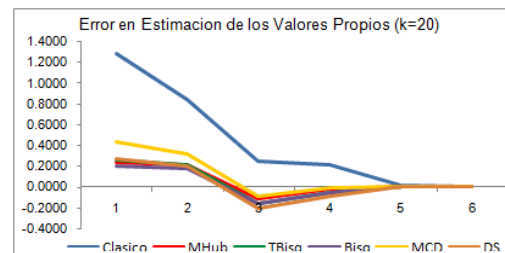
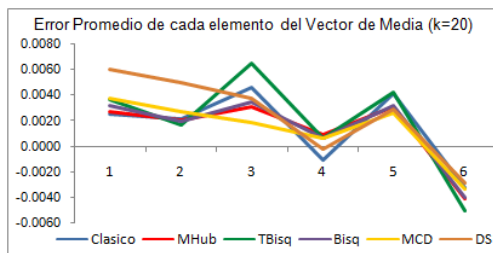
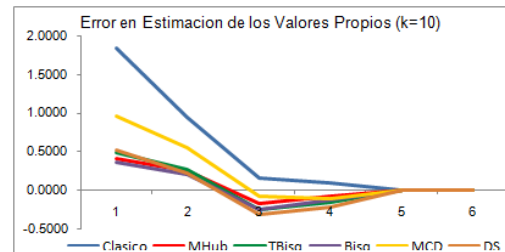
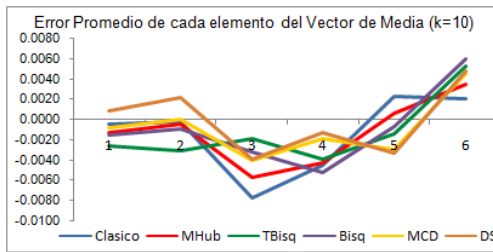
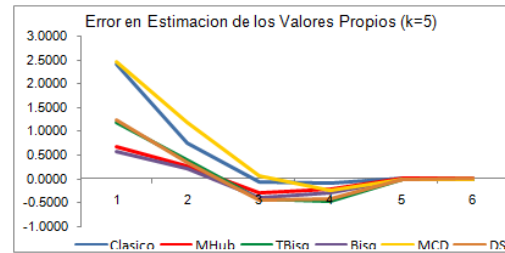
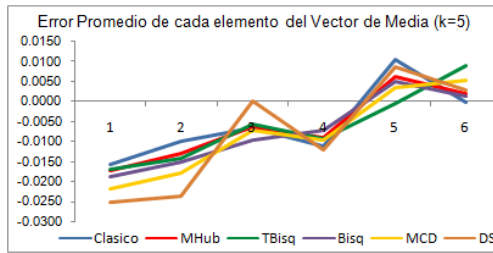
Cuadro 4.5a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_1)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.025 ± 0.745	1.614	0.795	0.009 ± 0.479	1.161	0.550	0.008 ± 0.348	0.833	0.396			
	Mhuber	0.025 ± 0.636	1.397	0.686	0.008 ± 0.426	0.971	0.468	0.007 ± 0.316	0.775	0.366			
	T-Bic	0.026 ± 0.792	1.844	0.887	0.008 ± 0.481	1.102	0.530	0.010 ± 0.332	0.787	0.376			
	Bic	0.027 ± 0.648	1.468	0.715	0.009 ± 0.448	0.980	0.479	0.007 ± 0.330	0.781	0.373			
	MCD	0.031 ± 0.754	1.757	0.848	0.007 ± 0.504	1.191	0.567	0.007 ± 0.342	0.788	0.379			
	DS	0.038 ± 0.789	1.766	0.864	0.007 ± 0.514	1.089	0.537	0.010 ± 0.347	0.810	0.389			
Valores Propios	Clásico	2.534 ± 1.997	9.518	4.683	2.075 ± 1.390	7.148	3.538	1.568 ± 0.969	3.970	2.169			
	Mhuber	0.807 ± 0.930	3.746	1.828	0.525 ± 0.658	2.188	1.124	0.343 ± 0.489	1.556	0.796			
	T-Bic	1.406 ± 1.433	7.429	3.422	0.622 ± 0.773	2.581	1.325	0.372 ± 0.586	1.726	0.895			
	Bic	0.791 ± 0.975	3.890	1.885	0.500 ± 0.715	2.250	1.155	0.321 ± 0.551	1.583	0.818			
	MCD	2.747 ± 1.775	8.255	4.259	1.110 ± 0.941	3.760	1.937	0.547 ± 0.622	2.259	1.143			
	DS	1.441 ± 1.561	12.618	5.207	0.673 ± 0.831	3.135	1.546	0.401 ± 0.609	1.779	0.930			
Matriz de Covarianzas	Clásico	1.389 ± 2.675	10.946	5.003	1.416 ± 2.008	8.770	4.064	1.200 ± 1.396	4.935	2.511			
	Mhuber	0.203 ± 1.495	4.178	1.959	0.203 ± 1.002	2.674	1.293	0.148 ± 0.704	1.875	0.909			
	T-Bic	0.322 ± 2.238	7.921	3.494	0.170 ± 1.198	3.144	1.504	0.119 ± 0.816	2.011	0.982			
	Bic	0.057 ± 1.514	4.401	1.991	0.105 ± 1.038	2.747	1.297	0.077 ± 0.760	1.962	0.933			
	MCD	1.628 ± 2.967	9.529	4.708	0.616 ± 1.512	4.325	2.151	0.303 ± 0.911	2.410	1.208			
	DS	0.330 ± 2.164	11.829	4.774	0.116 ± 1.295	3.434	1.615	0.085 ± 0.898	2.424	1.136			

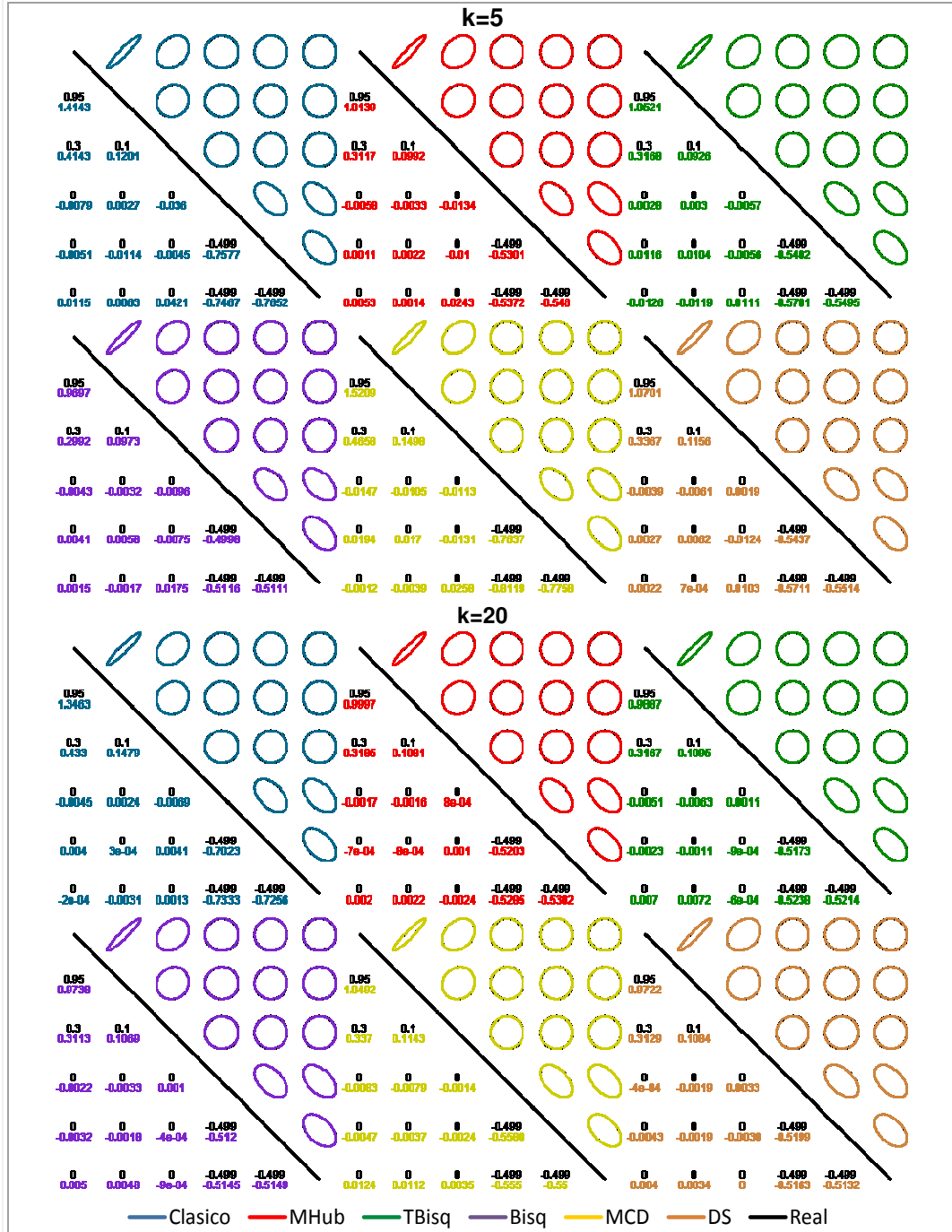


Elaborado por: Néstor Montaña P.

Cuadro 4.5b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = N(0, 9\Sigma_1)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_2)$

En el Cuadro 4.6a parte inferior se puede observar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador DS para $k=10$ subestima la primer componente en 0.0009 unidades, la segunda, tercera, cuarta y sexta por 0.0074, 0.0029, 0.0017 y 0.0058 respectivamente, además sobreestima la quinta componente por 0.0038 unidades, valores que proporcionan un Sesgo Promedio Total de 0.011 ± 0.514 .

El comportamiento de los Estimadores es muy parecido al caso anterior, así, el Estimador M de Huber y Bicuadrático brindan, en todos los tamaños muestrales considerados, la mejor estimación de la Matriz de Varianzas y Covarianzas, en particular, para $k=20$ el índice alcanzado por el Estimador Bicuadrático es 0.888, resultado del Sesgo Promedio igual a 0.089 ± 0.726 y Sesgo Máximo de 1.848.

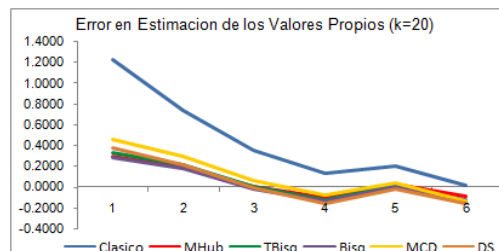
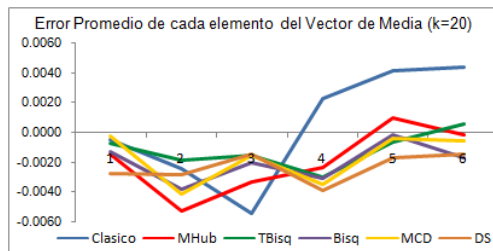
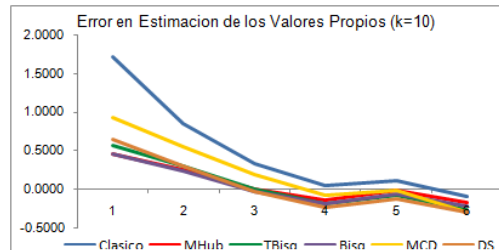
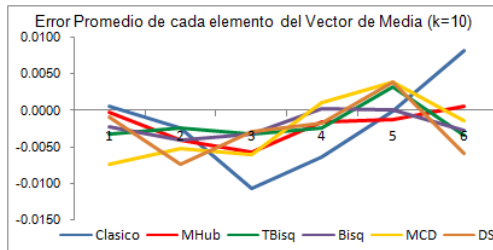
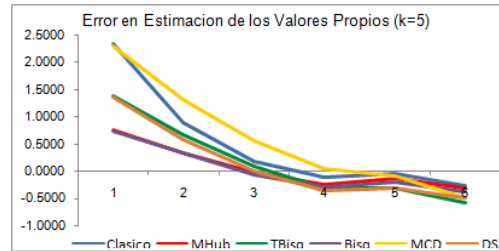
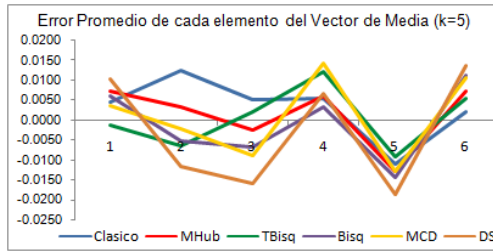
Cuadro 4.6a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_2)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.019	± 0.712	1.586	0.772	0.015	± 0.498	1.113	0.542	0.009	± 0.343	0.765	0.373
	Mhuber	0.018	± 0.656	1.454	0.709	0.007	± 0.441	1.017	0.488	0.007	± 0.308	0.786	0.367
	T-Bic	0.017	± 0.816	1.975	0.936	0.007	± 0.462	1.023	0.498	0.004	± 0.321	0.734	0.353
	Bic	0.021	± 0.688	1.555	0.755	0.006	± 0.435	1.018	0.487	0.006	± 0.316	0.785	0.369
	MCD	0.024	± 0.789	1.893	0.902	0.012	± 0.506	1.142	0.553	0.006	± 0.337	0.850	0.398
	DS	0.033	± 0.811	1.768	0.870	0.011	± 0.514	1.126	0.550	0.006	± 0.339	0.813	0.386
Valores Propios	Clásico	2.526	± 1.534	6.357	3.472	1.943	± 1.078	4.896	2.639	1.496	± 0.675	3.331	1.834
	Mhuber	0.930	± 0.700	3.019	1.550	0.579	± 0.458	1.637	0.891	0.384	± 0.331	1.162	0.625
	T-Bic	1.699	± 1.107	5.814	2.874	0.717	± 0.529	1.964	1.070	0.431	± 0.378	1.207	0.672
	Bic	0.955	± 0.744	2.819	1.506	0.593	± 0.493	1.529	0.872	0.385	± 0.380	1.147	0.637
	MCD	2.749	± 1.339	7.832	3.974	1.138	± 0.678	2.920	1.579	0.580	± 0.410	1.527	0.839
	DS	1.637	± 1.262	7.052	3.317	0.819	± 0.603	2.075	1.166	0.491	± 0.426	1.406	0.774
Matriz de Covarianzas	Clásico	1.235	± 2.568	10.936	4.913	1.214	± 1.915	7.384	3.505	1.112	± 1.400	5.085	2.532
	Mhuber	0.178	± 1.412	3.934	1.842	0.157	± 0.971	2.480	1.203	0.146	± 0.677	1.835	0.886
	T-Bic	0.387	± 2.274	8.180	3.614	0.130	± 1.169	3.129	1.476	0.119	± 0.772	2.009	0.967
	Bic	0.060	± 1.483	3.916	1.820	0.080	± 1.029	2.727	1.279	0.089	± 0.726	1.848	0.888
	MCD	1.493	± 2.973	10.055	4.841	0.537	± 1.499	4.439	2.158	0.283	± 0.881	2.550	1.238
	DS	0.321	± 2.113	9.310	3.915	0.095	± 1.292	3.570	1.652	0.102	± 0.859	2.241	1.067

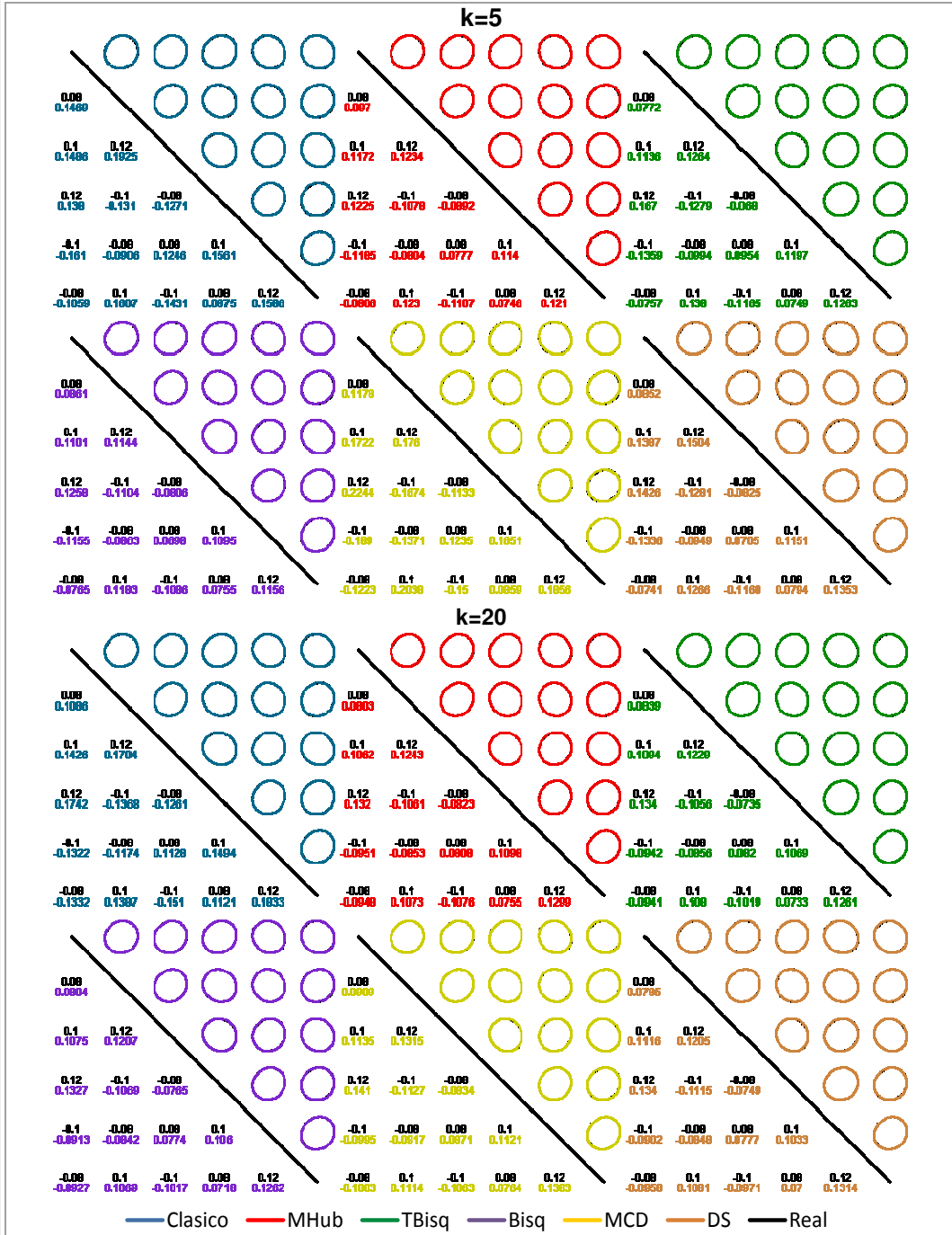


Elaborado por: Néstor Montaña P.

Cuadro 4.6b

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = N(\mathbf{0}, 9\Sigma_2)$



Elaborado por: Néstor Montaña P.

**El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada
bajo el modelo (2.1) con $\varepsilon = 0.05$ y $H = Uesf(d = 5)$**

Como se muestra en el Cuadro 4.7a, para $k=20$ el promedio de la estimación del Vector de Medias resultante al utilizar el Método M de Huber se encuentra a una distancia 0.016 con respecto al Vector de Medias Real; y la estimación más alejada está a 0.649 unidades de distancia; además este método presenta una Desviación Total de 0.314, por lo tanto el Índice correspondiente es 0.326.

Al considerar el Índice planteado se encuentra que para $k=5$, los Estimadores Clásico y M de Huber son los que mejor se aproximan al valor real del Vector de Medias, mientras que para $k = 10$ y 20 el Estimador M de Huber y Bicuadrático son los que presentan menor Índice.

Además, en todos los casos el Estimador M de Huber brinda una mejor estimación de la Matriz de Covarianzas y los Valores Propios asociados a la misma.

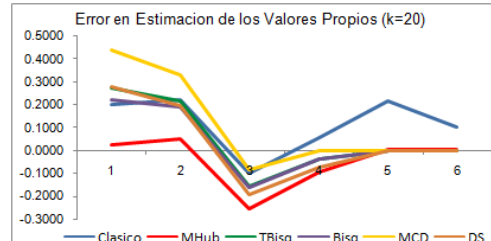
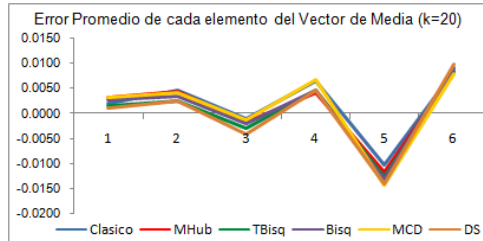
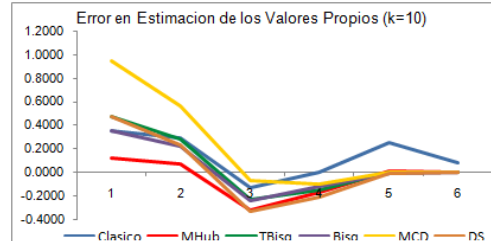
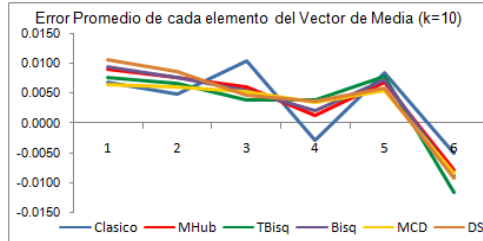
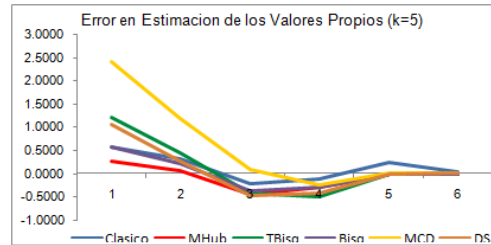
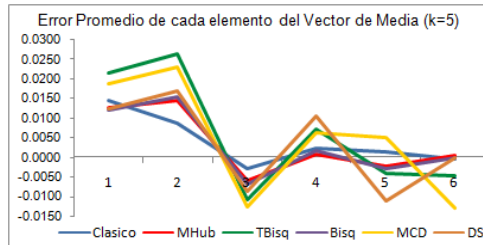
Cuadro 4.7a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = Uesf (d = 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.017 ± 0.612	1.318	0.649	0.017 ± 0.443	1.148	0.536	0.016 ± 0.309	0.689	0.338			
	Mhuber	0.020 ± 0.625	1.391	0.679	0.017 ± 0.437	1.140	0.531	0.016 ± 0.314	0.649	0.326			
	T-Bic	0.037 ± 0.799	1.834	0.890	0.018 ± 0.485	1.215	0.573	0.016 ± 0.320	0.693	0.343			
	Bic	0.021 ± 0.649	1.492	0.721	0.018 ± 0.441	1.133	0.531	0.017 ± 0.318	0.663	0.333			
	MCD	0.036 ± 0.763	1.741	0.846	0.015 ± 0.516	1.181	0.571	0.018 ± 0.326	0.780	0.375			
	DS	0.027 ± 0.778	1.834	0.880	0.019 ± 0.505	1.207	0.577	0.019 ± 0.338	0.711	0.356			
Valores Propios	Clásico	0.750 ± 0.873	3.009	1.544	0.535 ± 0.610	2.067	1.070	0.396 ± 0.461	1.417	0.758			
	Mhuber	0.586 ± 0.615	2.493	1.231	0.390 ± 0.594	1.442	0.809	0.280 ± 0.435	0.978	0.564			
	T-Bic	1.436 ± 1.356	5.992	2.928	0.620 ± 0.804	2.491	1.305	0.385 ± 0.526	1.570	0.827			
	Bic	0.772 ± 0.964	3.496	1.744	0.500 ± 0.723	2.177	1.133	0.337 ± 0.512	1.437	0.762			
	MCD	2.703 ± 1.710	7.750	4.054	1.112 ± 0.985	3.966	2.021	0.556 ± 0.570	2.024	1.050			
	DS	1.282 ± 1.330	16.955	6.522	0.659 ± 0.835	2.434	1.309	0.401 ± 0.569	1.714	0.895			
Matriz de Covarianzas	Clásico	0.356 ± 1.519	3.989	1.955	0.347 ± 1.057	2.640	1.348	0.292 ± 0.743	1.946	0.994			
	Mhuber	0.194 ± 0.946	3.135	1.425	0.147 ± 0.912	2.092	1.051	0.135 ± 0.654	1.422	0.737			
	T-Bic	0.348 ± 2.193	6.947	3.163	0.178 ± 1.234	2.993	1.468	0.136 ± 0.813	2.037	0.995			
	Bic	0.081 ± 1.512	4.072	1.888	0.108 ± 1.086	2.671	1.289	0.098 ± 0.767	1.865	0.910			
	MCD	1.598 ± 2.966	9.035	4.533	0.624 ± 1.578	4.527	2.243	0.318 ± 0.896	2.563	1.259			
	DS	0.194 ± 1.985	13.680	5.286	0.090 ± 1.316	3.410	1.605	0.098 ± 0.872	2.214	1.061			

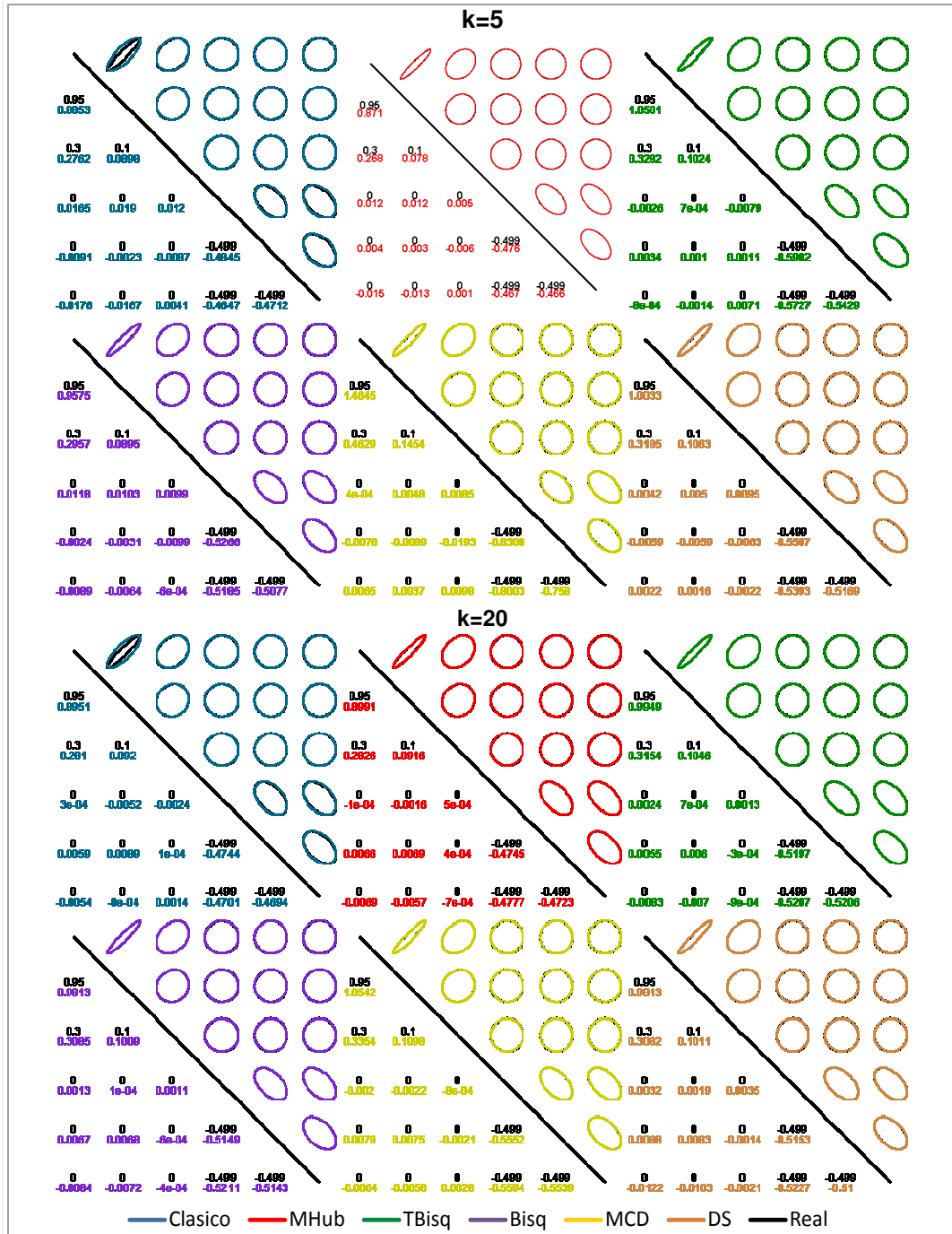


Elaborado por: Néstor Montaña P.

Cuadro 4.7b

Estimadores Robustos

Población $N(0, \Sigma = \Sigma_1)$, Contaminada con $\varepsilon = 0.05$ y $H = Uesf(d = 5)$



El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.05$ y $H = Uesf(d = 5)$

En el cuadro 4.8a se puede observar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador Bicuadrático para $k=10$ subestima la primer componente en 0.0086 unidades, la segunda, tercera y quinta por 0.0009, 0.006 y 0.0034 respectivamente, además sobreestima la cuarta y sexta componente por 0.0021 y 0.0026 unidades, valores que proporcionan un Sesgo Promedio Total de 0.012

La estimación de la Matriz de Covarianzas correspondiente al Estimador DS y $k=5$ converge a los valores 0.113, 0.121, 0.143, -0.116, -0.101 en su primer columna, mientras que para $k=20$ esos valores son 0.085, 0.105, 0.124, -0.112, -0.075 (Véase el Cuadro 4.8b)

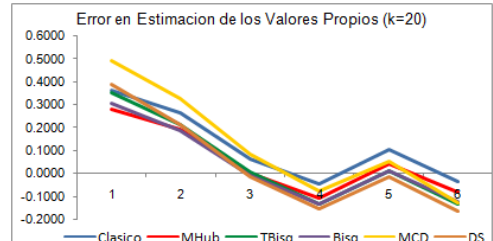
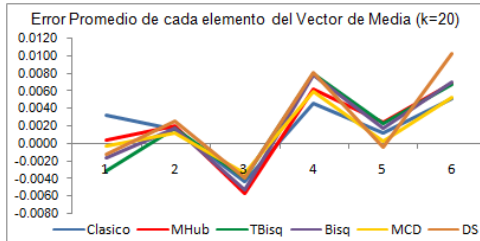
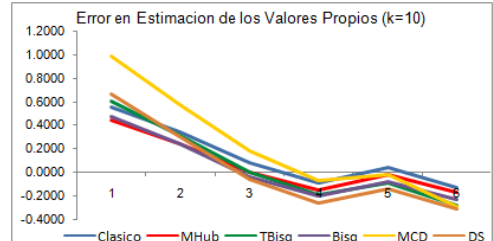
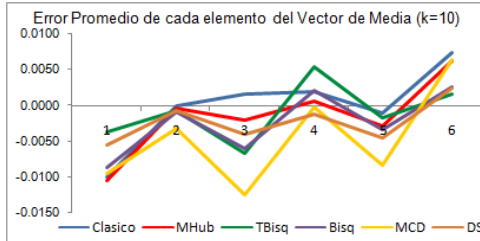
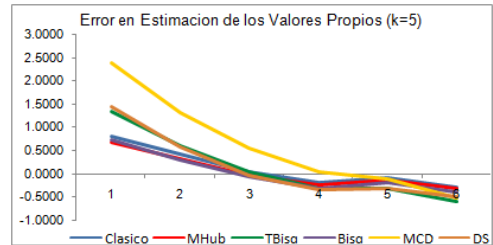
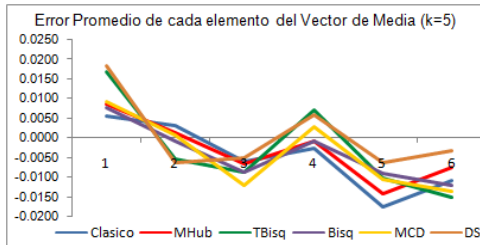
Cuadro 4.8a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = Uesf (d = 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.023	± 0.619	1.432	0.691	0.013	± 0.450	1.021	0.494	0.009	± 0.308	0.726	0.348
	Mhuber	0.020	± 0.618	1.437	0.692	0.013	± 0.445	0.978	0.478	0.011	± 0.306	0.737	0.352
	T-Bic	0.028	± 0.811	1.885	0.908	0.010	± 0.483	1.090	0.527	0.012	± 0.325	0.770	0.369
	Bic	0.019	± 0.657	1.478	0.718	0.012	± 0.456	1.020	0.496	0.012	± 0.314	0.763	0.363
	MCD	0.023	± 0.765	1.943	0.910	0.019	± 0.527	1.141	0.563	0.009	± 0.336	0.845	0.397
	DS	0.022	± 0.808	1.943	0.925	0.009	± 0.516	1.167	0.564	0.014	± 0.338	0.813	0.388
Valores Propios	Clásico	0.983	± 0.612	2.354	1.316	0.681	± 0.446	1.779	0.969	0.470	± 0.319	1.354	0.715
	Mhuber	0.847	± 0.462	2.203	1.171	0.560	± 0.436	1.665	0.887	0.365	± 0.313	1.149	0.609
	T-Bic	1.653	± 1.107	4.966	2.575	0.765	± 0.585	2.080	1.144	0.457	± 0.410	1.291	0.719
	Bic	0.949	± 0.704	2.439	1.364	0.618	± 0.517	1.922	1.019	0.404	± 0.378	1.325	0.702
	MCD	2.818	± 1.387	7.089	3.765	1.197	± 0.705	3.340	1.747	0.618	± 0.416	1.756	0.930
	DS	1.689	± 1.230	11.680	4.866	0.848	± 0.630	2.223	1.234	0.500	± 0.436	1.460	0.798
Matriz de Covarianzas	Clásico	0.314	± 1.453	4.123	1.963	0.332	± 1.035	2.714	1.360	0.297	± 0.728	1.942	0.989
	Mhuber	0.126	± 1.028	3.820	1.658	0.144	± 0.955	2.478	1.193	0.135	± 0.681	1.746	0.854
	T-Bic	0.330	± 2.255	7.117	3.234	0.140	± 1.215	3.380	1.578	0.134	± 0.792	2.090	1.005
	Bic	0.058	± 1.515	4.049	1.874	0.085	± 1.058	2.840	1.328	0.100	± 0.736	1.904	0.913
	MCD	1.511	± 3.051	10.483	5.015	0.573	± 1.539	4.550	2.221	0.305	± 0.920	2.425	1.217
	DS	0.349	± 2.162	10.207	4.239	0.091	± 1.300	3.700	1.697	0.109	± 0.879	2.254	1.081

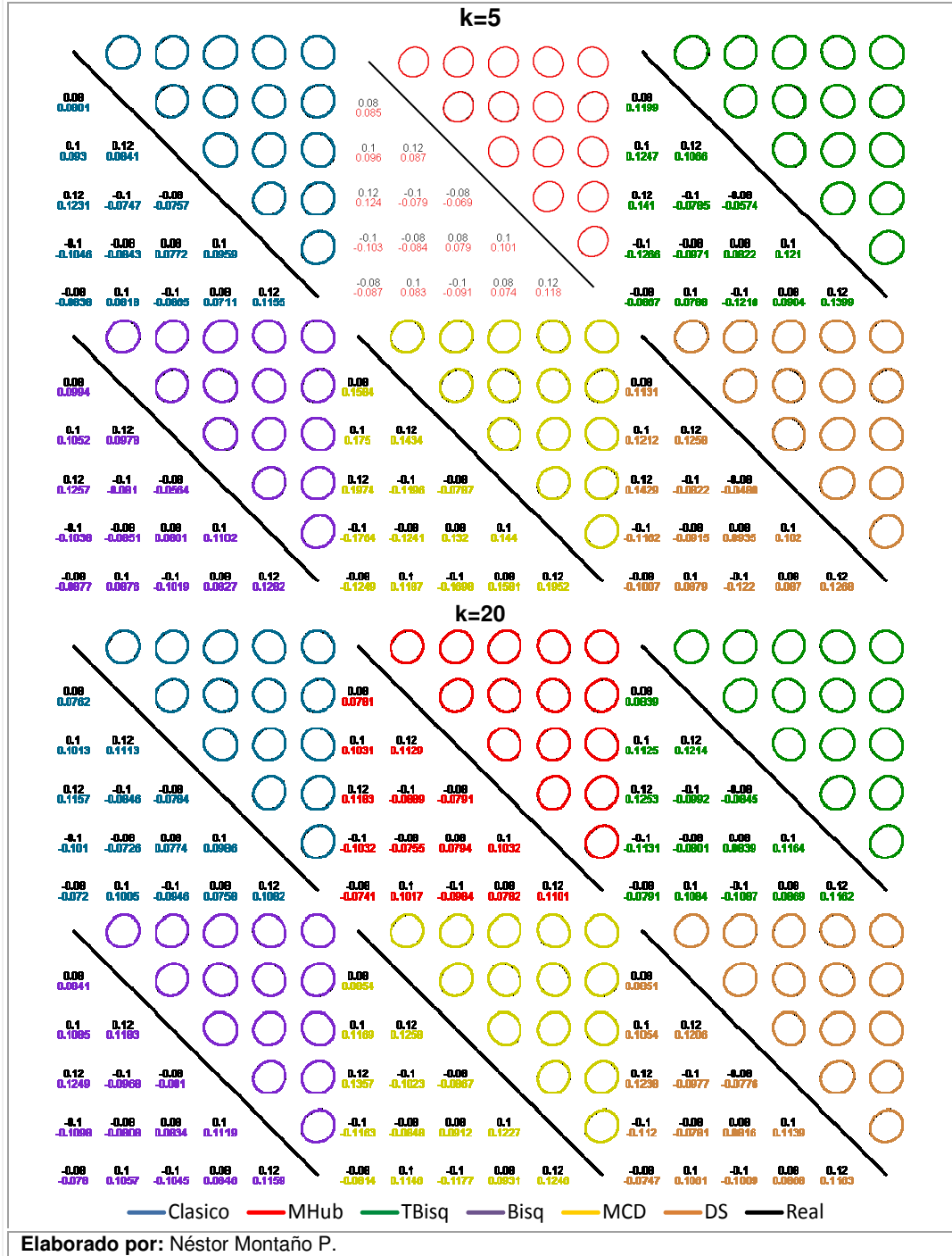


Elaborado por: Néstor Montaña P.

Cuadro 4.8b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = Uesf(d = 5)$



**El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada
bajo el modelo (2.1) con $\varepsilon = 0.05$ y $H = N(0.537\alpha_p, 9\Sigma_1)$**

De acuerdo al Cuadro 4.9a, el Estimador M de Huber brinda, en todos los tamaños muestrales considerados, la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para $k=5$ obtiene un Índice de 1.213 resultado del Sesgo Promedio igual a 0.613, Sesgo Máximo de 2.381 y Desviación Total de 0.646.

Realizando un análisis similar para la estimación del Vector de Medias, se aprecia que los Estimadores M de Huber y Bicuadrático son los que presentan una mejor balance entre su estimación Promedio, “peor” estimación y dispersión.

Además, como se muestra en la segunda parte del Cuadro 4.9a se mantiene el comportamiento de los Estimadores pues estos sobreestiman el mayor valor propio, subestiman los 4 siguientes para finalmente aproximar con poco error el menor valor propio; y se puede observar que para $k=10$ y 20 el Estimador Clásico sobreestima el mayor Valor Propio con error promedio mucho mayor a los Estimadores Robustos

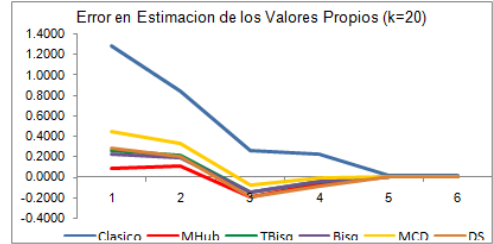
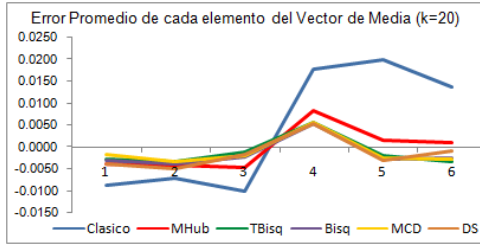
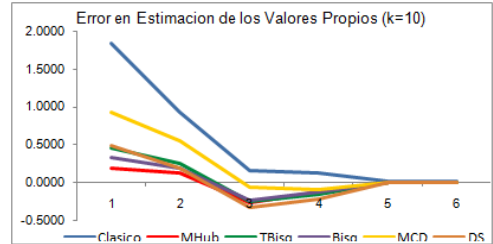
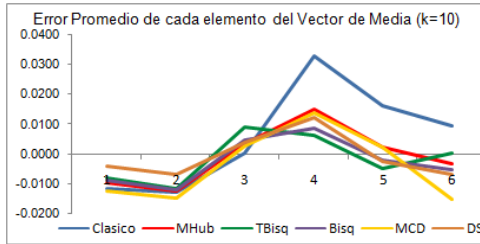
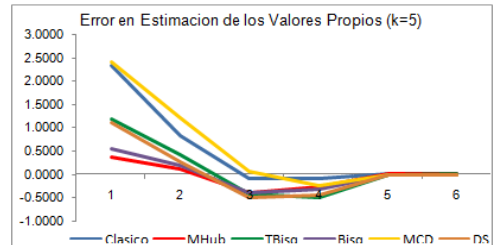
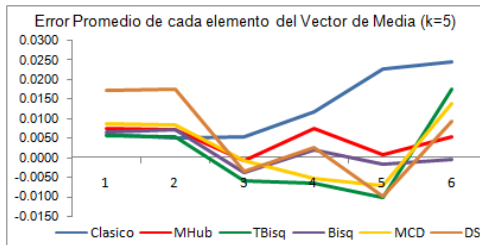
Cuadro 4.9a

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		k											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.037	± 0.702	1.624	0.787	0.042	± 0.516	1.168	0.575	0.034	± 0.348	0.858	0.413
	Mhuber	0.014	± 0.616	1.408	0.679	0.022	± 0.456	0.989	0.489	0.011	± 0.304	0.767	0.361
	T-Bic	0.024	± 0.783	1.686	0.831	0.019	± 0.488	1.139	0.549	0.008	± 0.326	0.766	0.367
	Bic	0.011	± 0.629	1.444	0.694	0.019	± 0.464	1.051	0.512	0.009	± 0.316	0.768	0.364
	MCD	0.021	± 0.760	1.660	0.814	0.028	± 0.520	1.233	0.594	0.008	± 0.341	0.794	0.381
	DS	0.029	± 0.798	1.760	0.862	0.017	± 0.521	1.229	0.589	0.009	± 0.334	0.803	0.382
Valores Propios	Clásico	2.487	± 2.029	9.092	4.536	2.058	± 1.392	6.101	3.184	1.576	± 0.948	4.330	2.285
	Mhuber	0.613	± 0.646	2.381	1.213	0.373	± 0.574	1.569	0.839	0.252	± 0.468	1.168	0.629
	T-Bic	1.412	± 1.371	6.091	2.958	0.594	± 0.742	2.445	1.260	0.374	± 0.573	1.696	0.881
	Bic	0.769	± 0.987	3.004	1.587	0.475	± 0.681	2.271	1.142	0.331	± 0.535	1.741	0.869
	MCD	2.709	± 1.805	7.501	4.005	1.075	± 0.950	4.213	2.079	0.565	± 0.573	1.962	1.033
	DS	1.330	± 1.349	11.336	4.672	0.655	± 0.802	2.774	1.410	0.402	± 0.583	2.026	1.004
Matriz de Covarianzas	Clásico	1.401	± 2.591	11.291	5.095	1.404	± 2.034	7.859	3.766	1.219	± 1.404	4.854	2.492
	Mhuber	0.094	± 0.979	3.120	1.398	0.061	± 0.899	2.174	1.045	0.041	± 0.646	1.556	0.748
	T-Bic	0.290	± 2.275	6.855	3.140	0.134	± 1.193	2.945	1.424	0.132	± 0.809	2.114	1.018
	Bic	0.029	± 1.525	3.927	1.827	0.076	± 1.034	2.888	1.333	0.102	± 0.755	2.014	0.957
	MCD	1.606	± 3.024	9.223	4.618	0.605	± 1.513	4.512	2.210	0.322	± 0.904	2.472	1.233
	DS	0.213	± 2.112	11.092	4.473	0.072	± 1.305	3.565	1.647	0.094	± 0.898	2.372	1.121

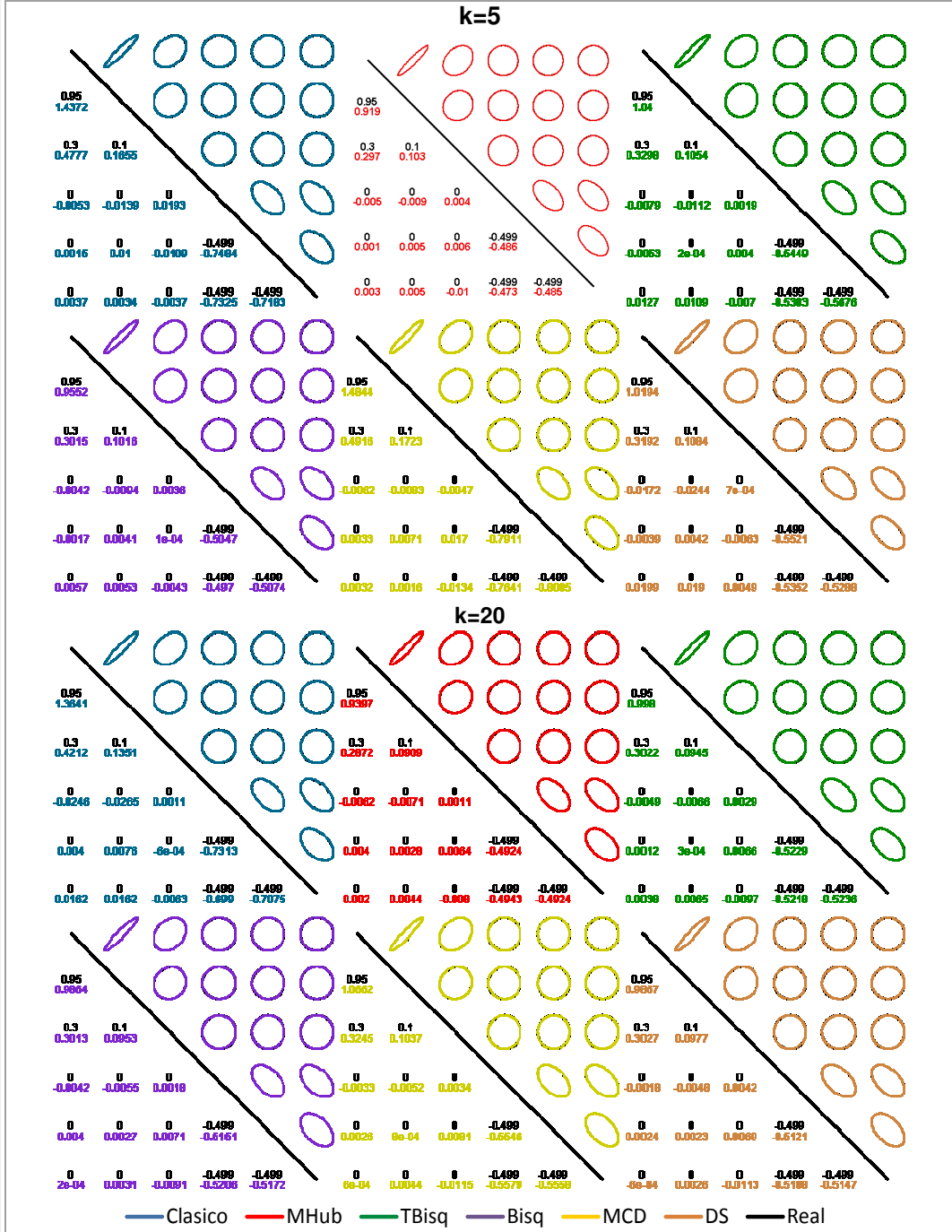


Elaborado por: Néstor Montaña P.

Cuadro 4.9 b

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$



Elaborado por: Néstor Montaña P.

**El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada
bajo el modelo (2.1) con $\varepsilon = 0.05$ y $H = N(0.537\alpha_p, 9\Sigma_2)$**

Como se muestra en el Cuadro 4.10a, para $k=10$ el promedio de la estimación del Vector de Medias resultante al utilizar el Método T-Bicuadrático se encuentra a una distancia 0.011 con respecto al Vector de Medias Real; y la estimación más alejada está a 1.056 unidades de distancia; además este método presenta una Desviación Total de 0.485, por lo tanto el Índice correspondiente es 0.517.

Al considerar el Índice planteado se aprecia que para $k=5$ y 10, los Estimadores M de Huber y Bicuadrático son los que mejor se aproximan al valor real del Vector de Medias, mientras que para $k=20$ el Estimador M de Huber y T-Bicuadrático son los que presentan menor Índice.

Además, en todos los tamaños muestrales el Estimador M de Huber y Bicuadrático brindan una mejor estimación de la Matriz de Varianzas y Covarianzas así como a los Valores Propios asociados a la misma.

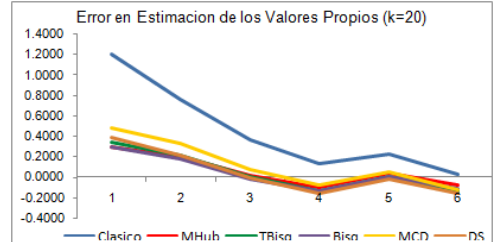
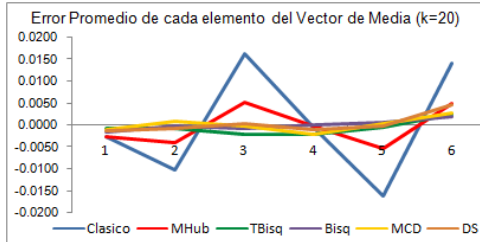
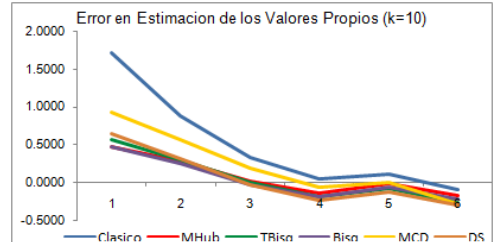
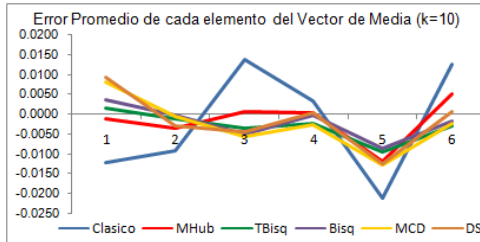
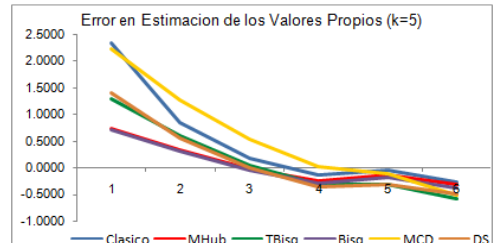
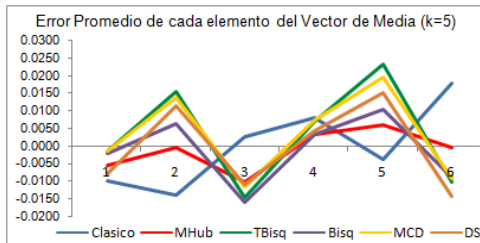
Cuadro 4.10a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.05$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5						k						20							
		Prom		Desv		Max		Prom		Desv		Max		Prom		Desv		Max		Indice	
Vector de Medias	Clásico	0.026	± 0.695	1.488	0.737	0.032	± 0.507	1.196	0.578	0.029	± 0.337	0.798	0.388								
	Mhuber	0.013	± 0.630	1.401	0.681	0.014	± 0.446	1.083	0.514	0.010	± 0.302	0.687	0.333								
	T-Bic	0.034	± 0.789	1.808	0.877	0.011	± 0.485	1.056	0.517	0.004	± 0.316	0.650	0.324								
	Bic	0.022	± 0.649	1.458	0.710	0.011	± 0.467	1.060	0.513	0.003	± 0.312	0.694	0.336								
	MCD	0.029	± 0.787	1.773	0.863	0.016	± 0.531	1.116	0.555	0.004	± 0.336	0.735	0.359								
	DS	0.028	± 0.779	1.766	0.857	0.017	± 0.504	1.155	0.559	0.005	± 0.337	0.723	0.355								
Valores Propios	Clásico	2.514	± 1.617	7.573	3.901	1.961	± 1.001	6.234	3.065	1.486	± 0.708	3.705	1.967								
	Mhuber	0.896	± 0.665	2.448	1.336	0.591	± 0.469	1.579	0.880	0.387	± 0.335	1.130	0.617								
	T-Bic	1.606	± 1.107	5.807	2.840	0.725	± 0.561	1.906	1.064	0.441	± 0.368	1.111	0.640								
	Bic	0.935	± 0.707	2.806	1.483	0.601	± 0.532	1.602	0.912	0.394	± 0.365	1.133	0.630								
	MCD	2.680	± 1.297	6.189	3.389	1.144	± 0.670	2.861	1.558	0.601	± 0.396	1.670	0.889								
	DS	1.664	± 1.202	9.358	4.075	0.816	± 0.601	2.130	1.182	0.495	± 0.397	1.469	0.787								
Matriz de Covarianzas	Clásico	1.219	± 2.497	10.952	4.889	1.235	± 1.944	7.753	3.644	1.113	± 1.375	5.064	2.518								
	Mhuber	0.155	± 1.364	3.836	1.785	0.178	± 0.989	2.564	1.244	0.147	± 0.696	1.886	0.910								
	T-Bic	0.291	± 2.197	7.649	3.379	0.149	± 1.191	3.178	1.506	0.125	± 0.784	1.984	0.964								
	Bic	0.058	± 1.464	4.205	1.909	0.101	± 1.028	2.654	1.261	0.092	± 0.747	1.909	0.916								
	MCD	1.430	± 2.908	9.830	4.722	0.554	± 1.481	4.163	2.066	0.295	± 0.900	2.392	1.196								
	DS	0.328	± 2.145	11.278	4.584	0.104	± 1.275	3.500	1.626	0.102	± 0.877	2.195	1.058								

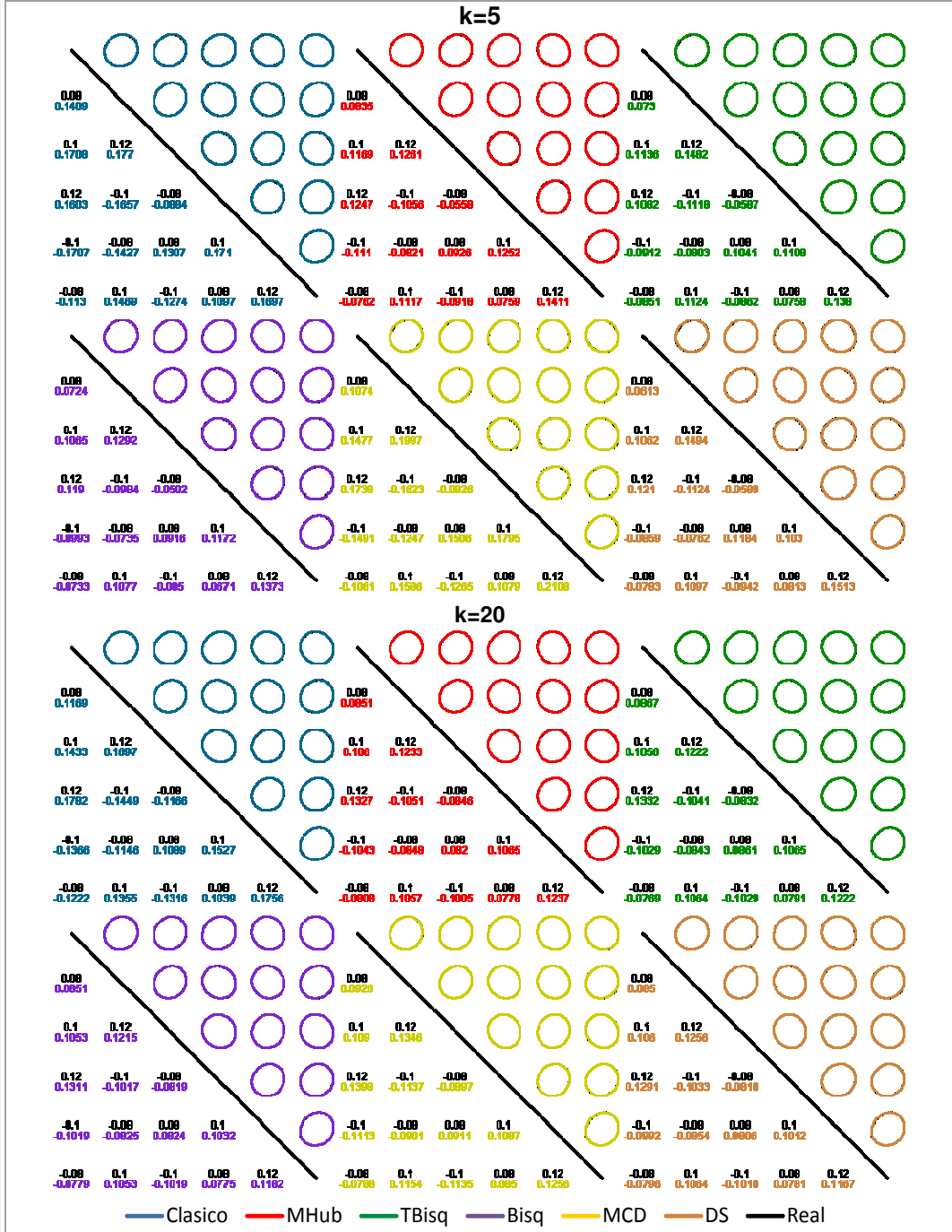


Elaborado por: Néstor Montaña P.

Cuadro 4.10b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.05$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada

Por Variable con $\varepsilon = 0.05$ y $H = N(0,9)$

En el cuadro 4.11a se puede apreciar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador MCD para $k=20$ sobreestima la primer componente por 0.0026 unidades, sobreestima también la segunda, cuarta y quinta componentes por 0.0036, 0.0023 y 0.0016 respectivamente, además subestima la tercer y sexta componente por 0.0035 y 0.0038 unidades, valores que proporcionan un Sesgo Promedio Total de 0.006

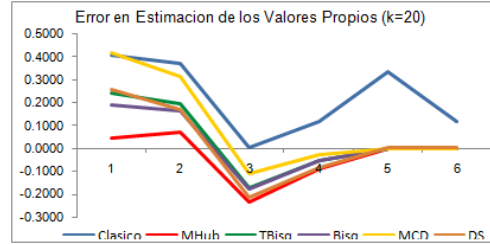
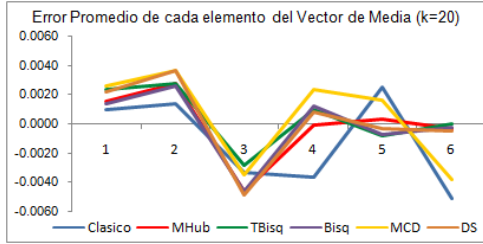
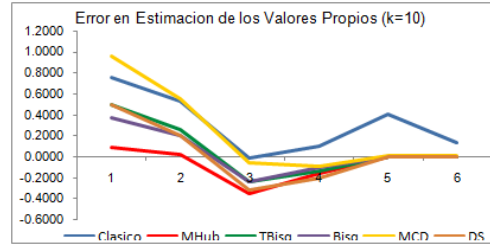
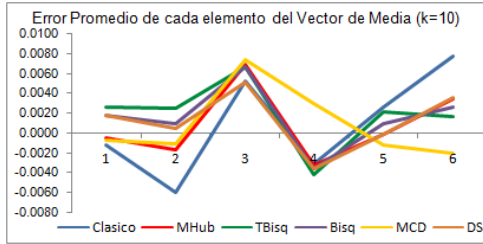
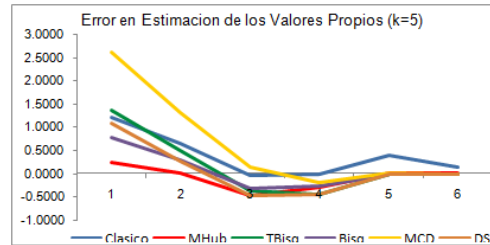
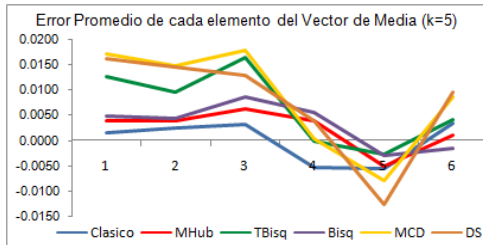
A pesar de que el promedio de la Matriz de Varianzas y Covarianzas resultante al utilizar el Estimador DS es para todos los tamaños muestrales la mas cercana a la Matriz de Varianzas y Covarianzas Real, la Dispersión observada en este Estimador causa que tenga índices mayores que los demás Estimadores.

La estimación de la Matriz de Covarianzas correspondiente al Estimador M de Huber y $k=5$ converge a los valores 0.855, 0.261, -0.006, -0.002, 0.007 en su primer columna, mientras que para $k=20$

Cuadro 4.11a
Estimadores Robustos
 Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = N(0, 9)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.009 ± 0.692	1.520	0.740	0.012 ± 0.485	1.087	0.528	0.008 ± 0.319	0.706	0.344			
	Mhuber	0.011 ± 0.633	1.483	0.709	0.009 ± 0.434	1.023	0.488	0.006 ± 0.307	0.707	0.340			
	T-Bic	0.023 ± 0.773	1.805	0.867	0.009 ± 0.468	1.054	0.477	0.005 ± 0.326	0.723	0.351			
	Bic	0.013 ± 0.663	1.594	0.757	0.008 ± 0.445	1.011	0.488	0.006 ± 0.314	0.719	0.346			
	MCD	0.031 ± 0.769	1.665	0.822	0.008 ± 0.499	1.078	0.528	0.007 ± 0.338	0.719	0.355			
	DS	0.030 ± 0.781	1.789	0.867	0.007 ± 0.502	1.120	0.543	0.007 ± 0.333	0.781	0.374			
Valores Propios	Clásico	1.439 ± 1.187	6.681	3.102	1.013 ± 0.945	3.510	1.823	0.669 ± 0.592	2.746	1.336			
	Mhuber	0.618 ± 0.579	2.102	1.100	0.404 ± 0.567	2.016	0.996	0.263 ± 0.450	1.112	0.608			
	T-Bic	1.552 ± 1.485	6.096	3.044	0.624 ± 0.792	2.873	1.430	0.356 ± 0.538	1.537	0.810			
	Bic	0.939 ± 0.979	3.487	1.802	0.497 ± 0.739	2.480	1.239	0.310 ± 0.530	1.464	0.768			
	MCD	2.922 ± 1.855	7.546	4.107	1.114 ± 0.989	4.175	2.092	0.534 ± 0.575	1.967	1.025			
	DS	1.283 ± 1.302	6.986	3.190	0.650 ± 0.861	2.861	1.457	0.383 ± 0.595	1.575	0.851			
Matriz de Covarianzas	Clásico	0.961 ± 2.091	9.666	4.239	0.782 ± 1.537	6.422	2.913	0.557 ± 1.008	4.068	1.878			
	Mhuber	0.248 ± 0.937	2.924	1.369	0.194 ± 0.885	2.161	1.080	0.095 ± 0.627	1.423	0.715			
	T-Bic	0.498 ± 2.307	7.954	3.586	0.180 ± 1.206	3.203	1.530	0.101 ± 0.790	1.894	0.929			
	Bic	0.241 ± 1.645	4.591	2.159	0.109 ± 1.066	2.912	1.362	0.064 ± 0.724	1.825	0.871			
	MCD	1.793 ± 3.138	9.300	4.744	0.626 ± 1.534	4.569	2.243	0.282 ± 0.899	2.390	1.190			
	DS	0.226 ± 2.041	7.716	3.328	0.083 ± 1.301	3.616	1.667	0.063 ± 0.861	2.079	1.001			

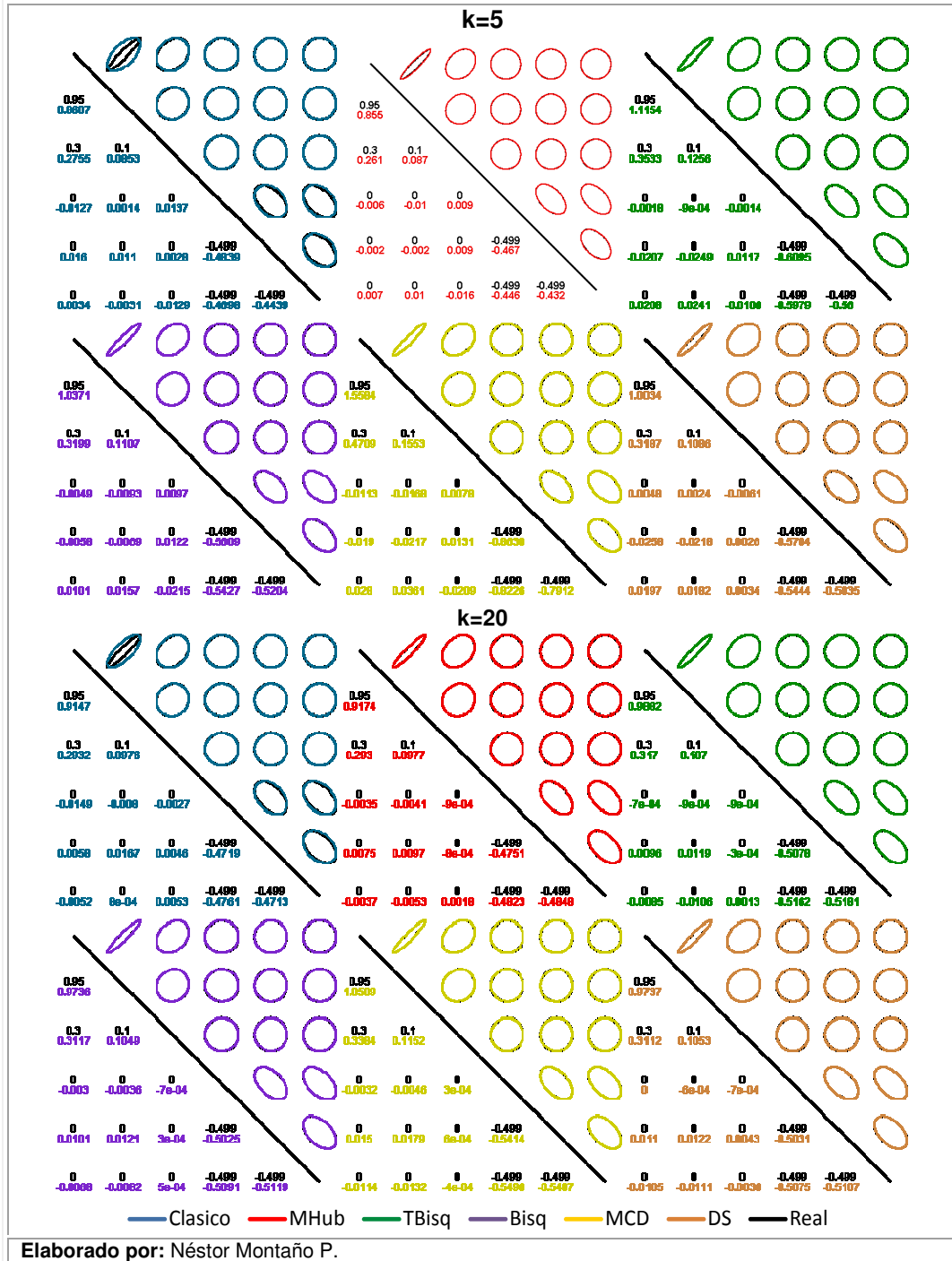


Elaborado por: Néstor Montaña P.

Cuadro 4.11 b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.05$ y $H = N(0, 9)$



esos valores son 0.917, 0.293, -0.004, 0.008, -0.004 (Véase el Cuadro 4.11b).

**El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada
Por Variable con $\varepsilon = 0.05$ y $H = N(0,9)$**

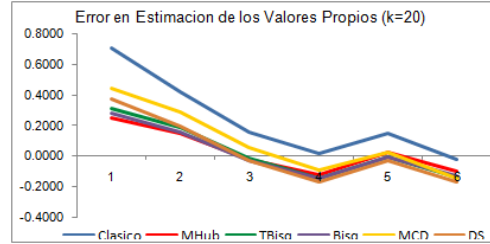
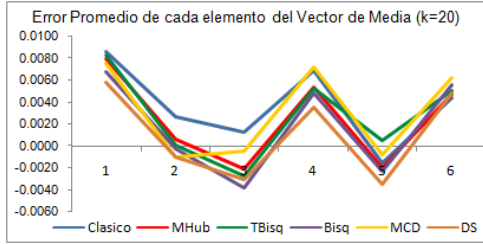
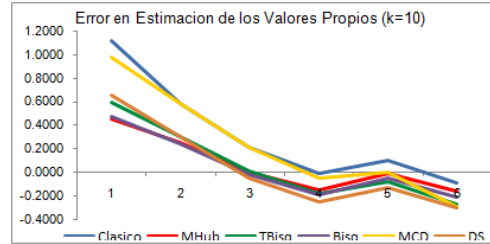
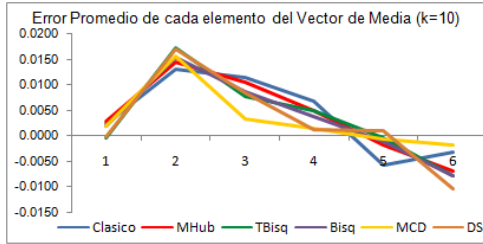
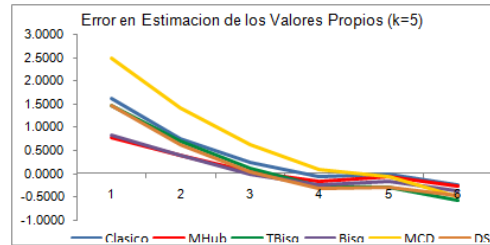
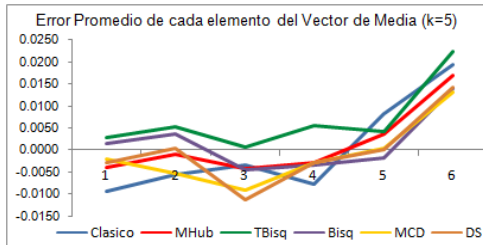
De acuerdo al Cuadro 4.12a, el Estimador M de Huber brinda, en todos los tamaños muestrales considerados, la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para $k=10$ obtiene un Índice de 1.225 resultado del Sesgo Promedio igual a 0.160, Sesgo Máximo de 2.543 y Desviación Total de 0.972; este Estimador también presenta mejor balance entre su estimación Promedio, “peor” estimación y dispersión en los resultados correspondientes la estimación del Vector de Medias y Valores Propios,

Además, como se muestra en la segunda parte del Cuadro 4.12a con respecto a cada Valor Propio, todos los estimadores sobreestiman los dos mayores Valores Propios así también subestiman el menor valor propio; se puede observar además que las estimaciones brindadas por el Estimador Clásico y el Estimador MCD en la mayoría de las ocasiones son mayores que las dadas por los demás Estimadores.

Cuadro 4.12a
Estimadores Robustos
 Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = N(0, 9)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.025 ± 0.691	1.603	0.773	0.020 ± 0.456	1.101	0.526	0.012 ± 0.337	0.709	0.353			
	Mhuber	0.018 ± 0.633	1.480	0.710	0.020 ± 0.434	1.082	0.512	0.011 ± 0.315	0.678	0.334			
	T-Bic	0.024 ± 0.813	2.145	0.994	0.021 ± 0.468	1.182	0.557	0.011 ± 0.329	0.719	0.353			
	Bic	0.016 ± 0.671	1.597	0.761	0.020 ± 0.447	1.087	0.518	0.011 ± 0.326	0.685	0.341			
	MCD	0.017 ± 0.801	1.915	0.911	0.016 ± 0.507	1.259	0.594	0.012 ± 0.348	0.763	0.374			
	DS	0.018 ± 0.804	1.940	0.921	0.022 ± 0.510	1.287	0.606	0.010 ± 0.351	0.756	0.372			
Valores Propios	Clásico	1.833 ± 1.232	6.967	3.344	1.280 ± 0.823	3.857	1.987	0.847 ± 0.509	2.383	1.247			
	Mhuber	0.939 ± 0.508	2.612	1.353	0.558 ± 0.456	1.544	0.853	0.334 ± 0.326	1.016	0.559			
	T-Bic	1.769 ± 1.185	6.240	3.065	0.747 ± 0.576	2.238	1.187	0.414 ± 0.356	1.123	0.631			
	Bic	1.025 ± 0.758	2.884	1.556	0.607 ± 0.515	1.763	0.962	0.375 ± 0.363	1.080	0.606			
	MCD	2.973 ± 1.420	8.225	4.206	1.191 ± 0.726	3.201	1.706	0.565 ± 0.390	1.486	0.814			
	DS	1.731 ± 1.290	10.502	4.508	0.832 ± 0.599	2.067	1.166	0.486 ± 0.387	1.188	0.687			
Matriz de Covarianzas	Clásico	0.965 ± 2.029	11.450	4.815	0.785 ± 1.480	6.883	3.050	0.583 ± 1.021	4.095	1.900			
	Mhuber	0.277 ± 1.112	4.336	1.909	0.160 ± 0.972	2.543	1.225	0.085 ± 0.663	1.724	0.824			
	T-Bic	0.465 ± 2.335	9.136	3.978	0.160 ± 1.203	3.198	1.520	0.083 ± 0.777	1.844	0.901			
	Bic	0.167 ± 1.584	4.345	2.032	0.104 ± 1.070	2.618	1.264	0.058 ± 0.726	1.735	0.840			
	MCD	1.679 ± 3.100	11.701	5.493	0.586 ± 1.545	4.702	2.278	0.245 ± 0.893	2.492	1.210			
	DS	0.434 ± 2.199	11.992	4.875	0.111 ± 1.295	3.300	1.569	0.071 ± 0.880	2.103	1.018			

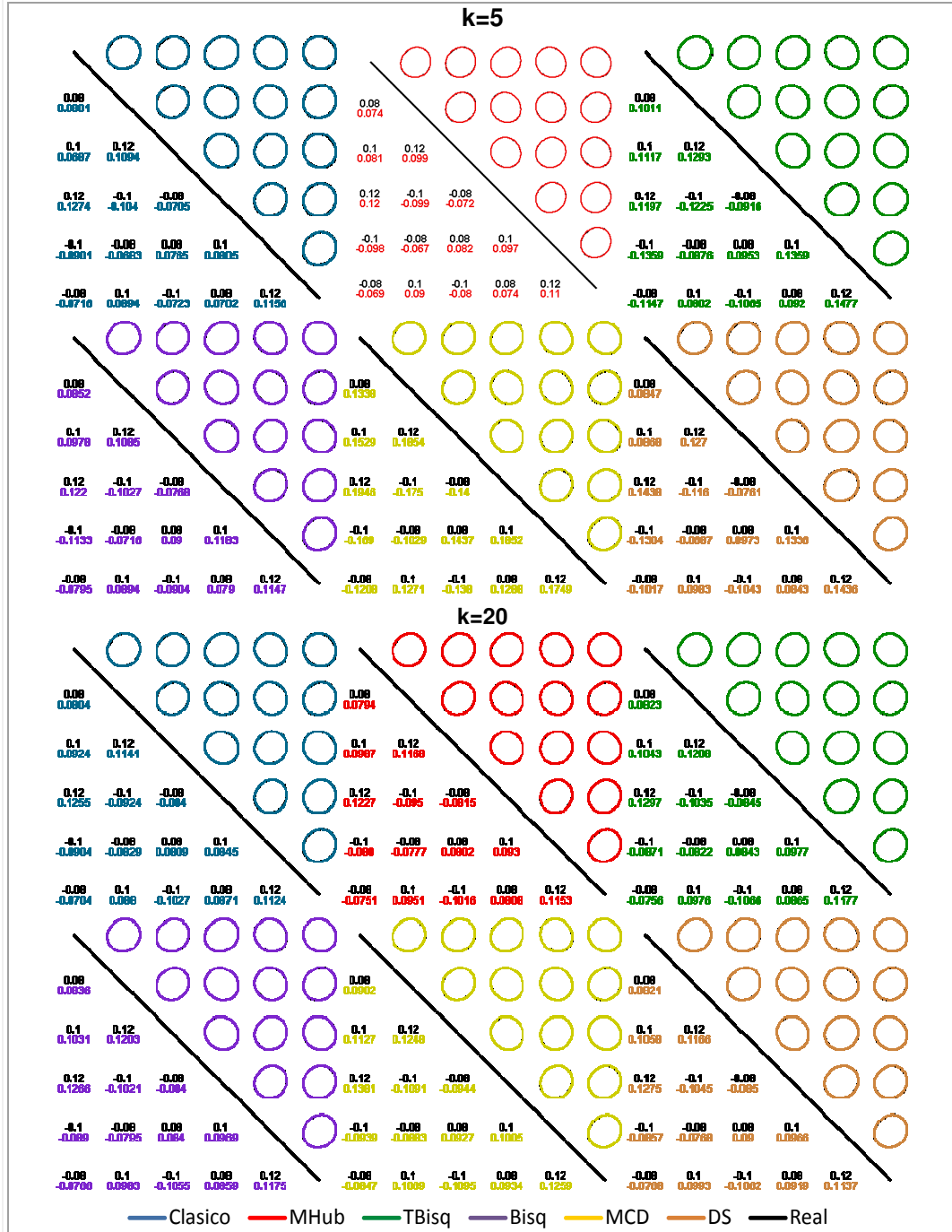


Elaborado por: Néstor Montaña P.

Cuadro 4.12b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada con $\varepsilon = 0.05$ y $H = N(0, 9)$



El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada

Por Variable con $\varepsilon = 0.05$ y $H = U(-5, 5)$

Como se muestra en el Cuadro 4.13a, para $k=5$ el promedio de la estimación del Vector de Medias resultante al utilizar el Método Bicuadrático se encuentra a una distancia 0.029 con respecto al Vector de Medias Real; y la estimación más alejada está a 1.384 unidades de distancia; además este método presenta una Desviación Total de 0.485, por lo tanto el Índice correspondiente es 0.659.

Al considerar el Índice planteado se aprecia que para todos los valores de k los Estimadores M de Huber y Bicuadrático son los que mejor se aproximan al valor real del Vector de Medias, Matriz de Varianzas y Covarianzas así como a los Valores Propios asociados a la misma.

En este escenario el Estimador Clásico no es el estimador que más sobrestima el mayor Valor Propio, sin embargo, si ocurre esto para el menor valor propio, evento que puede causar que se pierda una oportunidad de reducir el número de variables, pues el último Valor Propio asociado a Σ_1 es 0.002.

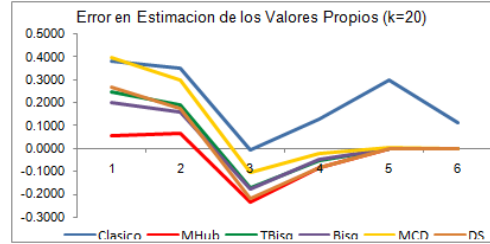
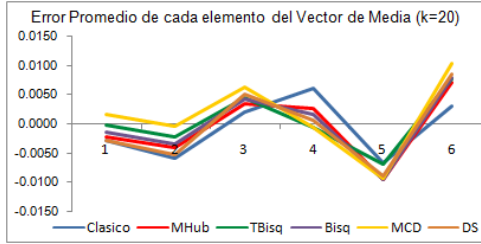
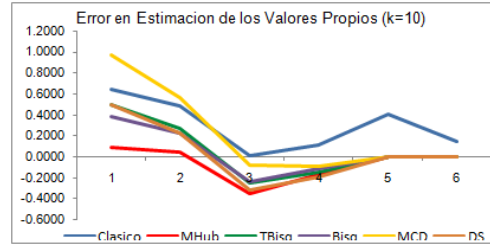
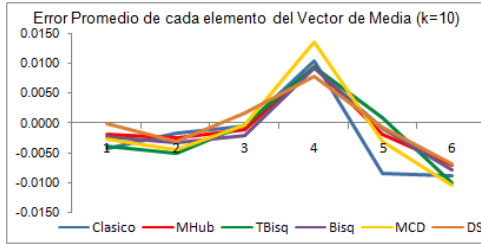
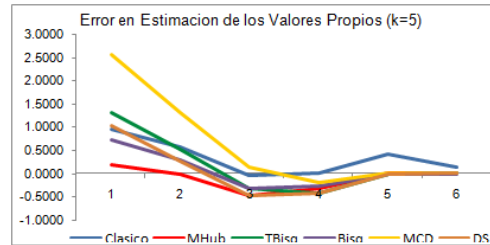
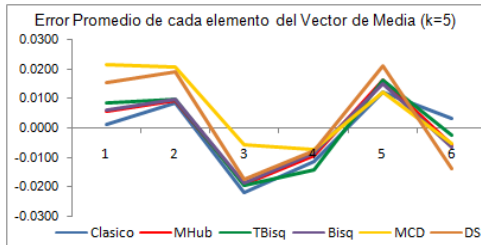
Cuadro 4.13a

Estimadores Robustos

Caso Población Normal $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y $H = U(-5,5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.029 ± 0.674	1.452	0.718	0.017 ± 0.464	1.064	0.515	0.012 ± 0.322	0.726	0.353			
	Mhuber	0.030 ± 0.632	1.337	0.666	0.012 ± 0.458	0.995	0.488	0.013 ± 0.305	0.713	0.344			
	T-Bic	0.032 ± 0.775	1.722	0.843	0.015 ± 0.482	1.065	0.521	0.011 ± 0.326	0.742	0.360			
	Bic	0.029 ± 0.659	1.384	0.690	0.013 ± 0.469	1.023	0.501	0.014 ± 0.317	0.708	0.346			
	MCD	0.034 ± 0.730	1.699	0.821	0.018 ± 0.522	1.172	0.571	0.015 ± 0.332	0.770	0.373			
	DS	0.040 ± 0.767	1.658	0.822	0.011 ± 0.506	1.147	0.555	0.015 ± 0.333	0.725	0.357			
Valores Propios	Clásico	1.196 ± 1.009	3.600	1.935	0.925 ± 0.772	2.829	1.509	0.620 ± 0.572	2.013	1.068			
	Mhuber	0.601 ± 0.581	2.281	1.154	0.402 ± 0.595	1.782	0.927	0.263 ± 0.437	1.183	0.628			
	T-Bic	1.515 ± 1.404	6.759	3.226	0.641 ± 0.794	3.233	1.556	0.357 ± 0.523	1.616	0.832			
	Bic	0.895 ± 1.094	4.115	2.034	0.516 ± 0.726	2.811	1.351	0.319 ± 0.523	1.517	0.786			
	MCD	2.876 ± 1.896	9.029	4.600	1.124 ± 0.970	3.990	2.028	0.507 ± 0.555	1.909	0.991			
	DS	1.237 ± 1.387	10.051	4.225	0.655 ± 0.784	3.132	1.524	0.395 ± 0.556	1.656	0.869			
Matriz de Covarianzas	Clásico	0.853 ± 1.911	5.721	2.828	0.740 ± 1.389	4.116	2.082	0.518 ± 0.933	2.706	1.386			
	Mhuber	0.280 ± 0.931	3.283	1.498	0.177 ± 0.877	2.138	1.064	0.092 ± 0.656	1.458	0.735			
	T-Bic	0.495 ± 2.299	8.461	3.752	0.179 ± 1.226	3.480	1.628	0.098 ± 0.802	1.968	0.956			
	Bic	0.217 ± 1.607	4.816	2.213	0.129 ± 1.082	2.870	1.361	0.067 ± 0.757	1.790	0.871			
	MCD	1.761 ± 3.018	10.940	5.240	0.640 ± 1.544	4.447	2.210	0.264 ± 0.886	2.289	1.146			
	DS	0.198 ± 2.007	8.630	3.611	0.098 ± 1.294	3.517	1.636	0.065 ± 0.878	2.140	1.028			

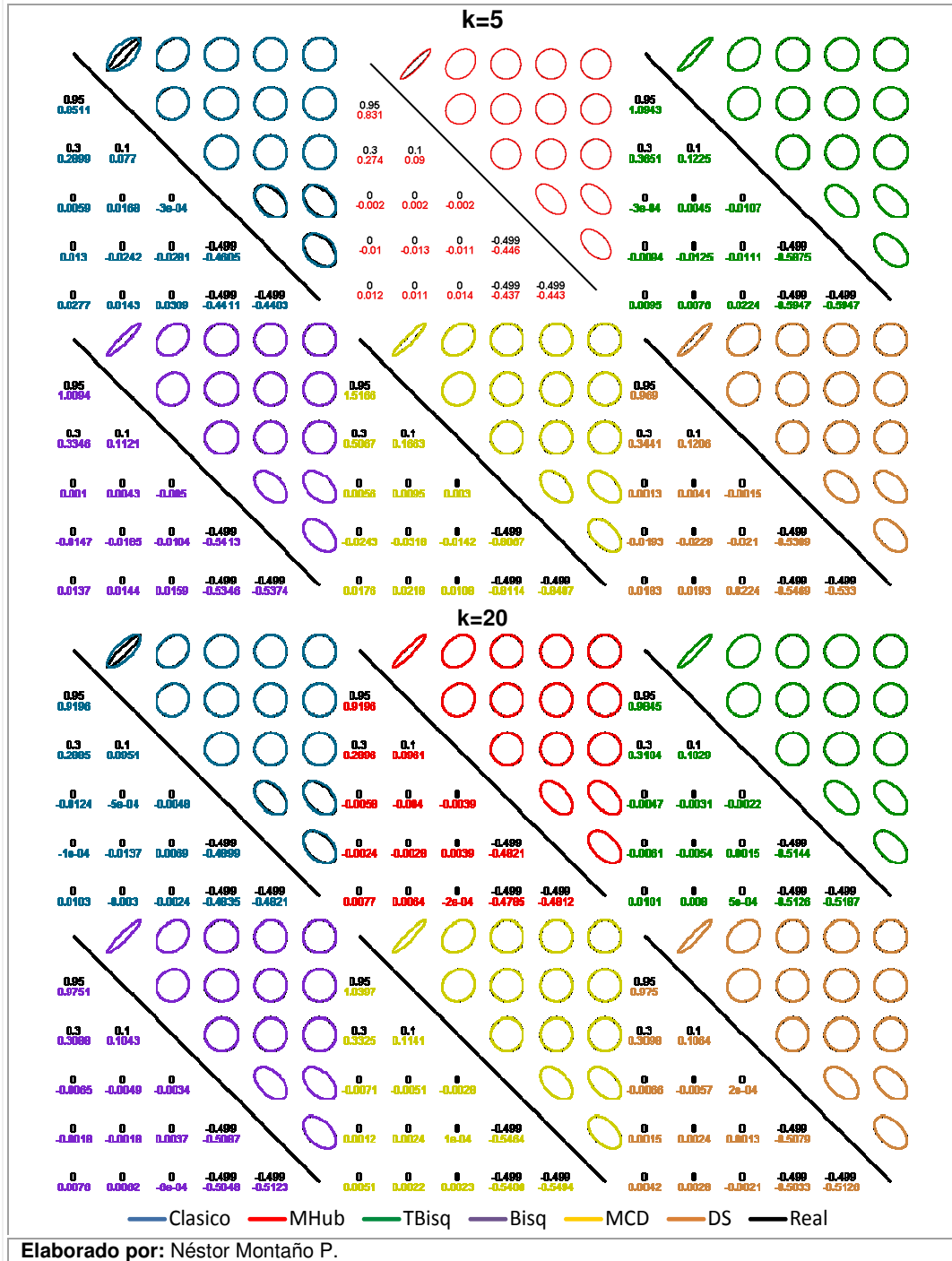


Elaborado por: Néstor Montaña P.

Cuadro 4.13b

Estimadores Robustos

Caso Población Normal $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y $H = U(-5, 5)$



El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada

Por Variable con $\varepsilon = 0.05$ y $H = U(-5,5)$

En el cuadro 4.14a se puede observar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador DS para $k=10$ subestima la primer componente por 0.002 unidades, además sobreestima la segunda, tercer y sexta componentes por 0.0042, 0.0048 y 0.0067 respectivamente y subestima la cuarta y quinta componente por 0.0103 y 0.0094 unidades, valores que proporcionan un Sesgo Promedio Total de 0.017.

Como en casos anteriores, el promedio de la Matriz de Varianzas y Covarianzas resultante al utilizar el Estimador DS es siempre la mas cercana a Matriz de Varianzas y Covarianzas Real, así como su "peor" estimación puede estar incluso mas alejada que la peor estimación del Estimador Clásico, nótese los valores correspondientes a $k=5$.

La estimación de la Matriz de Covarianzas correspondiente al Estimador Bicuadrático y $k=5$ converge a los valores 0.061, 0.101, 0.114, -0.102, -0.085 en su primer columna, mientras que para $k=20$

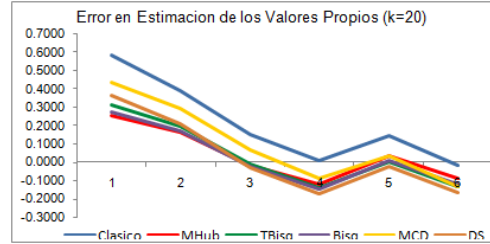
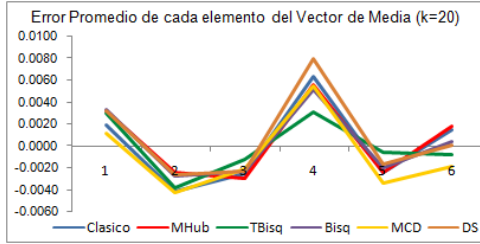
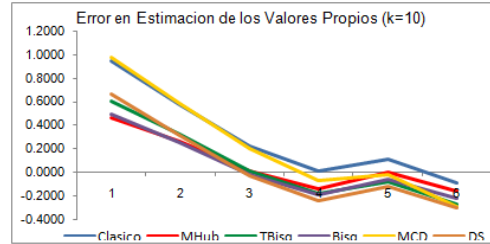
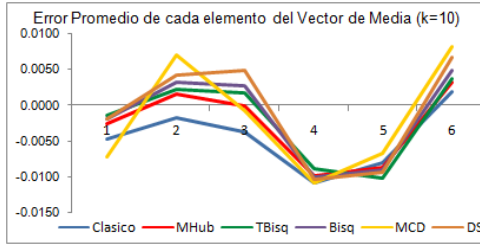
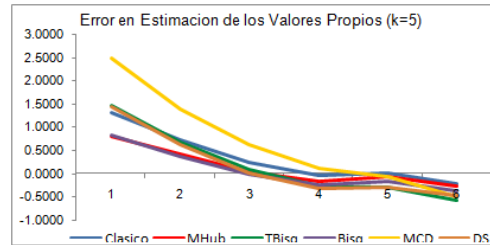
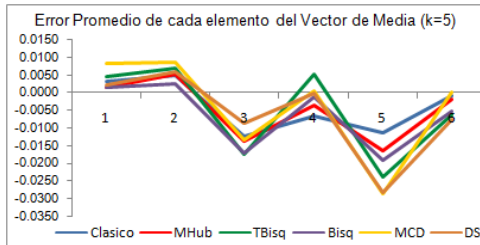
Cuadro 4.14a

Estimadores Robustos

Caso Población Normal $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.05$ y $H = U(-5, 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.019 ± 0.660	1.620	0.766	0.015 ± 0.471	1.132	0.539	0.009 ± 0.338	0.668	0.338			
	Mhuber	0.022 ± 0.636	1.541	0.733	0.014 ± 0.444	1.063	0.507	0.008 ± 0.312	0.697	0.339			
	T-Bic	0.032 ± 0.779	1.759	0.856	0.014 ± 0.488	1.126	0.543	0.006 ± 0.324	0.712	0.347			
	Bic	0.027 ± 0.650	1.617	0.765	0.015 ± 0.464	1.051	0.510	0.007 ± 0.317	0.702	0.342			
	MCD	0.034 ± 0.753	1.768	0.852	0.018 ± 0.521	1.198	0.579	0.008 ± 0.341	0.784	0.377			
	DS	0.031 ± 0.763	1.830	0.875	0.017 ± 0.524	1.185	0.575	0.009 ± 0.337	0.776	0.374			
Valores Propios	Clásico	1.553 ± 0.886	3.951	2.130	1.141 ± 0.636	2.889	1.555	0.736 ± 0.437	1.790	0.988			
	Mhuber	0.957 ± 0.689	2.825	1.490	0.575 ± 0.457	1.544	0.859	0.339 ± 0.323	1.114	0.592			
	T-Bic	1.746 ± 1.146	5.876	2.923	0.763 ± 0.558	1.873	1.064	0.417 ± 0.358	1.276	0.684			
	Bic	1.031 ± 0.793	2.831	1.552	0.628 ± 0.519	1.689	0.945	0.375 ± 0.347	1.205	0.642			
	MCD	2.962 ± 1.532	7.475	3.990	1.198 ± 0.705	2.938	1.614	0.553 ± 0.403	1.550	0.836			
	DS	1.700 ± 1.174	21.252	8.042	0.839 ± 0.604	2.336	1.260	0.486 ± 0.402	1.338	0.742			
Matriz de Covarianzas	Clásico	0.840 ± 1.956	5.678	2.824	0.733 ± 1.406	4.033	2.057	0.521 ± 0.917	2.701	1.379			
	Mhuber	0.297 ± 1.446	4.503	2.082	0.180 ± 0.992	2.699	1.290	0.096 ± 0.666	1.748	0.837			
	T-Bic	0.448 ± 2.242	8.297	3.662	0.166 ± 1.211	3.353	1.577	0.095 ± 0.765	1.902	0.921			
	Bic	0.165 ± 1.553	5.047	2.255	0.112 ± 1.065	2.811	1.329	0.067 ± 0.730	1.823	0.873			
	MCD	1.672 ± 3.063	11.412	5.382	0.585 ± 1.550	4.680	2.272	0.249 ± 0.871	2.370	1.163			
	DS	0.401 ± 2.158	16.901	6.487	0.115 ± 1.287	3.618	1.673	0.080 ± 0.873	2.334	1.096			

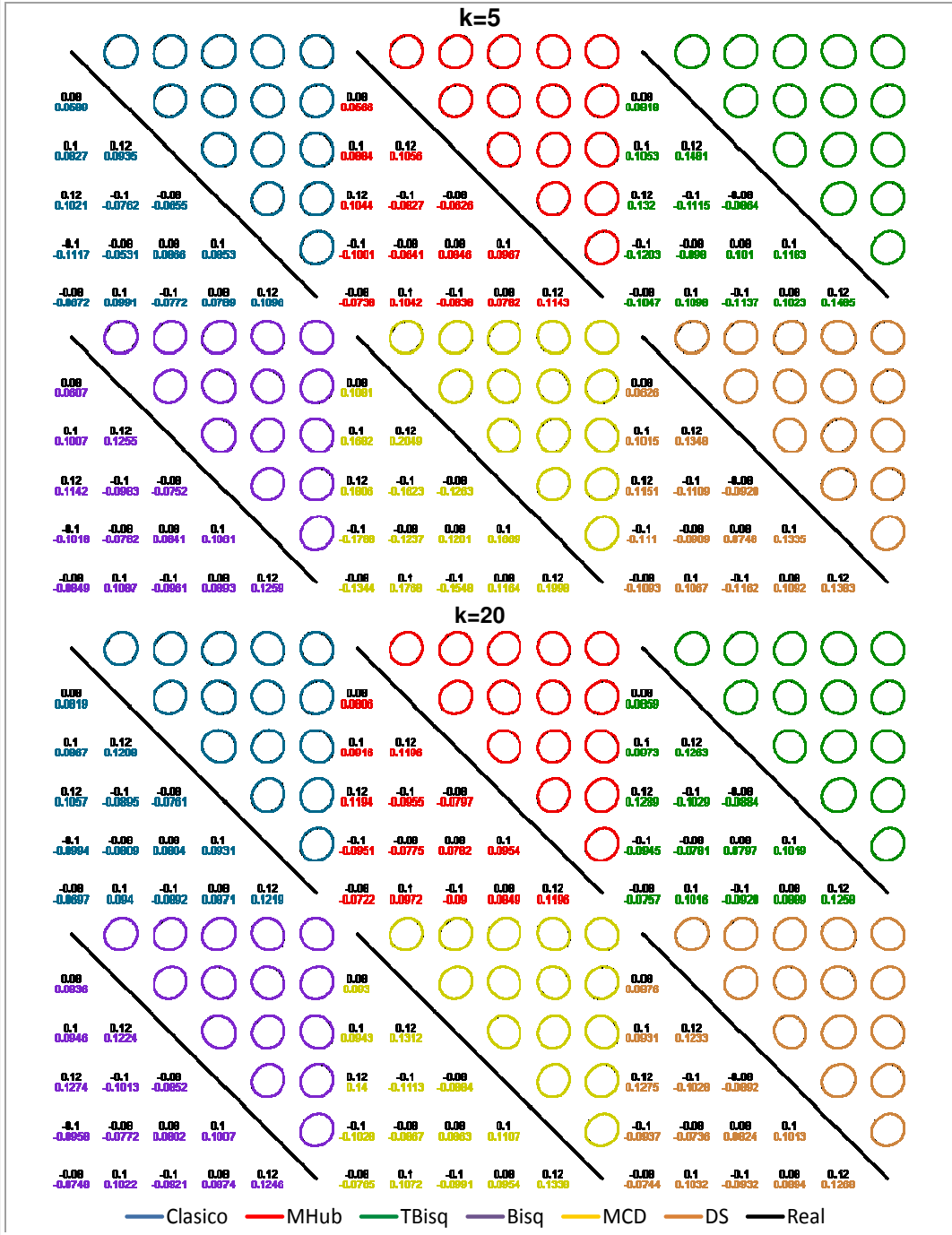


Elaborado por: Néstor Montaña P.

Cuadro 4.14 b

Estimadores Robustos

Caso Población Normal $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.05$ y $H = U(-5,5)$



Elaborado por: Néstor Montaña P.

esos valores son 0.084, 0.095, 0.127, -0.096, -0.075 (Véase el Cuadro 4.14b).

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_1)$

De acuerdo al Cuadro 4.15a, el Estimador M de Huber brinda, para $k=10$ y 20 , la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para $k=20$ obtiene un Índice de 0.877 resultado del Sesgo Promedio igual a 0.465, Sesgo Máximo de 1.646 y Desviación Total de 0.521, mientras que para $k=5$ el Estimador Bicuadrático el que mejor balance entre su estimación Promedio, “peor” estimación y dispersión.

En lo referente al Vector de Medias los resultados varían, el Estimador Bicuadrático brinda la “mejor” estimación para $k=5$ y 10 , mientras que para $k=20$ es el Estimador T-Bicuadrático el que presenta menor índice.

Nótese además que, con respecto a cada Valor Propio, todos los estimadores sobreestiman los dos mayores Valores Propios; para $k=20$ se puede observar que el promedio de las estimaciones brindadas por el Estimador Clásico sobreestima el mayor Valor Propio con error que cuadriplica el error correspondiente a las

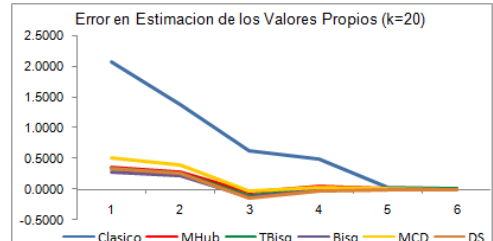
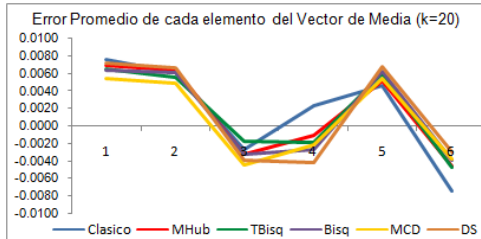
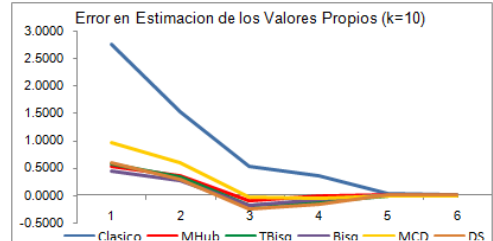
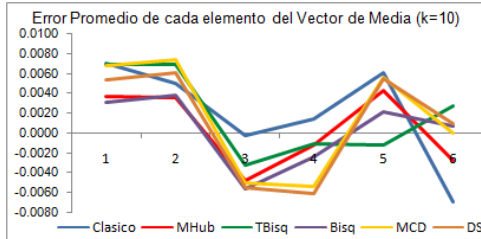
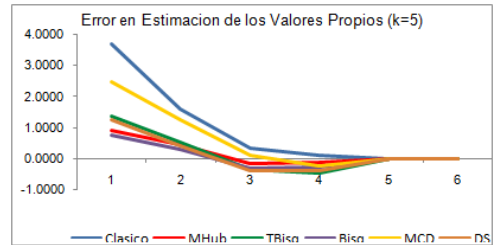
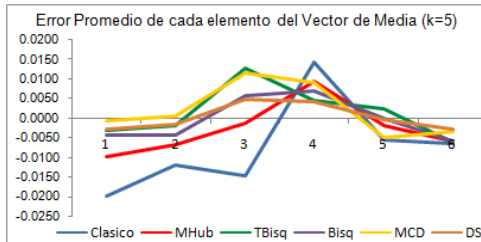
Cuadro 4.15a

Estimadores Robustos

Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_1)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

	5				10				20			
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.032 ± 0.792	1.648	0.824	0.013 ± 0.572	1.241	0.609	0.014 ± 0.387	0.926	0.442		
	Mhuber	0.016 ± 0.649	1.455	0.707	0.009 ± 0.457	1.061	0.509	0.012 ± 0.321	0.809	0.380		
	T-Bic	0.015 ± 0.793	1.707	0.838	0.011 ± 0.455	1.078	0.515	0.012 ± 0.322	0.758	0.364		
	Bic	0.012 ± 0.651	1.424	0.696	0.008 ± 0.448	1.022	0.493	0.012 ± 0.318	0.770	0.367		
	MCD	0.016 ± 0.762	1.664	0.814	0.014 ± 0.507	1.187	0.569	0.011 ± 0.329	0.830	0.390		
	DS	0.008 ± 0.756	1.611	0.792	0.013 ± 0.505	1.119	0.546	0.014 ± 0.328	0.764	0.368		
Valores Propios	Clásico	4.043 ± 2.526	12.767	6.445	3.213 ± 1.742	13.298	6.085	2.619 ± 1.166	5.937	3.241		
	Mhuber	1.057 ± 0.792	4.072	1.974	0.651 ± 0.676	2.274	1.200	0.465 ± 0.521	1.646	0.877		
	T-Bic	1.566 ± 1.446	8.265	3.759	0.685 ± 0.755	2.378	1.273	0.441 ± 0.571	1.798	0.936		
	Bic	0.919 ± 0.976	3.372	1.756	0.550 ± 0.698	2.600	1.282	0.390 ± 0.571	1.682	0.881		
	MCD	2.794 ± 1.807	9.425	4.675	1.142 ± 0.901	4.077	2.040	0.638 ± 0.624	2.249	1.170		
	DS	1.447 ± 1.442	10.816	4.568	0.727 ± 0.837	3.643	1.736	0.445 ± 0.583	1.735	0.921		
Matriz de Covarianzas	Clásico	2.666 ± 3.716	13.479	6.620	2.394 ± 2.619	11.817	5.610	2.129 ± 1.748	6.455	3.444		
	Mhuber	0.511 ± 1.217	5.035	2.254	0.361 ± 1.042	2.799	1.401	0.298 ± 0.734	1.870	0.967		
	T-Bic	0.530 ± 2.297	8.458	3.762	0.273 ± 1.219	3.272	1.588	0.228 ± 0.802	2.096	1.042		
	Bic	0.269 ± 1.553	4.529	2.117	0.200 ± 1.085	3.012	1.432	0.190 ± 0.763	1.950	0.968		
	MCD	1.725 ± 2.884	11.164	5.258	0.692 ± 1.514	4.683	2.296	0.422 ± 0.892	2.630	1.315		
	DS	0.465 ± 2.095	9.738	4.099	0.214 ± 1.347	4.117	1.893	0.198 ± 0.852	2.205	1.085		

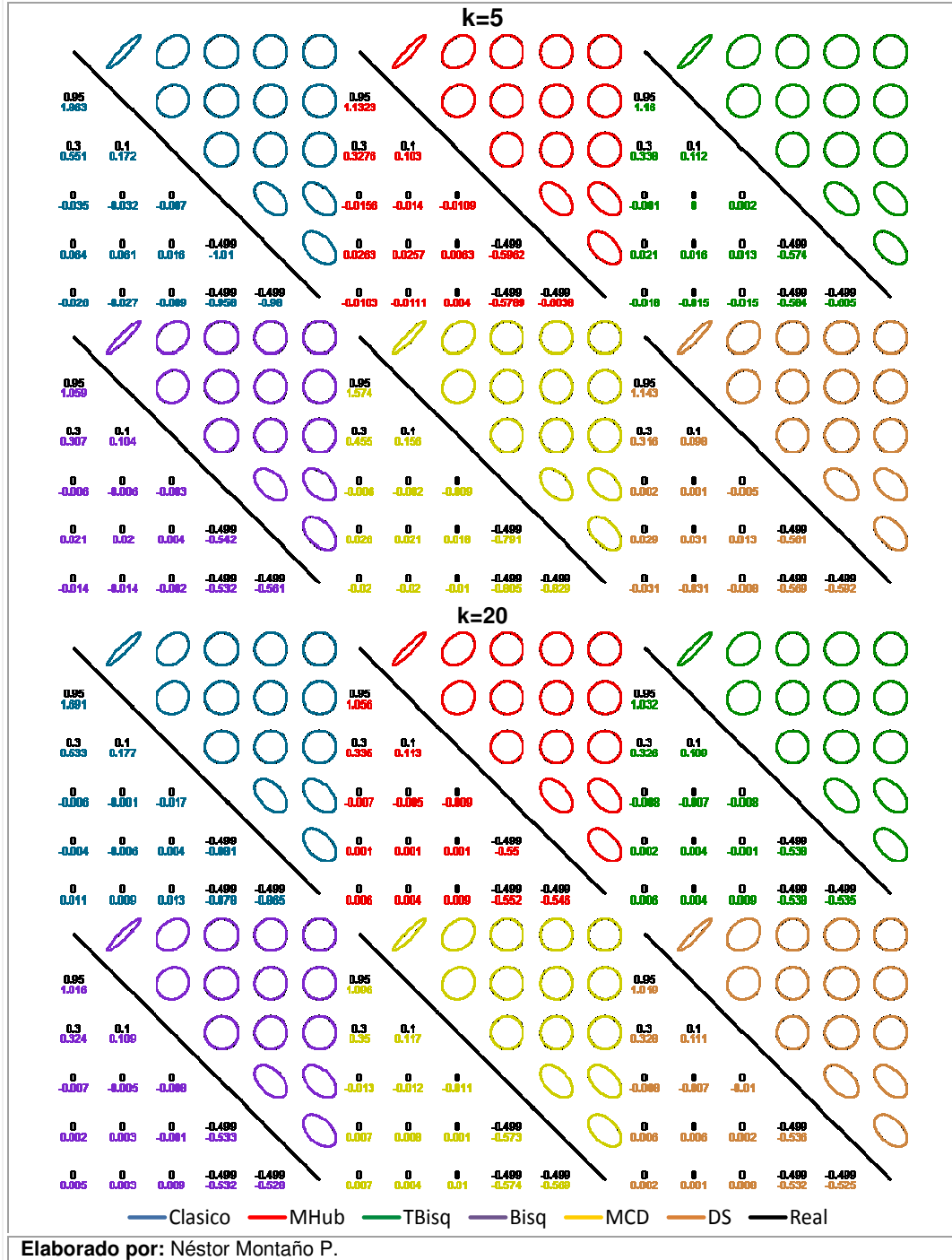


Elaborado por: Néstor Montaña P.

Cuadro 4.15b

Estimadores Robustos

Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_1)$



estimaciones dadas por los demás Estimadores.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_2)$

Como se muestra en el Cuadro 4.16a, para $k=5$ el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador MCD se encuentra a una distancia 0.022 con respecto al Vector de Medias Real; y la estimación más alejada está a 1.753 unidades de distancia; además este método presenta una Desviación Total de 0.762, por lo tanto el Índice correspondiente es 0.844.

Al considerar el Índice planteado se aprecia que para todos los valores de k los Estimadores M de Huber y Bicuadrático son los que mejor se aproximan al valor real del Vector de Medias así como la Matriz de Varianzas y Covarianzas.

En este escenario, al igual que el anterior, para la Matriz de Covarianzas y los Valores Propios asociados a la misma el Estimador Clásico presenta, en todos los tamaños muestrales, el mayor Índice haciendo indudable la necesidad usar Estimadores Robustos.

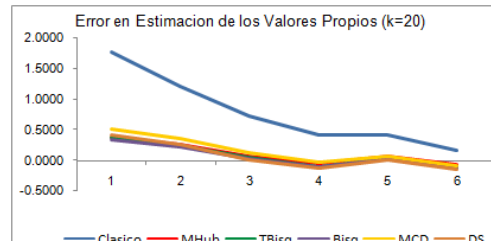
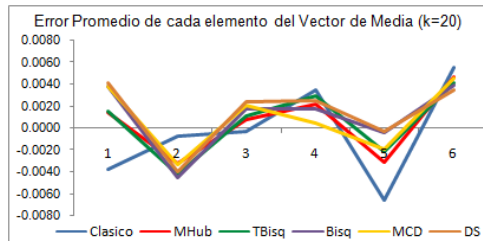
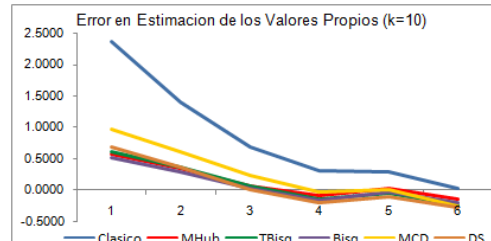
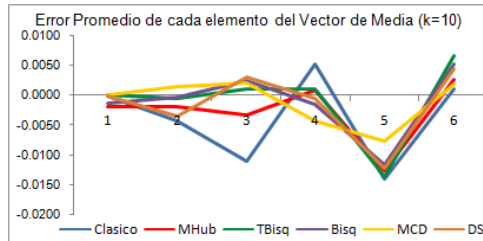
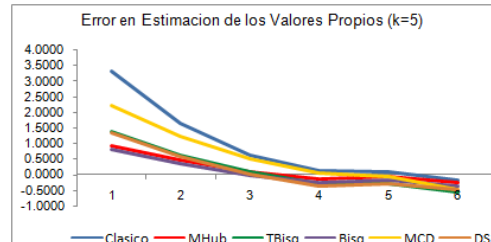
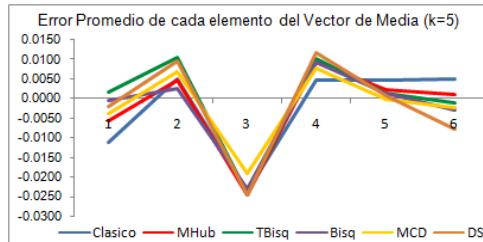
Cuadro 4.16a

Estimadores Robustos

Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_2)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.028 ± 0.772	1.795	0.865	0.019 ± 0.558	1.205	0.594	0.010 ± 0.369	0.779	0.386			
	Mhuber	0.027 ± 0.622	1.448	0.699	0.014 ± 0.462	0.991	0.489	0.007 ± 0.315	0.702	0.341			
	T-Bic	0.028 ± 0.778	1.808	0.871	0.015 ± 0.495	1.018	0.509	0.007 ± 0.322	0.709	0.346			
	Bic	0.025 ± 0.634	1.575	0.745	0.013 ± 0.474	0.976	0.488	0.007 ± 0.316	0.701	0.341			
	MCD	0.022 ± 0.771	1.753	0.849	0.009 ± 0.516	1.099	0.541	0.007 ± 0.330	0.767	0.368			
	DS	0.030 ± 0.762	1.740	0.844	0.014 ± 0.527	1.075	0.539	0.008 ± 0.329	0.730	0.356			
Valores Propios	Clásico	3.775 ± 1.844	10.563	5.394	2.859 ± 1.283	6.134	3.426	2.345 ± 0.851	4.823	2.673			
	Mhuber	1.090 ± 0.720	3.928	1.913	0.700 ± 0.494	1.728	0.974	0.475 ± 0.330	1.493	0.766			
	T-Bic	1.656 ± 1.063	4.903	2.540	0.774 ± 0.558	2.012	1.115	0.480 ± 0.380	1.348	0.736			
	Bic	1.023 ± 0.786	3.313	1.707	0.654 ± 0.523	1.711	0.963	0.434 ± 0.364	1.179	0.659			
	MCD	2.653 ± 1.384	7.067	3.701	1.187 ± 0.679	3.396	1.754	0.648 ± 0.398	1.530	0.859			
	DS	1.597 ± 1.175	10.837	4.536	0.851 ± 0.585	2.475	1.304	0.512 ± 0.398	1.316	0.742			
Matriz de Covarianzas	Clásico	2.339 ± 3.433	13.963	6.578	2.082 ± 2.488	9.153	4.574	1.936 ± 1.771	6.716	3.474			
	Mhuber	0.408 ± 1.494	4.423	2.108	0.329 ± 1.019	2.713	1.354	0.270 ± 0.709	2.027	1.002			
	T-Bic	0.420 ± 2.181	7.586	3.396	0.255 ± 1.202	3.247	1.568	0.205 ± 0.804	2.080	1.029			
	Bic	0.191 ± 1.487	4.521	2.067	0.197 ± 1.075	2.758	1.343	0.172 ± 0.755	1.966	0.964			
	MCD	1.462 ± 2.815	9.664	4.647	0.637 ± 1.499	4.375	2.170	0.381 ± 0.889	2.383	1.218			
	DS	0.364 ± 2.069	11.823	4.752	0.196 ± 1.325	3.658	1.726	0.182 ± 0.851	2.212	1.082			

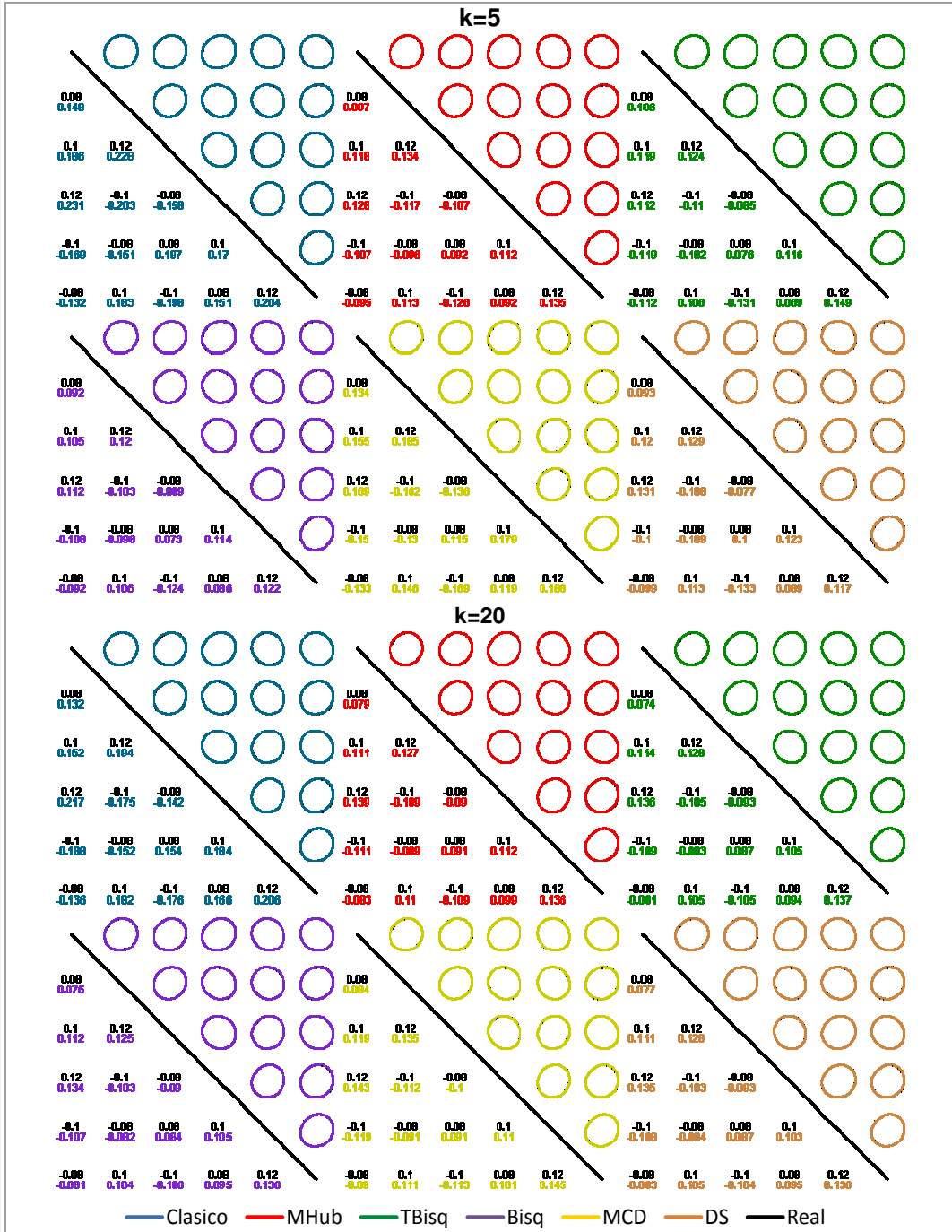


Elaborado por: Néstor Montaña P.

Cuadro 4.16b

Estimadores Robustos

Caso Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(\mathbf{0}, 9\Sigma_2)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.10$ y $H = Uesf(d = 5)$

En el cuadro 4.17a se puede apreciar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador M de Huber para $k=5$ subestima la primer componente por 0.0058 unidades, además sobreestima la segunda, cuarta y quinta componentes por 0.0022, 0.0072 y 0.0009 respectivamente y subestima también la tercer y sexta componente por 0.0062 y 0.0095 unidades, valores que proporcionan un Sesgo Promedio Total de 0.014.

El promedio de la Matriz de Varianzas y Covarianzas resultante al utilizar el Estimador DS es siempre la mas cercana a Matriz de Varianzas y Covarianzas Real, así como su “peor” estimación puede estar incluso mas alejada que la peor estimación del Estimador Clásico, nótese los valores correspondientes a $k=5$.

La estimación de la Matriz de Varianzas Covarianzas correspondiente al Estimador DS y $k=5$ converge a los valores 1.005, 0.298, -0.021, 0.006, 0.014 en su primer columna, mientras que para $k=20$ converge a 1.017, 0.321, 0.007, -0.007, 0. (Véase Cuadro 4.17b)

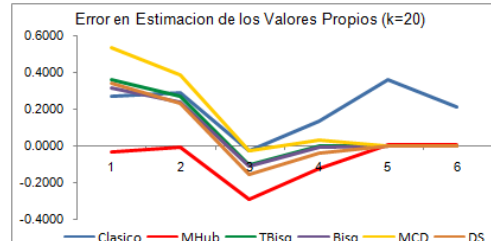
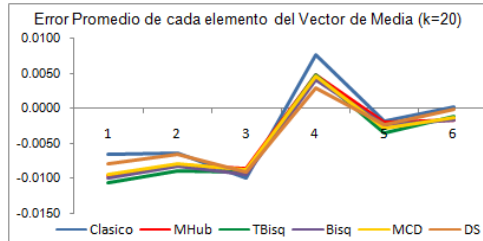
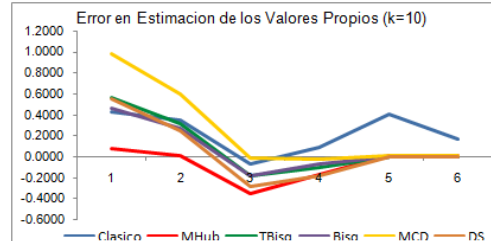
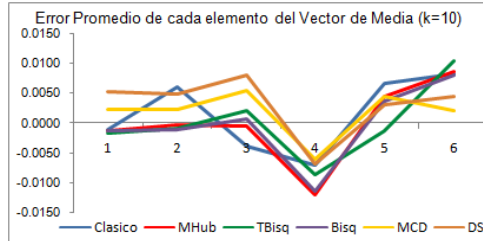
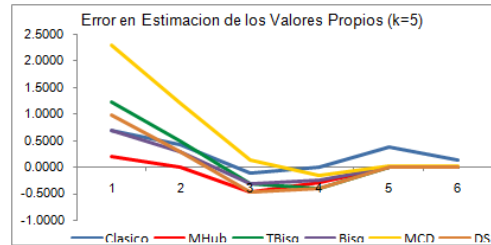
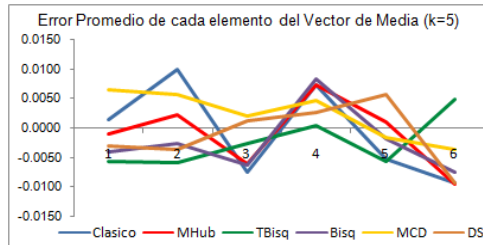
Cuadro 4.17a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada con $\varepsilon = 0.10$ y $H = Uesf(d = 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.018 ± 0.640	1.378	0.679	0.015 ± 0.453	0.972	0.480	0.016 ± 0.320	0.690	0.342			
	Mhuber	0.014 ± 0.650	1.451	0.705	0.016 ± 0.437	0.975	0.476	0.016 ± 0.304	0.659	0.326			
	T-Bic	0.012 ± 0.812	1.846	0.890	0.014 ± 0.465	1.114	0.531	0.018 ± 0.315	0.764	0.366			
	Bic	0.014 ± 0.677	1.475	0.722	0.015 ± 0.449	0.991	0.485	0.017 ± 0.316	0.695	0.343			
	MCD	0.011 ± 0.789	1.761	0.854	0.010 ± 0.467	1.189	0.555	0.016 ± 0.331	0.768	0.371			
	DS	0.012 ± 0.785	1.751	0.849	0.014 ± 0.500	1.119	0.544	0.014 ± 0.326	0.759	0.366			
Valores Propios	Clásico	0.926 ± 0.836	3.875	1.879	0.710 ± 0.642	2.263	1.205	0.590 ± 0.487	1.666	0.914			
	Mhuber	0.586 ± 0.563	2.708	1.286	0.404 ± 0.566	1.260	0.743	0.320 ± 0.413	0.923	0.552			
	T-Bic	1.420 ± 1.309	6.877	3.202	0.679 ± 0.799	2.562	1.346	0.459 ± 0.569	1.888	0.972			
	Bic	0.847 ± 0.966	3.598	1.804	0.574 ± 0.696	2.368	1.213	0.413 ± 0.549	1.648	0.870			
	MCD	2.594 ± 1.731	8.948	4.425	1.150 ± 0.886	4.466	2.167	0.662 ± 0.571	2.210	1.148			
	DS	1.182 ± 1.209	5.750	2.714	0.690 ± 0.839	2.690	1.406	0.440 ± 0.582	1.760	0.927			
Matriz de Covarianzas	Clásico	0.649 ± 1.631	5.001	2.427	0.575 ± 1.109	3.226	1.637	0.523 ± 0.794	2.170	1.162			
	Mhuber	0.254 ± 0.901	3.279	1.478	0.209 ± 0.866	1.944	1.006	0.213 ± 0.614	1.367	0.732			
	T-Bic	0.447 ± 2.246	7.270	3.321	0.276 ± 1.214	3.272	1.587	0.244 ± 0.841	2.208	1.098			
	Bic	0.205 ± 1.554	4.320	2.026	0.223 ± 1.087	2.791	1.367	0.207 ± 0.775	2.078	1.020			
	MCD	1.596 ± 2.825	10.143	4.855	0.712 ± 1.452	4.369	2.178	0.436 ± 0.922	2.580	1.312			
	DS	0.177 ± 1.972	6.563	2.904	0.162 ± 1.327	3.471	1.653	0.178 ± 0.872	2.147	1.066			

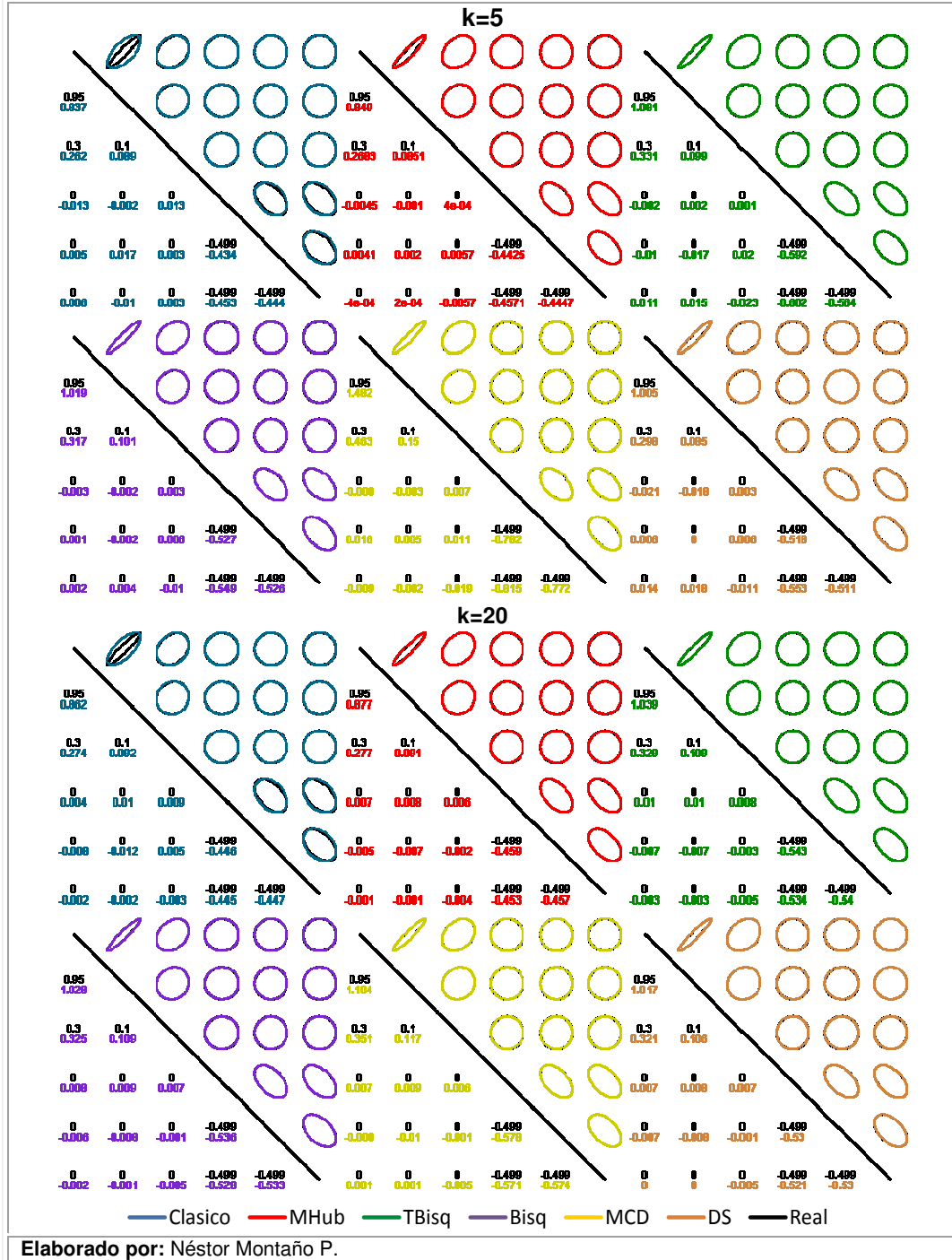


Elaborado por: Néstor Montaña P.

Cuadro 4.17 b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = Uesf(d = 5)$



El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.10$ y $H = Uesf(d = 5)$

De acuerdo al Cuadro 4.18a, el Estimador M de Huber brinda, para todos los valores de k, la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para k=10 obtiene un Índice de 0.885 resultado del Sesgo Promedio igual a 0.612, Sesgo Máximo de 1.579 y Desviación Total de 0.465.

En lo referente al Vector de Medias los resultados varían, el Estimador M de Huber brinda la “mejor” estimación para k=5 y 10, mientras que para k=20 su índice es igual al del Estimador Clásico presentando ambos el mejor balance entre su estimación Promedio, “peor” estimación y dispersión, sin embargo, en la estimación de los Valores Propios y la Matriz de Varianzas y Covarianzas el Estimador Clásico es superado por el estimador M de Huber, Bicuadrático y T-Bicuadrático.

Nótese nuevamente que, con respecto a cada Valor Propio, todos los estimadores sobreestiman los dos mayores Valores Propios; y es este escenario se puede observar que el estimador MCD es el que sobreestima con mayor error.

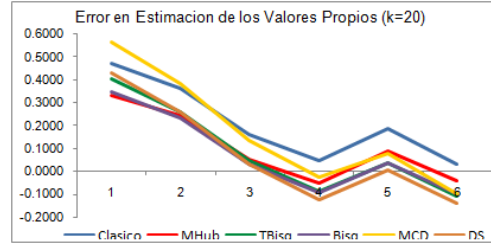
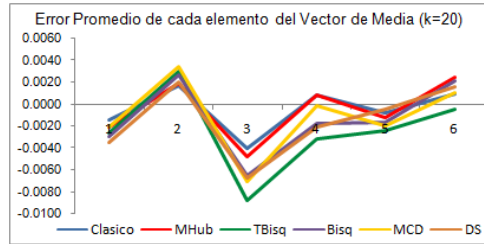
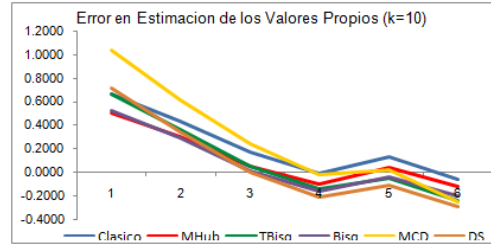
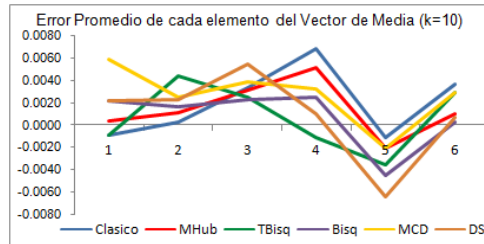
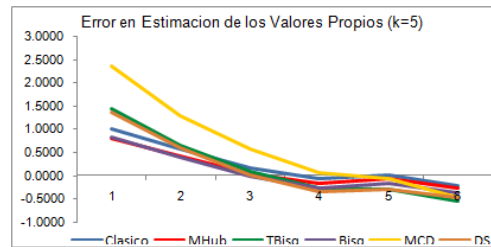
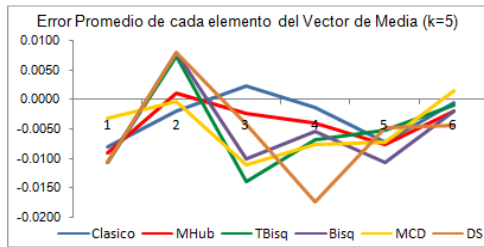
Cuadro 4.18a

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = Uesf(d = 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.011	± 0.634	1.465	0.703	0.008	± 0.442	1.005	0.485	0.005	± 0.319	0.682	0.335
	Mhuber	0.013	± 0.621	1.374	0.669	0.007	± 0.447	0.988	0.480	0.006	± 0.322	0.676	0.335
	T-Bic	0.021	± 0.793	1.766	0.860	0.007	± 0.492	1.059	0.519	0.010	± 0.338	0.705	0.351
	Bic	0.021	± 0.634	1.438	0.698	0.006	± 0.462	1.026	0.498	0.008	± 0.332	0.698	0.346
	MCD	0.016	± 0.746	1.887	0.883	0.009	± 0.522	1.238	0.590	0.008	± 0.345	0.771	0.375
	DS	0.023	± 0.768	1.822	0.871	0.009	± 0.526	1.351	0.629	0.008	± 0.350	0.713	0.357
Valores Propios	Clásico	1.199	± 0.656	2.889	1.581	0.824	± 0.456	1.820	1.033	0.645	± 0.331	1.438	0.805
	Mhuber	0.968	± 0.658	2.913	1.513	0.612	± 0.465	1.579	0.885	0.430	± 0.347	1.218	0.665
	T-Bic	1.711	± 1.135	6.645	3.164	0.810	± 0.578	2.277	1.222	0.507	± 0.401	1.372	0.760
	Bic	1.038	± 0.747	3.246	1.677	0.651	± 0.512	1.645	0.936	0.444	± 0.381	1.375	0.733
	MCD	2.798	± 1.390	7.408	3.865	1.250	± 0.709	3.141	1.700	0.706	± 0.416	1.705	0.943
	DS	1.624	± 1.175	9.764	4.188	0.875	± 0.614	2.354	1.281	0.536	± 0.417	1.509	0.821
Matriz de Covarianzas	Clásico	0.608	± 1.572	4.485	2.222	0.546	± 1.101	3.011	1.553	0.516	± 0.771	2.424	1.237
	Mhuber	0.326	± 1.432	4.050	1.936	0.276	± 0.999	2.716	1.330	0.256	± 0.702	2.080	1.013
	T-Bic	0.453	± 2.209	7.943	3.535	0.257	± 1.206	3.475	1.646	0.230	± 0.820	2.203	1.085
	Bic	0.198	± 1.548	4.317	2.021	0.183	± 1.079	2.846	1.370	0.189	± 0.765	2.159	1.038
	MCD	1.546	± 2.921	10.005	4.824	0.676	± 1.545	4.749	2.323	0.426	± 0.937	2.808	1.390
	DS	0.400	± 2.103	11.775	4.759	0.191	± 1.350	3.726	1.756	0.198	± 0.880	2.489	1.189

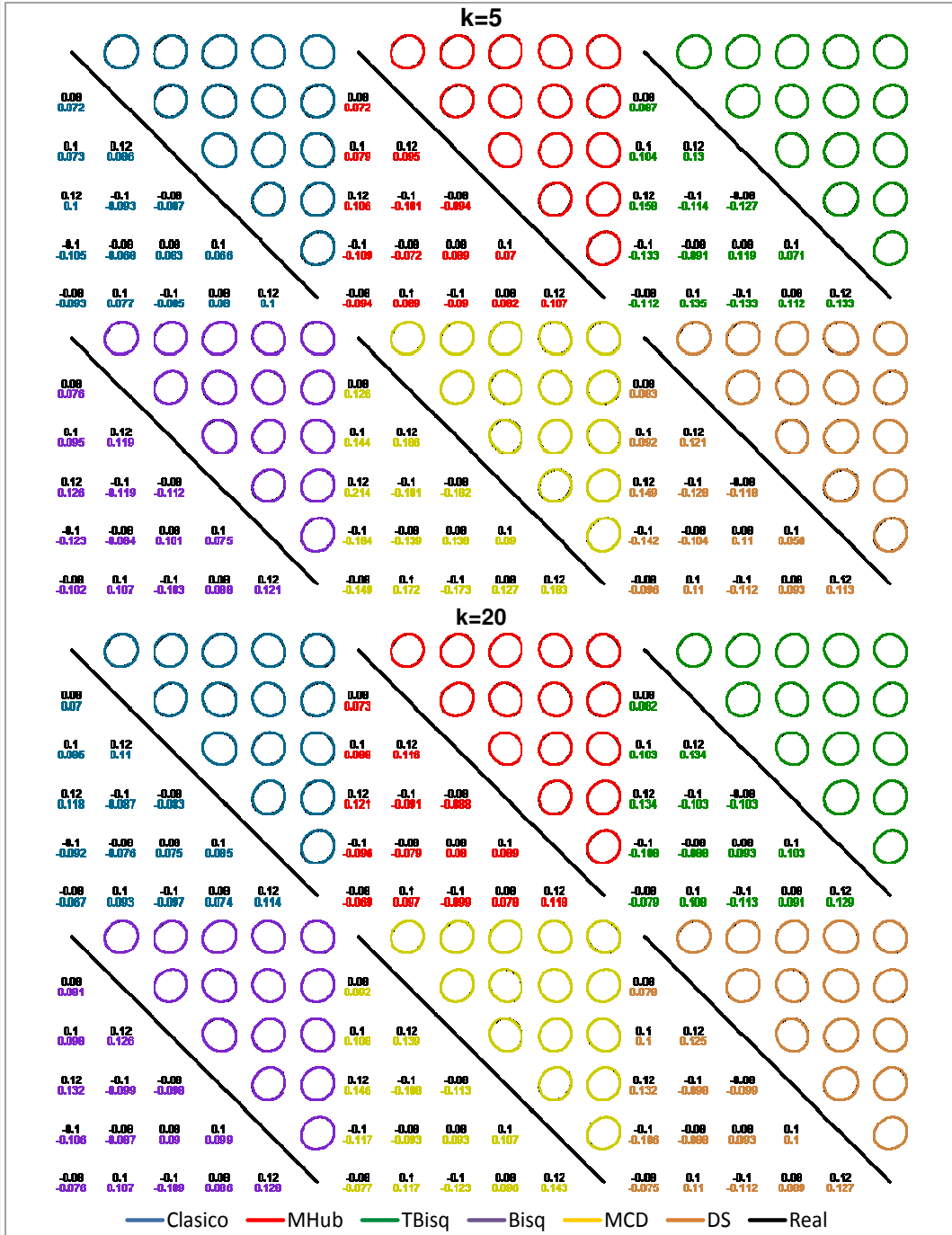


Elaborado por: Néstor Montaña P.

Cuadro 4.18b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = Uesf(d = 5)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada bajo el modelo (2.1) con $\varepsilon = 0.10$ y $H = N(0.537\alpha_p, 9\Sigma_1)$

Como se muestra en el Cuadro 4.19a, para $k=20$ el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador T-Bicuadrático se encuentra a una distancia 0.018 con respecto al Vector de Medias Real; y la estimación más alejada está a 1.111 unidades de distancia; además este método presenta una Desviación Total de 0.453, por lo tanto el Índice correspondiente es 0.527.

Al considerar el Índice planteado se observa que para $k=5$ y 10 los Estimadores M de Huber y Bicuadrático son los que mejor se aproximan al valor real del Vector de Medias mientras que para $k=20$ los estimadores Bicuadrático y T-Bicuadrático son los que alcanzan menor índice.

En este escenario, al igual que el anterior, para la Matriz de Covarianzas y los Valores Propios asociados a la misma el Estimador Clásico presenta, en todos los tamaños muestrales, el mayor Índice haciendo indudable la necesidad usar Estimadores Robustos.

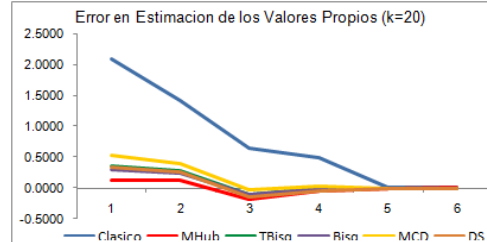
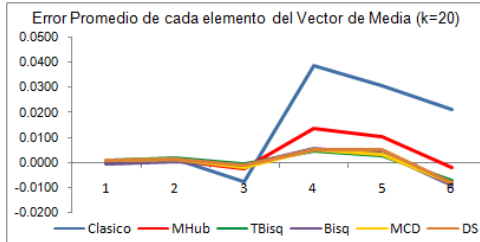
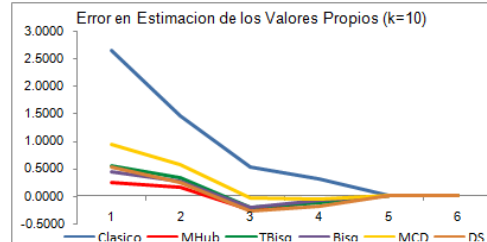
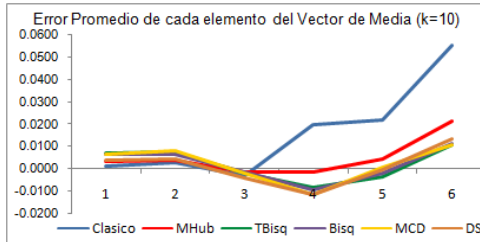
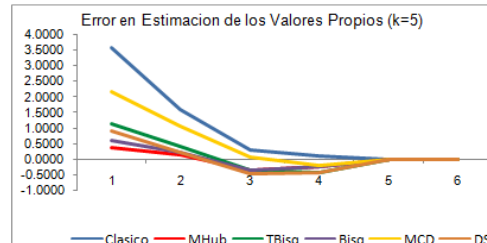
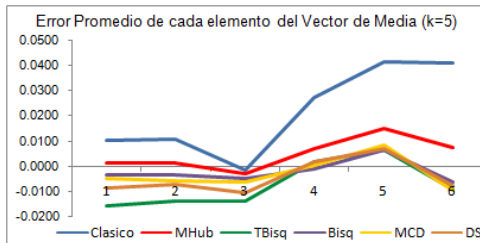
Cuadro 4.19a

Estimadores Robustos

Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.066	± 0.757	1.833	0.885	0.063	± 0.534	1.243	0.613	0.054	± 0.377	0.919	0.450
	Mhuber	0.018	± 0.614	1.384	0.672	0.022	± 0.432	0.964	0.473	0.018	± 0.313	0.690	0.340
	T-Bic	0.027	± 0.760	1.714	0.834	0.018	± 0.453	1.111	0.527	0.009	± 0.317	0.684	0.337
	Bic	0.012	± 0.641	1.449	0.701	0.017	± 0.433	1.044	0.498	0.011	± 0.313	0.680	0.335
	MCD	0.016	± 0.740	1.626	0.794	0.019	± 0.475	1.227	0.574	0.010	± 0.329	0.688	0.343
	DS	0.019	± 0.765	1.649	0.811	0.019	± 0.514	1.175	0.569	0.011	± 0.337	0.685	0.345
Valores Propios	Clásico	3.924	± 2.694	13.365	6.661	3.097	± 1.703	8.199	4.333	2.665	± 1.184	6.464	3.438
	Mhuber	0.596	± 0.810	2.477	1.294	0.392	± 0.609	1.587	0.863	0.254	± 0.492	1.167	0.638
	T-Bic	1.341	± 1.332	5.705	2.793	0.675	± 0.739	2.608	1.341	0.456	± 0.574	1.918	0.982
	Bic	0.774	± 1.003	3.152	1.643	0.552	± 0.693	2.241	1.162	0.409	± 0.559	1.718	0.895
	MCD	2.426	± 1.629	7.117	3.724	1.097	± 0.919	3.485	1.834	0.669	± 0.628	2.909	1.402
	DS	1.132	± 1.291	12.580	5.001	0.676	± 0.814	2.528	1.339	0.452	± 0.603	1.819	0.958
Matriz de Covarianzas	Clásico	2.600	± 3.590	13.959	6.716	2.322	± 2.547	9.080	4.650	2.180	± 1.834	6.691	3.568
	Mhuber	0.044	± 1.275	3.384	1.568	0.032	± 0.900	2.257	1.063	0.021	± 0.665	1.530	0.739
	T-Bic	0.352	± 2.177	6.547	3.025	0.256	± 1.198	3.113	1.522	0.244	± 0.845	2.104	1.064
	Bic	0.136	± 1.539	4.379	2.018	0.190	± 1.086	2.845	1.374	0.210	± 0.790	2.061	1.020
	MCD	1.460	± 2.742	9.062	4.421	0.653	± 1.473	3.940	2.022	0.444	± 0.943	3.006	1.464
	DS	0.136	± 1.927	10.865	4.309	0.150	± 1.290	3.350	1.597	0.202	± 0.884	2.197	1.094

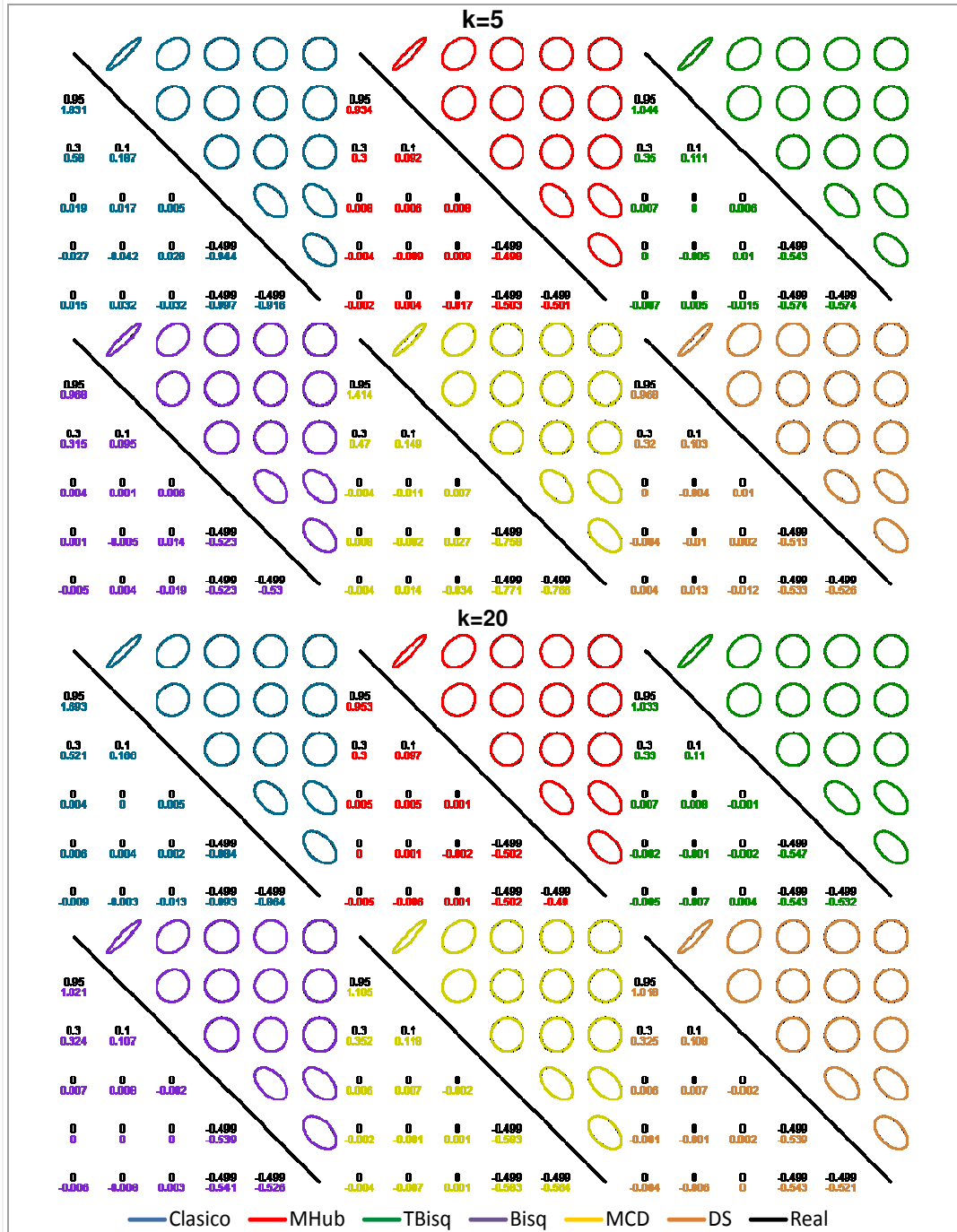


Elaborado por: Néstor Montaña P.

Cuadro 4.19b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$



Elaborado por: Néstor Montaña P.

**El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada
bajo el modelo (2.1) con $\varepsilon = 0.10$ y $H = N(0.537\alpha_p, 9\Sigma_2)$**

En el cuadro 4.20a se puede observar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador Bicuadrático para $k=5$ sobreestima la primer componente por 0.0055 unidades, también sobreestima la tercer, cuarta y quinta componentes por 0.0258, 0.0009 y 0.0082 respectivamente y subestima también la segunda y sexta componente por 0.0016 y 0.0135 unidades, valores que proporcionan un Sesgo Promedio Total de 0.012.

Se puede observar que el promedio de la estimación de cada componente del Vector de Medias correspondiente al Estimador Clásico se aparta notoriamente de las estimaciones brindadas por los demás estimadores.

La estimación de la Matriz de Varianzas Varizanzas y Covarianzas correspondiente al Estimador M de Huber y $k=5$ converge a los valores 0.07, 0.101, 0.145, -0.113, -0.098 en su primer columna, mientras que para $k=20$ converge a 0.093, 0.1, 0.136, -0.114, -0.093. (Véase Cuadro 4.20b)

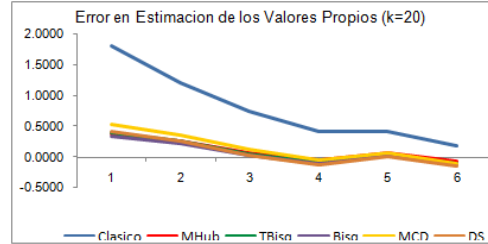
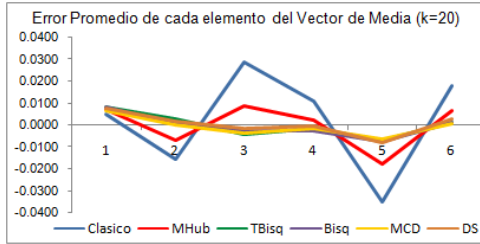
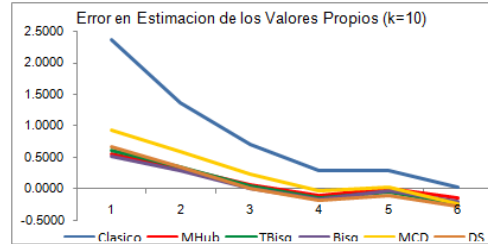
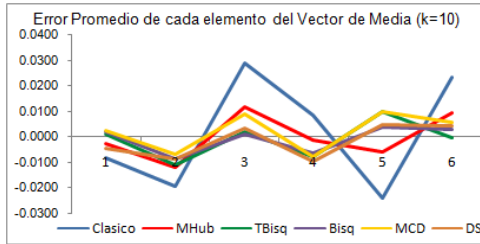
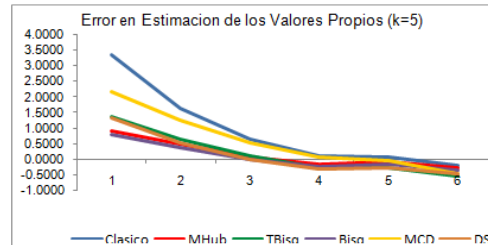
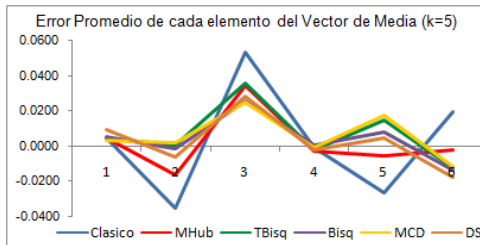
Cuadro 4.20a

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		k											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.072	± 0.781	1.711	0.855	0.050	± 0.560	1.312	0.641	0.053	± 0.394	0.856	0.434
	Mhuber	0.039	± 0.645	1.522	0.735	0.020	± 0.458	1.043	0.507	0.023	± 0.328	0.677	0.343
	T-Bic	0.041	± 0.740	1.893	0.892	0.017	± 0.469	1.001	0.496	0.012	± 0.330	0.717	0.353
	Bic	0.031	± 0.640	1.649	0.773	0.012	± 0.451	0.999	0.487	0.012	± 0.327	0.697	0.345
	MCD	0.033	± 0.745	1.741	0.840	0.018	± 0.491	1.045	0.518	0.010	± 0.334	0.722	0.356
	DS	0.035	± 0.764	1.703	0.834	0.015	± 0.509	1.063	0.529	0.011	± 0.339	0.756	0.369
Valores Propios	Clásico	3.793	± 1.916	10.184	5.298	2.835	± 1.242	6.622	3.566	2.372	± 0.836	4.670	2.626
	Mhuber	1.074	± 0.716	3.138	1.643	0.679	± 0.482	2.073	1.078	0.482	± 0.357	1.314	0.718
	T-Bic	1.656	± 1.076	10.567	4.433	0.760	± 0.551	2.246	1.186	0.490	± 0.387	1.396	0.758
	Bic	1.002	± 0.751	3.028	1.594	0.641	± 0.505	1.971	1.039	0.441	± 0.378	1.373	0.730
	MCD	2.609	± 1.338	6.996	3.647	1.155	± 0.689	3.260	1.701	0.658	± 0.423	1.687	0.923
	DS	1.576	± 1.112	6.108	2.932	0.835	± 0.593	2.357	1.262	0.526	± 0.403	1.566	0.832
Matriz de Covarianzas	Clásico	2.349	± 3.511	14.310	6.723	2.063	± 2.480	9.570	4.704	1.961	± 1.809	6.697	3.489
	Mhuber	0.402	± 1.523	4.513	2.146	0.306	± 1.008	3.081	1.465	0.277	± 0.724	2.025	1.009
	T-Bic	0.428	± 2.133	10.863	4.475	0.238	± 1.182	3.127	1.516	0.216	± 0.801	2.086	1.034
	Bic	0.182	± 1.494	4.417	2.031	0.184	± 1.059	2.762	1.335	0.179	± 0.759	1.960	0.966
	MCD	1.455	± 2.768	9.909	4.717	0.615	± 1.462	4.450	2.176	0.390	± 0.904	2.459	1.251
	DS	0.356	± 2.075	8.510	3.647	0.182	± 1.292	3.493	1.656	0.191	± 0.865	2.293	1.116

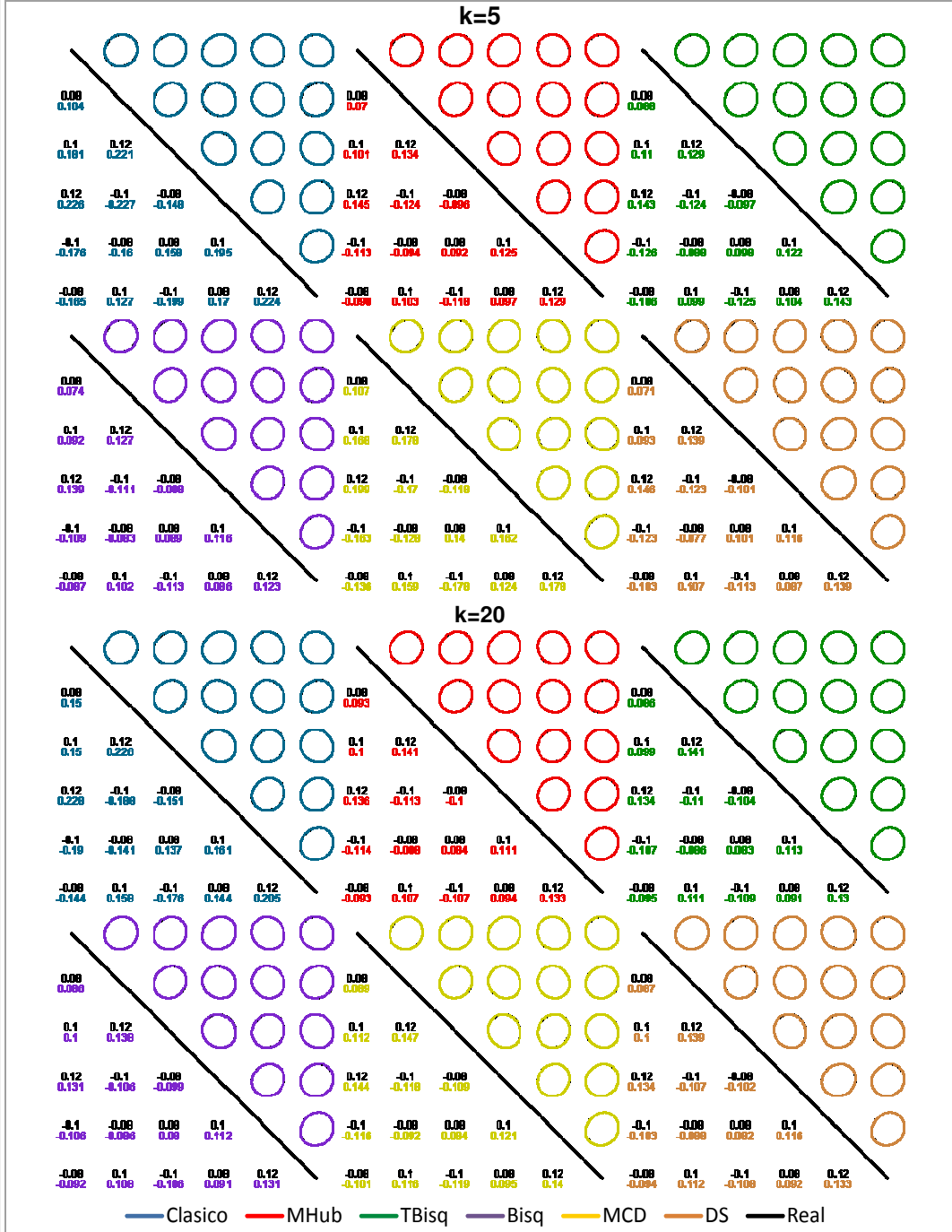


Elaborado por: Néstor Montaña P.

Cuadro 4.20 b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada

Por Variable con $\varepsilon = 0.10$ y $H = N(0, 9)$

De acuerdo al Cuadro 4.21a, el Estimador M de Huber brinda, para todos los valores de k, la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para k=20 obtiene un Índice de 0.550 resultado del Sesgo Promedio igual a 0.258, Sesgo Máximo igual a 0.953 y Desviación Total de 0.440.

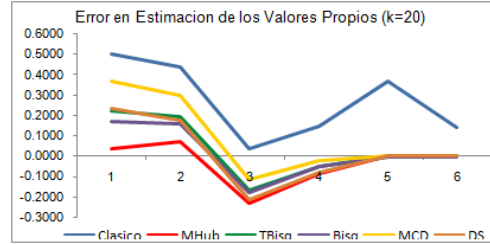
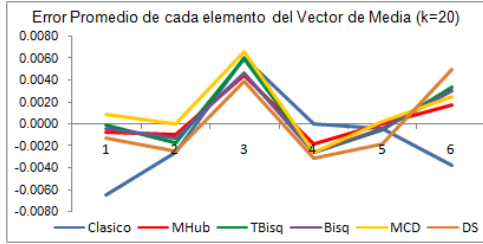
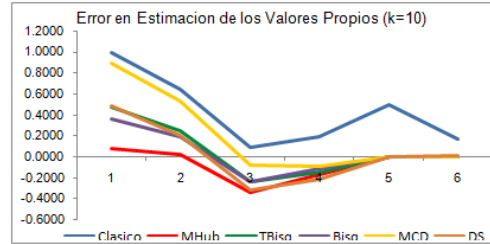
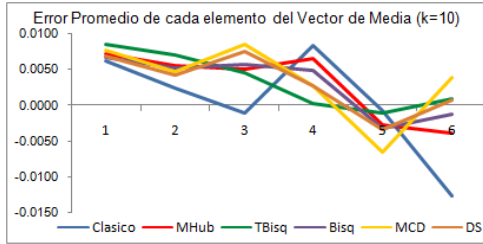
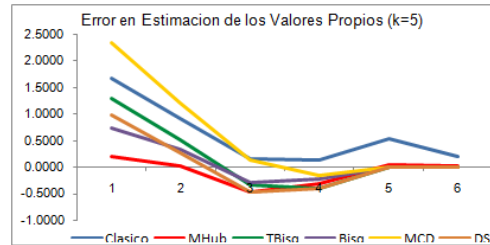
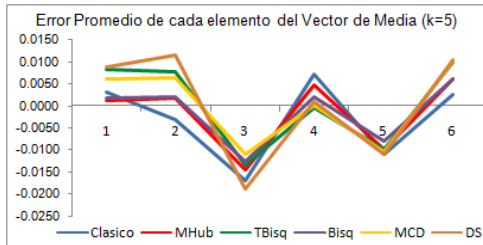
En lo referente al Vector de Medias y la Matriz de Varianzas y Covarianzas los resultados se mantienen, el Estimador M de Huber brinda la “mejor” estimación para todos los tamaños muestrales considerados.

En este escenario el Estimador Clásico es el estimador que más sobrestima el menor Valor Propio, evento que puede causar que se pierda una oportunidad de reducir el número de variables, pues el último Valor Propio asociado a Σ_1 es 0.002, además para k=10 y 20 este estimador se aleja claramente de los demás estimadores.

Cuadro 4.21a
Estimadores Robustos
 Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(0, 9)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.022 ± 0.717	1.619	0.786	0.017 ± 0.480	0.955	0.484	0.010 ± 0.341	0.771	0.374			
	Mhuber	0.019 ± 0.625	1.362	0.669	0.013 ± 0.424	0.902	0.446	0.005 ± 0.305	0.692	0.334			
	T-Bic	0.022 ± 0.787	1.853	0.887	0.012 ± 0.450	1.022	0.495	0.008 ± 0.317	0.738	0.354			
	Bic	0.016 ± 0.637	1.424	0.692	0.012 ± 0.432	0.942	0.462	0.006 ± 0.316	0.727	0.350			
	MCD	0.020 ± 0.762	1.727	0.836	0.015 ± 0.484	1.120	0.540	0.008 ± 0.342	0.747	0.366			
	DS	0.028 ± 0.766	1.696	0.830	0.012 ± 0.477	1.096	0.528	0.008 ± 0.334	0.804	0.382			
Valores Propios	Clásico	2.011 ± 1.547	7.019	3.526	1.307 ± 0.986	4.504	2.266	0.785 ± 0.611	3.440	1.612			
	Mhuber	0.593 ± 0.782	1.945	1.107	0.394 ± 0.566	1.349	0.769	0.258 ± 0.440	0.953	0.550			
	T-Bic	1.478 ± 1.457	5.258	2.731	0.596 ± 0.737	2.467	1.267	0.342 ± 0.532	1.486	0.787			
	Bic	0.885 ± 1.122	3.583	1.863	0.490 ± 0.706	2.043	1.080	0.300 ± 0.514	1.394	0.736			
	MCD	2.632 ± 1.782	7.962	4.125	1.048 ± 0.868	3.332	1.749	0.490 ± 0.553	1.608	0.884			
	DS	1.196 ± 1.298	6.337	2.944	0.648 ± 0.808	2.658	1.371	0.371 ± 0.563	1.517	0.817			
Matriz de Covarianzas	Clásico	1.494 ± 2.524	11.520	5.179	1.055 ± 1.744	6.868	3.222	0.665 ± 1.074	3.755	1.831			
	Mhuber	0.274 ± 1.202	3.051	1.509	0.195 ± 0.847	1.988	1.010	0.098 ± 0.611	1.321	0.676			
	T-Bic	0.496 ± 2.221	6.894	3.204	0.162 ± 1.133	3.209	1.501	0.090 ± 0.759	1.823	0.891			
	Bic	0.249 ± 1.590	4.526	2.122	0.100 ± 1.036	2.584	1.240	0.052 ± 0.712	1.722	0.829			
	MCD	1.631 ± 2.822	9.665	4.706	0.588 ± 1.450	4.124	2.054	0.250 ± 0.848	2.208	1.102			
	DS	0.189 ± 1.961	7.293	3.148	0.078 ± 1.293	3.244	1.538	0.059 ± 0.833	2.023	0.972			

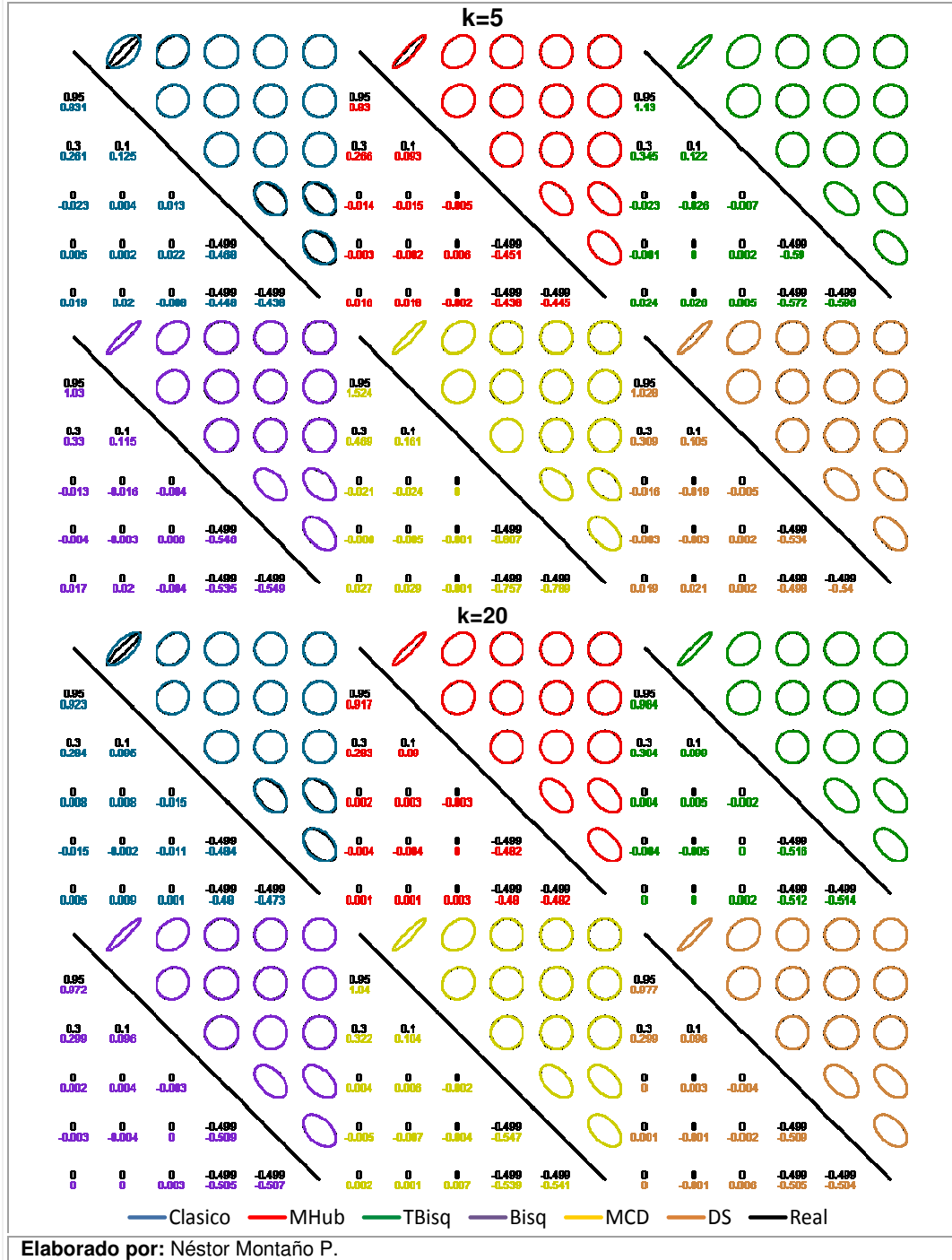


Elaborado por: Néstor Montaña P.

Cuadro 4.21b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = N(0, 9)$



El caso de una Población Normal $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada

Por Variable con $\varepsilon = 0.10$ y $H = N(0, 9)$

Como se muestra en el Cuadro 4.22a, para $k=10$ el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador DS se encuentra a una distancia 0.016 con respecto al Vector de Medias Real; y la estimación más alejada está a 1.126 unidades de distancia; además este método presenta una Desviación Total de 0.502, por lo tanto el Índice correspondiente es 0.548.

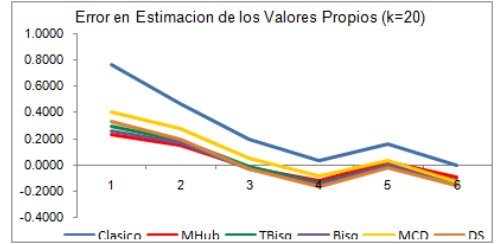
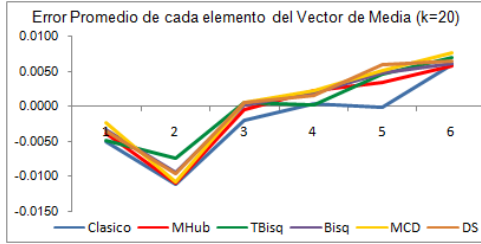
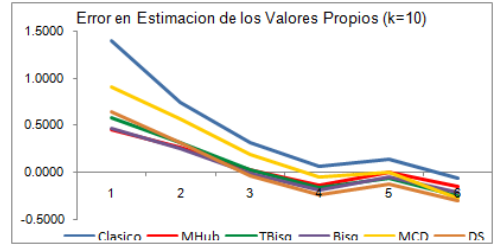
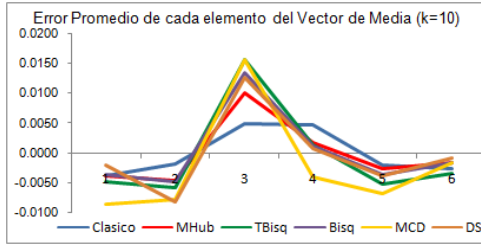
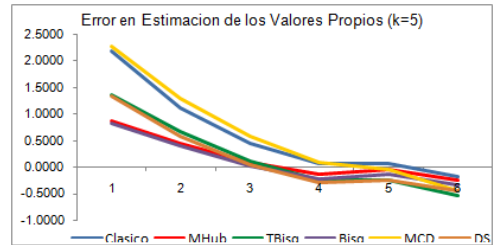
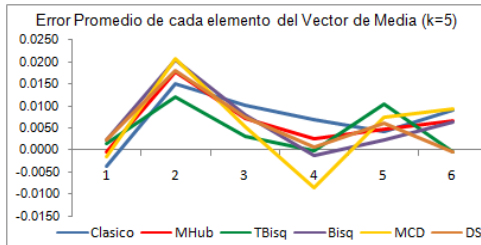
Al considerar el Índice planteado se aprecia que para todos los valores de k los Estimadores M de Huber y Bicuadrático son los que mejor se aproximan al valor real del Vector de Medias, situación que se repite para la Matriz de Covarianzas y los vectores propios asociados a la misma.

Nótese que, con respecto a cada Valor Propio, todos los estimadores sobreestiman los dos mayores Valores Propios; y para $k=10$ y 20 el estimador Clásico es el que sobreestima con mayor error los tres primeros valores propios, lo que es confirmado con los valores de Sesgo promedio donde en ambos casos el estimador Clásico es el que mayor Sesgo presenta.

Cuadro 4.22a
Estimadores Robustos
 Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(0, 9)$

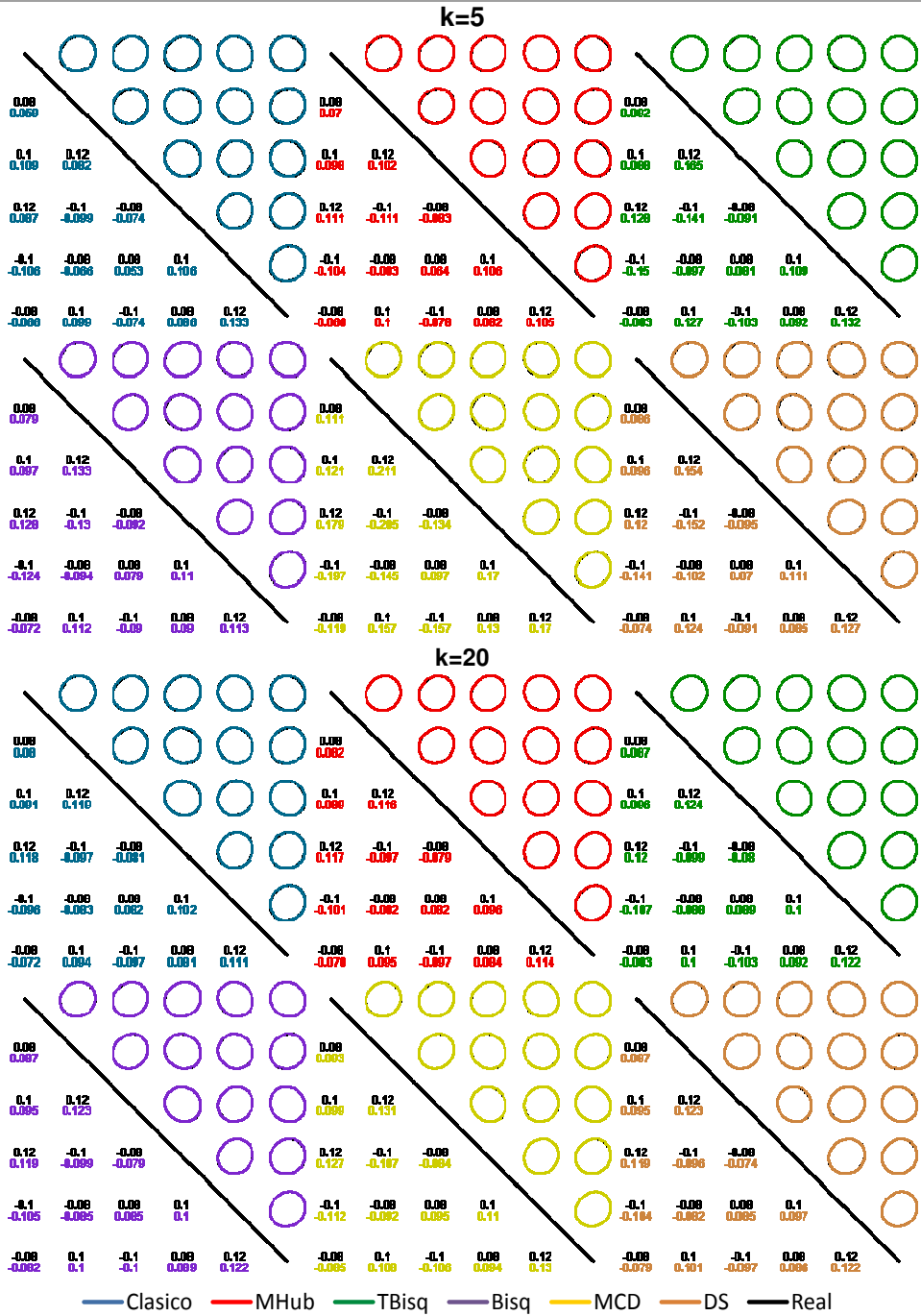
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.022 ± 0.709	1.663	0.798	0.009 ± 0.489	1.069	0.523	0.014 ± 0.323	0.715	0.351			
	Mhuber	0.021 ± 0.630	1.401	0.684	0.012 ± 0.444	0.974	0.477	0.014 ± 0.303	0.705	0.340			
	T-Bic	0.016 ± 0.766	1.757	0.846	0.018 ± 0.471	1.045	0.512	0.012 ± 0.316	0.708	0.345			
	Bic	0.023 ± 0.640	1.458	0.707	0.015 ± 0.455	0.987	0.486	0.013 ± 0.306	0.693	0.337			
	MCD	0.026 ± 0.759	1.683	0.823	0.021 ± 0.503	1.016	0.514	0.014 ± 0.327	0.760	0.367			
	DS	0.021 ± 0.772	1.807	0.867	0.016 ± 0.502	1.126	0.548	0.013 ± 0.331	0.755	0.367			
Valores Propios	Clásico	2.493 ± 1.514	7.004	3.670	1.626 ± 0.891	4.697	2.405	0.935 ± 0.551	2.383	1.290			
	Mhuber	1.020 ± 0.693	3.059	1.591	0.561 ± 0.441	1.600	0.867	0.322 ± 0.314	0.986	0.541			
	T-Bic	1.660 ± 1.142	4.611	2.471	0.729 ± 0.544	2.206	1.159	0.403 ± 0.366	1.190	0.653			
	Bic	1.022 ± 0.720	3.023	1.588	0.598 ± 0.474	1.969	1.014	0.359 ± 0.341	1.167	0.623			
	MCD	2.714 ± 1.320	7.383	3.806	1.119 ± 0.664	3.071	1.618	0.519 ± 0.384	1.370	0.758			
	DS	1.585 ± 1.069	8.169	3.608	0.820 ± 0.578	2.272	1.223	0.454 ± 0.383	1.344	0.727			
Matriz de Covarianzas	Clásico	1.514 ± 2.573	11.663	5.250	1.068 ± 1.686	6.733	3.162	0.665 ± 1.072	3.808	1.848			
	Mhuber	0.392 ± 1.503	5.073	2.323	0.188 ± 0.967	2.558	1.238	0.077 ± 0.651	1.713	0.814			
	T-Bic	0.452 ± 2.182	7.897	3.510	0.178 ± 1.179	3.188	1.515	0.080 ± 0.764	1.907	0.917			
	Bic	0.222 ± 1.555	4.817	2.198	0.114 ± 1.045	2.621	1.260	0.052 ± 0.699	1.793	0.848			
	MCD	1.533 ± 2.887	9.955	4.792	0.556 ± 1.452	4.366	2.125	0.219 ± 0.837	2.155	1.071			
	DS	0.389 ± 2.068	9.554	4.004	0.124 ± 1.293	3.673	1.697	0.059 ± 0.828	2.126	1.004			



Elaborado por: Néstor Montaña P.

Cuadro 4.22b
 Estimadores Robustos
 Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = N(0, 9)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada

Por Variable con $\varepsilon = 0.10$ y $H = U(-5,5)$

En el cuadro 4.23a se puede apreciar el error correspondiente a cada componente del promedio de la estimación del Vector de Medias para cada estimador, así por ejemplo se tiene que el Estimador M Huber para $k=10$ subestima la primer componente por 0.0068 unidades, también subestima la segunda, tercer y sexta componentes por 0.0054, 0.0084 y 0.0059 respectivamente y sobreestima la cuarta y quinta componente por 0.0044 y 0.0006 unidades, valores que proporcionan un Sesgo Promedio Total de 0.012.

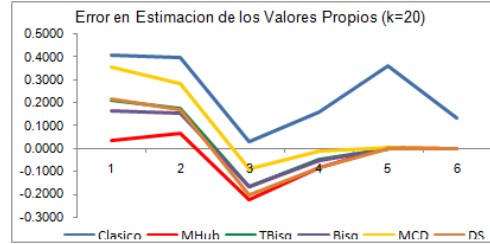
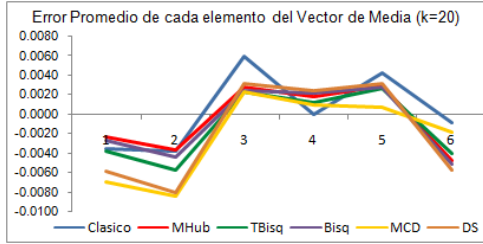
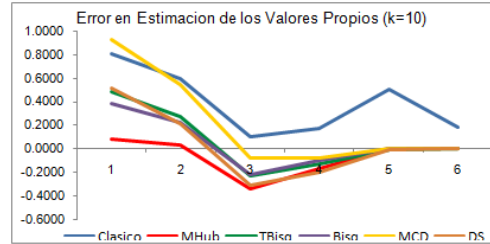
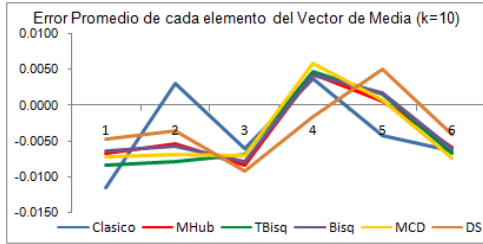
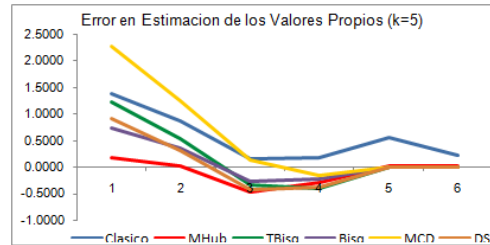
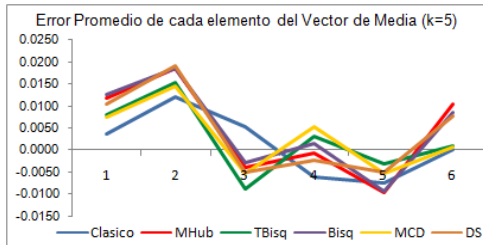
Se puede observar que para $k=20$ el promedio de la estimación de cada Valor Propio correspondiente al Estimador Clásico se aparta notoriamente de las estimaciones brindadas por los demás estimadores.

La estimación de la Matriz de Varianzas Varizanzas y Covarianzas brindada por el Estimador T-Bicuadrático converge para $k=5$ a los valores 1.097, 0.356, -0.013, -0.014, 0.025 en su primer columna, mientras que para $k=20$ converge a 0.97, 0.313, 0.004, -0.001 y -0.003 (Véase Cuadro 4.23b)

Cuadro 4.23a
Estimadores Robustos
 Caso Población $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = U(-5, 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.017 ± 0.695	1.464	0.725	0.016 ± 0.470	1.034	0.507	0.009 ± 0.321	0.714	0.348			
	Mhuber	0.026 ± 0.606	1.320	0.651	0.014 ± 0.432	1.008	0.485	0.008 ± 0.302	0.678	0.329			
	T-Bic	0.020 ± 0.748	1.521	0.763	0.016 ± 0.464	1.095	0.525	0.009 ± 0.316	0.708	0.344			
	Bic	0.026 ± 0.629	1.370	0.675	0.014 ± 0.443	0.987	0.481	0.008 ± 0.308	0.685	0.334			
	MCD	0.019 ± 0.733	1.533	0.761	0.015 ± 0.494	1.130	0.547	0.011 ± 0.324	0.687	0.341			
	DS	0.024 ± 0.725	1.663	0.804	0.013 ± 0.498	1.159	0.557	0.013 ± 0.335	0.724	0.357			
Valores Propios	Clásico	1.742 ± 1.166	4.799	2.569	1.153 ± 0.820	3.408	1.794	0.706 ± 0.582	2.063	1.117			
	Mhuber	0.575 ± 0.779	2.221	1.192	0.394 ± 0.538	1.736	0.889	0.248 ± 0.440	1.186	0.625			
	T-Bic	1.433 ± 1.374	6.355	3.054	0.618 ± 0.733	2.773	1.375	0.326 ± 0.537	1.594	0.819			
	Bic	0.888 ± 1.023	3.869	1.927	0.504 ± 0.674	2.731	1.303	0.284 ± 0.529	1.585	0.799			
	MCD	2.607 ± 1.594	8.126	4.109	1.087 ± 0.905	3.623	1.872	0.464 ± 0.568	1.892	0.975			
	DS	1.136 ± 1.136	6.041	2.771	0.665 ± 0.783	2.598	1.348	0.351 ± 0.579	1.803	0.911			
Matriz de Covarianzas	Clásico	1.372 ± 2.299	7.008	3.560	0.973 ± 1.579	4.792	2.448	0.612 ± 0.991	2.753	1.452			
	Mhuber	0.269 ± 1.173	2.868	1.437	0.191 ± 0.858	2.099	1.049	0.101 ± 0.624	1.546	0.757			
	T-Bic	0.481 ± 2.186	7.273	3.313	0.194 ± 1.167	3.352	1.571	0.082 ± 0.774	2.000	0.952			
	Bic	0.275 ± 1.568	4.397	2.080	0.133 ± 1.038	3.018	1.396	0.051 ± 0.727	1.955	0.911			
	MCD	1.629 ± 2.817	8.813	4.419	0.622 ± 1.461	4.111	2.064	0.248 ± 0.847	2.446	1.180			
	DS	0.190 ± 1.886	7.079	3.052	0.122 ± 1.256	3.475	1.618	0.054 ± 0.860	2.311	1.075			

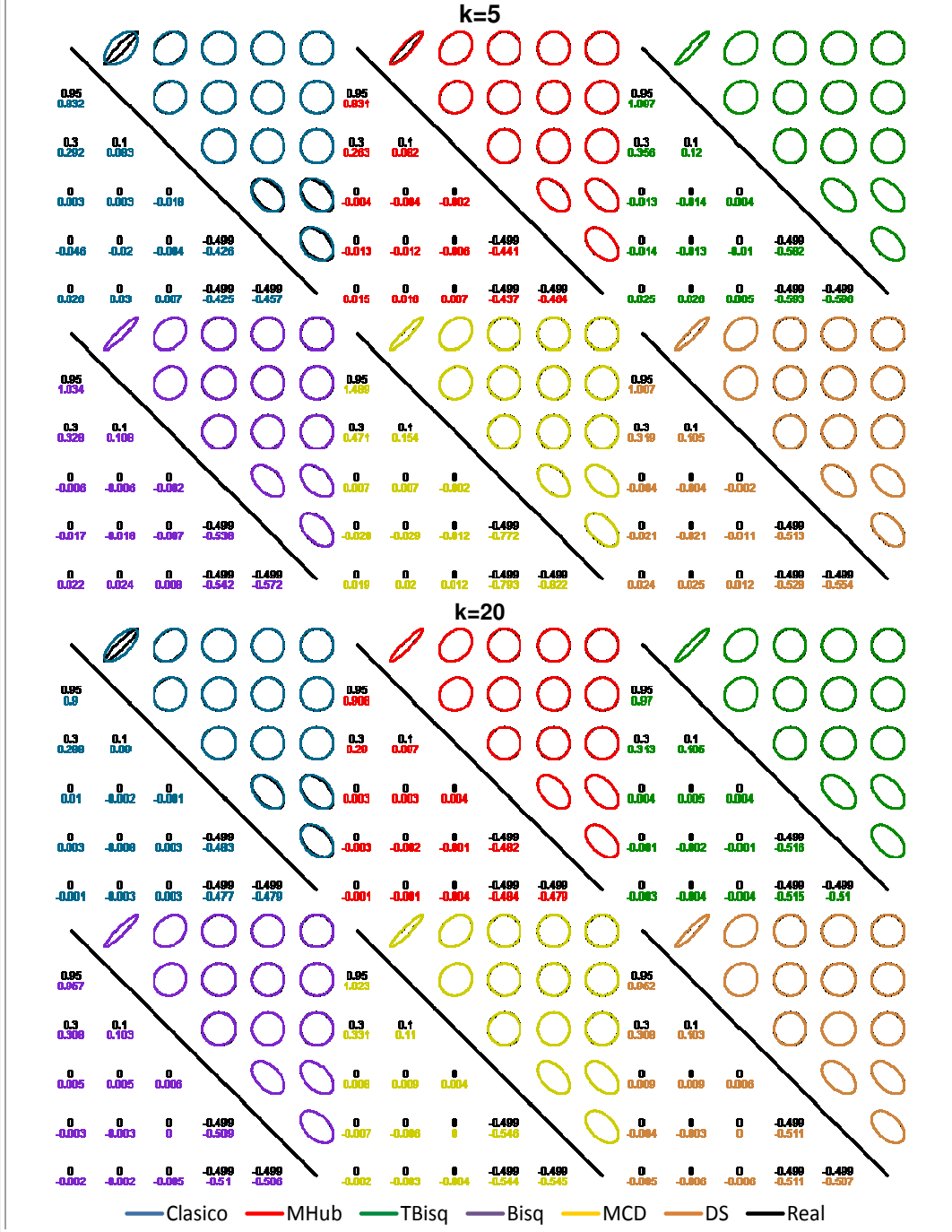


Elaborado por: Néstor Montaña P.

Cuadro 4.23 b

Estimadores Robustos

Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.10$ y $H = U(-5, 5)$



Elaborado por: Néstor Montaña P.

El caso de una Población Normal $N(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada

Por Variable con $\varepsilon = 0.10$ y $H = U(-5, 5)$

De acuerdo al Cuadro 4.24a, el Estimador M de Huber brinda, para todos los valores de k, la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas, en particular, para k=5 obtiene un Índice de 0.683 resultado del Sesgo Promedio igual a 0.019, Sesgo Máximo igual a 1.375 y Desviación Total de 0.655.

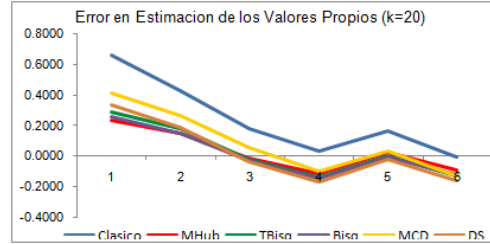
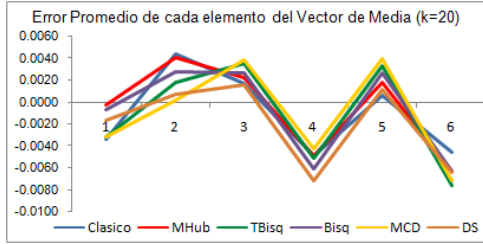
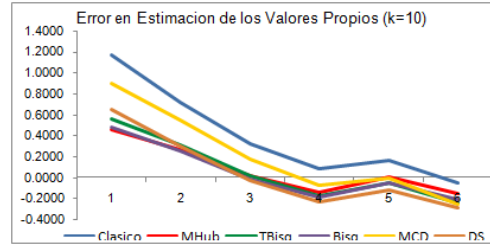
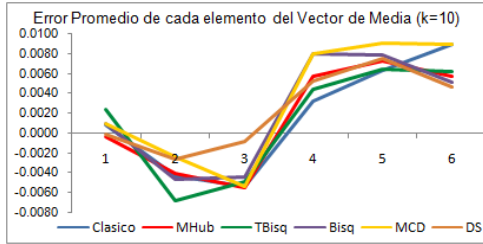
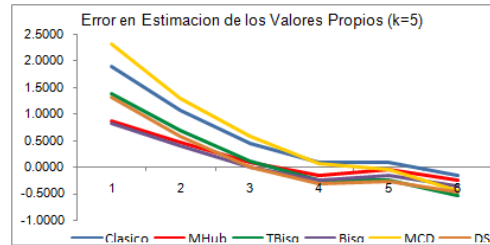
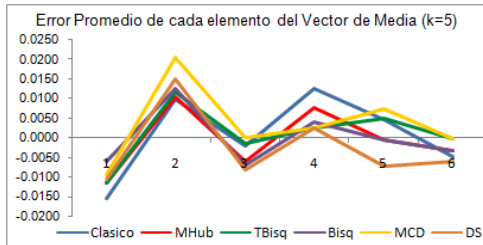
En lo referente al Vector de Medias y la Matriz de Varianzas y Covarianzas, el Estimador M de Huber y el Estimador Bicuadrático brindan la “mejor” mejor balance entre su estimación Promedio, “peor” estimación y dispersión para todos los tamaños muestrales considerados.

Se mantiene el comportamiento de los Estimadores para muestras generadas a partir de Σ_2 , sobreestimando los dos mayores Valores Propios para finalmente subestimar (con menor error) el último valor propio; además se puede observar que para k=10 y 20 el Estimador Clásico sobreestima el mayor Valor Propio con error promedio mayor al alcanzado por los Estimadores Robustos.

Cuadro 4.24a
Estimadores Robustos
 Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = U(-5,5)$

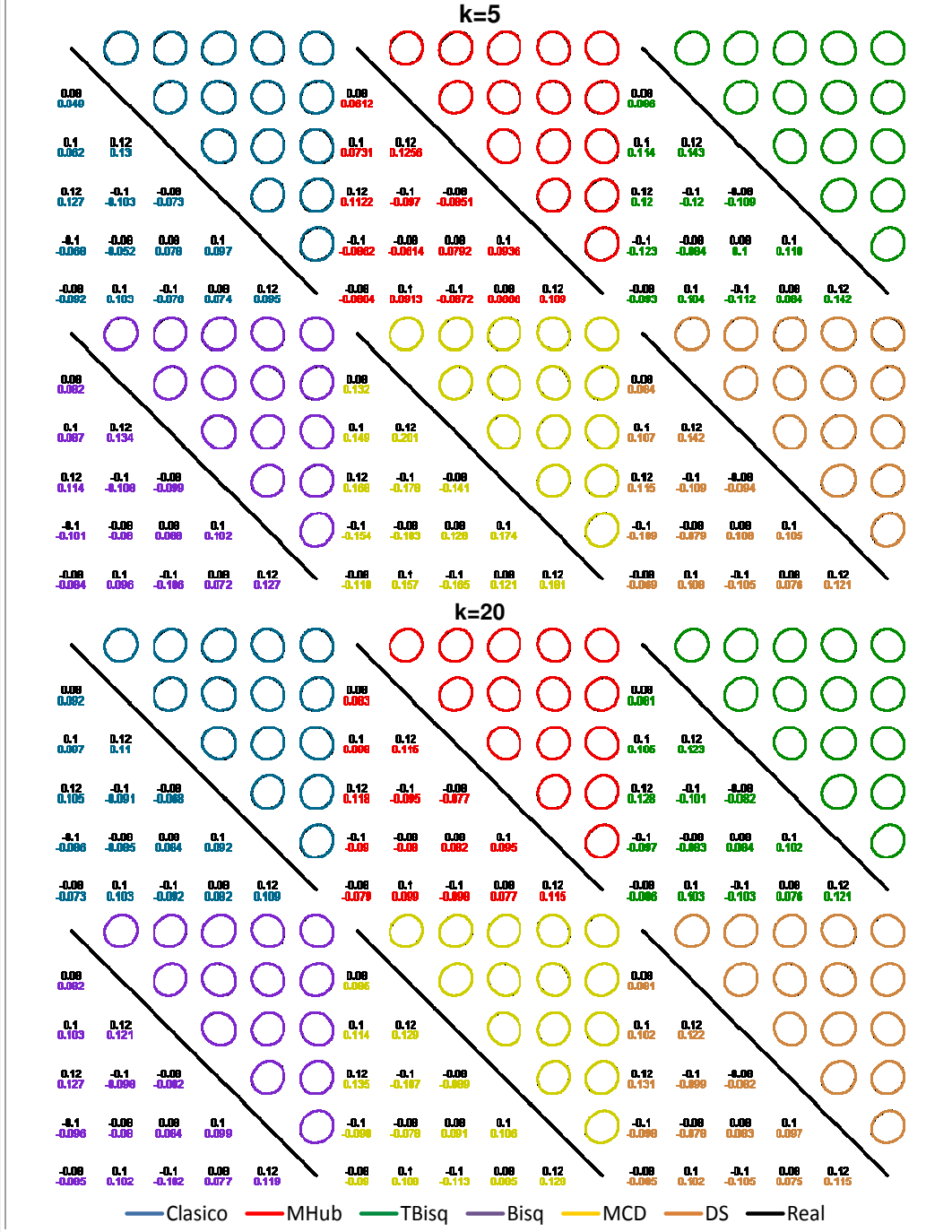
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5						k						20					
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice		
Vector de Medias	Clásico	0.023	± 0.732	1.522	0.759	0.013	± 0.465	1.031	0.503	0.009	± 0.315	0.716	0.347						
	Mhuber	0.019	± 0.655	1.375	0.683	0.013	± 0.451	1.013	0.492	0.009	± 0.294	0.716	0.340						
	T-Bic	0.017	± 0.775	1.726	0.839	0.013	± 0.456	1.060	0.510	0.011	± 0.299	0.716	0.342						
	Bic	0.017	± 0.655	1.339	0.670	0.014	± 0.454	1.036	0.501	0.010	± 0.297	0.712	0.340						
	MCD	0.024	± 0.750	1.716	0.830	0.016	± 0.506	1.310	0.611	0.011	± 0.313	0.724	0.349						
	DS	0.022	± 0.768	1.663	0.818	0.011	± 0.509	1.238	0.586	0.010	± 0.325	0.742	0.359						
Valores Propios	Clásico	2.220	± 1.026	4.908	2.718	1.417	± 0.696	3.177	1.763	0.828	± 0.430	1.947	1.068						
	Mhuber	1.029	± 0.507	2.649	1.395	0.564	± 0.455	1.571	0.863	0.320	± 0.308	0.932	0.520						
	T-Bic	1.668	± 1.138	8.686	3.831	0.715	± 0.544	1.896	1.052	0.396	± 0.359	1.056	0.604						
	Bic	1.026	± 0.754	2.827	1.536	0.609	± 0.529	1.829	0.989	0.356	± 0.360	1.024	0.580						
	MCD	2.749	± 1.400	7.812	3.987	1.103	± 0.653	2.797	1.518	0.517	± 0.370	1.501	0.796						
	DS	1.565	± 1.115	8.909	3.863	0.819	± 0.594	2.176	1.196	0.455	± 0.376	1.187	0.673						
Matriz de Covarianzas	Clásico	1.385	± 2.374	6.748	3.502	0.974	± 1.536	4.496	2.335	0.603	± 0.979	2.846	1.476						
	Mhuber	0.402	± 1.134	4.554	2.030	0.193	± 0.950	2.514	1.219	0.080	± 0.646	1.709	0.811						
	T-Bic	0.459	± 2.210	8.620	3.763	0.171	± 1.123	3.123	1.472	0.073	± 0.737	1.969	0.927						
	Bic	0.212	± 1.567	4.478	2.086	0.122	± 1.043	2.989	1.384	0.046	± 0.697	1.889	0.878						
	MCD	1.543	± 2.918	10.579	5.013	0.544	± 1.426	4.271	2.080	0.219	± 0.832	2.375	1.142						
	DS	0.359	± 2.053	10.475	4.296	0.121	± 1.275	3.474	1.623	0.056	± 0.825	2.227	1.036						



Elaborado por: Néstor Montaña P.

Cuadro 4.24b
 Estimadores Robustos
 Caso Población $N(0, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada $\varepsilon = 0.10$ y $H = U(-5,5)$



Elaborado por: Néstor Montaña P.

**El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminada bajo el modelo (2.1) , con $\varepsilon = 0.05$ y 0.10 y
 $H = N(\mathbf{0}, 9\Sigma_1)$**

En el Cuadro 4.25a se pueden observar los resultados de la Simulación; permitiendo efectuar inferencias sobre los mejores estimadores robustos cuando la población tiene un 5% y 10% de contaminación para diversos valores de k en cada escenario.

Así para k=5, 10 y 20 con el 5% de contaminación, el Estimador DS es el estimador que se encuentra más próximo al Vector de Medias verdadero; siendo la distancia promedio, cuando k=5, 0.036 ± 0.872 y el Índice asociado es 1.007; cuando k=10 el mismo estimador se encuentra a una Distancia Promedio de 0.026 ± 0.594 con respecto al Vector de Medias real. Mientras que, el mejor estimador, para la Matriz de Varianzas y Covarianzas cuando k=5, es el Bicuadrático con una Distancia Media de 5.226 ± 6.466 .

Al incrementar el porcentaje de contaminación a 10% el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador DS es el que menor índice tiene para cuando k toma los valores de 5 y 20 siendo el que mejor se aproxima al Vector real de Medias. Se tiene además, que el Estimador Bicuadrático es el más próximo a la

Cuadro 4.25a

Estimadores Robustos

Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(\mathbf{0}, 9\Sigma_1)$

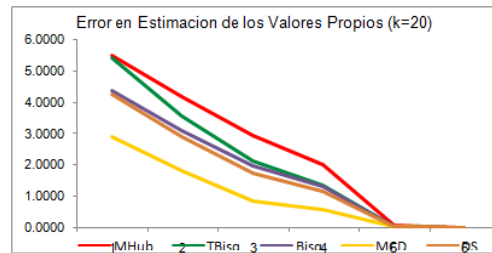
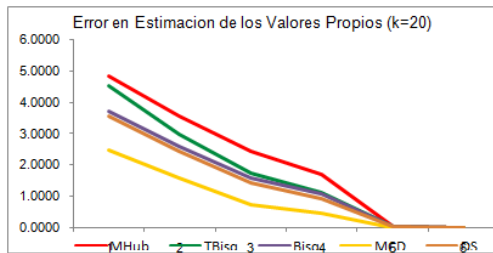
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	42.608 ± 4.657	---	---	---	53.602 ± 4.508	---	---	---	3.249 ± 4.584	---	---	---
	Mhuber	0.032 ± 0.933	2.372	1.112	0.020 ± 0.607	1.462	0.696	0.014 ± 0.428	1.058	0.500			
	T-Bic	0.037 ± 1.044	2.527	1.203	0.020 ± 0.683	1.772	0.825	0.013 ± 0.484	1.077	0.525			
	Bic	0.031 ± 0.923	2.231	1.062	0.013 ± 0.591	1.443	0.682	0.010 ± 0.413	0.960	0.461			
	MCD	0.028 ± 0.966	2.533	1.176	0.024 ± 0.629	1.516	0.723	0.012 ± 0.439	0.978	0.476			
	DS	0.036 ± 0.872	2.113	1.007	0.026 ± 0.594	1.254	0.625	0.010 ± 0.416	0.957	0.461			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	9.269 ± 6.097	50.477	21.948	7.838 ± 4.812	26.539	13.063	6.693 ± 3.022	21.498	10.404			
	T-Bic	9.039 ± 7.113	56.119	24.090	7.419 ± 4.940	33.058	15.139	5.800 ± 2.964	17.077	8.614			
	Bic	7.379 ± 5.840	41.751	18.323	5.908 ± 4.162	27.178	12.416	4.956 ± 2.688	13.763	7.136			
	MCD	8.906 ± 6.567	53.802	23.092	4.382 ± 3.134	23.802	10.439	3.052 ± 1.874	9.259	4.728			
	DS	9.623 ± 9.557	105.761	41.647	5.598 ± 3.898	35.926	15.141	4.645 ± 2.602	13.886	7.044			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.341 ± 6.577	49.284	21.068	6.616 ± 5.042	24.952	12.203	5.839 ± 3.285	17.984	9.036			
	T-Bic	5.917 ± 7.436	49.249	20.867	5.706 ± 5.599	31.859	14.388	4.822 ± 3.646	15.772	8.080			
	Bic	5.226 ± 6.466	40.432	17.375	4.733 ± 4.577	25.836	11.716	4.217 ± 2.977	12.887	6.694			
	MCD	5.992 ± 7.280	49.476	20.916	3.180 ± 3.817	22.059	9.685	2.431 ± 2.366	9.333	4.710			
	DS	6.793 ± 8.344	90.529	35.222	4.371 ± 4.392	34.685	14.482	3.898 ± 2.939	13.863	6.900			

		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	---	---	---	---	8.800 ± 4.569	---	---	---	3.044 ± 4.389	---	---	---
	Mhuber	0.028 ± 0.932	2.277	1.079	0.028 ± 0.645	1.717	0.797	0.014 ± 0.433	1.004	0.484			
	T-Bic	0.035 ± 1.014	2.727	1.259	0.025 ± 0.692	1.688	0.802	0.023 ± 0.495	1.118	0.545			
	Bic	0.027 ± 0.932	2.375	1.111	0.024 ± 0.627	1.495	0.715	0.020 ± 0.415	0.998	0.478			
	MCD	0.032 ± 0.989	2.753	1.258	0.024 ± 0.647	1.618	0.763	0.024 ± 0.441	1.021	0.495			
	DS	0.022 ± 0.847	2.337	1.069	0.021 ± 0.622	1.575	0.739	0.024 ± 0.419	0.924	0.456			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	10.470 ± 6.441	61.047	25.986	8.990 ± 4.926	23.711	12.542	7.781 ± 3.169	20.076	10.342			
	T-Bic	10.572 ± 8.583	50.235	23.130	9.087 ± 5.835	44.292	19.738	6.956 ± 3.367	19.590	9.971			
	Bic	8.651 ± 6.752	46.478	20.627	7.016 ± 4.388	22.338	11.247	5.894 ± 2.888	17.437	8.740			
	MCD	9.898 ± 8.133	51.189	23.073	5.148 ± 3.626	18.358	9.044	3.573 ± 2.068	10.757	5.466			
	DS	11.441 ± 11.383	82.544	35.123	6.648 ± 4.630	48.542	19.940	5.570 ± 2.860	15.451	7.960			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.397 ± 6.931	54.289	23.206	7.626 ± 5.369	22.542	11.845	6.821 ± 3.507	17.724	9.351			
	T-Bic	7.139 ± 8.672	60.206	25.339	7.052 ± 6.710	37.776	17.180	5.806 ± 4.169	18.596	9.524			
	Bic	6.297 ± 7.059	47.554	20.303	5.651 ± 5.061	22.453	11.055	5.042 ± 3.335	15.814	8.064			
	MCD	6.773 ± 8.168	52.902	22.615	3.808 ± 4.311	18.716	8.945	2.868 ± 2.657	10.873	5.466			
	DS	8.276 ± 9.683	89.998	35.986	5.263 ± 4.960	39.261	16.494	4.698 ± 3.328	15.181	7.736			

$\varepsilon = 0.05$

$\varepsilon = 0.10$

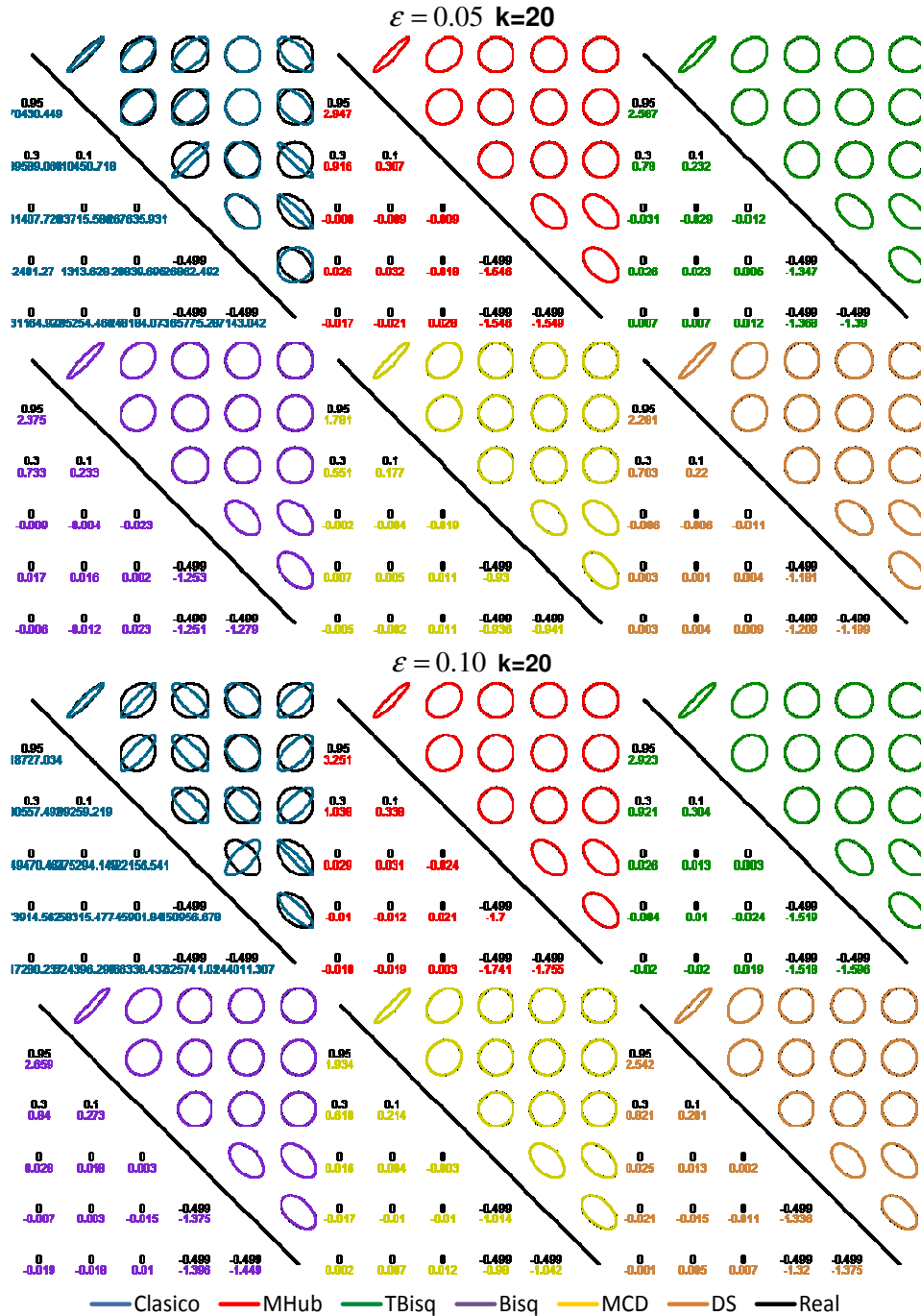


Elaborado por: Néstor Montaña P.

Cuadro 4.25b

Estimadores Robustos

Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(\mathbf{0}, 9\Sigma_1)$



Elaborado por: Néstor Montaña P.

Matriz de Varianzas y Covarianzas para $k=5$, con distancia promedio de 6.297 ± 7.059 ; y cuando el valor de k se incrementa a 10 el mejor estimador es el MCD con una distancia de 3.808 ± 4.311 .

El error de las estimaciones obtenidas al utilizar el Estimador Clásico es “muy grande” en todos los casos, por lo que no es presentado, además, para $k=5$ y $\varepsilon=0.05$ en 38 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon=0.10$ el algoritmo no converge en 21 ocasiones

**El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
Contaminada bajo el modelo (2.1), con $\varepsilon = 0.05$ y 0.10 y
 $H = N(\mathbf{0}, 9\Sigma_1)$**

Para el caso de muestras generadas a partir de la Distribución Cauchy, con una contaminación del 5% y con un valor de $k=5$ se tiene que el estimador que más se aproxima al valor real del Vector de Medias, en relación a los demás estimadores, es el Estimador DS donde la distancia promedio es 0.020 ± 0.855 , con un Índice asociado de 0.981. Bajo el mismo razonamiento, cuando el valor de k varía a 20 DS sigue siendo el mejor estimador con distancia promedio de 0.018 ± 0.425 y con 0.458 como Índice. Mientras que cuando $k=10$, el Estimador más próximo al Vector de Medias es el Bicuadrático con

distancia promedio de 0.020 ± 0.604 . El mejor estimador para la Matriz de Covarianzas y sus Valores Propios correspondientes, cuando $k=5$ es el Bicuadrático. Para $k=10$ y 20 el Estimador MCD es el que se encuentra a una distancia más cercana con respecto a la Matriz real (Véase Cuadro 4.26)

Por otro lado, cuando la contaminación se incrementa al 10%, para $k=5, 10$ y 20 el estimador más próximo al Vector real de Medias es el Bicuadrático. Se tiene además que el mejor estimador de la Matriz de Covarianza es: para $k=5$ el Bicuadrático mientras que para $k=10$ y $k=20$ es el MCD.

En este caso, para $k=5$ y $\varepsilon = 0.05$ en 12 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon = 0.10$ el algoritmo no converge en 14 ocasiones.

Cuadro 4.26

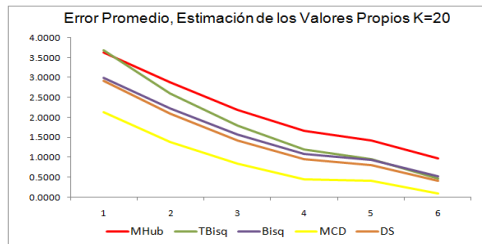
Estimadores Robustos

Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(\mathbf{0}, 9\Sigma_2)$

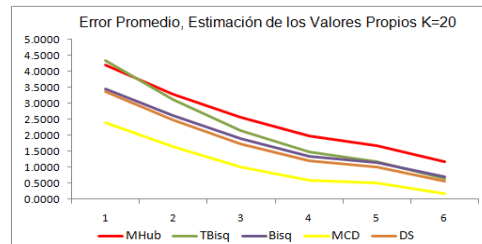
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		Eps=0,05												
		5				10				20				
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	4.758	± 4.935	---	---	7.626	± 4.612	---	---	6.119	3.826	± 4.460	---	4.143
	Mhuber	0.032	± 0.914	2.330	1.092	0.022	± 0.621	1.423	0.689	0.021	± 0.433	0.995	0.483	
	T-Bic	0.031	± 0.958	2.340	1.110	0.025	± 0.689	1.530	0.748	0.020	± 0.483	1.085	0.529	
	Bic	0.035	± 0.896	2.270	1.067	0.020	± 0.604	1.400	0.674	0.016	± 0.429	0.947	0.464	
	MCD	0.027	± 0.953	2.407	1.129	0.030	± 0.652	1.654	0.779	0.020	± 0.451	0.987	0.486	
	DS	0.020	± 0.855	2.068	0.981	0.026	± 0.612	1.540	0.726	0.018	± 0.425	0.931	0.458	
Valores Propios	Clásico	---	± ---	---	---	---	± ---	---	---	---	± ---	---	---	
	Mhuber	8.205	± 4.968	37.976	17.050	6.553	± 3.611	21.378	10.514	5.662	± 2.318	14.046	7.342	
	T-Bic	8.802	± 6.749	64.015	26.522	6.382	± 3.934	22.064	10.793	5.122	± 2.496	13.896	7.171	
	Bic	6.654	± 5.072	28.441	13.389	5.062	± 3.272	16.415	8.250	4.314	± 2.101	12.142	6.186	
	MCD	8.064	± 5.613	32.257	15.311	3.876	± 2.414	14.284	6.858	2.758	± 1.451	7.506	3.905	
	DS	9.529	± 8.762	88.870	35.720	4.741	± 3.071	23.898	10.570	4.077	± 2.046	11.421	5.848	
Matriz de Covarianzas	Clásico	---	± ---	---	---	---	± ---	---	---	---	± ---	---	---	
	Mhuber	6.552	± 6.111	40.964	17.875	5.777	± 4.499	22.668	10.982	5.276	± 3.095	15.256	7.876	
	T-Bic	5.442	± 7.384	64.489	25.772	4.866	± 5.078	27.268	12.404	4.426	± 3.522	16.608	8.185	
	Bic	4.555	± 5.935	33.339	14.610	4.105	± 4.055	19.862	9.341	3.849	± 2.872	13.339	6.687	
	MCD	5.200	± 6.806	38.356	16.787	2.698	± 3.486	17.933	8.039	2.201	± 2.332	9.162	4.565	
	DS	6.591	± 7.830	100.841	38.421	3.710	± 3.951	25.958	11.206	3.552	± 2.878	12.062	6.164	
		Eps=0,10												
		5				10				20				
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	8.190	± 4.307	---	---	4.569	± 4.443	---	---	1.680	± 4.206	---	---	
	Mhuber	0.027	± 0.920	1.994	0.981	0.015	± 0.620	1.564	0.733	0.022	± 0.436	1.114	0.524	
	T-Bic	0.018	± 1.001	2.632	1.217	0.017	± 0.689	1.659	0.788	0.021	± 0.505	1.152	0.559	
	Bic	0.022	± 0.903	1.954	0.960	0.011	± 0.610	1.495	0.705	0.017	± 0.426	0.987	0.477	
	MCD	0.027	± 0.979	2.241	1.082	0.011	± 0.645	1.699	0.785	0.020	± 0.444	1.018	0.494	
	DS	0.027	± 0.858	2.267	1.050	0.017	± 0.576	1.528	0.707	0.020	± 0.427	1.049	0.499	
Valores Propios	Clásico	---	± ---	---	---	---	± ---	---	---	---	± ---	---	---	
	Mhuber	8.849	± 4.950	34.853	16.217	7.403	± 3.795	20.570	10.589	6.592	± 2.545	15.170	8.103	
	T-Bic	10.017	± 7.301	63.464	26.927	7.992	± 4.930	31.312	14.744	6.078	± 2.778	15.458	8.105	
	Bic	7.699	± 5.482	34.257	15.812	5.945	± 3.617	19.321	9.627	5.107	± 2.357	12.254	6.572	
	MCD	8.678	± 5.866	37.280	17.275	4.478	± 2.688	14.379	7.182	3.178	± 1.673	7.767	4.206	
	DS	10.782	± 9.889	81.659	34.110	5.563	± 3.730	25.240	11.511	4.831	± 2.270	11.794	6.298	
Matriz de Covarianzas	Clásico	---	± ---	---	---	---	± ---	---	---	---	± ---	---	---	
	Mhuber	7.267	± 6.166	41.780	18.405	6.590	± 4.813	23.868	11.757	6.177	± 3.366	17.129	8.891	
	T-Bic	6.415	± 8.275	73.127	29.272	6.206	± 6.224	36.365	16.265	5.307	± 4.017	18.829	9.385	
	Bic	5.511	± 6.648	42.440	18.200	4.889	± 4.646	21.737	10.424	4.618	± 3.247	14.977	7.614	
	MCD	5.735	± 7.316	49.371	20.807	3.206	± 3.854	19.165	8.741	2.611	± 2.555	10.200	5.122	
	DS	7.698	± 9.102	90.917	35.906	4.429	± 4.553	27.775	12.253	4.287	± 3.192	14.300	7.260	

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$,
Contaminada bajo el modelo (2.1), con $\varepsilon = 0.05$ y 0.10 y
 $H = Uesf(d = 5)$

Para el caso de muestras generadas a partir de la Distribución Cauchy, con una contaminación del 5% y con un valor de $k=5$ se tiene que el estimador que más se aproxima al valor real del Vector de Medias, en relación a los demás estimadores, es el Estimador M-Huber donde la distancia promedio es 0.037 ± 0.898 , con un Índice asociado de 0.991. Bajo el mismo razonamiento, cuando el valor de k varía a 10 el Estimador DS es el que presenta mejor estimación con distancia promedio de 0.027 ± 0.574 y con 0.717 como Índice. Mientras que cuando $k=20$, el Estimador más próximo al Vector de Medias es el Bicuadrático con distancia promedio de 0.009 ± 0.425 . El mejor estimador para la Matriz de Covarianzas y sus Valores Propios correspondientes, es el MCD para $k=5, 10$ y 20 puesto que es el que se encuentra a una distancia más cercana con respecto a la Matriz real en los tres casos (Véase Cuadro 4.27)

Por otro lado, cuando la contaminación se incrementa al 10%, para $k=5$ y 10 el estimador más próximo al Vector real de Medias es el Estimador DS. Se tiene además que el “mejor” estimador de la Matriz de Varianzas y Covarianza para $k=5$ es el M-Huber y el estimador

Cuadro 4.27

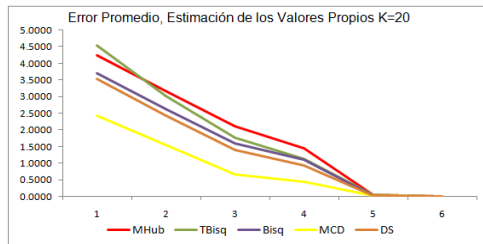
Estimadores Robustos

Caso Población $CAU(0, \Sigma)$, $\Sigma = \Sigma_1$, contaminada $\varepsilon = 0.05$ y 0.10 $H = Uesf(d = 5)$

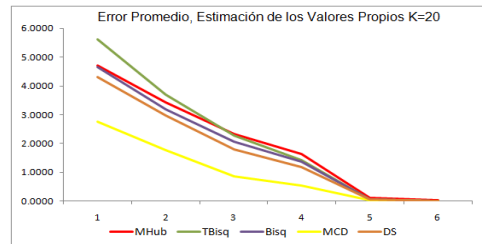
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		Eps=0,05											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	10.395 ± 4.343	---	---	7.369	10.818 ± 4.331	---	---	7.575	1.649 ± 4.753	---	---	3.201
	Mhuber	0.039 ± 0.865	1.931	0.945	0.014 ± 0.608	1.377	0.666	0.013 ± 0.420	0.937	0.457			
	T-Bic	0.042 ± 0.949	2.266	1.085	0.026 ± 0.667	1.738	0.810	0.012 ± 0.479	1.125	0.539			
	Bic	0.046 ± 0.860	2.137	1.014	0.019 ± 0.601	1.427	0.683	0.009 ± 0.417	0.969	0.465			
	MCD	0.050 ± 0.916	2.465	1.144	0.025 ± 0.611	1.476	0.704	0.013 ± 0.440	1.085	0.512			
	DS	0.033 ± 0.851	2.181	1.022	0.021 ± 0.593	1.383	0.665	0.014 ± 0.426	1.094	0.511			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.049 ± 3.841	28.160	13.017	5.363 ± 2.889	15.954	8.069	5.013 ± 1.965	12.193	6.391			
	T-Bic	7.257 ± 5.053	53.384	21.898	5.323 ± 3.011	17.681	8.672	4.683 ± 1.996	13.056	6.578			
	Bic	5.728 ± 4.043	33.334	14.368	4.303 ± 2.589	14.719	7.204	4.014 ± 1.832	10.457	5.435			
	MCD	7.446 ± 4.461	40.069	17.325	3.766 ± 2.261	11.958	5.995	2.849 ± 1.431	7.044	3.775			
	DS	7.581 ± 6.355	51.039	21.658	4.214 ± 2.595	14.401	7.070	3.943 ± 1.926	9.814	5.228			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	5.581 ± 5.068	31.994	14.214	4.721 ± 3.679	19.455	9.285	4.675 ± 2.659	13.642	6.992			
	T-Bic	4.481 ± 6.307	51.551	20.780	4.108 ± 4.454	21.512	10.025	4.098 ± 3.171	14.623	7.297			
	Bic	3.920 ± 5.220	36.951	15.364	3.643 ± 3.435	18.093	8.357	3.620 ± 2.593	11.876	6.030			
	MCD	4.802 ± 6.395	45.080	18.759	2.644 ± 3.563	15.169	7.125	2.299 ± 2.396	9.290	4.662			
	DS	5.103 ± 6.672	58.928	23.567	3.248 ± 3.663	17.159	8.023	3.444 ± 2.735	12.183	6.121			
		Eps=0,10											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	3.505 ± 4.198	---	---	3.852	2.711 ± 4.442	---	---	3.577	1.808 ± 4.462	---	---	3.135
	Mhuber	0.040 ± 0.892	2.070	1.001	0.024 ± 0.623	1.534	0.727	0.019 ± 0.414	0.914	0.449			
	T-Bic	0.040 ± 1.016	2.535	1.197	0.024 ± 0.674	1.920	0.872	0.016 ± 0.485	0.973	0.491			
	Bic	0.048 ± 0.890	2.226	1.055	0.030 ± 0.577	1.594	0.734	0.012 ± 0.415	0.900	0.442			
	MCD	0.051 ± 0.953	2.326	1.110	0.025 ± 0.630	1.855	0.837	0.013 ± 0.446	1.037	0.499			
	DS	0.041 ± 0.822	2.125	0.996	0.016 ± 0.578	1.424	0.673	0.014 ± 0.421	0.944	0.460			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	9.582 ± 6.587	43.948	20.039	7.493 ± 4.708	27.500	13.234	6.482 ± 3.254	17.400	9.045			
	T-Bic	12.259 ± 10.792	76.325	33.125	9.534 ± 6.318	45.240	20.364	7.223 ± 4.081	21.066	10.790			
	Bic	9.566 ± 8.146	44.409	20.707	7.193 ± 4.668	34.804	15.555	6.169 ± 3.284	16.623	8.692			
	MCD	9.998 ± 8.153	49.449	22.533	4.872 ± 3.419	20.503	9.598	3.430 ± 2.095	9.593	5.039			
	DS	12.252 ± 11.559	175.424	66.412	6.410 ± 4.386	34.035	14.944	5.644 ± 3.094	16.727	8.488			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.496 ± 7.047	42.089	18.877	6.297 ± 4.837	25.575	12.237	5.649 ± 3.419	15.909	8.326			
	T-Bic	8.309 ± 9.707	74.977	30.998	7.426 ± 6.976	42.062	18.821	6.062 ± 4.330	21.100	10.497			
	Bic	6.854 ± 7.928	48.585	21.122	5.793 ± 5.150	30.059	13.667	5.270 ± 3.606	15.475	8.117			
	MCD	6.861 ± 8.168	53.034	22.688	3.609 ± 4.188	20.054	9.284	2.762 ± 2.524	10.187	5.158			
	DS	8.770 ± 9.950	159.610	59.443	5.071 ± 4.894	35.552	15.173	4.775 ± 3.411	15.217	7.801			

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

más cercano a la Matriz real y para $k=10$ y 20 es el MCD

Para $k=5$ y $\varepsilon = 0.05$ en 11 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon = 0.10$ el algoritmo no converge en 8 ocasiones

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada bajo el modelo (2.1), con $\varepsilon = 0.05$ y 0.10 y $H = Uesf(d = 5)$

De acuerdo a los datos que se presentan en el Cuadro 4.28 se determina que el mejor estimador para el Vector de Medias, cuando existe un 5% de contaminación y con $k=5$, es el M-Huber donde la distancia promedio al Vector real es 0.039 ± 0.865 y para $k=10$ es el DS ubicado a una distancia de 0.021 ± 0.893 del Vector real; mientras que para la Matriz de Varianzas y Covarianzas los Estimadores M-Huber y MCD son los que mejor se aproximan al valor real de la Matriz así como a los Valores Propios asociados a la misma.

Para una población contaminada en un 10%; se tiene que las mejores estimaciones para el Vector de Medias para los diferentes valores de k , están dadas por el Estimador M-Huber ubicándose, para el caso $k=5$, a una distancia promedio de 0.026 ± 0.850 con respecto al Vector Real de Medias. Por otro lado, para la Matriz de

Cuadro 4.28

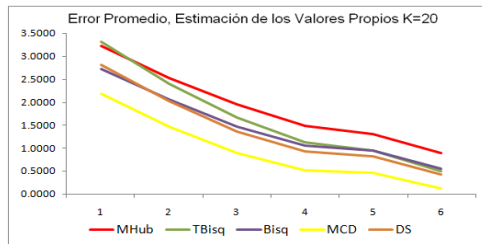
Estimadores Robustos

Caso Población $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, contaminada $\varepsilon = 0.05$ y 0.10 $H = Uesf(d = 5)$

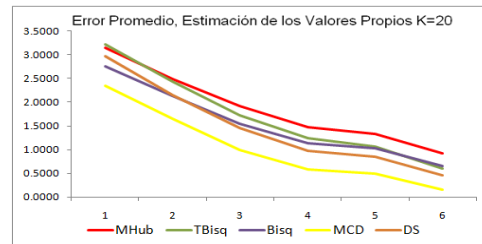
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

Eps=0,05													
k													
5													
10													
20													
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	10.395 ± 4.343	---	7.369	10.818 ± 4.331	---	7.575	1.649 ± 4.753	---	3.201			
	Mhuber	0.039 ± 0.865	1.931	0.945	0.014 ± 0.608	1.377	0.666	0.013 ± 0.420	0.937	0.457			
	T-Bic	0.042 ± 0.949	2.266	1.085	0.026 ± 0.667	1.738	0.810	0.012 ± 0.479	1.125	0.539			
	Bic	0.046 ± 0.860	2.137	1.014	0.019 ± 0.601	1.427	0.683	0.009 ± 0.417	0.969	0.465			
	MCD	0.050 ± 0.916	2.465	1.144	0.025 ± 0.611	1.476	0.704	0.013 ± 0.440	1.085	0.512			
	DS	0.033 ± 0.851	2.181	1.022	0.021 ± 0.593	1.383	0.665	0.014 ± 0.426	1.094	0.511			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.049 ± 3.841	28.160	13.017	5.363 ± 2.889	15.954	8.069	5.013 ± 1.965	12.193	6.391			
	T-Bic	7.257 ± 5.053	53.384	21.898	5.323 ± 3.011	17.681	8.672	4.683 ± 1.996	13.056	6.578			
	Bic	5.728 ± 4.043	33.334	14.368	4.303 ± 2.589	14.719	7.204	4.014 ± 1.832	10.457	5.435			
	MCD	7.446 ± 4.461	40.069	17.325	3.766 ± 2.261	11.958	5.995	2.849 ± 1.431	7.044	3.775			
	DS	7.581 ± 6.355	51.039	21.658	4.214 ± 2.595	14.401	7.070	3.943 ± 1.926	9.814	5.228			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	5.581 ± 5.068	31.994	14.214	4.721 ± 3.679	19.455	9.285	4.675 ± 2.659	13.642	6.992			
	T-Bic	4.481 ± 6.307	51.551	20.780	4.108 ± 4.454	21.512	10.025	4.098 ± 3.171	14.623	7.297			
	Bic	3.920 ± 5.220	36.951	15.364	3.543 ± 3.435	18.093	8.357	3.620 ± 2.593	11.876	6.030			
	MCD	4.802 ± 6.395	45.080	18.759	2.644 ± 3.563	15.169	7.125	2.299 ± 2.396	9.290	4.662			
	DS	5.103 ± 6.672	58.928	23.567	3.248 ± 3.663	17.159	8.023	3.444 ± 2.735	12.183	6.121			
Eps=0,10													
k													
5													
10													
20													
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	9.894 ± 4.157	---	7.025	2.428 ± 4.623	---	3.526	0.960 ± 3.999	---	2.479			
	Mhuber	0.026 ± 0.850	2.072	0.983	0.015 ± 0.591	1.519	0.708	0.014 ± 0.421	0.985	0.473			
	T-Bic	0.027 ± 0.943	2.697	1.222	0.018 ± 0.692	1.527	0.746	0.010 ± 0.465	1.074	0.516			
	Bic	0.017 ± 0.834	2.124	0.991	0.013 ± 0.594	1.405	0.670	0.009 ± 0.398	0.883	0.430			
	MCD	0.019 ± 0.915	2.610	1.181	0.021 ± 0.681	1.599	0.767	0.012 ± 0.458	0.989	0.486			
	DS	0.028 ± 0.878	2.111	1.006	0.019 ± 0.622	1.405	0.682	0.012 ± 0.421	0.932	0.455			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	6.665 ± 3.746	25.412	11.941	5.560 ± 2.645	14.672	7.626	4.959 ± 1.852	9.975	5.595			
	T-Bic	6.833 ± 4.357	36.795	15.995	5.722 ± 2.830	16.834	8.462	4.737 ± 1.946	9.186	5.290			
	Bic	5.402 ± 3.663	27.589	12.218	4.593 ± 2.477	11.942	6.337	4.170 ± 1.855	8.657	4.894			
	MCD	6.927 ± 4.038	25.361	12.109	4.202 ± 2.223	10.813	5.746	3.147 ± 1.483	7.472	4.034			
	DS	7.415 ± 6.048	71.988	28.484	4.721 ± 2.673	15.773	7.723	4.193 ± 1.907	9.287	5.129			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	5.350 ± 4.813	30.266	13.476	4.941 ± 3.669	16.584	8.398	4.648 ± 2.547	11.845	6.346			
	T-Bic	4.351 ± 6.121	38.668	16.380	4.564 ± 4.312	19.449	9.442	4.237 ± 3.005	11.908	6.383			
	Bic	3.853 ± 4.769	29.523	12.715	3.846 ± 3.447	13.771	7.021	3.809 ± 2.583	10.364	5.585			
	MCD	4.581 ± 6.172	32.211	14.321	3.077 ± 3.692	14.307	7.025	2.577 ± 2.525	8.841	4.648			
	DS	5.091 ± 6.459	71.140	27.563	3.681 ± 3.899	19.580	9.053	3.672 ± 2.805	11.194	5.890			

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

Varianzas y Covarianzas, el Estimador que más se aproxima a la Matriz real es el Bicuadrático cuando $k=5$ y 10 ; situación que difiere para $k=5$ donde el Estimador MCD es el que mejor estima la Matriz real de Varianzas y Covarianzas con una distancia promedio de 2.577 ± 2.525 .

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$,
Contaminada Por Variable con $\varepsilon = 0.05$ y 0.10 y
 $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$

En el Cuadro 4.29 se pueden observar los resultados de la Simulación; permitiendo efectuar inferencias sobre los mejores estimadores robustos cuando la población tiene un 5% y 10% de contaminación para diversos valores de k en cada escenario.

Así para $k=5$ con el 5% de contaminación, el Estimador DS es el estimador que se encuentra más próximo al Vector real de Medias cuya distancia promedio es 0.024 ± 0.811 y el Índice asociado es 0.967 ; para $k=10$ se tiene que el Estimador Bicuadrático se encuentra a una Distancia Promedio de 0.019 ± 0.585 con respecto al Vector de Medias real. Mientras que, el mejor estimador, para la Matriz de Covarianzas cuando $k=5$, es el Bicuadrático con una Distancia Media de 5.376 ± 6.581 .

Al incrementar el porcentaje de contaminación a 10% el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador DS es el que menor índice tiene para los diferentes valores de k siendo el que mejor se aproxima al Vector real de Medias. Se tiene además, que el Estimador M-Huber es el más próximo a la Matriz Real de Varianzas y Covarianzas para $k=5$, cuya distancia promedio es 8.242 ± 7.255 ; y cuando el valor de k se incrementa a 10 el mejor estimador es el Bicuadrático con una distancia de 5.869 ± 4.836 .

En este tipo de contaminación, para $k=5$ y $\varepsilon = 0.05$ en 37 de las 500 repeticiones el algoritmo del Estimador M de Huber no converge, así también con $k=5$ y $\varepsilon = 0.10$ el algoritmo no converge en 35 ocasiones; mientras que considerando $k=10$ para $\varepsilon = 0.05$ y $\varepsilon = 0.10$ el del Estimador M de Huber no converge en 3 y 2 ocasiones respectivamente.

Cuadro 4.29

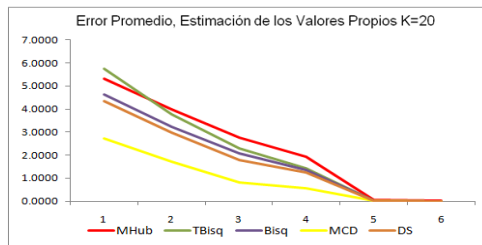
Estimadores Robustos

Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.05$ y 0.10 $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$

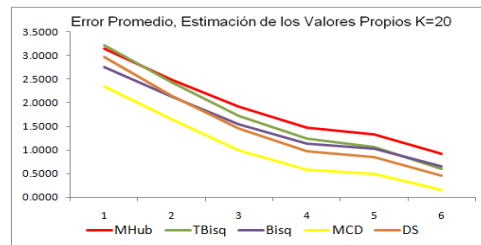
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

Eps=0,05												
k												
5												
10												
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	7.529 ± 4.778	---	6.153	1.955 ± 4.782	---	3.369	37.005 ± 4.133	---	20.569		
	Mhuber	0.026 ± 0.928	2.389	1.115	0.025 ± 0.590	1.531	0.715	0.016 ± 0.458	0.967	0.481		
	T-Bic	0.023 ± 0.982	2.472	1.159	0.026 ± 0.659	1.551	0.745	0.011 ± 0.475	1.088	0.525		
	Bic	0.018 ± 0.894	2.217	1.043	0.019 ± 0.585	1.344	0.649	0.010 ± 0.426	0.951	0.462		
	MCD	0.021 ± 0.938	2.849	1.269	0.027 ± 0.619	1.479	0.708	0.018 ± 0.451	0.989	0.486		
	DS	0.024 ± 0.811	2.066	0.967	0.022 ± 0.590	1.393	0.668	0.012 ± 0.439	0.960	0.470		
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	9.028 ± 6.163	43.113	19.434	7.316 ± 3.759	24.634	11.903	6.404 ± 3.204	15.518	8.375		
	T-Bic	9.186 ± 8.197	45.993	21.125	7.291 ± 4.916	27.895	13.367	6.123 ± 3.519	17.258	8.967		
	Bic	7.598 ± 6.707	37.423	17.243	5.910 ± 4.006	20.344	10.087	5.167 ± 2.838	15.186	7.730		
	MCD	8.837 ± 7.068	48.226	21.377	4.219 ± 3.096	14.825	7.380	3.034 ± 1.888	10.421	5.114		
	DS	10.175 ± 9.430	126.488	48.698	5.380 ± 3.847	30.366	13.198	4.800 ± 2.844	13.937	7.193		
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.149 ± 6.579	42.091	18.606	6.173 ± 4.029	23.666	11.289	5.583 ± 3.465	15.637	8.228		
	T-Bic	6.107 ± 7.718	52.080	21.969	5.648 ± 5.588	28.393	13.209	5.110 ± 3.984	17.761	8.951		
	Bic	5.376 ± 6.581	36.982	16.313	4.759 ± 4.535	20.231	9.842	4.399 ± 3.187	14.250	7.279		
	MCD	6.000 ± 7.449	46.431	19.960	3.092 ± 3.846	15.110	7.349	2.407 ± 2.485	10.047	4.980		
	DS	7.260 ± 8.309	115.032	43.534	4.209 ± 4.312	29.968	12.830	4.029 ± 3.268	13.463	6.920		
Eps=0,10												
k												
5												
10												
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	2.787 ± 4.624	---	---	0.590 ± 4.224	---	---	3.366 ± 4.563	---	---	---	---
	Mhuber	0.047 ± 0.913	2.651	1.204	0.029 ± 0.629	1.532	0.730	0.024 ± 0.446	1.030	0.500		
	T-Bic	0.049 ± 0.978	2.968	1.331	0.017 ± 0.699	1.706	0.807	0.007 ± 0.481	1.274	0.587		
	Bic	0.034 ± 0.876	2.608	1.173	0.016 ± 0.600	1.478	0.698	0.004 ± 0.426	1.026	0.485		
	MCD	0.036 ± 0.924	2.591	1.184	0.008 ± 0.631	1.604	0.748	0.012 ± 0.436	1.061	0.503		
	DS	0.025 ± 0.851	2.279	1.052	0.021 ± 0.610	1.374	0.668	0.009 ± 0.426	0.931	0.455		
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	10.311 ± 6.858	43.361	20.177	8.128 ± 4.097	36.811	16.345	7.459 ± 3.631	18.958	10.016		
	T-Bic	11.725 ± 10.490	99.338	40.517	8.837 ± 5.835	39.295	17.989	7.401 ± 3.919	22.790	11.370		
	Bic	9.185 ± 7.462	63.888	26.845	6.647 ± 4.265	24.823	11.912	6.171 ± 3.166	17.574	8.971		
	MCD	9.716 ± 7.953	59.332	25.667	4.731 ± 3.317	34.589	14.212	3.359 ± 1.931	9.675	4.988		
	DS	13.163 ± 13.001	194.815	73.660	6.286 ± 4.311	45.617	18.738	5.695 ± 2.873	17.212	8.593		
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.242 ± 7.255	41.177	18.891	6.888 ± 4.354	32.220	14.487	6.525 ± 3.726	16.383	8.878		
	T-Bic	7.871 ± 9.278	95.077	37.409	6.920 ± 6.390	35.327	16.212	6.200 ± 4.499	21.365	10.688		
	Bic	6.583 ± 7.455	65.320	26.453	5.369 ± 4.836	22.454	10.886	5.278 ± 3.578	15.542	8.133		
	MCD	6.686 ± 7.742	57.740	24.056	3.497 ± 3.981	27.729	11.736	2.711 ± 2.464	9.748	4.975		
	DS	9.635 ± 11.001	165.656	62.097	4.960 ± 4.706	39.968	16.545	4.825 ± 3.310	16.234	8.123		

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$,
Contaminada Por Variable con $\varepsilon = 0.05$ y 0.10 y
 $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$

De acuerdo a los datos que se presentan en el Cuadro 4.30 se determina que el mejor estimador para el Vector de Medias, cuando existe un 5% de contaminación y con $k=5$, es el DS donde la distancia promedio al Vector real es 0.035 ± 0.865 ; mientras que para la Matriz de Varianzas y Covarianzas los Estimadores Bicuadrático y MCD son los que mejor se aproximan al valor real de la Matriz mientras que para estimar los Valores Propios son los Estimadores M-Huber y MCD los que mejor se aproximan al valor real de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas.

Para una población contaminada en un 10%; se tiene que las mejores estimaciones para el Vector de Medias para $k=5$ y 10 , están dadas por el Estimador DS ubicándose, para el caso $k=5$, a una distancia promedio de 0.022 ± 0.832 con respecto al Vector Real de Medias. Por otro lado, para la Matriz de Varianzas y Covarianzas, el Estimador que más se aproxima a la Matriz real es el Bicuadrático cuando $k=5$ con una distancia promedio de 5.351 ± 6.386 ; situación que difiere para valores de $k=10$ y 20 donde el Estimador MCD es el que mejor estima la Matriz real de Varianzas y Covarianzas.

Cuadro 4.30

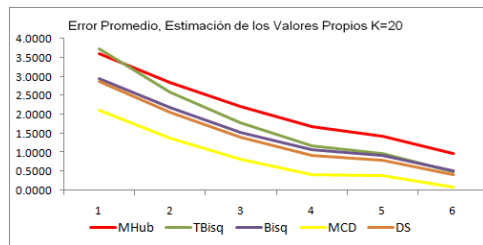
Estimadores Robustos

Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.05$ y 0.10 $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$

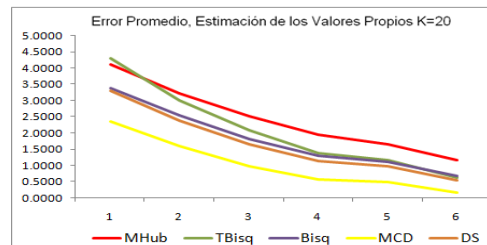
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				10								
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Eps=0,5	Vector de Medias	Clásico	8.819 ±	4.541	---	---	6.816 ±	4.591	---	---	2.579 ±	4.580	---	---
		Mhuber	0.047 ±	0.898	2.127	1.024	0.024 ±	0.650	1.428	0.701	0.022 ±	0.431	1.047	0.500
		T-Bic	0.042 ±	1.005	2.676	1.241	0.026 ±	0.687	1.694	0.802	0.024 ±	0.472	1.136	0.544
		Bic	0.045 ±	0.892	2.098	1.012	0.023 ±	0.616	1.330	0.656	0.018 ±	0.412	1.032	0.487
		MCD	0.044 ±	0.948	2.521	1.171	0.019 ±	0.628	1.530	0.725	0.024 ±	0.432	1.130	0.529
		DS	0.037 ±	0.850	2.044	0.977	0.011 ±	0.600	1.436	0.682	0.022 ±	0.415	1.013	0.483
Eps=0,5	Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
		Mhuber	8.379 ±	5.421	48.199	20.666	6.692 ±	3.677	22.404	10.924	5.622 ±	2.458	14.453	7.511
		T-Bic	8.748 ±	6.662	67.400	27.604	6.610 ±	4.066	24.072	11.582	5.135 ±	2.664	15.256	7.685
		Bic	6.975 ±	5.370	52.158	21.501	5.201 ±	3.242	17.740	8.728	4.260 ±	2.113	14.043	6.805
		MCD	8.133 ±	5.772	48.134	20.680	3.959 ±	2.659	14.317	6.979	2.709 ±	1.517	8.566	4.264
		DS	9.376 ±	7.708	79.331	32.139	4.875 ±	3.208	27.497	11.860	4.015 ±	2.067	14.937	7.006
Eps=0,5	Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
		Mhuber	6.676 ±	6.533	51.635	21.615	5.908 ±	4.662	23.724	11.431	5.242 ±	3.133	15.117	7.831
		T-Bic	5.434 ±	7.455	72.597	28.495	5.065 ±	5.292	30.655	13.671	4.435 ±	3.539	16.903	8.292
		Bic	4.767 ±	6.100	53.905	21.591	4.234 ±	4.261	21.557	10.018	3.803 ±	2.950	14.236	6.996
		MCD	5.263 ±	6.949	56.931	23.048	2.789 ±	3.676	17.970	8.145	2.151 ±	2.340	9.832	4.774
		DS	6.529 ±	7.638	90.614	34.927	3.822 ±	4.187	31.015	13.008	3.499 ±	2.833	14.647	6.993
Eps=0,10	Vector de Medias	Clásico	1.353 ±	4.711	---	---	3.989 ±	4.807	---	---	3.734 ±	4.606	---	---
		Mhuber	0.051 ±	0.875	2.266	1.064	0.042 ±	0.615	1.676	0.778	0.031 ±	0.431	0.979	0.481
		T-Bic	0.042 ±	0.988	2.822	1.284	0.034 ±	0.733	1.748	0.838	0.019 ±	0.486	1.204	0.570
		Bic	0.042 ±	0.879	2.473	1.131	0.025 ±	0.621	1.458	0.701	0.020 ±	0.414	0.962	0.465
		MCD	0.040 ±	0.913	2.730	1.228	0.027 ±	0.663	1.572	0.754	0.015 ±	0.437	1.067	0.506
		DS	0.022 ±	0.832	2.048	0.967	0.020 ±	0.608	1.466	0.698	0.018 ±	0.423	0.993	0.478
Eps=0,10	Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
		Mhuber	8.864 ±	4.615	35.134	16.204	7.436 ±	3.902	29.403	13.580	6.461 ±	2.592	12.408	7.154
		T-Bic	9.746 ±	6.453	151.160	55.787	7.691 ±	4.472	34.247	15.470	5.975 ±	2.843	13.716	7.511
		Bic	7.534 ±	5.022	34.635	15.731	5.803 ±	3.490	26.269	11.854	4.988 ±	2.355	12.044	6.462
		MCD	8.496 ±	6.037	31.007	15.180	4.480 ±	2.809	18.315	8.535	3.104 ±	1.605	9.134	4.614
		DS	9.457 ±	7.690	53.318	23.488	5.679 ±	3.704	25.075	11.486	4.697 ±	2.238	11.490	6.142
Eps=0,10	Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
		Mhuber	7.207 ±	6.012	39.090	17.436	6.637 ±	4.820	29.339	13.599	6.057 ±	3.276	15.309	8.214
		T-Bic	6.206 ±	7.884	107.326	40.472	6.015 ±	6.041	40.741	17.599	5.203 ±	4.006	18.023	9.077
		Bic	5.351 ±	6.386	38.429	16.722	4.823 ±	4.520	29.851	13.064	4.508 ±	3.141	13.721	7.123
		MCD	5.561 ±	7.164	40.384	17.703	3.231 ±	3.851	23.501	10.194	2.537 ±	2.477	11.078	5.364
		DS	6.733 ±	7.871	69.064	27.889	4.570 ±	4.562	29.546	12.893	4.155 ±	3.029	14.703	7.296

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

Para $k=5$ y $\varepsilon=0.05$ en 11 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon=0.10$ el algoritmo no converge en 15 ocasiones

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada Por Variable con $\varepsilon=0.05$ y 0.10 y $H = N(0,9)$

En el Cuadro 4.31 se muestran los resultados de la Simulación; permitiendo efectuar inferencias sobre los mejores estimadores robustos cuando la población tiene un 5% y 10% de contaminación para diversos valores de k en cada escenario.

Así para $k=5$ con el 5% de contaminación, el Estimador DS es el estimador que se encuentra más próximo al Vector real de Medias cuya distancia promedio es 0.020 ± 0.923 y el Índice asociado es 1.007; para $k=10$ se tiene que el Estimador Bicuadrático que se encuentra a una Distancia Promedio de 0.021 ± 0.593 con respecto al Vector de Medias real y para $k=20$ es el M-Huber con un índice de 0.478 y ubicado a una distancia de 0.019 ± 0.414 . Mientras que, el mejor estimador, para la Matriz de Covarianzas cuando $k=5$, es el M-Huber con una Distancia Media de 7.806 ± 8.243 .

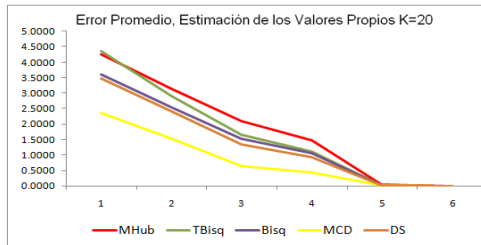
Al incrementar el porcentaje de contaminación al 10% el promedio de la estimación del Vector de Medias resultante al utilizar el

Cuadro 4.31
Estimadores Robustos
 Caso $CAU(0, \Sigma)$, $\Sigma = \Sigma_1$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(0,9)$

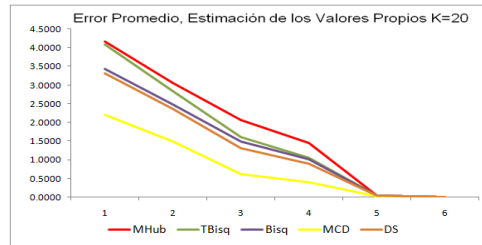
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

Eps=0,05		5				10				k			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	3.163 ± 4.738	---	---	---	1.161 ± 4.830	---	---	---	2.927 ± 4.524	---	---	---
	Mhuber	0.021 ± 0.945	2.401	1.122	0.031 ± 0.652	1.730	0.804	0.019 ± 0.414	1.002	0.478	0.019 ± 0.414	1.002	0.478
	T-Bic	0.019 ± 1.021	3.205	1.415	0.020 ± 0.704	1.820	0.848	0.013 ± 0.444	1.135	0.531	0.013 ± 0.444	1.135	0.531
	Bic	0.023 ± 0.953	2.686	1.221	0.021 ± 0.593	1.607	0.740	0.014 ± 0.391	1.028	0.477	0.014 ± 0.391	1.028	0.477
	MCD	0.018 ± 1.016	2.841	1.292	0.015 ± 0.640	1.730	0.795	0.019 ± 0.410	1.132	0.520	0.019 ± 0.410	1.132	0.520
	DS	0.020 ± 0.923	2.079	1.007	0.017 ± 0.620	1.645	0.761	0.017 ± 0.392	1.108	0.506	0.017 ± 0.392	1.108	0.506
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	10.097 ± 7.955	60.074	26.042	7.152 ± 4.597	47.117	19.622	5.895 ± 2.844	17.424	8.721	5.895 ± 2.844	17.424	8.721
	T-Bic	12.262 ± 9.560	84.031	35.284	7.887 ± 5.992	68.044	27.308	5.610 ± 2.994	18.549	9.051	5.610 ± 2.994	18.549	9.051
	Bic	9.826 ± 7.823	57.590	25.080	6.321 ± 4.536	39.282	16.713	4.789 ± 2.642	16.217	7.882	4.789 ± 2.642	16.217	7.882
	MCD	10.964 ± 8.628	66.821	28.804	4.371 ± 3.055	24.547	10.658	2.914 ± 1.821	9.055	4.597	2.914 ± 1.821	9.055	4.597
	DS	13.875 ± 11.939	187.614	71.143	5.806 ± 3.858	58.536	22.733	4.522 ± 2.538	14.242	7.101	4.522 ± 2.538	14.242	7.101
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.806 ± 8.243	59.407	25.152	5.993 ± 4.838	42.631	17.821	5.129 ± 3.129	15.566	7.941	5.129 ± 3.129	15.566	7.941
	T-Bic	8.175 ± 9.502	87.803	35.160	6.069 ± 5.828	50.813	20.903	4.677 ± 3.638	16.800	8.372	4.677 ± 3.638	16.800	8.372
	Bic	6.945 ± 7.771	64.238	26.318	5.070 ± 4.766	33.109	14.315	4.077 ± 2.984	14.143	7.068	4.077 ± 2.984	14.143	7.068
	MCD	7.531 ± 8.696	71.894	29.374	3.217 ± 3.718	20.615	9.184	2.312 ± 2.329	8.646	4.429	2.312 ± 2.329	8.646	4.429
	DS	9.989 ± 10.300	147.724	56.004	4.535 ± 4.402	46.574	18.504	3.797 ± 2.964	12.827	6.529	3.797 ± 2.964	12.827	6.529
Eps=0,10		5				10				k			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	2.476 ± 4.795	---	---	---	1.907 ± 4.841	---	---	---	3.374 ± 4.698	---	---	3.189
	Mhuber	0.024 ± 0.906	2.241	1.057	0.026 ± 0.602	1.355	0.661	0.021 ± 0.425	0.883	0.443	0.021 ± 0.425	0.883	0.443
	T-Bic	0.024 ± 1.077	2.428	1.176	0.029 ± 0.662	1.593	0.761	0.018 ± 0.439	1.046	0.501	0.018 ± 0.439	1.046	0.501
	Bic	0.019 ± 1.004	2.408	1.144	0.026 ± 0.578	1.245	0.616	0.018 ± 0.386	0.853	0.419	0.018 ± 0.386	0.853	0.419
	MCD	0.017 ± 1.047	2.556	1.207	0.028 ± 0.599	1.407	0.678	0.017 ± 0.406	0.937	0.454	0.017 ± 0.406	0.937	0.454
	DS	0.017 ± 0.884	2.240	1.047	0.023 ± 0.563	1.366	0.651	0.018 ± 0.406	0.905	0.443	0.018 ± 0.406	0.905	0.443
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	11.104 ± 9.408	87.560	36.024	6.892 ± 4.204	23.904	11.667	5.761 ± 2.813	16.663	8.412	5.761 ± 2.813	16.663	8.412
	T-Bic	13.952 ± 10.687	133.054	52.564	7.824 ± 5.321	32.752	15.299	5.347 ± 2.873	15.517	7.912	5.347 ± 2.873	15.517	7.912
	Bic	11.281 ± 9.080	137.662	52.675	6.035 ± 4.131	22.488	10.884	4.617 ± 2.468	14.281	7.122	4.617 ± 2.468	14.281	7.122
	MCD	11.685 ± 8.730	119.244	46.553	4.232 ± 2.939	16.648	7.939	2.767 ± 1.821	9.296	4.628	2.767 ± 1.821	9.296	4.628
	DS	15.349 ± 12.168	198.923	75.480	5.352 ± 3.763	31.113	13.409	4.361 ± 2.547	12.626	6.511	4.361 ± 2.547	12.626	6.511
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.704 ± 9.441	88.694	35.613	5.753 ± 4.615	23.373	11.247	5.018 ± 3.004	14.235	7.419	5.018 ± 3.004	14.235	7.419
	T-Bic	9.484 ± 10.274	153.100	57.619	6.084 ± 5.684	33.393	15.054	4.477 ± 3.410	14.592	7.493	4.477 ± 3.410	14.592	7.493
	Bic	8.178 ± 8.842	111.845	42.955	4.856 ± 4.402	22.569	10.609	3.935 ± 2.810	12.308	6.351	3.935 ± 2.810	12.308	6.351
	MCD	8.091 ± 8.769	103.680	40.180	3.109 ± 3.539	17.562	8.070	2.203 ± 2.204	8.756	4.388	2.203 ± 2.204	8.756	4.388
	DS	11.222 ± 11.070	187.458	69.917	4.218 ± 4.082	26.086	11.462	3.666 ± 2.834	12.260	6.253	3.666 ± 2.834	12.260	6.253

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

Estimador DS se encuentra a una distancia promedio de 0.017 ± 0.884 con respecto al Vector de Medias Real; para el caso en que $k=5$. Se tiene además, que el Estimador M-Huber es el más próximo a la Matriz Real de Covarianzas para $k=5$; con distancia promedio de 8.704 ± 9.441 ; y para $k=10$ y 20 es el Estimador MCD

En este caso, cuando $k=5$ y $\varepsilon = 0.05$ en 19 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon = 0.10$ el algoritmo no converge en 21 ocasiones

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada Por Variable con $\varepsilon = 0.05$ y 0.10 y $H = N(0,9)$

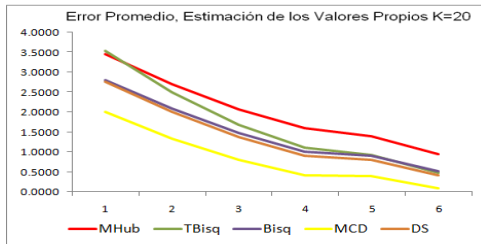
Para el caso de muestras generadas a partir de la Distribución Cauchy, con una contaminación del 5% y con un valor de $k=5$ se tiene que el estimador que más se aproxima al valor real del Vector de Medias, en relación a los demás estimadores, es el Estimador DS donde la distancia promedio es 0.026 ± 0.845 , con un Índice asociado de 1.064. Bajo el mismo razonamiento, cuando el valor de k varía a 10, el Estimador Bicuadrático es el mejor de los estimadores con distancia promedio de 0.027 ± 0.595 y con 0.689 como Índice. Mientras que cuando $k=20$, el Estimador Bicuadrático continua siendo el estimador más próximo al Vector de Medias con distancia promedio de 0.010 ± 0.409 . El mejor estimador para la Matriz de

Cuadro 4.32
Estimadores Robustos
 Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = N(0,9)$

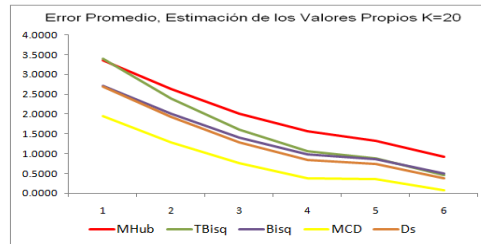
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

	Eps=0,10											
	5				10				k			
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	4.944 ± 4.911	---	---	63.671 ± 4.686	---	---	---	1.150 ± 5.097	---	---	---
	Mhuber	0.042 ± 0.943	2.883	1.289	0.019 ± 0.615	1.633	0.756	0.007 ± 0.433	0.974	0.471	---	---
	T-Bic	0.037 ± 0.986	2.799	1.274	0.033 ± 0.670	1.489	0.731	0.010 ± 0.465	1.055	0.510	---	---
	Bic	0.039 ± 0.863	2.485	1.129	0.027 ± 0.595	1.444	0.689	0.010 ± 0.409	0.907	0.442	---	---
	MCD	0.035 ± 0.962	2.750	1.249	0.026 ± 0.615	1.455	0.699	0.008 ± 0.429	1.005	0.481	---	---
	DS	0.026 ± 0.845	2.321	1.064	0.023 ± 0.612	1.461	0.699	0.010 ± 0.427	0.960	0.466	---	---
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.470 ± 5.177	47.289	20.312	6.337 ± 3.243	19.119	9.566	5.379 ± 2.394	13.941	7.298	---	---
	T-Bic	8.928 ± 6.626	92.030	35.861	6.210 ± 3.592	26.344	12.716	4.883 ± 2.440	15.457	7.593	---	---
	Bic	7.039 ± 5.117	32.613	14.923	4.926 ± 2.868	16.233	8.009	4.054 ± 2.051	12.810	6.305	---	---
	MCD	8.534 ± 5.472	45.723	19.910	3.851 ± 2.245	11.951	6.016	2.615 ± 1.342	7.361	3.773	---	---
	DS	9.394 ± 8.029	90.188	35.870	4.557 ± 2.986	18.785	8.776	3.885 ± 1.969	11.558	5.804	---	---
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	6.806 ± 6.366	49.479	20.884	5.596 ± 4.286	22.414	10.765	5.017 ± 3.071	14.916	7.668	---	---
	T-Bic	5.626 ± 7.406	78.491	30.507	4.761 ± 5.064	28.860	12.895	4.213 ± 3.462	16.500	8.059	---	---
	Bic	4.935 ± 6.207	40.400	17.181	4.020 ± 4.035	19.370	9.142	3.623 ± 2.772	13.562	6.652	---	---
	MCD	5.646 ± 7.199	46.916	19.920	2.714 ± 3.472	16.880	7.689	2.087 ± 2.246	9.209	4.514	---	---
	DS	6.569 ± 7.763	116.387	43.573	3.566 ± 3.910	20.419	9.298	3.394 ± 2.753	12.545	6.231	---	---

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

Covarianzas y sus Valores Propios correspondientes, para los valores de $k=5, 10$ y 20 es el Bicuadrático. (Véase Cuadro 4.32)

Por otro lado, cuando la contaminación se incrementa al 10%, para $k=5$ y 10 el estimador más próximo al Vector real de Medias es el DS; y siendo para $k=20$ el Bicuadrático con 0.024 ± 0.388 de distancia al Vector real. Se tiene además que los mejores estimadores para la Matriz de Covarianza para $k=5$ sigue siendo el Bicuadrático el estimador cercano a la Matriz real y para $k=10$ y 20 es el MCD.

Para $k=5$ y $\varepsilon = 0.05$ en 15 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon = 0.10$ el algoritmo no converge en 12 ocasiones.

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$,

Contaminada Por Variable con $\varepsilon = 0.05$ y 0.10 y $H = U(-5, 5)$

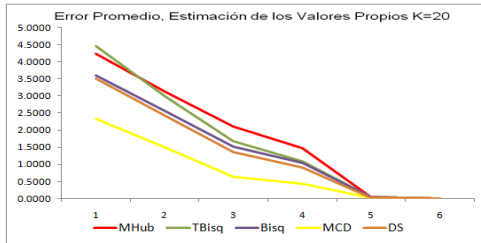
De acuerdo a los datos que se presentan en el Cuadro 4.33 se determina que el mejor estimador para el Vector de Medias, cuando existe un 5% de contaminación y con $k=5$, es el DS donde la distancia promedio al Vector real es 0.035 ± 0.865 ; mientras que para la Matriz de Varianzas y Covarianzas los Estimadores M de Huber y MCD son los que mejor se aproximan al valor real de la Matriz así como a los Valores Propios asociados a la misma.

Cuadro 4.33
Estimadores Robustos
 Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$, Contaminada $\varepsilon = 0.05$ y 0.10 y $H = U(-5,5)$

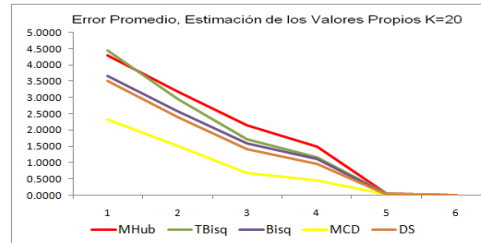
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

Eps=0,05													
k													
5													
10													
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	4.053 ±	4.977	---	---	2.557 ±	4.290	---	---	1.482 ±	5.021	---	---
	Mhuber	0.044 ±	0.964	2.648	1.218	0.014 ±	0.625	1.530	0.723	0.017 ±	0.430	1.057	0.501
	T-Bic	0.070 ±	1.043	3.275	1.463	0.013 ±	0.668	1.719	0.800	0.012 ±	0.462	1.168	0.547
	Bic	0.045 ±	0.958	2.653	1.219	0.006 ±	0.594	1.423	0.674	0.012 ±	0.407	0.967	0.462
	MCD	0.058 ±	1.029	2.956	1.348	0.013 ±	0.615	1.596	0.741	0.016 ±	0.431	0.951	0.466
	DS	0.035 ±	0.864	2.165	1.021	0.016 ±	0.594	1.401	0.670	0.014 ±	0.424	0.955	0.464
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	10.278 ±	7.783	52.328	23.463	7.091 ±	3.788	26.531	12.470	5.846 ±	2.829	14.787	7.821
	T-Bic	12.754 ±	10.760	78.584	34.033	7.681 ±	5.217	38.645	17.181	5.723 ±	3.078	17.453	8.751
	Bic	10.193 ±	8.780	62.416	27.130	6.260 ±	4.172	29.072	13.168	4.802 ±	2.614	14.119	7.179
	MCD	11.424 ±	9.105	76.052	32.194	4.233 ±	2.835	15.311	7.460	2.874 ±	1.866	9.201	4.647
	DS	15.627 ±	14.327	381.452	137.135	5.411 ±	3.803	28.458	12.557	4.554 ±	2.563	13.504	6.874
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.061 ±	8.104	50.260	22.142	5.945 ±	4.083	24.733	11.587	5.089 ±	3.033	13.971	7.364
	T-Bic	8.572 ±	10.122	80.266	32.987	5.911 ±	5.785	33.496	15.064	4.763 ±	3.719	17.392	8.625
	Bic	7.265 ±	8.501	61.005	25.590	5.008 ±	4.612	28.585	12.735	4.080 ±	2.966	13.709	6.918
	MCD	7.864 ±	9.237	69.216	28.772	3.100 ±	3.677	16.189	7.655	2.271 ±	2.313	10.144	4.909
	DS	11.112 ±	11.336	307.910	110.119	4.227 ±	4.258	26.541	11.675	3.813 ±	2.965	12.891	6.557
Eps=0,10													
k													
5													
10													
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	1.593 ±	4.373	---	---	0.900 ±	4.641	---	---	13.297 ±	4.872	---	---
	Mhuber	0.019 ±	0.891	2.896	1.269	0.012 ±	0.595	1.432	0.680	0.011 ±	0.407	0.966	0.461
	T-Bic	0.024 ±	0.980	3.350	1.451	0.016 ±	0.666	1.515	0.732	0.006 ±	0.432	1.064	0.501
	Bic	0.022 ±	0.877	2.598	1.166	0.011 ±	0.579	1.381	0.657	0.005 ±	0.386	0.980	0.457
	MCD	0.033 ±	0.975	2.577	1.195	0.018 ±	0.615	1.439	0.691	0.012 ±	0.393	1.064	0.490
	DS	0.031 ±	0.845	2.471	1.115	0.010 ±	0.585	1.292	0.629	0.007 ±	0.393	0.990	0.463
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	10.394 ±	7.812	65.090	27.765	6.993 ±	4.036	25.424	12.151	5.970 ±	2.821	16.493	8.428
	T-Bic	13.363 ±	10.924	157.130	60.473	7.917 ±	4.772	30.836	14.509	5.746 ±	3.023	19.665	9.478
	Bic	10.848 ±	8.056	95.845	38.249	6.128 ±	3.751	25.633	11.838	4.878 ±	2.654	14.361	7.298
	MCD	11.199 ±	8.316	65.713	28.409	4.316 ±	2.926	15.212	7.485	2.885 ±	1.743	8.820	4.483
	DS	12.946 ±	11.144	225.422	83.171	5.564 ±	3.884	40.292	16.580	4.596 ±	2.473	14.892	7.320
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.230 ±	8.063	60.465	25.586	5.874 ±	4.587	24.155	11.539	5.204 ±	3.052	15.647	7.968
	T-Bic	9.121 ±	10.155	139.023	52.766	6.180 ±	5.699	28.504	13.461	4.811 ±	3.554	17.524	8.630
	Bic	7.889 ±	8.295	86.085	34.090	4.960 ±	4.503	24.305	11.256	4.162 ±	2.959	13.597	6.906
	MCD	7.771 ±	8.533	77.515	31.273	3.190 ±	3.747	16.872	7.936	2.304 ±	2.250	9.232	4.596
	DS	9.409 ±	9.652	194.914	71.325	4.393 ±	4.393	37.390	15.392	3.873 ±	2.921	13.530	6.775

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

Para una población contaminada en un 10%; se tiene que las mejores estimaciones para el Vector de Medias para los diferentes valores de k , están dadas por el Estimador DS ubicándose, para el caso $k=5$, a una distancia promedio de 0.035 ± 0.864 con respecto al Vector Real de Medias. Por otro lado, para la Matriz de Varianzas y Covarianzas, el Estimador que más se aproxima a la Matriz real es el M-Huber cuando $k=5$ con una distancia promedio de 8.23 ± 8.063 ; situación que difiere para valores de $k=10$ y 20 donde el Estimador MCD es el que mejor estima la Matriz real de Varianzas y Covarianzas.

En el presente caso, cuando $k=5$ y $\varepsilon=0.05$ en 30 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon=0.10$ el algoritmo no converge en 18 ocasiones.

El caso de una Población Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$, Contaminada Por Variable con $\varepsilon=0.05$ y 0.10 y $H = U(-5,5)$

En el Cuadro 4.34 se muestran los resultados de la Simulación; permitiendo efectuar inferencias sobre los mejores estimadores robustos cuando la población tiene un 5% y 10% de contaminación para diversos valores de k en cada escenario.

Así para $k=5$ con el 5% de contaminación, el Estimador DS es el estimador que se encuentra más próximo al Vector real de Medias cuya distancia promedio es 0.041 ± 0.882 y distancia máxima de 2.090, el Índice asociado es 1.004; para $k=10$ se tiene que el Estimador Bicuadrático se encuentra a una Distancia Promedio de 0.030 ± 0.602 con respecto al Vector de Medias real. Mientras que, el mejor estimador, para la Matriz de Covarianzas cuando $k=5$, es el Bicuadrático con una Distancia Media de 4.871 ± 6.032 .

Al incrementar el porcentaje de contaminación a 10% el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador DS se encuentra a una distancia promedio de 0.028 ± 0.812 con respecto al Vector de Medias Real; para $k=5$. Se tiene además, que el Estimador MCD es el más próximo a la Matriz Real de Covarianzas para $k=5, 10$ y 20 ; con distancia promedio de 5.613 ± 7.297 ; 2.842 ± 3.588 y 2.051 ± 2.215 respectivamente.

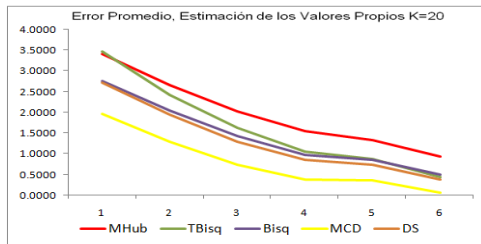
Para $k=5$ y $\varepsilon = 0.05$ en 19 de las 500 repeticiones el algoritmo del estimador M de Huber no converge, mientras que para mismo $k=5$ y $\varepsilon = 0.10$ el algoritmo no converge en 15 ocasiones.

Cuadro 4.34
Estimadores Robustos
 Caso $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$ Contaminada $\varepsilon = 0.05$ y 0.10 y $H = U(-5,5)$

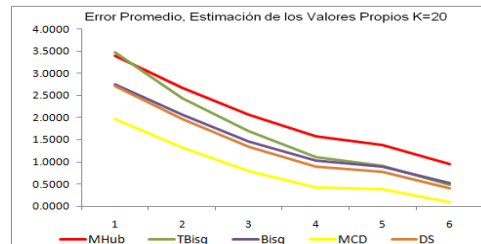
Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

Eps=0,05													
Vector de Medias	5				10				10				
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
	Clásico	3.626 ±	4.494	---	---	3.615 ±	4.636	---	---	3.507 ±	4.254	---	---
Mhuber	0.044 ±	0.947	2.131	1.041	0.024 ±	0.620	1.432	0.692	0.014 ±	0.417	0.957	0.463	
T-Bic	0.041 ±	1.062	2.607	1.237	0.028 ±	0.678	1.529	0.745	0.015 ±	0.444	1.085	0.508	
Bic	0.041 ±	0.941	2.281	1.088	0.030 ±	0.602	1.258	0.630	0.014 ±	0.408	0.965	0.462	
MCD	0.041 ±	1.030	2.614	1.228	0.027 ±	0.637	1.474	0.713	0.015 ±	0.422	0.966	0.468	
DS	0.041 ±	0.882	2.090	1.004	0.029 ±	0.593	1.342	0.655	0.017 ±	0.407	0.961	0.462	
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	
	Mhuber	8.465 ±	5.419	50.638	21.507	6.548 ±	2.771	19.089	9.469	5.273 ±	2.280	14.528	7.360
	T-Bic	8.848 ±	6.161	54.446	23.152	6.490 ±	3.698	29.898	13.362	4.770 ±	2.375	13.217	6.787
	Bic	6.919 ±	4.771	33.468	15.053	5.198 ±	3.070	16.301	8.189	3.968 ±	1.980	11.461	5.803
	MCD	8.662 ±	5.417	36.574	16.885	4.038 ±	2.376	13.007	6.474	2.532 ±	1.465	6.693	3.563
	DS	9.476 ±	8.177	77.427	31.693	4.947 ±	3.122	26.612	11.561	3.780 ±	1.941	10.858	5.526
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	
	Mhuber	6.811 ±	6.553	54.763	22.709	5.798 ±	3.667	21.950	10.472	4.904 ±	3.000	14.592	7.499
	T-Bic	5.605 ±	7.491	69.649	27.582	4.984 ±	5.174	29.607	13.255	4.092 ±	3.406	16.458	7.985
	Bic	4.871 ±	6.032	40.504	17.136	4.246 ±	4.149	20.200	9.531	3.527 ±	2.696	12.236	6.153
	MCD	5.789 ±	7.400	48.326	20.506	2.865 ±	3.712	16.740	7.772	1.995 ±	2.210	9.943	4.716
	DS	6.605 ±	7.831	79.186	31.207	3.871 ±	4.119	27.766	11.919	3.280 ±	2.711	13.322	6.438
Eps=0,10													
Vector de Medias	5				10				10				
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
	Clásico	4.101 ±	4.222	---	---	2.429 ±	4.432	---	---	2.537 ±	4.787	---	---
Mhuber	0.042 ±	0.916	2.028	0.995	0.024 ±	0.598	1.505	0.709	0.008 ±	0.423	1.015	0.482	
T-Bic	0.034 ±	1.012	2.342	1.129	0.021 ±	0.673	1.645	0.780	0.008 ±	0.470	1.085	0.521	
Bic	0.032 ±	0.884	2.188	1.035	0.017 ±	0.589	1.386	0.664	0.009 ±	0.410	0.957	0.459	
MCD	0.028 ±	0.985	2.318	1.110	0.020 ±	0.633	1.492	0.715	0.007 ±	0.437	0.967	0.470	
DS	0.028 ±	0.812	1.922	0.921	0.023 ±	0.587	1.385	0.665	0.008 ±	0.426	0.958	0.464	
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	
	Mhuber	8.630 ±	4.324	30.013	14.322	6.505 ±	3.604	17.782	9.297	5.329 ±	2.224	15.686	7.746
	T-Bic	9.410 ±	6.397	54.562	23.456	6.683 ±	4.181	20.700	10.521	4.800 ±	2.429	15.417	7.549
	Bic	7.371 ±	5.315	42.325	18.337	5.153 ±	3.154	16.763	8.357	4.004 ±	2.030	14.901	6.978
	MCD	8.511 ±	5.762	32.264	15.512	3.959 ±	2.516	13.901	6.792	2.562 ±	1.363	7.842	3.922
	DS	9.783 ±	8.773	93.367	37.307	4.856 ±	3.133	23.513	10.501	3.826 ±	1.991	12.358	6.058
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	
	Mhuber	7.089 ±	5.777	40.113	17.660	5.817 ±	4.397	20.527	10.247	4.981 ±	2.961	16.815	8.252
	T-Bic	6.056 ±	8.030	64.263	26.116	5.241 ±	5.368	27.110	12.573	4.160 ±	3.374	16.289	7.941
	Bic	5.287 ±	6.499	42.188	17.992	4.257 ±	4.076	21.118	9.817	3.590 ±	2.724	15.214	7.176
	MCD	5.613 ±	7.297	39.305	17.405	2.842 ±	3.588	18.214	8.215	2.051 ±	2.215	8.949	4.405
	DS	7.001 ±	8.476	88.979	34.819	3.848 ±	4.135	25.524	11.169	3.353 ±	2.727	13.275	6.451

$\varepsilon = 0.05$



$\varepsilon = 0.10$



Elaborado por: Néstor Montaña P.

Población Normal $N(\mathbf{0}, \Sigma)$ **y Cauchy** $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminadas bajo el modelo (2.1), con $\varepsilon = 0.30$ **y** $H = N(\mathbf{0}, 9\Sigma_1)$

En el Cuadro 4.35 se observa los resultados de la Simulación; permitiendo efectuar inferencias sobre los mejores estimadores robustos cuando la muestra sigue la Distribución Normal y cuando sigue la Distribución Cauchy, ambas 30% contaminada y clasificada para diferentes tamaños muestrales k .

Así, al trabajar con datos que provienen de una Distribución Normal se tiene que para $k=5$ y 20 , el Estimador Bicuadrático es el estimador que se encuentra más próximo al Vector real de Medias; siendo la distancia promedio, cuando $k=5$, 0.024 ± 0.639 y el Índice asociado es 0.689 ; cuando $k=10$ el Estimador T-Bicuadrático se encuentra a una Distancia Promedio de 0.017 ± 0.446 con respecto al Vector de Medias real. Mientras que, el mejor estimador, para la Matriz de Varianzas y Covarianzas con sus respectivos Valores Propios, cuando $k=5$, notoriamente es el Bicuadrático, puesto que el índice relacionado es 2.684 , con una Distancia Media de 0.715 ± 1.783 de la Matriz real; siendo el mismo estimador el que se aproxima a la Matriz real cuando $k=10$; para $k=20$ el Estimador DS es el mejor de los estimadores que conforman este estudio, debido a que la

distancia promedio entre la estimación y el valor real es 0.682 ± 0.962 .

Para el caso de la generación de datos a partir de una Distribución Cauchy, el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador DS es el que menor índice presenta para cuando k toma los valores de 5 y 20 siendo el que mejor se aproxima al Vector real de Medias; sin embargo para $k=10$ es estimador que más se aproxima es el Bicuadrático con una distancia promedio de 0.040 ± 0.634 . Se tiene además, que el Estimador Bicuadrático es el más próximo a la Matriz Real de Varianzas y Covarianzas para $k=5$, cuya distancia promedio es 9.166 ± 8.677 ; y cuando el valor de k se incrementa a 10 y a 20 el mejor estimador es el MCD con una distancia de 5.679 ± 5.566 y 4.871 ± 3.619 respectivamente.

Cuando la muestras son generadar por $CAU(\mathbf{0}, \Sigma)$, para $k=5$ en 25 de las 500 repeticiones el algoritmo del estimador M de Huber no converge.

Cuadro 4.35

Estimadores Robustos

Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$

Contaminada $\varepsilon = 0.30$ y $H = N(\mathbf{0}, 9\Sigma_1)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		k											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.026 ± 0.897	2.203	1.042	0.016 ± 0.643	1.333	0.664	0.016 ± 0.449	0.989	0.485			
	Mhuber	0.022 ± 0.661	1.517	0.733	0.017 ± 0.465	0.959	0.480	0.011 ± 0.325	0.754	0.363			
	T-Bic	0.023 ± 0.699	1.599	0.774	0.017 ± 0.454	0.921	0.464	0.007 ± 0.312	0.699	0.339			
	Bic	0.024 ± 0.639	1.405	0.689	0.017 ± 0.446	0.968	0.477	0.007 ± 0.308	0.674	0.329			
	MCD	0.024 ± 0.699	1.511	0.745	0.016 ± 0.449	0.976	0.480	0.006 ± 0.316	0.730	0.351			
	DS	0.019 ± 0.697	1.603	0.773	0.016 ± 0.453	0.976	0.481	0.008 ± 0.313	0.695	0.339			
Valores Propios	Clásico	7.060 ± 2.988	20.090	10.046	6.518 ± 2.287	13.513	7.439	6.085 ± 1.701	11.076	6.287			
	Mhuber	1.936 ± 1.280	6.088	3.101	1.508 ± 0.927	3.861	2.099	1.256 ± 0.626	3.008	1.630			
	T-Bic	1.704 ± 1.360	7.935	3.666	1.164 ± 0.929	3.545	1.879	0.938 ± 0.656	2.926	1.507			
	Bic	1.301 ± 1.166	4.538	2.335	1.047 ± 0.866	3.219	1.711	0.880 ± 0.642	2.674	1.398			
	MCD	2.628 ± 1.697	7.700	4.008	1.545 ± 1.027	5.159	2.577	1.176 ± 0.709	2.988	1.624			
	DS	1.472 ± 1.229	5.453	2.718	1.061 ± 0.870	3.701	1.877	0.886 ± 0.658	2.642	1.395			
Matriz de Covarianzas	Clásico	5.439 ± 4.891	19.370	9.900	5.387 ± 3.635	13.288	7.436	5.260 ± 2.549	10.884	6.231			
	Mhuber	1.358 ± 1.934	6.567	3.286	1.173 ± 1.348	4.159	2.227	1.052 ± 0.892	3.142	1.695			
	T-Bic	0.886 ± 2.169	9.246	4.100	0.800 ± 1.394	4.003	2.065	0.734 ± 0.947	3.057	1.579			
	Bic	0.715 ± 1.783	5.554	2.684	0.722 ± 1.311	3.676	1.903	0.689 ± 0.874	2.866	1.477			
	MCD	1.744 ± 2.655	8.654	4.351	1.123 ± 1.589	5.416	2.709	0.945 ± 1.000	3.094	1.680			
	DS	0.683 ± 2.054	7.024	3.254	0.681 ± 1.374	4.278	2.111	0.682 ± 0.962	2.783	1.475			
Vector de Medias	Clásico	5.106 ± 3.766	---	---	0.969 ± 3.407	---	---	1.462 ± 3.898	---	---			
	Mhuber	0.043 ± 0.902	2.588	1.178	0.039 ± 0.637	1.487	0.721	0.014 ± 0.462	1.046	0.507			
	T-Bic	0.064 ± 1.068	2.975	1.369	0.046 ± 0.763	1.776	0.862	0.019 ± 0.546	1.371	0.645			
	Bic	0.055 ± 0.945	2.782	1.261	0.040 ± 0.634	1.447	0.707	0.015 ± 0.452	1.052	0.506			
	MCD	0.066 ± 0.988	2.900	1.318	0.042 ± 0.707	1.926	0.892	0.003 ± 0.491	1.075	0.523			
	DS	0.060 ± 0.874	2.384	1.106	0.041 ± 0.612	1.542	0.732	0.011 ± 0.462	1.026	0.499			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	13.435 ± 5.573	41.530	20.179	12.264 ± 4.749	33.144	16.719	11.519 ± 3.584	21.024	12.042			
	T-Bic	14.276 ± 8.978	63.392	28.882	12.899 ± 6.121	39.063	19.361	11.040 ± 4.366	23.177	12.861			
	Bic	11.904 ± 7.269	43.934	21.036	10.351 ± 4.842	35.720	16.971	9.683 ± 3.918	21.015	11.539			
	MCD	11.581 ± 7.525	51.066	23.390	7.400 ± 4.195	29.523	13.706	5.888 ± 2.819	16.063	8.257			
	DS	15.642 ± 12.397	79.120	35.720	10.560 ± 5.979	45.279	20.606	9.204 ± 3.703	21.802	11.569			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	11.098 ± 6.662	41.275	19.679	10.565 ± 5.822	30.459	15.615	10.176 ± 4.069	20.092	11.445			
	T-Bic	10.142 ± 10.821	61.540	27.501	10.324 ± 8.141	39.531	19.332	9.406 ± 5.604	23.673	12.895			
	Bic	9.166 ± 8.677	44.689	20.844	8.611 ± 6.160	32.667	15.813	8.433 ± 4.266	20.560	11.086			
	MCD	8.243 ± 9.165	48.852	22.087	5.679 ± 5.566	27.895	13.046	4.871 ± 3.619	15.351	7.947			
	DS	11.780 ± 12.012	77.526	33.773	8.635 ± 6.828	39.933	18.466	7.903 ± 4.358	21.995	11.418			

Elaborado por: Néstor Montaña P.

Población Normal $N(\mathbf{0}, \Sigma)$ **y Cauchy** $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$

Contaminadas bajo el modelo (2.1), con $\varepsilon = 0.30$ y $H = Uesf(d = 5)$

Para el caso de muestras generadas a partir de la Distribución Normal, cuando k toma el valor de 5 se obtiene que el estimador que más se aproxima al verdadero valor del Vector de Medias, en relación a los demás estimadores, es el Estimador M-Huber cuya distancia promedio al vector real es 0.021 ± 0.609 con un Índice asociado de 0.683. Bajo el mismo razonamiento, el Estimador M-Huber sigue siendo el mejor estimador con distancia promedio de 0.007 ± 0.433 y con 0.463 como Índice cuando k toma el valor de 10 mientras que cuando $k=20$ el estimador MCD es el mejor, su distancia promedio es 0.005 ± 0.314 . El mejor estimador para la Matriz de Covarianzas y sus Valores Propios correspondientes, para todos los valores de k es el Estimador M-Huber, además es el que se encuentra a una distancia más cercana con respecto a la Matriz de Covarianzas poblacional (Véase Cuadro 4.36)

Por otro lado, cuando se tienen datos a partir de una Distribución Cauchy, los estimadores que presentan menor Sesgo Promedio Total con respecto a los Valores Propios son el Estimador M de Huber y el Bicuadrático. Se tiene además que el mejor estimador

para la Matriz de Covarianza para $k=5$ es el estimador M de Huber mientras que para $k=10$ y $k=20$ es el Estimador MCD

Cuadro 4.36
Estimadores Robustos
Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminada $\varepsilon = 0.30$ y $H = Uesf(d = 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		k											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.021 ± 0.659	1.300	0.660	0.009 ± 0.469	1.040	0.506	0.011 ± 0.320	0.684	0.338			
	Mhuber	0.021 ± 0.609	1.419	0.683	0.007 ± 0.433	0.949	0.463	0.007 ± 0.310	0.677	0.331			
	T-Bic	0.035 ± 0.656	1.721	0.804	0.012 ± 0.446	1.163	0.540	0.006 ± 0.316	0.705	0.343			
	Bic	0.025 ± 0.613	1.478	0.705	0.010 ± 0.441	1.074	0.508	0.005 ± 0.314	0.712	0.344			
	MCD	0.022 ± 0.657	1.508	0.729	0.010 ± 0.463	1.108	0.527	0.005 ± 0.325	0.750	0.360			
	DS	0.028 ± 0.668	1.635	0.777	0.010 ± 0.470	1.080	0.520	0.005 ± 0.314	0.712	0.343			
Valores Propios	Clásico	1.459 ± 0.819	3.560	1.946	1.328 ± 0.619	3.031	1.660	1.294 ± 0.469	2.364	1.376			
	Mhuber	0.558 ± 0.767	2.667	1.331	0.368 ± 0.563	1.724	0.885	0.289 ± 0.397	1.069	0.585			
	T-Bic	1.595 ± 1.266	6.000	2.954	1.085 ± 0.877	3.485	1.815	0.926 ± 0.600	2.632	1.386			
	Bic	1.197 ± 1.072	3.917	2.062	0.962 ± 0.847	3.403	1.737	0.858 ± 0.581	2.543	1.327			
	MCD	2.463 ± 1.503	7.387	3.784	1.403 ± 0.925	3.853	2.060	1.125 ± 0.642	2.947	1.571			
	DS	1.272 ± 1.274	5.837	2.794	0.908 ± 0.858	2.955	1.574	0.797 ± 0.582	2.523	1.301			
Matriz de Covarianzas	Clásico	1.346 ± 1.811	5.831	2.996	1.281 ± 1.275	4.106	2.221	1.264 ± 0.890	3.074	1.743			
	Mhuber	0.254 ± 1.273	3.880	1.802	0.195 ± 0.879	2.463	1.179	0.189 ± 0.602	1.463	0.751			
	T-Bic	0.819 ± 2.114	7.142	3.358	0.744 ± 1.340	3.901	1.995	0.735 ± 0.905	2.789	1.476			
	Bic	0.653 ± 1.683	4.856	2.398	0.658 ± 1.254	3.848	1.920	0.679 ± 0.857	2.743	1.426			
	MCD	1.623 ± 2.529	7.820	3.991	1.016 ± 1.472	4.378	2.289	0.912 ± 0.964	3.170	1.682			
	DS	0.464 ± 1.932	6.847	3.081	0.529 ± 1.324	3.634	1.829	0.609 ± 0.891	2.724	1.408			
Vector de Medias	Clásico	3.577 ± 4.034	---	---	2.300 ± 3.959	---	---	4.446 ± 3.668	---	---			
	Mhuber	0.018 ± 0.904	2.243	1.055	0.015 ± 0.599	1.380	0.665	0.007 ± 0.412	0.947	0.455			
	T-Bic	0.026 ± 1.083	2.752	1.287	0.012 ± 0.757	1.726	0.832	0.007 ± 0.525	1.318	0.617			
	Bic	0.019 ± 0.976	2.757	1.251	0.016 ± 0.662	1.412	0.696	0.002 ± 0.463	1.088	0.518			
	MCD	0.026 ± 1.037	2.717	1.260	0.017 ± 0.716	1.692	0.808	0.006 ± 0.468	1.095	0.523			
	DS	0.041 ± 0.852	2.316	1.069	0.019 ± 0.628	1.289	0.645	0.005 ± 0.436	0.927	0.456			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	10.567 ± 6.030	46.356	20.984	8.952 ± 4.511	37.307	16.923	8.405 ± 3.260	21.313	10.993			
	T-Bic	19.203 ± 13.917	101.350	44.823	17.454 ± 10.151	67.820	31.809	14.420 ± 6.674	55.136	25.410			
	Bic	14.960 ± 10.828	81.751	35.846	12.678 ± 7.201	42.285	20.721	11.778 ± 5.238	43.956	20.324			
	MCD	13.166 ± 8.902	55.982	26.016	7.685 ± 4.848	33.731	15.421	5.547 ± 2.669	16.613	8.276			
	DS	27.565 ± 21.373	255.311	101.416	14.287 ± 9.803	74.584	32.891	11.282 ± 5.831	39.784	18.966			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	8.664 ± 6.709	43.840	19.738	7.796 ± 5.181	35.738	16.238	7.530 ± 3.716	19.385	10.211			
	T-Bic	13.869 ± 13.562	102.928	43.453	14.093 ± 10.898	72.784	32.592	12.298 ± 7.615	45.818	21.910			
	Bic	11.348 ± 10.523	83.602	35.158	10.544 ± 7.721	41.015	19.760	10.202 ± 5.847	36.497	17.515			
	MCD	9.510 ± 9.728	57.001	25.413	6.030 ± 5.449	29.620	13.700	4.609 ± 3.437	16.264	8.103			
	DS	21.075 ± 19.602	270.729	103.802	11.798 ± 9.599	70.072	30.490	9.745 ± 6.146	34.810	16.900			

Elaborado por: Néstor Montaña P.

Población Normal $N(\mathbf{0}, \Sigma)$ **y Cauchy** $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$

Contaminadas bajo el modelo (2.1), con $\varepsilon = 0.30$ y $H = Uesf(d = 5)$

De acuerdo a los resultados que se presentan en el Cuadro 4.37 se determina que el mejor estimador para el Vector de Medias, cuando se trabaja con datos que siguen una Distribución Normal, con $k=5$, es el Bicuadrático cuya distancia promedio al Vector de Medias verdadero es 0.028 ± 0.671 ; cuando k varía a 10 el estimador más cercano es el T-Bicuadrático ubicado a una distancia promedio de 0.013 ± 0.481 del Vector real y es el Estimador MCD el que obtiene menor distancia promedio, 0.005 ± 0.275 , cuando $k=20$. Mientras que para la Matriz de Varianzas y Covarianzas el Estimador MCD junto con el Bicuadrático son los que “mejor” se aproximan a $\Sigma = \Sigma_2$.

Cuando se trabaja con una Muestra con datos que se generan a partir de una Distribución Cauchy, se observa que las mejores estimaciones para el Vector de Medias para los diferentes valores de k , están dadas por los Estimadores M de Huber y Bicuadrático. Por otro lado, para la Matriz de Varianzas y Covarianzas, el Estimador que más se aproxima al verdadero valor de $\Sigma = \Sigma_2$ es el Bicuadrático cuando $k=5$ y el MCD para los dos valores restantes de k .

Cuadro 4.37

Estimadores Robustos

Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$

Contaminada $\varepsilon = 0.30$ y $H = Uesf(d = 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		k											
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.025 ± 0.664	1.445	0.711	0.016 ± 0.456	0.946	0.473	0.011 ± 0.315	0.728	0.351			
	Mhuber	0.026 ± 0.639	1.431	0.699	0.016 ± 0.434	0.951	0.467	0.009 ± 0.306	0.679	0.331			
	T-Bic	0.034 ± 0.754	1.594	0.794	0.013 ± 0.481	1.020	0.505	0.009 ± 0.337	0.737	0.361			
	Bic	0.028 ± 0.671	1.360	0.686	0.015 ± 0.443	1.021	0.493	0.009 ± 0.318	0.695	0.340			
	MCD	0.031 ± 0.769	1.511	0.770	0.015 ± 0.499	1.129	0.548	0.009 ± 0.344	0.746	0.366			
	DS	0.030 ± 0.757	1.658	0.815	0.020 ± 0.476	1.059	0.518	0.010 ± 0.327	0.748	0.362			
Valores Propios	Clásico	1.776 ± 0.660	3.324	1.920	1.436 ± 0.488	2.614	1.513	1.302 ± 0.336	2.274	1.304			
	Mhuber	1.465 ± 0.747	3.202	1.804	1.083 ± 0.540	2.457	1.360	0.905 ± 0.358	1.871	1.044			
	T-Bic	1.991 ± 1.234	5.704	2.976	1.231 ± 0.751	2.998	1.660	0.909 ± 0.448	1.976	1.111			
	Bic	1.393 ± 0.859	3.432	1.895	0.961 ± 0.583	2.502	1.348	0.744 ± 0.401	1.792	0.979			
	MCD	2.989 ± 1.389	7.226	3.868	1.643 ± 0.802	3.695	2.047	1.204 ± 0.492	2.520	1.405			
	DS	1.752 ± 1.219	4.959	2.643	1.085 ± 0.680	3.216	1.660	0.801 ± 0.434	1.878	1.038			
Matriz de Covarianzas	Clásico	1.309 ± 1.799	5.518	2.876	1.243 ± 1.249	3.819	2.104	1.232 ± 0.886	3.130	1.750			
	Mhuber	0.958 ± 1.680	5.176	2.605	0.854 ± 1.165	3.597	1.872	0.814 ± 0.804	2.749	1.456			
	T-Bic	0.918 ± 2.331	8.655	3.968	0.746 ± 1.437	4.576	2.253	0.670 ± 0.990	3.135	1.598			
	Bic	0.697 ± 1.789	5.683	2.723	0.617 ± 1.191	3.690	1.832	0.588 ± 0.828	2.660	1.359			
	MCD	1.838 ± 3.006	10.828	5.224	1.135 ± 1.674	5.709	2.839	0.965 ± 1.093	3.625	1.894			
	DS	0.733 ± 2.167	7.757	3.552	0.576 ± 1.389	4.302	2.089	0.587 ± 0.906	2.796	1.430			
Vector de Medias	Clásico	1.071 ± 3.644	---	---	1.912 ± 3.930	---	---	1.781 ± 3.920	---	---			
	Mhuber	0.038 ± 0.856	1.762	0.886	0.019 ± 0.587	1.354	0.653	0.008 ± 0.407	0.863	0.426			
	T-Bic	0.044 ± 0.972	2.278	1.098	0.019 ± 0.639	1.421	0.693	0.010 ± 0.439	0.941	0.463			
	Bic	0.038 ± 0.832	1.822	0.897	0.016 ± 0.577	1.369	0.654	0.008 ± 0.393	0.844	0.415			
	MCD	0.042 ± 0.932	2.126	1.033	0.013 ± 0.670	1.420	0.701	0.015 ± 0.455	1.038	0.503			
	DS	0.032 ± 0.914	1.990	0.979	0.016 ± 0.637	1.404	0.686	0.011 ± 0.450	0.931	0.464			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	6.387 ± 2.421	17.037	8.615	5.682 ± 1.867	13.006	6.852	5.221 ± 1.241	8.590	5.018			
	T-Bic	7.063 ± 3.155	24.952	11.724	5.772 ± 1.860	11.041	6.224	5.307 ± 1.332	8.706	5.115			
	Bic	5.649 ± 2.497	15.515	7.887	5.274 ± 1.886	10.917	6.026	5.060 ± 1.415	8.091	4.855			
	MCD	7.205 ± 2.933	18.052	9.397	5.043 ± 1.904	10.043	5.663	4.271 ± 1.295	7.423	4.330			
	DS	7.477 ± 3.904	29.158	13.513	5.684 ± 2.107	17.398	8.396	5.206 ± 1.454	8.701	5.120			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	5.411 ± 3.679	22.484	10.525	5.211 ± 3.150	15.245	7.869	4.995 ± 2.154	10.499	5.883			
	T-Bic	4.992 ± 5.872	28.863	13.242	5.048 ± 3.772	14.850	7.890	5.001 ± 2.539	10.855	6.132			
	Bic	4.500 ± 4.387	19.777	9.555	4.722 ± 3.310	13.257	7.096	4.789 ± 2.382	10.496	5.889			
	MCD	5.162 ± 6.100	23.232	11.498	4.081 ± 3.865	13.199	7.048	3.807 ± 2.656	10.164	5.542			
	DS	5.392 ± 6.072	36.751	16.072	4.718 ± 4.057	19.082	9.286	4.753 ± 2.882	11.751	6.462			

Elaborado por: Néstor Montaña P.

Población Normal $N(\mathbf{0}, \Sigma)$ **y Cauchy** $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminadas bajo el modelo (2.1), con $\varepsilon = 0.30$ **y**
 $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$

De la simulación se concluye que cuando la Muestra es generada a través de la Distribución Normal se tiene que para $k=5$ y 10 , el Estimador Bicuadrático es el que mejor estima al Vector de Medias, en particular, cuando $k=5$, su Distancia Promedio es igual a 0.019 ± 0.607 , mientras que su peor estimación se encuentra a una distancia de 1.323 , obteniendo así Índice de 0.649 ; cuando $k=10$ su Distancia Promedio es igual a 0.023 ± 0.435 , mientras que su peor estimación se encuentra a una distancia de 0.944 , obteniendo así Índice de 0.467 ; para los Valores Propios asociados a $\Sigma = \Sigma_1$, el Estimador M-Huber es el mejor estimador para todos los tamaños muestrales estudiados (véase el Cuadro 4.38)

Para el caso de la generación de datos a partir de una Distribución Cauchy, el Estimador M-Huber es el que menor índice presenta para cuando k toma el valor de 5 mientras que para $k=10$ y 20 el estimador que más se aproxima es el DS con una distancia promedio de 0.026 ± 0.628 y 0.017 ± 0.445 respectivamente. Se tiene además, para $k=10$ y $k=20$ el Estimador MCD es el más próximo al verdadero valor de la Matriz de Covarianzas y los Valores Propios.

Cuadro 4.38

Estimadores Robustos

Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$

Contaminada $\varepsilon = 0.30$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_1)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.135 ± 0.889	2.338	1.120	0.132 ± 0.600	1.481	0.737	0.127 ± 0.469	1.160	0.585							
	Mhuber	0.052 ± 0.648	1.573	0.757	0.052 ± 0.431	1.036	0.507	0.044 ± 0.319	0.748	0.370							
	T-Bic	0.013 ± 0.671	1.469	0.718	0.022 ± 0.451	0.991	0.488	0.011 ± 0.306	0.674	0.331							
	Bic	0.019 ± 0.607	1.323	0.649	0.023 ± 0.435	0.944	0.467	0.010 ± 0.306	0.699	0.338							
	MCD	0.020 ± 0.649	1.450	0.706	0.023 ± 0.448	0.976	0.482	0.010 ± 0.305	0.691	0.335							
	DS	0.020 ± 0.680	1.487	0.729	0.022 ± 0.459	0.983	0.488	0.011 ± 0.313	0.695	0.340							
Valores Propios	Clásico	7.174 ± 3.146	15.910	8.743	6.303 ± 2.273	11.974	6.850	6.079 ± 1.733	10.291	6.034							
	Mhuber	1.386 ± 1.137	4.652	2.392	0.913 ± 0.719	2.803	1.478	0.768 ± 0.556	2.046	1.123							
	T-Bic	1.575 ± 1.424	6.350	3.116	1.066 ± 0.891	3.031	1.663	0.929 ± 0.637	2.801	1.455							
	Bic	1.191 ± 1.149	4.443	2.261	0.940 ± 0.775	3.123	1.613	0.871 ± 0.610	2.435	1.305							
	MCD	2.459 ± 1.549	7.516	3.841	1.385 ± 0.897	4.026	2.103	1.127 ± 0.680	2.742	1.516							
	DS	1.373 ± 1.252	5.763	2.796	0.942 ± 0.828	3.130	1.633	0.859 ± 0.630	2.420	1.303							
Matriz de Covarianzas	Clásico	5.531 ± 5.061	18.995	9.862	5.236 ± 3.501	12.798	7.178	5.253 ± 2.575	10.718	6.182							
	Mhuber	0.895 ± 1.691	5.754	2.780	0.667 ± 1.073	3.281	1.673	0.610 ± 0.756	2.157	1.174							
	T-Bic	0.827 ± 2.086	6.505	3.139	0.735 ± 1.294	3.749	1.926	0.729 ± 0.899	2.931	1.520							
	Bic	0.647 ± 1.697	4.946	2.430	0.650 ± 1.179	3.742	1.857	0.684 ± 0.858	2.711	1.418							
	MCD	1.651 ± 2.530	7.524	3.902	1.011 ± 1.411	4.828	2.417	0.908 ± 0.942	2.931	1.594							
	DS	0.609 ± 1.972	6.858	3.146	0.583 ± 1.285	3.842	1.903	0.659 ± 0.900	2.655	1.405							

		5				k				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	6.348 ± 3.902	---	---	5.133 ± 3.680	---	---	5.230 ± 3.806	---	---							
	Mhuber	0.098 ± 0.896	2.292	1.095	0.082 ± 0.637	1.575	0.764	0.073 ± 0.460	1.258	0.597							
	T-Bic	0.074 ± 1.096	3.130	1.433	0.053 ± 0.790	2.077	0.973	0.033 ± 0.581	1.454	0.689							
	Bic	0.070 ± 0.968	2.737	1.258	0.063 ± 0.665	1.556	0.761	0.059 ± 0.447	1.163	0.556							
	MCD	0.058 ± 1.044	2.986	1.363	0.033 ± 0.719	1.554	0.769	0.017 ± 0.477	1.277	0.590							
	DS	0.047 ± 0.839	2.602	1.163	0.026 ± 0.628	1.456	0.704	0.017 ± 0.445	1.070	0.511							
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---							
	Mhuber	14.180 ± 6.259	40.493	20.310	13.232 ± 5.564	31.719	16.838	12.477 ± 3.994	26.799	14.423							
	T-Bic	17.478 ± 10.609	117.340	48.476	15.567 ± 8.290	49.586	24.481	13.591 ± 5.236	32.565	17.131							
	Bic	12.732 ± 8.582	40.318	20.544	10.813 ± 5.418	29.393	15.208	10.087 ± 4.161	23.274	12.508							
	MCD	12.604 ± 9.198	59.783	27.195	7.827 ± 4.512	24.892	12.410	5.843 ± 2.883	15.074	7.934							
	DS	24.623 ± 23.300	144.045	63.990	14.828 ± 10.302	51.198	25.443	12.219 ± 6.091	30.627	16.312							
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---							
	Mhuber	11.730 ± 7.262	37.973	18.988	11.411 ± 6.293	30.351	16.018	11.016 ± 4.629	24.787	13.477							
	T-Bic	12.450 ± 12.894	124.419	49.921	12.554 ± 9.540	53.262	25.119	11.570 ± 6.598	30.099	16.089							
	Bic	9.754 ± 9.556	42.007	20.439	9.052 ± 6.271	29.192	14.838	8.767 ± 4.672	21.877	11.772							
	MCD	8.979 ± 9.891	59.179	26.016	6.116 ± 5.625	25.124	12.288	4.861 ± 3.555	14.042	7.486							
	DS	18.805 ± 19.963	132.009	56.926	12.245 ± 10.136	50.137	24.173	10.532 ± 6.508	29.629	15.556							

Elaborado por: Néstor Montaña P.

Población Normal $N(\mathbf{0}, \Sigma)$ **y Cauchy** $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
Contaminadas bajo el modelo (2.1), con $\varepsilon = 0.30$ **y**
 $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$

De acuerdo a los resultados que se presentan en el Cuadro 4.39 se determina que el mejor estimador para el Vector de Medias, cuando se trabaja con datos que siguen una Distribución Normal, con $k=5$ y 10 , es el Bicuadrático cuya distancia promedio al Vector de Medias verdadero es 0.024 ± 0.628 y 0.006 ± 0.446 respectivamente; cuando k varía a 20 el Estimador T-Bicuadrático es el que obtiene menor distancia promedio, 0.014 ± 0.309 , sin embargo son los Estimadores MCD y Bicuadrático los que alcanzan menor Índice. Mientras que para la Matriz de Varianzas y Covarianzas el Estimador MCD junto con el Bicuadrático son los que “mejor” se aproximan a $\Sigma = \Sigma_2$ y a los Valores Propios asociados a la misma.

Cuando se trabaja con una Muestra con datos que se generan a partir de una Distribución Cauchy, se observa que la mejor estimación de los Valores Propios asociados a la Matriz de Varianzas y Covarianzas para los diferentes valores de k , está dada por el Estimador MCD, por otro lado, para la Matriz de Varianzas y Covarianzas, el Estimador que más se aproxima al verdadero valor

de $\Sigma = \Sigma_2$ es el Bicuadrático cuando $k=5$ y el MCD para los dos valores restantes de k .

Cuadro 4.39
Estimadores Robustos
Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
Contaminada $\varepsilon = 0.30$ y $H = N(0.537\mathbf{a}_p, 9\Sigma_2)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice													
	5				k				20				
					10								
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
Vector de Medias	Clásico	0.135 ± 0.889	2.338	1.120	0.132 ± 0.600	1.481	0.737	0.127 ± 0.469	1.160	0.585			
	Mhuber	0.052 ± 0.648	1.573	0.757	0.052 ± 0.431	1.036	0.507	0.044 ± 0.319	0.748	0.370			
	T-Bic	0.013 ± 0.671	1.469	0.718	0.022 ± 0.451	0.991	0.488	0.011 ± 0.306	0.674	0.331			
	Bic	0.019 ± 0.607	1.323	0.649	0.023 ± 0.435	0.944	0.467	0.010 ± 0.306	0.699	0.338			
	MCD	0.020 ± 0.649	1.450	0.706	0.023 ± 0.448	0.976	0.482	0.010 ± 0.305	0.691	0.335			
	DS	0.020 ± 0.680	1.487	0.729	0.022 ± 0.459	0.983	0.488	0.011 ± 0.313	0.695	0.340			
Valores Propios	Clásico	7.174 ± 3.146	15.910	8.743	6.303 ± 2.273	11.974	6.850	6.079 ± 1.733	10.291	6.034			
	Mhuber	1.386 ± 1.137	4.652	2.392	0.913 ± 0.719	2.803	1.478	0.768 ± 0.556	2.046	1.123			
	T-Bic	1.575 ± 1.424	6.350	3.116	1.066 ± 0.891	3.031	1.663	0.929 ± 0.637	2.801	1.455			
	Bic	1.191 ± 1.149	4.443	2.261	0.940 ± 0.775	3.123	1.613	0.871 ± 0.610	2.435	1.305			
	MCD	2.459 ± 1.549	7.516	3.841	1.385 ± 0.897	4.026	2.103	1.127 ± 0.680	2.742	1.516			
	DS	1.373 ± 1.252	5.763	2.796	0.942 ± 0.828	3.130	1.633	0.859 ± 0.630	2.420	1.303			
Matriz de Covarianzas	Clásico	5.531 ± 5.061	18.995	9.862	5.236 ± 3.501	12.798	7.178	5.253 ± 2.575	10.718	6.182			
	Mhuber	0.895 ± 1.691	5.754	2.780	0.667 ± 1.073	3.281	1.673	0.610 ± 0.756	2.157	1.174			
	T-Bic	0.827 ± 2.086	6.505	3.139	0.735 ± 1.294	3.749	1.926	0.729 ± 0.899	2.931	1.520			
	Bic	0.647 ± 1.697	4.946	2.430	0.650 ± 1.179	3.742	1.857	0.684 ± 0.858	2.711	1.418			
	MCD	1.651 ± 2.530	7.524	3.902	1.011 ± 1.411	4.828	2.417	0.908 ± 0.942	2.931	1.594			
	DS	0.609 ± 1.972	6.858	3.146	0.583 ± 1.285	3.842	1.903	0.659 ± 0.900	2.655	1.405			
Vector de Medias	5				k				10				
	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	
	Clásico	6.348 ± 3.902	---	---	5.133 ± 3.680	---	---	5.230 ± 3.806	---	---	---	---	
	Mhuber	0.098 ± 0.896	2.292	1.095	0.082 ± 0.637	1.575	0.764	0.073 ± 0.460	1.258	0.597			
	T-Bic	0.074 ± 1.096	3.130	1.433	0.053 ± 0.790	2.077	0.973	0.033 ± 0.581	1.454	0.689			
Bic	0.070 ± 0.968	2.737	1.258	0.063 ± 0.665	1.556	0.761	0.059 ± 0.447	1.163	0.556				
MCD	0.058 ± 1.044	2.986	1.363	0.033 ± 0.719	1.554	0.769	0.017 ± 0.477	1.277	0.590				
DS	0.047 ± 0.839	2.602	1.163	0.026 ± 0.628	1.456	0.704	0.017 ± 0.445	1.070	0.511				
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	
	Mhuber	14.180 ± 6.259	40.493	20.310	13.232 ± 5.564	31.719	16.838	12.477 ± 3.994	26.799	14.423			
	T-Bic	17.478 ± 10.609	117.340	48.476	15.567 ± 8.290	49.586	24.481	13.591 ± 5.236	32.565	17.131			
	Bic	12.732 ± 8.582	40.318	20.544	10.813 ± 5.418	29.393	15.208	10.087 ± 4.161	23.274	12.508			
	MCD	12.604 ± 9.198	59.783	27.195	7.827 ± 4.512	24.892	12.410	5.843 ± 2.883	15.074	7.934			
	DS	24.623 ± 23.300	144.045	63.990	14.828 ± 10.302	51.198	25.443	12.219 ± 6.091	30.627	16.312			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	
	Mhuber	11.730 ± 7.262	37.973	18.988	11.411 ± 6.293	30.351	16.018	11.016 ± 4.629	24.787	13.477			
	T-Bic	12.450 ± 12.894	124.419	49.921	12.554 ± 9.540	53.262	25.119	11.570 ± 6.598	30.099	16.089			
	Bic	9.754 ± 9.556	42.007	20.439	9.052 ± 6.271	29.192	14.838	8.767 ± 4.672	21.877	11.772			
	MCD	8.979 ± 9.891	59.179	26.016	6.116 ± 5.625	25.124	12.288	4.861 ± 3.555	14.042	7.486			
	DS	18.805 ± 19.963	132.009	56.926	12.245 ± 10.136	50.137	24.173	10.532 ± 6.508	29.629	15.556			

Elaborado por: Néstor Montaña P.

Población Normal $N(\mathbf{0}, \Sigma)$ **y Cauchy** $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$

Contaminadas Por Variable, con $\varepsilon = 0.30$ **y** $H = N(0, 9)$

De la simulación se concluye que cuando la Muestra es generada a través de la Distribución Normal se tiene que para $k=5$ y 20 , el Estimador M-Huber y el Bicuadrático son los estimadores que se encuentra más próximos al Vector real de Medias; siendo la distancia promedio, para el primero cuando $k=5$, 0.010 ± 0.561 , y el Índice asociado es 0.611 para ambos; cuando $k=10$ el Estimador M-Huber se encuentra a una Distancia Promedio de 0.019 ± 0.409 con respecto al Vector de Medias real. Mientras que, el mejor estimador, para la Matriz de Varianzas y Covarianzas con sus respectivos Valores Propios, cuando $k=5$, notoriamente es el M-Huber, puesto que el índice relacionado es 1.381 , con una Distancia Media de 0.248 ± 1.114 de la Matriz real; siendo el mismo estimador el que se aproxima a la Matriz real cuando $k=10$ y 20 . (véase Cuadro 4.40)

Para el caso de la generación de datos a partir de una Distribución Cauchy, el promedio de la estimación del Vector de Medias resultante al utilizar el Estimador M-Huber es el que menor índice presenta para cuando k toma los valores de 5 y 20 siendo el que mejor se aproxima al Vector real de Medias; sin embargo para $k=10$

es estimador que más se aproxima es el Bicuadrático con una distancia promedio de 0.015 ± 0.544 .

Cuadro 4.40
Estimadores Robustos
Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminada $\varepsilon = 0.30$ y $H = N(0, 9)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom		Desv		Prom		Desv		Prom		Desv	
		Max	Indice	Max	Indice	Max	Indice	Max	Indice	Max	Indice		
Vector de Medias	Clásico	0.015 ±	0.707	1.600	0.774	0.022 ±	0.450	1.011	0.494	0.008 ±	0.298	0.648	0.318
	Mhuber	0.010 ±	0.561	1.263	0.611	0.019 ±	0.409	0.878	0.435	0.006 ±	0.281	0.676	0.321
	T-Bic	0.010 ±	0.680	1.445	0.711	0.016 ±	0.408	1.018	0.480	0.006 ±	0.296	0.693	0.332
	Bic	0.009 ±	0.572	1.253	0.611	0.018 ±	0.408	0.906	0.444	0.007 ±	0.287	0.687	0.327
	MCD	0.009 ±	0.683	1.471	0.721	0.019 ±	0.442	1.018	0.493	0.007 ±	0.295	0.729	0.344
	DS	0.013 ±	0.648	1.498	0.720	0.019 ±	0.454	0.967	0.480	0.006 ±	0.299	0.741	0.349
Valores Propios	Clásico	2.707 ±	1.662	7.523	3.964	1.440 ±	1.002	4.475	2.305	0.742 ±	0.605	2.432	1.260
	Mhuber	0.535 ±	0.659	2.125	1.106	0.341 ±	0.532	1.359	0.744	0.227 ±	0.402	1.054	0.561
	T-Bic	1.215 ±	1.118	4.742	2.358	0.503 ±	0.690	2.624	1.272	0.300 ±	0.465	1.430	0.732
	Bic	0.806 ±	0.910	3.645	1.787	0.415 ±	0.632	1.891	0.979	0.266 ±	0.449	1.398	0.704
	MCD	2.124 ±	1.383	6.376	3.294	0.855 ±	0.785	2.898	1.513	0.414 ±	0.507	1.805	0.909
	DS	1.006 ±	1.063	3.649	1.906	0.543 ±	0.716	2.118	1.126	0.325 ±	0.510	1.558	0.797
Matriz de Covarianzas	Clásico	2.126 ±	3.026	12.146	5.766	1.208 ±	1.779	6.416	3.134	0.644 ±	1.013	3.617	1.758
	Mhuber	0.248 ±	1.114	2.782	1.381	0.165 ±	0.796	1.879	0.947	0.078 ±	0.582	1.397	0.686
	T-Bic	0.408 ±	1.919	5.704	2.677	0.140 ±	1.059	2.910	1.370	0.082 ±	0.690	1.807	0.860
	Bic	0.259 ±	1.466	4.011	1.912	0.084 ±	0.948	2.493	1.175	0.053 ±	0.657	1.712	0.807
	MCD	1.287 ±	2.433	7.607	3.776	0.470 ±	1.267	3.661	1.800	0.217 ±	0.758	2.115	1.030
	DS	0.171 ±	1.767	4.622	2.186	0.069 ±	1.137	3.086	1.430	0.057 ±	0.765	1.917	0.913
Vector de Medias	Clásico	1.995 ±	4.455	---	---	14.870 ±	4.349	---	---	2.676 ±	4.476	---	---
	Mhuber	0.014 ±	0.846	2.148	1.003	0.021 ±	0.571	1.435	0.676	0.005 ±	0.390	0.803	0.399
	T-Bic	0.037 ±	0.947	2.629	1.204	0.012 ±	0.611	1.298	0.640	0.009 ±	0.410	1.039	0.486
	Bic	0.035 ±	0.832	2.407	1.091	0.015 ±	0.544	1.212	0.590	0.007 ±	0.355	0.889	0.417
	MCD	0.024 ±	0.896	2.424	1.115	0.022 ±	0.583	1.311	0.638	0.010 ±	0.388	0.895	0.431
	DS	0.023 ±	0.809	2.251	1.028	0.021 ±	0.553	1.218	0.597	0.008 ±	0.383	0.820	0.404
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	9.344 ±	6.439	44.991	20.258	6.523 ±	4.048	27.164	12.578	5.707 ±	2.365	14.565	7.546
	T-Bic	12.185 ±	9.831	88.331	36.782	7.270 ±	4.966	32.603	14.946	5.307 ±	2.573	14.377	7.419
	Bic	9.607 ±	8.111	43.365	20.361	5.746 ±	3.901	25.963	11.870	4.614 ±	2.233	13.508	6.785
	MCD	8.459 ±	6.133	48.850	21.147	3.724 ±	2.571	13.445	6.580	2.673 ±	1.588	8.140	4.134
	DS	9.668 ±	8.542	93.567	37.259	5.128 ±	3.708	41.091	16.642	4.412 ±	2.216	13.026	6.552
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.480 ±	6.801	46.852	20.378	5.541 ±	4.238	23.389	11.056	5.010 ±	2.625	13.384	7.007
	T-Bic	8.577 ±	9.307	85.596	34.493	5.779 ±	5.164	30.050	13.665	4.486 ±	3.124	14.308	7.306
	Bic	7.156 ±	7.751	47.097	20.668	4.699 ±	4.127	22.201	10.343	3.961 ±	2.539	12.574	6.358
	MCD	5.955 ±	6.810	53.378	22.048	2.791 ±	3.248	13.893	6.644	2.168 ±	2.042	8.306	4.172
	DS	7.133 ±	7.785	81.086	32.001	4.105 ±	3.916	33.740	13.920	3.752 ±	2.550	11.162	5.821

Elaborado por: Néstor Montaña P.

Se tiene además, que el Estimador Bicuadrático es el más próximo a la Matriz Real de Varianzas y Covarianzas para $k=5$, cuya distancia promedio es 7.156 ± 7.751 ; y cuando el valor de k se incrementa a 10 y a 20 el mejor estimador es el MCD con una distancia de 2.791 ± 3.248 y 2.168 ± 2.042 respectivamente

**Población Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
Contaminadas Por Variable, con $\varepsilon = 0.30$ y $H = N(0, 9)$**

De acuerdo a los datos que se presentan en el Cuadro 4.41 se determina que el mejor estimador para el Vector de Medias, cuando se trabaja con datos que siguen una Distribución Normal, con $k=5$, es el M-Huber donde la distancia promedio al Vector real es 0.032 ± 0.601 ; cuando k varía a 10 el mejor estimador es el T-Bicuadrático ubicado, en este caso, a una distancia promedio de 0.015 ± 0.427 del Vector real y es el Estimador M-Huber el que se destaca por mejor estimación del vector real de medias con una distancia promedio de 0.012 ± 0.278 cuando $k=20$. Mientras que para la Matriz de Varianzas y Covarianzas es el Estimadores Bicuadrático el que mejor se aproximan al valor real de la Matriz así como a los Valores Propios asociados a la misma para $k=5$.

Cuadro 4.41
Estimadores Robustos
Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
Contaminada $\varepsilon = 0.30$ y $H = N(0,9)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice													
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.033 ±	0.762	1.620	0.805	0.014 ±	0.450	1.135	0.533	0.012 ±	0.294	0.621	0.309
	Mhuber	0.032 ±	0.601	1.399	0.677	0.013 ±	0.406	1.067	0.495	0.012 ±	0.278	0.603	0.298
	T-Bic	0.031 ±	0.674	1.667	0.791	0.015 ±	0.427	0.983	0.475	0.013 ±	0.285	0.640	0.313
	Bic	0.027 ±	0.599	1.448	0.691	0.014 ±	0.415	0.997	0.476	0.012 ±	0.284	0.630	0.309
	MCD	0.032 ±	0.673	1.715	0.807	0.017 ±	0.457	1.065	0.513	0.013 ±	0.295	0.639	0.316
	DS	0.029 ±	0.720	1.549	0.766	0.018 ±	0.466	1.066	0.517	0.013 ±	0.302	0.683	0.333
Valores Propios	Clásico	3.163 ±	1.522	7.507	4.064	1.701 ±	0.879	4.611	2.397	0.868 ±	0.489	2.300	1.219
	Mhuber	0.940 ±	0.650	3.029	1.540	0.481 ±	0.393	1.748	0.874	0.289 ±	0.289	0.943	0.507
	T-Bic	1.376 ±	0.937	4.652	2.322	0.609 ±	0.444	1.621	0.891	0.351 ±	0.317	1.048	0.572
	Bic	0.922 ±	0.722	2.530	1.391	0.522 ±	0.440	1.426	0.796	0.320 ±	0.323	0.957	0.533
	MCD	2.155 ±	1.078	6.018	3.084	0.922 ±	0.554	2.321	1.266	0.457 ±	0.332	1.288	0.692
	DS	1.290 ±	0.923	5.139	2.451	0.712 ±	0.510	1.922	1.048	0.397 ±	0.341	1.156	0.632
Matriz de Covarianzas	Clásico	2.096 ±	2.983	12.334	5.805	1.170 ±	1.717	7.192	3.360	0.637 ±	1.002	3.449	1.696
	Mhuber	0.398 ±	1.400	4.787	2.195	0.147 ±	0.869	2.249	1.088	0.078 ±	0.597	1.505	0.727
	T-Bic	0.376 ±	1.897	5.602	2.625	0.133 ±	1.026	2.519	1.226	0.076 ±	0.680	1.773	0.843
	Bic	0.216 ±	1.427	4.288	1.977	0.084 ±	0.936	2.386	1.136	0.052 ±	0.647	1.632	0.777
	MCD	1.185 ±	2.379	8.705	4.090	0.442 ±	1.260	3.724	1.809	0.196 ±	0.758	1.995	0.983
	DS	0.262 ±	1.833	6.771	2.956	0.097 ±	1.175	3.161	1.478	0.057 ±	0.758	1.918	0.911
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	2.282 ±	4.508	368.896	125.229	1.979 ±	4.565	930.958	312.501	2.187 ±	4.767	484.081	163.678
	Mhuber	0.030 ±	0.851	2.100	0.993	0.021 ±	0.559	1.453	0.677	0.017 ±	0.374	0.889	0.426
	T-Bic	0.029 ±	0.920	2.354	1.101	0.024 ±	0.599	1.647	0.757	0.017 ±	0.408	0.957	0.460
	Bic	0.025 ±	0.838	1.846	0.903	0.024 ±	0.532	1.358	0.638	0.017 ±	0.365	0.856	0.413
	MCD	0.038 ±	0.862	1.964	0.955	0.025 ±	0.576	1.370	0.657	0.018 ±	0.387	0.854	0.420
	DS	0.030 ±	0.782	1.783	0.865	0.021 ±	0.540	1.336	0.632	0.017 ±	0.374	0.908	0.433
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.500 ±	4.531	35.893	16.308	6.095 ±	3.178	19.690	9.654	5.063 ±	1.957	10.439	5.819
	T-Bic	9.077 ±	6.688	46.764	20.843	6.105 ±	3.444	23.997	11.182	4.456 ±	2.022	11.164	5.881
	Bic	7.262 ±	5.246	43.502	18.670	4.744 ±	2.830	16.565	8.046	3.812 ±	1.753	9.533	5.033
	MCD	7.056 ±	4.786	25.832	12.558	3.436 ±	2.093	12.315	5.948	2.378 ±	1.142	6.437	3.319
	DS	8.611 ±	6.807	64.202	26.540	4.452 ±	2.799	29.210	12.154	3.678 ±	1.637	9.439	4.918
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.120 ±	5.750	38.422	17.097	5.508 ±	3.922	20.545	9.992	4.767 ±	2.605	11.621	6.331
	T-Bic	6.096 ±	7.561	62.670	25.442	4.902 ±	4.724	28.535	12.720	3.916 ±	3.031	13.643	6.863
	Bic	5.381 ±	6.055	43.711	18.382	3.991 ±	3.654	18.498	8.714	3.454 ±	2.461	10.824	5.580
	MCD	4.719 ±	6.091	36.967	15.926	2.495 ±	3.099	15.285	6.960	1.954 ±	1.976	7.603	3.844
	DS	6.301 ±	7.012	70.413	27.909	3.602 ±	3.647	27.820	11.689	3.281 ±	2.500	10.687	5.489

Elaborado por: Néstor Montaña P.

Cuando se trabaja con una Población con datos que se generan a partir de una Distribución Cauchy, se observa que las mejores

estimaciones para el Vector de Medias para los diferentes valores de $k=5$ y 10 , están dadas por el Estimador DS ubicándose, para el caso $k=20$, es el Estimadote Bicuadrático el que mejor se aproxima a la Matriz real a una distancia promedio de 0.017 ± 0.365 con respecto al Vector Real de Medias. Por otro lado, para la Matriz de Varianzas y Covarianzas, el Estimador que más se aproxima a la Matriz real es el MCD para cualquiera de los tres valores que toma k .

**Población Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminadas Por Variable, con $\varepsilon = 0.30$ y $H = U(-5, 5)$**

Para el caso de muestras generadas a partir de la Distribución Normal, cuando k toma el valor de 5 se obtiene que el estimador que más se aproxima al valor real del Vector de Medias, en relación a los demás estimadores, es el Estimador M-Huber donde la distancia promedio al vector real es 0.018 ± 0.572 con un Índice asociado de 0.632 . Bajo el mismo razonamiento, cuando el valor de k varía M-Huber sigue siendo el mejor estimador con distancia promedio de 0.010 ± 0.388 y con 0.408 como Índice cuando k toma el valor de 10 y cuando $k=20$ la distancia es 0.017 ± 0.279 . El mejor estimador para la Matriz de Covarianzas y sus Valores Propios correspondientes, para todos los valores de k es el Estimador M-Huber debido a que es el

que se encuentra a una distancia más cercana con respecto a la Matriz real (Véase Cuadro 4.42)

Cuadro 4.42
Estimadores Robustos
Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_1$
Contaminada $\varepsilon = 0.30$ y $H = U(-5, 5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice

		5				k				20			
		Prom		Desv		Prom		Desv		Prom		Desv	
		Indice	Max	Indice	Max	Indice	Max	Indice	Max	Indice	Max		
Vector de Medias	Clásico	0.016 ± 0.682	1.633	0.777	0.006 ± 0.451	0.990	0.482	0.016 ± 0.299	0.669	0.328			
	Mhuber	0.018 ± 0.572	1.306	0.632	0.010 ± 0.388	0.826	0.408	0.017 ± 0.279	0.683	0.326			
	T-Bic	0.024 ± 0.640	1.490	0.718	0.015 ± 0.405	0.873	0.431	0.017 ± 0.292	0.680	0.330			
	Bic	0.019 ± 0.565	1.330	0.638	0.014 ± 0.401	0.846	0.420	0.018 ± 0.275	0.710	0.334			
	MCD	0.022 ± 0.669	1.542	0.745	0.013 ± 0.436	0.937	0.462	0.019 ± 0.299	0.711	0.343			
	DS	0.019 ± 0.657	1.496	0.724	0.016 ± 0.444	0.931	0.464	0.019 ± 0.300	0.777	0.365			
Valores Propios	Clásico	2.309 ± 1.314	5.947	3.190	1.275 ± 0.768	3.440	1.828	0.665 ± 0.530	2.025	1.073			
	Mhuber	0.537 ± 0.711	2.176	1.141	0.343 ± 0.533	1.257	0.711	0.239 ± 0.406	0.954	0.533			
	T-Bic	1.161 ± 1.290	5.246	2.566	0.506 ± 0.668	2.129	1.101	0.308 ± 0.483	1.443	0.745			
	Bic	0.749 ± 0.968	3.157	1.625	0.425 ± 0.632	2.190	1.082	0.272 ± 0.484	1.330	0.695			
	MCD	2.082 ± 1.476	7.299	3.619	0.849 ± 0.791	2.863	1.501	0.411 ± 0.504	1.614	0.843			
	DS	0.987 ± 1.121	5.496	2.535	0.548 ± 0.745	2.682	1.325	0.324 ± 0.518	1.449	0.764			
Matriz de Covarianzas	Clásico	1.872 ± 2.706	8.073	4.217	1.102 ± 1.574	4.437	2.371	0.579 ± 0.916	2.439	1.311			
	Mhuber	0.281 ± 1.103	2.720	1.368	0.160 ± 0.805	1.771	0.912	0.084 ± 0.577	1.329	0.663			
	T-Bic	0.373 ± 1.940	6.383	2.899	0.147 ± 1.048	2.484	1.226	0.078 ± 0.719	1.826	0.874			
	Bic	0.195 ± 1.424	3.893	1.837	0.099 ± 0.951	2.538	1.196	0.049 ± 0.675	1.613	0.779			
	MCD	1.247 ± 2.405	7.949	3.867	0.469 ± 1.262	3.627	1.786	0.198 ± 0.779	2.038	1.005			
	DS	0.140 ± 1.733	5.871	2.581	0.077 ± 1.147	3.160	1.461	0.049 ± 0.785	1.799	0.878			
Vector de Medias	Clásico	9.859 ± 4.324	---	---	2.197 ± 4.665	---	---	55.482 ± 4.911	---	---			
	Mhuber	0.031 ± 0.825	1.947	0.934	0.015 ± 0.543	1.277	0.612	0.009 ± 0.386	0.870	0.422			
	T-Bic	0.028 ± 0.910	2.208	1.049	0.015 ± 0.598	1.343	0.652	0.011 ± 0.414	0.898	0.441			
	Bic	0.033 ± 0.810	2.049	0.964	0.016 ± 0.523	1.166	0.568	0.011 ± 0.364	0.827	0.400			
	MCD	0.037 ± 0.859	2.218	1.038	0.019 ± 0.568	1.280	0.622	0.010 ± 0.397	0.871	0.426			
	DS	0.036 ± 0.773	1.823	0.877	0.015 ± 0.540	1.216	0.590	0.012 ± 0.375	0.835	0.407			
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	9.206 ± 6.249	42.463	19.306	6.592 ± 3.889	21.457	10.646	5.759 ± 2.522	14.302	7.528			
	T-Bic	11.753 ± 9.421	65.290	28.821	7.200 ± 4.470	23.079	11.583	5.246 ± 2.682	15.053	7.660			
	Bic	9.352 ± 6.973	57.392	24.572	5.653 ± 3.698	17.479	8.943	4.586 ± 2.362	11.287	6.078			
	MCD	8.072 ± 5.896	47.096	20.355	3.731 ± 2.632	15.056	7.140	2.684 ± 1.590	7.359	3.878			
	DS	9.739 ± 7.804	85.718	34.420	5.028 ± 3.445	19.964	9.479	4.357 ± 2.280	10.848	5.828			
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---			
	Mhuber	7.393 ± 6.584	39.490	17.822	5.604 ± 4.153	20.144	9.967	5.054 ± 2.739	12.860	6.884			
	T-Bic	8.236 ± 9.095	68.459	28.596	5.722 ± 5.068	23.383	11.391	4.442 ± 3.101	13.994	7.179			
	Bic	6.993 ± 7.295	53.261	22.517	4.619 ± 4.014	18.504	9.046	3.952 ± 2.597	10.976	5.842			
	MCD	5.613 ± 6.447	42.320	18.127	2.791 ± 3.160	14.477	6.809	2.192 ± 2.007	7.648	3.949			
	DS	7.195 ± 7.753	89.374	34.774	4.036 ± 3.745	18.741	8.841	3.716 ± 2.587	10.915	5.739			

Elaborado por: Néstor Montaña P.

Por otro lado, cuando se tienen datos a partir de una Distribución Cauchy, los estimador más próximos al Vector real de Medias son el Estimador DS y el Bicuadrático. Se tiene además que el mejor estimador para la Matriz de Covarianza para los tres valores que toma k es el MCD.

**Población Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
Contaminadas Por Variable, con $\varepsilon = 0.30$ y $H = U(-5, 5)$**

De acuerdo a los datos que se presentan en el Cuadro 4.43 se determina que el mejor estimador para el Vector de Medias, cuando se trabaja con datos que siguen una Distribución Normal, con $k=5$, es el Bicuadrático donde la distancia promedio al Vector verdadero es 0.017 ± 0.591 ; cuando k varía a 10 continua siendo el Bicuadrático ubicado, en este caso, a una distancia promedio de 0.015 ± 0.413 del Vector de Medias poblacional y es el Estimador M-Huber el que se destaca por mejor estimación del verdadero Vector de Medias con una distancia promedio de 0.005 ± 0.275 cuando $k=20$. Mientras que para la Matriz de Varianzas y Covarianzas es el Estimadores M-Huber el que mejor se aproximan a $\Sigma = \Sigma_2$ así como a los Valores Propios asociados a la misma.

Cuadro 4.43
Estimadores Robustos
 Caso Normal $N(\mathbf{0}, \Sigma)$ y Cauchy $CAU(\mathbf{0}, \Sigma)$, $\Sigma = \Sigma_2$
 Contaminada $\varepsilon = 0.30$ y $H = U(-5,5)$

Sesgo Promedio Total, Sesgo Máximo Total, Desviación Total e Índice													
		5				10				20			
		Prom	Desv	Max	Indice	Prom	Desv	Max	Indice	Prom	Desv	Max	Indice
Vector de Medias	Clásico	0.016 ±	0.684	1.592	0.764	0.014 ±	0.457	1.028	0.500	0.006 ±	0.288	0.629	0.308
	Mhuber	0.014 ±	0.579	1.381	0.658	0.015 ±	0.419	0.901	0.445	0.005 ±	0.275	0.616	0.299
	T-Bic	0.021 ±	0.665	1.463	0.716	0.017 ±	0.414	0.923	0.451	0.006 ±	0.279	0.638	0.308
	Bic	0.017 ±	0.591	1.338	0.649	0.015 ±	0.413	0.887	0.438	0.006 ±	0.282	0.626	0.305
	MCD	0.025 ±	0.677	1.537	0.746	0.015 ±	0.437	0.970	0.474	0.005 ±	0.292	0.722	0.340
	DS	0.020 ±	0.703	1.546	0.756	0.015 ±	0.459	0.975	0.483	0.006 ±	0.303	0.698	0.336
Valores Propios	Clásico	2.754 ±	1.110	6.101	3.322	1.520 ±	0.658	3.439	1.872	0.786 ±	0.390	1.690	0.955
	Mhuber	0.901 ±	0.475	2.659	1.345	0.496 ±	0.415	1.306	0.739	0.301 ±	0.293	0.921	0.505
	T-Bic	1.315 ±	0.888	3.863	2.022	0.631 ±	0.477	1.589	0.899	0.362 ±	0.358	1.032	0.584
	Bic	0.888 ±	0.688	2.307	1.294	0.537 ±	0.455	1.445	0.813	0.330 ±	0.341	1.091	0.587
	MCD	2.068 ±	1.115	5.678	2.954	0.917 ±	0.561	2.300	1.259	0.459 ±	0.356	1.201	0.672
	DS	1.257 ±	0.893	4.841	2.330	0.707 ±	0.518	1.771	0.998	0.407 ±	0.367	1.209	0.661
Matriz de Covarianzas	Clásico	1.885 ±	2.655	7.816	4.119	1.098 ±	1.541	4.292	2.310	0.583 ±	0.919	2.490	1.330
	Mhuber	0.375 ±	1.016	3.897	1.763	0.155 ±	0.867	2.101	1.041	0.076 ±	0.602	1.524	0.734
	T-Bic	0.335 ±	1.826	6.650	2.937	0.149 ±	1.049	2.468	1.222	0.076 ±	0.695	1.696	0.822
	Bic	0.196 ±	1.377	4.020	1.865	0.093 ±	0.955	2.310	1.120	0.052 ±	0.643	1.604	0.766
	MCD	1.122 ±	2.310	8.077	3.837	0.439 ±	1.248	3.712	1.800	0.194 ±	0.757	2.013	0.988
	DS	0.234 ±	1.747	6.632	2.871	0.090 ±	1.163	3.037	1.430	0.059 ±	0.760	1.945	0.921

k													
		5								5			
		Prom	Desv		Indice	Prom	Desv		Indice	Prom	Desv	Max	Prom
Vector de Medias	Clásico	6.524 ±	4.306	---	---	1.818 ±	4.537	---	---	1.821 ±	4.930	---	---
	Mhuber	0.026 ±	2.261	0.872	1.053	0.021 ±	1.376	0.562	0.653	0.012 ±	0.382	0.943	0.446
	T-Bic	0.021 ±	2.463	0.924	1.136	0.010 ±	1.480	0.603	0.698	0.012 ±	0.411	1.039	0.488
	Bic	0.023 ±	2.311	0.828	1.054	0.006 ±	1.141	0.534	0.560	0.010 ±	0.370	0.907	0.429
	MCD	0.019 ±	2.514	0.873	1.135	0.017 ±	1.279	0.566	0.621	0.010 ±	0.381	0.918	0.436
	DS	0.014 ±	2.035	0.786	0.945	0.015 ±	1.277	0.530	0.607	0.009 ±	0.366	0.907	0.427
Valores Propios	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	8.710 ±	31.947	4.307	14.988	6.177 ±	19.234	3.128	9.513	5.190 ±	1.942	11.144	6.092
	T-Bic	9.629 ±	37.595	6.421	17.882	6.217 ±	20.702	3.862	10.260	4.604 ±	2.059	11.664	6.109
	Bic	7.448 ±	31.429	4.816	14.564	4.723 ±	17.229	2.868	8.273	3.900 ±	1.739	9.514	5.051
	MCD	7.469 ±	28.025	4.914	13.469	3.398 ±	11.464	2.087	5.650	2.421 ±	1.171	6.130	3.241
	DS	9.302 ±	73.115	7.882	30.100	4.295 ±	15.206	2.738	7.413	3.762 ±	1.682	8.392	4.612
Matriz de Covarianzas	Clásico	---	---	---	---	---	---	---	---	---	---	---	---
	Mhuber	7.328 ±	36.795	5.584	16.569	5.564 ±	21.300	4.035	10.300	4.901 ±	2.615	13.025	6.847
	T-Bic	6.464 ±	50.059	8.065	21.529	4.980 ±	24.887	4.826	11.564	4.037 ±	3.069	13.615	6.907
	Bic	5.564 ±	36.788	6.221	16.191	3.983 ±	18.053	3.701	8.579	3.537 ±	2.462	10.638	5.546
	MCD	4.959 ±	36.517	6.428	15.968	2.477 ±	16.257	3.072	7.269	1.990 ±	2.028	7.484	3.834
	DS	6.821 ±	70.948	7.820	28.530	3.510 ±	18.351	3.567	8.476	3.364 ±	2.463	10.196	5.341

Elaborado por: Néstor Montaña P.

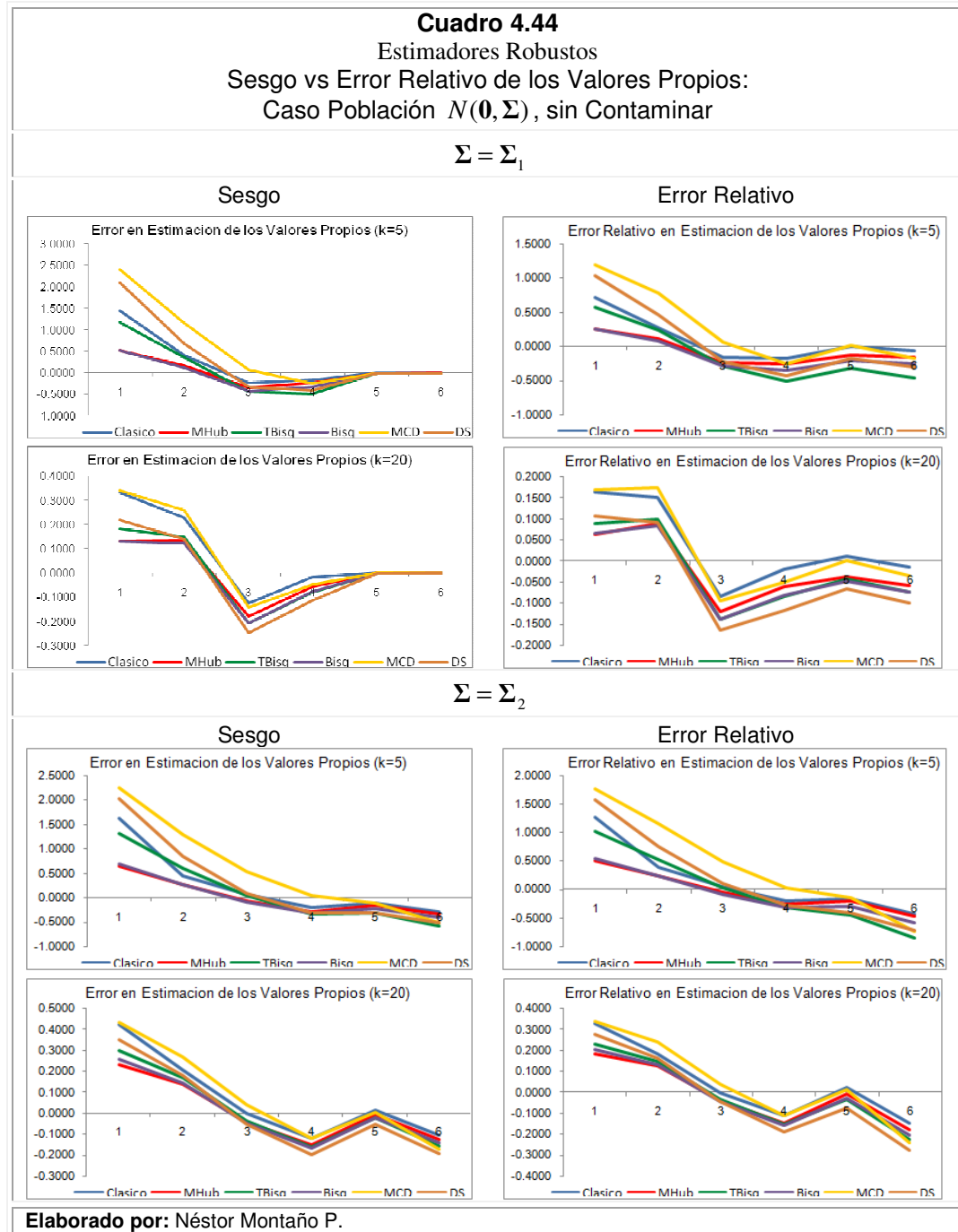
Quando se trabaja con una Población con datos que se generan a partir de una Distribución Cauchy, se observa que las mejores

estimaciones para el Vector de Medias para los diferentes valores de k , están dadas por el Estimador DS ubicándose, para el caso $k=5$, a una distancia promedio de 0.014 ± 2.035 con respecto al Vector Real de Medias. Por otro lado, para la Matriz de Varianzas y Covarianzas, el Estimador que más se aproxima a la Matriz real es el MCD para cualquiera de los tres valores que toma k

En lo correspondiente a los valores propios, se ha mostrado los resultados en términos del sesgo promedio; un análisis alternativo se puede realizar utilizando el error relativo, por ello a continuación se presentan los resultados correspondientes a los casos de muestras generadas a partir de una Distribución Normal sin contaminar, y una Distribución Cauchi sin contaminar.

El Cuadro 4.44 presenta en el lado izquierdo el Sesgo promedio mientras que en el derecho el Error Relativo correspondiente a cada valor propio, todo esto para muestras generadas a partir de una Distribución Normal sin contaminar; se observa que, para $\Sigma = \Sigma_1$ y $k=5$, el Error Relativo correspondiente estimador M de Huber es 0.253 para el primer valor propio, 0.109, -0.231, -0.252, -0.126, -0.158 para los demás valores propios; mientras que para $\Sigma = \Sigma_2$ el comportamiento del Error Relativo es “parecido” al del Sesgo puesto

que los valores propios asociados a $\Sigma = \Sigma_2$ son más cercanos a uno y menos variables que los asociados a $\Sigma = \Sigma_1$



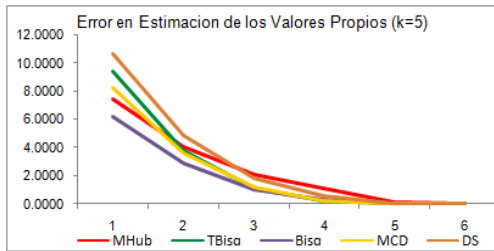
En el Cuadro 4.45 se observa el Sesgo y el Error Relativo correspondiente a la estimación de los valores propios, para el caso de muestra generadas a partir de una Distribución Cauchy sin contaminar.

Al igual que en el análisis anterior, para $\Sigma = \Sigma_2$ los resultados correspondientes al Error Relativo son “parecidos” a los del Sesgo, mientras que para $\Sigma = \Sigma_1$ se puede ver que el “pequeño” Sesgo promedio que se obtiene al estimar los dos últimos valores propios equivale a una sobreestimación de entre 38,4% y 156% para $k=5$ y de entre 52% y 164% para $k=20$

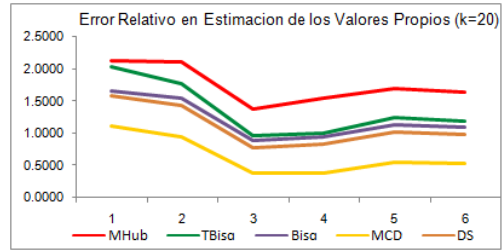
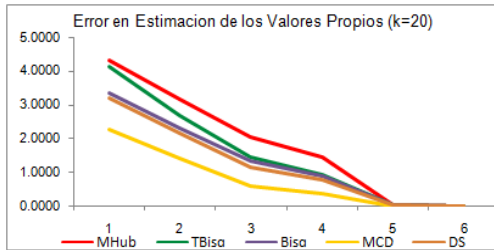
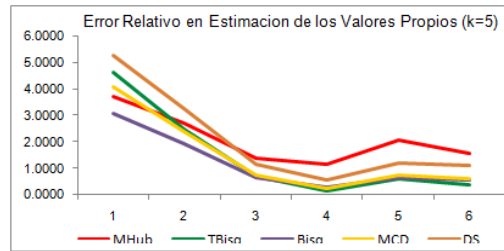
Cuadro 4.45
Estimadores Robustos
Sesgo vs Error Relativo de los Valores Propios:
Caso Población $CAU(0, \Sigma)$, sin Contaminar

$\Sigma = \Sigma_1$

Sesgo

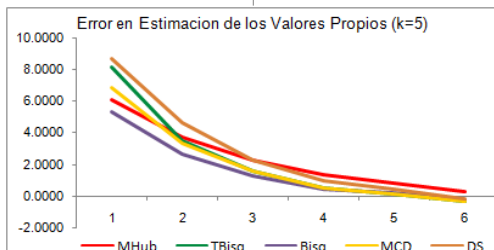


Error Relativo

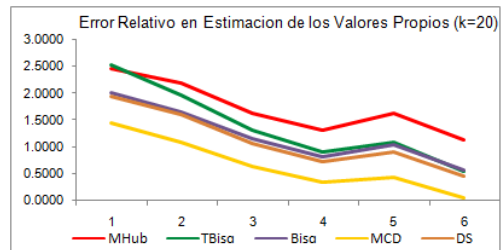
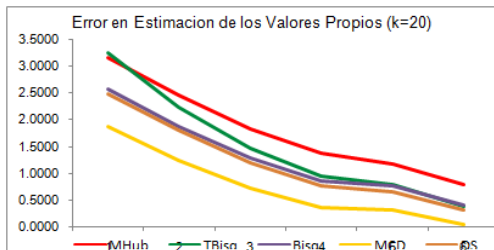
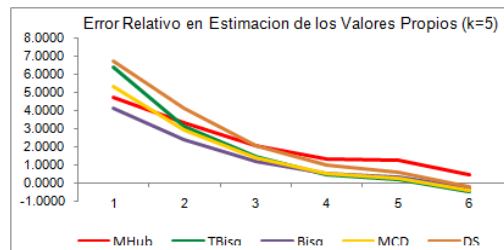


$\Sigma = \Sigma_2$

Sesgo



Error Relativo



Elaborado por: Néstor Montaña P.

4.4.2 Análisis en función del Tamaño de la Muestra, Sesgo y Curtosis Muestral

El objetivo en esta parte del estudio es tratar de establecer una “regla empírica” donde se utilice al tamaño de la Muestra, al Sesgo y la Curtosis Muestral como elementos de decisión sobre el estimador a utilizar.

En cada muestra generada se ha calculado el Sesgo y Curtosis Muestral, además la distancia Euclidiana entre los valores reales del Vector de Medias, Matriz de Covarianzas y Valores Propios asociados a la misma versus la estimación brindada por cada Método; a partir de ello se establece el estimador que menor distancia presenta en cada caso.

Se presenta a continuación los resultados clasificados por tamaño Muestral.

Tamaño de Muestra igual a 30

En el Cuadro 4.50a se muestra la tabla de resumen de los resultados obtenidos para estimadores estudiados en función de los valores de Sesgo y Curtosis, en la tabla se puede observar los tres estimadores que con mayor frecuencia presentan la menor distancia entre el valor del parámetro y la estimación brindada por los

mismos; para cada estimador se presenta la distancia Promedio y su Desviación Estándar, además de la Frecuencia Relativa de ser el estimador con “más cercano” al valor real.

Así, cuando el Sesgo Muestral es menor a 45 y la Curtosis Muestral se encuentra en el intervalo [20, 65) el 25,3% de las veces el estimador Clásico es el que presenta la menor distancia para la estimación del Vector de Medias, siendo su distancia Promedio igual a 0.503 ± 0.174 . Mientras que para los Valores Propios y la Matriz de Varianzas y Covarianzas el Estimador M de Huber en más del 50% de los casos brinda la estimación “más cercana”.

Para valores de Sesgo Muestral mayores a 45 y menores a 90, el Estimador DS es con mayor frecuencia el “mejor” estimador, sin embargo su Promedio y Desviación Estándar de la distancia para el caso de los Valores Propios y Matriz de Covarianzas es mayor a la presentada por los otros dos estimadores mostrados en cada caso, por ejemplo, cuando la Curtosis se encuentra en el intervalo [110, 155) la distancia promedio correspondiente al Estimador DS al estimar la Matriz de Varianzas y Covarianzas es 11.075 mientras que para el estimador Bicuadrático y MCD son 8.397 y 8.540 respectivamente, situación que puede ser causado por la alta dispersión observada en el Estimador DS.

Cuadro 4.46
Estimadores Robustos
Estimadores Robustos en función del Sesgo y la Curtosis para k=5

		Curtosis Muestral							
		[20, 65)				[65, 110)			
				Distancia	Frecuencia			Distancia	Frecuencia
[0, 45)	Localizacion	1 Clas	0.503 ±	0.174	0.253	MHub	0.502 ±	0.205	0.239
		2 MHub	0.450 ±	0.147	0.212	Bis	0.502 ±	0.211	0.205
		3 Bis	0.459 ±	0.151	0.210	DS	0.537 ±	0.202	0.174
	Valores Propios	1 MHub	1.154 ±	1.155	0.535	MHub	3.598 ±	4.828	0.389
		2 Bis	1.212 ±	1.191	0.323	Bis	3.510 ±	4.979	0.306
		3 DS	1.919 ±	2.271	0.068	DS	5.165 ±	10.074	0.164
	Matriz de Covarianzas	1 MHub	1.229 ±	1.020	0.548	MHub	3.414 ±	4.317	0.419
		2 Bis	1.266 ±	1.030	0.338	Bis	3.287 ±	4.334	0.299
		3 DS	1.871 ±	1.965	0.058	DS	4.739 ±	8.726	0.156
[45, 90)	Localizacion	1 DS	0.596 ±	0.231	0.304	DS	0.615 ±	0.238	0.339
		2 MHub	0.619 ±	0.247	0.264	MHub	0.655 ±	0.254	0.262
		3 Bis	0.620 ±	0.252	0.185	Bis	0.649 ±	0.254	0.178
	Valores Propios	1 DS	10.648 ±	13.286	0.321	DS	12.391 ±	16.178	0.366
		2 Bis	7.640 ±	6.283	0.279	Bis	9.334 ±	6.217	0.229
		3 MHub	8.372 ±	6.229	0.191	MCD	9.607 ±	5.930	0.195
	Matriz de Covarianzas	1 DS	9.536 ±	11.528	0.322	DS	11.075 ±	14.074	0.376
		2 Bis	6.899 ±	5.442	0.278	Bis	8.397 ±	5.385	0.220
		3 MHub	7.691 ±	5.541	0.184	MCD	8.540 ±	5.075	0.204
[90, 135)	Localizacion	1 DS	0.619 ±	0.236	0.334	DS	0.627 ±	0.216	0.296
		2 MHub	0.666 ±	0.255	0.251	MHub	0.661 ±	0.248	0.259
		3 Bis	0.654 ±	0.247	0.191	Bis	0.649 ±	0.240	0.222
	Valores Propios	1 DS	11.190 ±	12.920	0.422	DS	13.493 ±	10.911	0.333
		2 Bis	8.957 ±	6.306	0.238	MCD	10.479 ±	7.010	0.241
		3 MCD	9.368 ±	5.947	0.167	MHub	11.112 ±	6.280	0.204
	Matriz de Covarianzas	1 DS	10.044 ±	11.236	0.429	DS	12.128 ±	9.662	0.352
		2 Bis	8.102 ±	5.460	0.224	MCD	9.362 ±	5.914	0.259
		3 MCD	8.367 ±	5.090	0.180	Bis	9.249 ±	5.669	0.204
[135, 180)	Localizacion	1 DS	0.619 ±	0.262	0.371				
		2 Bis	0.638 ±	0.257	0.257				
		3 MHub	0.686 ±	0.256	0.143				
	Valores Propios	1 Bis	10.560 ±	7.907	0.314				
		2 DS	13.448 ±	13.000	0.257				
		3 MCD	10.199 ±	6.006	0.229				
	Matriz de Covarianzas	1 Bis	9.532 ±	6.750	0.371				
		2 DS	12.125 ±	11.515	0.257				
		3 MCD	9.142 ±	5.161	0.229				

Elaborado por: Néstor Montaña P.

Tamaño de Muestra igual a 60

De acuerdo al Cuadro 4.52, cuando el Sesgo Muestral es mayor a 90 y menor a 180 y la Curtosis Muestral se encuentra en el intervalo [130, 220) el estimador MCD tiene una frecuencia igual a 0.688 de brinda la estimación “más cercana” a los Valores Propios, seguido por el estimador DS y Bicuadrático con frecuencias 0.181 y 0.09 respectivamente.

En lo referente al Vector de Medias, El 28.9% de las veces el Estimador M de Huber es el más cercano al verdadero valor del Vector de Medias, al Estimador M de Huber le corresponde distancia Promedio igual a 0.435 mientras que la desviación de las distancias observadas es 0.174.

Nótese que el estimador que con mayor frecuencia es el mejor estimador de la Matriz de Covarianzas y los Valores Propios asociados a la misma es, para sesgo mayores a 90, el Estimador MCD.

Cuadro 4.47
Estimadores Robustos
Estimadores Robustos en función del Sesgo y la Curtosis para k=10

		Curtosis Muestral						
		[40, 130)			[130, 220)			
		Distancia	Frecuencia	Distancia	Frecuencia			
Sesgo Muestral	[0, 90)	Localizacion	1 Clas	0.383 ± 0.243	0.229	MHub	0.402 ± 0.156	0.265
			2 MHub	0.315 ± 0.110	0.208	DS	0.395 ± 0.147	0.252
			3 TBis	0.335 ± 0.120	0.176	Bis	0.397 ± 0.152	0.160
		Valores Propios	1 MHub	0.994 ± 1.569	0.494	MCD	4.004 ± 2.707	0.579
			2 Bis	0.985 ± 1.407	0.246	DS	5.346 ± 4.914	0.155
			3 DS	1.144 ± 1.631	0.107	MHub	6.076 ± 4.204	0.131
		Matriz de Covarianzas	1 MHub	1.053 ± 1.415	0.537	MCD	3.700 ± 2.366	0.152
			2 Bis	1.042 ± 1.248	0.257	MHub	5.657 ± 3.801	0.156
			3 DS	1.201 ± 1.432	0.083	DS	4.927 ± 4.347	0.152
	[90, 180)	Localizacion	[130, 220)			[220, 310)		
			1 MHub	0.435 ± 0.160	0.289	MHub	0.423 ± 0.148	0.276
			2 DS	0.420 ± 0.151	0.258	DS	0.414 ± 0.148	0.259
Valores Propios		3 Bis	0.423 ± 0.155	0.198	Bis	0.416 ± 0.153	0.212	
		1 MCD	4.593 ± 2.542	0.688	MCD	4.661 ± 2.488	0.753	
		2 DS	6.056 ± 4.611	0.181	DS	6.185 ± 4.328	0.158	
Matriz de Covarianzas		3 Bis	6.183 ± 3.609	0.090	Bis	6.357 ± 3.430	0.069	
		1 MCD	4.225 ± 2.212	0.706	MCD	4.301 ± 2.193	0.779	
		2 DS	5.567 ± 4.081	0.178	DS	5.707 ± 3.872	0.149	
3 Bis		5.705 ± 3.201	0.078	Bis	5.889 ± 3.075	0.005		
[180, 270)		Localizacion	[220, 310)			[310, 400]		
			1 MHub	0.441 ± 0.159	0.288	DS	0.417 ± 0.151	0.290
	2 DS		0.421 ± 0.153	0.258	MHub	0.429 ± 0.133	0.274	
	Valores Propios	3 Bis	0.425 ± 0.153	0.200	Bis	0.422 ± 0.138	0.194	
		1 MCD	4.890 ± 2.705	0.738	MCD	5.256 ± 2.946	0.790	
		2 DS	6.514 ± 4.988	0.163	DS	6.840 ± 4.262	0.113	
	Matriz de Covarianzas	3 Bis	6.647 ± 3.855	0.073	Bis	7.143 ± 3.938	0.081	
		1 MCD	4.507 ± 2.371	0.759	MCD	4.858 ± 2.555	0.855	
		2 DS	6.011 ± 4.426	0.157	DS	6.336 ± 3.745	0.097	
	3 Bis	6.164 ± 3.431	0.064	Bis	6.657 ± 3.458	0.016		
	[270, 360]	Localizacion	[310, 400]					
			1 DS	0.436 ± 0.150	0.319			
2 MHub			0.456 ± 0.152	0.278				
Valores Propios		3 Bis	0.447 ± 0.160	0.250				
		1 MCD	5.011 ± 2.528	0.722				
		2 DS	6.345 ± 3.730	0.208				
Matriz de Covarianzas		3 Bis	6.539 ± 3.277	0.056				
		1 MCD	4.707 ± 2.329	0.778				
		2 DS	5.994 ± 3.430	0.181				
3 Bis	6.201 ± 3.061	0.028						

Elaborado por: Néstor Montaña P.

Tamaño de Muestra igual a 120

Como se muestra en el Cuadro 4.52a para Curtosis mayores a 130 el Estimador M de Huber es con mayor frecuencia el estimador “más cercano” al Vector de Medias poblacional, seguido siempre por el Estimador Bicuadrático.

Para valores de Sesgo Muestral es menor a 190 y la Curtosis Muestral se encuentra en el intervalo $[40, 230)$ el 48.7% de las veces el estimador M de Huber es el estimador “más cercano” a los Valores Propios, seguido por el estimador DS y Bicuadrático con frecuencias 0.181 y 0.09 respectivamente.

Cuando la Curtosis es mayor a 230 el Estimador MCD es el que con mayor frecuencia brinda la mejor estimación de la Matriz de Covarianzas y los Valores Propios asociados a la misma, incluso para Curtosis mayor a 610 el Estimador MCD el 100% de las veces es el estimador más cercano a los verdaderos valores de la Matriz de Varianzas y Covarianzas.

Cuadro 4.48
Estimadores Robustos
Estimadores Robustos en función del Sesgo y la Curtosis para k=20

		Curtosis Muestral							
		[40, 230)			[230, 420)				
			Distancia	Frecuencia		Distancia	Frecuencia		
Sesgo Muestral	[0, 190)	Localizacion	1 Clas	0.278 ± 0.837	0.242	MHub	0.299 ± 0.104	0.307	
			2 TBis	0.227 ± 0.081	0.187	Bis	0.290 ± 0.106	0.212	
			3 MHub	0.221 ± 0.077	0.184	MCD	0.308 ± 0.113	0.185	
	Valores Propios	1 MHub	0.710 ± 1.215	0.487	MCD	3.168 ± 1.604	0.948		
		2 Bis	0.785 ± 1.345	0.208	Bis	5.144 ± 2.986	0.017		
		3 DS	0.757 ± 1.052	0.113	DS	5.007 ± 3.213	0.015		
	Matriz de Covarianzas	1 MHub	0.759 ± 1.099	0.552	MCD	2.967 ± 1.418	0.949		
		2 Bis	0.754 ± 0.980	0.217	Bis	4.775 ± 2.669	0.019		
		3 TBis	0.832 ± 1.197	0.085	MHub	5.668 ± 2.475	0.018		
	Sesgo Muestral	[190, 380)	Localizacion	1 MHub	0.300 ± 0.105	0.307	MHub	0.297 ± 0.102	0.284
				2 Bis	0.289 ± 0.102	0.219	Bis	0.284 ± 0.102	0.238
				3 DS	0.294 ± 0.103	0.200	DS	0.293 ± 0.104	0.197
Valores Propios		1 MCD	3.162 ± 1.451	0.976	MCD	3.223 ± 1.504	0.986		
		2 Bis	5.156 ± 2.580	0.013	Bis	5.125 ± 2.479	0.008		
		3 DS	4.975 ± 2.738	0.009	DS	5.028 ± 2.746	0.006		
Matriz de Covarianzas		1 MCD	2.963 ± 1.279	0.980	MCD	3.042 ± 1.351	0.989		
		2 Bis	4.795 ± 2.296	0.009	Bis	4.807 ± 2.232	0.006		
		3 DS	4.619 ± 2.431	0.008	DS	4.704 ± 2.461	0.005		
Sesgo Muestral		[380, 570)	Localizacion	1 MHub	0.296 ± 0.102	0.310	MHub	0.310 ± 0.130	0.333
				2 Bis	0.283 ± 0.099	0.228	Bis	0.301 ± 0.123	0.197
				3 DS	0.292 ± 0.101	0.192	DS	0.310 ± 0.116	0.182
	Valores Propios	1 MCD	3.317 ± 1.544	0.983	MCD	3.4084 ± 1.682	0.985		
		2 Bis	5.284 ± 2.580	0.009	Bis	5.36173 ± 2.654	0.015		
		3 DS	5.189 ± 2.827	0.007					
	Matriz de Covarianzas	1 MCD	3.131 ± 1.390	0.985	MCD	3.269 ± 1.540	1.000		
		2 Bis	4.962 ± 2.314	0.007					
		3 DS	4.858 ± 2.526	0.006					
	Sesgo Muestral	[570, 760]	Localizacion	1 MHub	0.276 ± 0.093	0.323			
				2 Bis	0.270 ± 0.084	0.306			
				3 DS	0.287 ± 0.081	0.161			
Valores Propios		1 MCD	3.595 ± 1.818	1.000					
		1 MCD	3.433 ± 1.690	1.000					
Matriz de Covarianzas									

Elaborado por: Néstor Montaña P.

CONCLUSIONES

A partir del Análisis Estadístico realizado a los resultados obtenidos de la Simulación Matemática desarrollada, se concluye lo siguiente:

1. Se confirma la sensibilidad de los Estimadores de Máxima Verosimilitud para el Vector de Medias y Matriz de Varianzas y Covarianzas ante desviaciones de la Distribución Normal Multivariada.
2. De acuerdo al Cuadro 4.1 y 4.2 los Estimadores Robustos son comparables a los Estimadores de Máxima Verosimilitud para el caso de muestras generadas a partir de una Normal Multivariada sin contaminar, es decir, se verifica que los estimadores Robustos Estudiados son Eficientes.
3. Al considerar el Sesgo Promedio Total como medida de la “bondad” del Estimador, se concluye que para los casos en los que se considera Población Normal Contaminada, (bajo cualquier tipo de contaminación) el Estimador Clásico presenta su “peor” rendimiento al tratar de estimar la Matriz de Covarianzas y los Valores Propios asociados a la misma.

4. El tamaño muestral influye en la estimación, pues en todos los estimadores se cumple que al aumentar el tamaño muestral se disminuye el Sesgo Promedio Total y la Desviación Total.
5. Para $\varepsilon = 0.30$ el comportamiento de los estimadores Robustos estudiados, es “parecido” al comportamiento observado por los mismos para $\varepsilon = 0.05$ y $\varepsilon = 0.10$, se concluye entonces que el Punto de Ruptura, para 6 variables, es mayor a 0.3.
6. El estimador T-Bicuadrático o T-Biponderado es casi siempre superado por los demás Estimadores Robustos, situación que era previsible pues este estimador fue construido para mejorar el comportamiento de los estimadores S para “grandes” número de variables.
7. A pesar de que la mayoría de los Estimadores Robustos considerados en el presente estudio han sido construidos bajo el supuesto de contaminación simétrica, al ser sometidos a contaminación asimétrica se comportan de manera similar a cuando la contaminación es simétrica.

8. Así también, los estimadores han sido construidos siguiendo el entorno de contaminación (2.1), sin embargo al contaminar cada variable independientemente, según lo explicado en el capítulo 4, los Estimadores brindan “buenas” estimaciones; sin embargo, en el presente estudio no se analiza las consecuencias que puede tener una contaminación por variable en las Técnicas de Análisis Multivariado.

9. Cuando $\Sigma = \Sigma_1$ y la muestra es generada a partir de una Población Normal, los estimadores en general sobreestiman el primer y segundo valor propio para luego estimar con error “pequeño” los valores propios restantes, para este caso, los Estimadores Clásico y MCD son los que sobreestiman con mayor error el primer valor propio, sin embargo, el Estimador MCD reduce considerablemente su error mientras aumenta el tamaño muestral.

10. Por otro lado, cuando $\Sigma = \Sigma_1$ y la muestra es generada a partir de una Población Norma, los Estimadores sobreestiman el primer y segundo valor propio, sin embargo, al final generalmente subestiman el menor valor propio, para esta matriz se puede observar que para $k=10$ y 20 que el estimador Clásico se aleja de los Estimadores Robustos presentando mayor error en si estimación.

- 11.** En todos los casos cuando las Muestras son generadas a partir de la Distribución Cauchy, el Estimador Clásico brinda las estimaciones más distantes de cada parámetro poblacional mientras que el algoritmo del Estimador M de Huber puede no converger a una solución; además, los valores propios son siempre sobreestimados. Algo que en parte confirma la idea de las tres esquinas propuesta por Andrews en [1] y mencionada en el capítulo 4.
- 12.** Cuando la muestra es generada a partir de la Distribución Normal Multivariada, para cualquier tamaño muestral los Estimadores M de Huber y Bicuadrático son generalmente los que alcanzan el menor índice al estimar el Vector de Medias, Matriz de Covarianzas y Valores Propios.
- 13.** Cuando se considera muestras generadas a partir de una Distribución Cauchy, el Estimador Bicuadrático presenta el mejor comportamiento para $k=5$, mientras que para $k=10$ y 20 el Estimador MCD es el que “mejor” estima la Matriz de Covarianzas y los Valores propios asociados a la misma, todo ello en base al índice planteado.
- 14.** A las estimaciones brindadas por el estimador DS en varias ocasiones le corresponde un Sesgo Promedio Total menor al Sesgo presentado por los demás estimadores, sin embargo su “peor” estimación puede incluso

encontrarse más alejada que la “peor” estimación utilizando el método Clásico; esto genera que no presente un Balance adecuado y no sea considerado el mejor estimador (de acuerdo al índice planteado) en los escenarios simulados.

RECOMENDACIONES

A partir del Análisis Estadístico realizado a los resultados obtenidos de la Simulación Matemática desarrollada, se plantean las siguientes recomendaciones, todas ellas consideran que se trabaja con seis variables:

1. Siendo la Matriz de Varianzas y Covarianzas el punto de partida de muchas de las técnicas Multivariadas, se recomienda, cuando exista contaminación de la Muestra, el uso de Estimadores Robustos, pues en ningún caso estudiado el Estimador de Máxima Verosimilitud fue el “mejor” estimador para dicho parámetro.
2. Cuando se considera que las variables aleatorias a estudiar se distribuyen de acuerdo a una Normal Multivariada, se recomienda el uso de los estimadores M de Huber y Bicuadrático.
3. Si se considera que las variables aleatorias a estudiar se distribuyen de acuerdo a una Cauchy Multivariada y se tiene tamaño muestral cercano a 30, entonces se recomienda el uso del estimador Bicuadrático, mientras que para tamaños muestrales mayores se recomienda el uso del Estimador MCD.

4. Para tamaño muestral igual a 30 y Sesgo menor a 90 se recomienda el uso del Estimador M de Huber o del Estimador Bicuadrático; no se recomienda el uso del estimador DS a causa de la desviación presentada por el mismo. Así también, para el mismo tamaño muestral y Sesgo mayor a 90 y menor a 135: cuando la Curtosis Muestral se encuentre en el intervalo [110, 155) se recomienda el uso del Estimador Bicuadrático, mientras que con Curtosis Muestral en el intervalo [155, 200) se recomienda el uso del Estimador MCD

5. Con el tamaño muestral igual a 60 y la Curtosis Muestral sea menor a 130 se recomienda el uso del Estimador M de Huber o Bicuadrático, caso contrario se recomienda el uso del estimador MCD o Bicuadrático

6. Para tamaño muestral igual a 120, cuando la Curtosis Muestral sea menor a 230 se recomienda el uso del estimador M de Huber o Bicuadrático, caso contrario se recomienda el uso del estimador MCD o Bicuadrático.

7. Se recomienda la realización de próximos estudios que complementen el presente trabajo, considerando por ejemplo: Mayor número de variables, Otros estimadores, etc.

BIBLIOGRAFÍA

- [1] ANDREWS, D. et al. (1972). *Robust Estimates of Location: Survey and Advances*. Princeton Univ. Press
- [2] BOENTE, G. & YOHAI, V., “*Notas de estadística*”, [en línea] Universidad de Buenos Aires, Buenos Aires – Argentina
- [3] CHOW, Y. & TEICHER H. (1978). “*Probability theory: independence, interchangeability, martingales*”, Springer-Verlag, New York – USA,
- [4] CIPRA, B (2000) “*The Best of the 20th Century: Editors Name Top 10 Algorithms*”, SIAM News, 33, 4, 1-2.
- [5] DEVLIN, S., GNANADESIKAN, R., Y KETTENRING J. (1981). “*Robust estimation of dispersion matrices and Principal Components*”, Journal of the American Statistical Association; 76, 354-362.
- [6] FISHMAN, G (2006) “*A First Course in Monte Carlo*”, Thompson, Belmont–USA
- [7] FLURY, B. (1997). “*A first course in multivariate Statistics*”, Springer texts in Statics, New York–USA
- [8] GARCÍA, G. (2007). “*Métodos cuantitativos II*”, Edición electrónica Universidad Autónoma Barcelona, Barcelona – España.
- [9] GENTO, P., ORTEGA, J. y GARCÍA, G. (2004). “*Alternativas estadísticas al cálculo del Valor en Riesgo*”, Estadística Española; Vol. 46, 119 - 148

- [10] GIRALDO, N. (2006). "*Procesos estocásticos*", [en línea] Universidad Nacional de Colombia.
- [11] HAMPEL, F. (1971). "*A general qualitative definition of robustness*", *The Annals of Mathematical Statistics*; 42, 1887–1896
- [12] HAMPEL, F. (1974). "*The influence curve and its role in robust estimation*", *Journal of the American Statistical Association*; 69, 383–393
- [13] HODGES, J. & LEHMAN, E. (1963). "*Estimates of Location based on rank test*", *The Annals of Mathematical Statistics*; 33, 1124–1147
- [14] HUBER, P. (1964). "*Robust estimation of a location parameter*", *The Annals of Mathematical Statistics*; 35, 73–101
- [15] HUBER, P. (1972). "*The 1972 wald lecture. Robust Statistics: A Review*", *The Annals of Mathematical Statistics*; 43, 1041–1067
- [16] HUBER, P. (1979). "*Robust Statistical Procedures*", Society for industrial and Applied Mathematics
- [17] JIMENEZ, D. (1963) "*El método de Montecarlo y sus aplicaciones*", *Estadística Española*; 19, 12–30
- [18] MARONNA, R., MARTIN, D. Y YOHAI, V. (2006). "*Robust Statistics: Theory and Methods*", John Willey & Sons, New York–USA
- [19] MARONNA, R. (1976). "*Robust M-Estimators of Multivariate Location and Scatter*", *The Annals of Statistics*; 4, 51–67

- [20] MATSUMOTO, M., NISHIMURA, T. (1998) "*Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator*", ACM Transactions on Modeling and Computer Simulation, 8, 3–30.
- [21] MENDENHALL, W., SCHEAFFER, R. Y WAKERLY, D. (2002). "*Estadística matemática con aplicaciones*", Thomson, México-México
- [22] MÉNDEZ, A. (2001). "*Estimación robusta: una aplicación informática con fines didácticos*", Estadística Española; 43, 105–123
- [23] PERELLÓ, M. (2002). "*Estadística y Probabilidades*" [en línea], Universidad Politécnica Cataluña UPC ediciones
- [24] R DEVELOPMENT CORE TEAM (2008). "*R: A language and environment for statistical computing. R Foundation for Statistical Computing*", Vienna, Austria.
- [25] RENCHER, A. (1998) "*Multivariate Statistical Inference and Applications*", John Willey & Sons, New York–USA
- [26] ROCKE D. (1996) "*Robustness properties of multivariate location and shape in high dimension*", The Annals of Statistics; 24, 1327–1345
- [27] TUKEY, J. (1960) "*A survey of sampling from contaminated distributions*", Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling", Olkin Ediciones Stanford University Press, Stanford. CA.
- [28] SEBER, G. (1984) "*Multivariate observation*", John Willey & Sons, New York–USA

[29] SHORACK, G. (2000) *“Probability for statisticians”* , Springer-Verlag, New York–USA

[30] YOHAJ, V.. (2004) *“Notas de probabilidad y estadística”*, [en línea] Universidad de Buenos Aires, Buenos Aires – Argentina

[31] ZAMAR, R. (1994) *“Estimación Robusta”* , Estadística Española; 36, 327–387

[32] ZUO, Y. (2006) “Robust location and scatter estimators in multivariate analysis” (invited book chapter to honor Peter Bickel on his 65th Birthday), The Frontiers in Statistics, Imperial College Press

[33] ZURITA, G. (2008) *“Probabilidad y Estadística: Fundamentos y Aplicaciones”* ICM-ESPOL, Guayaquil-Ecuador