

**ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL**

**Facultad de Ingeniería en Estadística Informática**

“Índice de natalidad del Ecuador mediante Series  
Temporales”

TESIS DE GRADO

Previa a la obtención del Título de:

**INGENIERO EN ESTADÍSTICA INFORMÁTICA**

Presentada por:

Luz Marina Basilio Gómez

GUAYAQUIL – ECUADOR

AÑO

2001

## AGRADECIMIENTO

En primer lugar a Dios por su bendición, al Mat. John Ramírez, Director de Tesis, por su valiosa colaboración, y a mis amigos: Ing. Samaniego e Ing. Rivas por su inmenso apoyo en la realización de este trabajo.

## DEDICATORIA

A mi tío Carlos Gómez

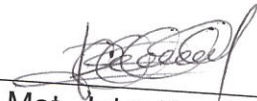
A mis padres

A mis hermanos

TRIBUNAL DE GRADUACIÓN



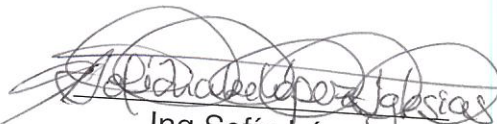
Ing. Félix Ramírez  
DIRECTOR DEL ICM



Mat. John Ramírez  
DIRECTOR DE TESIS



Mat. César Guerrero  
VOCAL



Ing. Sofía López  
VOCAL

## DECLARACIÓN EXPRESA

“La responsabilidad del contenido de esta Tesis de Grado, me corresponden exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”



---

Luz Marina Basilio G.

## RESUMEN

El presente trabajo desarrolla un análisis del número de nacimientos registrados en el país mensualmente, dentro del período 1990-1997.

Para un análisis exhaustivo de nacimientos en el Ecuador, se estudia a esta variable clasificada en: el total de nacimientos en el Ecuador, el número total de hombres nacidos vivos en el Ecuador, el total de mujeres nacidas vivas en el Ecuador, el total (hombres y mujeres) de nacimientos por provincia.

Cada una de estas clasificaciones, se las maneja como una serie, para las cuales se trata de hallar el modelo adecuado que mejor se ajuste a los datos, y se lo realiza mediante de series de tiempo, utilizando el método de Box y Jenkins, descrito en la primera parte del trabajo, y basándose en pruebas estadísticas para seleccionar el mejor modelo.

También se realiza un análisis del comportamiento que cada una de las series ha tenido durante los años que corresponden al período de estudio.

## ÍNDICE GENERAL

	<b>Pág.</b>
RESUMEN.....	.....III
ÍNDICE GENERAL.....	.....IV
ÍNDICE DE FIGURAS.....	.....VI
ÍNDICE DE TABLAS.....	.....VIII
INTRODUCCIÓN.....	.....1
I. LA NATALIDAD EN EL ECUADOR.....	.....3
1.1 Los componentes básicos del cambio de la población.....	.....3
1.2 Fertilidad.....	.....4
1.3 Distribución de la población.....	.....5
1.3.1 Distribución por edad.....	.....5
1.3.2 Distribución por sexo.....	.....7
1.4 Población en los países desde 1950.....	.....9
1.5 Poblaciones en países industrializados desde 1950.....	.....12
1.6 Proyecciones poblacionales.....	.....13
1.7 Población actual del mundo y sus estimaciones.....	.....15
1.8 Población del Ecuador.....	.....18
1.9 Fecundidad en el Ecuador.....	.....23

1.10 Crecimiento de la población del Ecuador.....	28
II. MARCO TEÓRICO .....	32
2.1 Series de tiempo y Procesos Estocásticos.....	32
2.1.1 Series de tiempo Determinísticas y Estadísticas.....	33
2.1.2 Procesos estocásticos.....	34
2.1.3 Procesos estocásticos estacionarios.....	35
2.2 Matrices definidas positivas y de autocovarianza.....	38
2.3 Tipos de Procesos Lineales.....	40
2.3.1 Ecuaciones de primer orden.....	42
2.3.2 Ecuaciones de Segundo orden.....	43
2.4 Modelos Lineales.....	44
2.4.1 Modelos Autorregresivos (AR).....	44
2.4.2 Modelos de Medias Móviles MA.....	49
2.4.3 Modelos Mixtos Autorregresivos Medias Móviles (ARMA).....	52
2.5 Procesos no estacionarios.....	57
2.5.1 Modelos ARIMA.....	60
2.6 Modelos estacionarios multiplicativos (SARIMA).....	61
III. ANÁLISIS DE SERIES DE TIEMPO.....	64
3.1 Método de Box y Jenkins.....	64
3.1.1 Elección de $D$ .....	65
3.1.2 Elección de $P$ Y $Q$ .....	66
3.1.3 Fase de verificación.....	67



3.1.4	Uso de los ruidos blancos.....	71
3.1.5	Elección del modelo.....	72
3.2	Clasificación de las series utilizadas para el análisis.....	75
3.3	Análisis de la Serie del Total de nacidos vivos en el país.....	77
3.4	Análisis de la serie del total de hombres nacidos en el país.....	93
3.5	Análisis de la serie del total de mujeres nacidas en el país.....	106
3.6	Análisis de serie de la provincia Pichincha.....	114
3.7	Análisis de serie de la provincia Chimborazo.....	123
3.8	Análisis de serie de la provincia Loja.....	133
3.9	Análisis de la serie de la provincia Azuay.....	140
3.10	Análisis de la serie de la provincia Manabí.....	152
3.11	Análisis de la serie de la provincia Guayas.....	161
3.12	Análisis de la serie de la provincia El Oro.....	172
3.13	Análisis de la serie de la provincia Napo.....	181
3.14	Análisis de la serie de la provincia Pastaza.....	191

## ÍNDICE DE FIGURAS

		<b>Pág.</b>
Figura 1.1	Tasas de natalidad del Ecuador.....	27
Figura 1.2	Población del Ecuador.....	29
Figura 2.1	Ejemplo del gráfico de una serie de tiempo.....	33
Figura 3.1	Total de niños nacidos vivos en el país.....	77
Figura 3.2	Producto Interno Bruto del Ecuador. Período 1990-1997.....	78
Figura 3.3	Autocorrelaciones de la serie Total.....	80
Figura 3.4	Autocorrelaciones estimadas de la serie Total con una diferenciación estacional.....	81
Figura 3.5	Gráfico de la Serie Total con una diferenciación estacionaria y una diferenciación estacional.....	82
Figura 3.6	Autocorrelaciones estimadas de la serie Total con una diferenciación estacionaria y una diferenciación estacional.....	82
Figura 3.7	Autocorrelaciones parciales estimadas de la serie Total con una diferenciación estacional y una diferenciación estacionaria.....	83
Figura 3.8	Autocorrelaciones de los residuos de la serie Total con el modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,1).....	87
Figura 3.9	Autocorrelaciones parciales de los residuos de la serie Total con el modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,1).....	88
Figura 3.10	Comparación de la serie real del Total con los pronósticos de los modelos propuestos.....	90
Figura 3.11	Gráfico de las predicciones de la serie del Total de nacimientos en el país.....	92
Figura 3.12	Total de hombres y mujeres nacidos en el Ecuador.....	93
Figura 3.13	Total de Hombres nacidos en el país.....	94
Figura 3.14	Autocorrelaciones del Total de Hombres nacidos en el Ecuador.....	95
Figura 3.15	Autocorrelaciones parciales del Total de Hombres nacidos.....	95
Figura 3.16	Gráfico de la serie del total de Hombres con una diferenciación estacional y una estacionaria.....	96

Figura 3.17	Autocorrelaciones del Total de Hombres con una diferenciación estacional y una estacionaria.....	97
Figura 3.18	Autocorrelaciones parciales del Total de Hombres con una diferenciación estacional y una estacionaria.....	98
Figura 3.19	Comparación de la serie real del Total de hombres con los pronósticos de los modelos propuestos.....	103
Figura 3.20	Gráfico de las predicciones de la serie Total de hombre con el modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,0).....	105
Figura 3.21	Total de Hombres nacidos en el país.....	106
Figura 3.22	Autocorrelaciones del Total de Mujeres con una diferencia estacional y una diferencia estacionaria.....	107
Figura 3.23	Autocorrelaciones parciales del Total de Mujeres con una diferencia estacional y estacionaria.....	108
Figura 3.24	Comparación de la serie real del Total de Mujeres con los pronósticos del modelo SARIMA <sub>12</sub> (0,1,0)x(0,1,1).....	111
Figura 3.25	Gráfico de las predicciones de la serie Total de Mujeres con el modelo SARIMA <sub>12</sub> (0,1,0)x(0,1,1).....	113
Figura 3.26	Gráfico de la serie de la Provincia del Pichincha.....	115
Figura 3.27	Comparación de la serie real de la provincia Pichincha con los pronósticos de los modelos propuestos.....	120
Figura 3.28	Gráfico de las predicciones de la serie de la Prov. Pichincha.....	122
Figura 3.29	Gráfico de la serie de la Provincia Chimborazo.....	123
Figura 3.30	Autocorrelaciones de la serie Prov. Chimborazo	
	Autocorrelaciones de la serie Prov. Chimborazo.....	124
Figura 3.31	Autocorrelaciones parciales de la serie Prov. Chimborazo..	124
Figura 3.32	Autocorrelaciones estimadas de la serie de la Prov. Chimborazo con una diferenciación estacional.....	125
Figura 3.33	Autocorrelaciones parciales estimadas de la serie de la Prov. Chimborazo con una diferenciación estacional.....	126
Figura 3.34	Comparación de la serie real de la provincia Chimborazo con los pronósticos del modelo SARIMA <sub>12</sub> (0,0,0)(1,1,1).....	129
Figura 3.35	Autocorrelaciones y Autocorrelaciones parciales de los Residuos del modelo SARIMA <sub>12</sub> (0,0,0)(1,1,1) de la serie de la Prov. Chimborazo.....	130
Figura 3.36	Gráfico de las predicciones de la serie de la Prov. Chimborazo con el modelo SARIMA <sub>12</sub> (0,0,0)x(1,1,1).....	132
Figura 3.37	Gráfico de la serie de la Provincia Loja.....	133
Figura 3.38	Autocorrelaciones estimadas de la serie de la Prov. Loja con una diferencia estacional.....	134
Figura 3.39	Autocorrelaciones parciales estimadas de la serie de la Prov. Loja con una diferencia estacional.....	134
Figura 3.40	Comparación de la serie real de la provincia Loja con los pronósticos del modelo SARIMA(1,0,1)x(1,1,0) <sub>12</sub> .....	136

Figura 3.41	Gráfico de las predicciones de la serie de la Prov. Loja con el modelo $SARIMA_{12}(1,0,1)x(1,1,0)$ .....	..138
Figura 3.42	Comparación de los años 95 y 96 con los pronósticos del 97 de la serie de la Prov. Loja.....	..139
Figura 3.43	Gráfico de la serie de la Provincia Azuay.....	..140
Figura 3.44	Autocorrelaciones de la serie Prov. Azuay.....	..140
Figura 3.45	Autocorrelaciones parciales de la serie Prov. Azuay.....	..141
Figura 3.46	Autocorrelaciones estimadas de la serie Prov. Azuay con una diferencia estacional.....	..142
Figura 3.47	Autocorrelaciones parciales estimadas de la serie Prov. Azuay con una diferencia estacional.....	..142
Figura 3.48	Autocorrelaciones estimadas de la serie Prov. Azuay con una diferencia estacionaria.....	..145
Figura 3.49	Autocorrelaciones parciales estimadas de la serie Prov. Azuay con una diferencia estacionaria.....	..146
Figura 3.50	Comparación de los datos reales con los pronósticos de los modelos de la serie de la Prov. Azuay.....	..149
Figura 3.51	Gráfico de las predicciones de la serie de la Prov. Azuay con el modelo $SARIMA_{12}(0,1,0)x(1,0,1)$ .....	..151
Figura 3.52	Gráfico de la serie de la Provincia Manabí.....	..152
Figura 3.53	Autocorrelaciones de la serie Prov. Manabí con una diferenciación estacional.....	..153
Figura 3.54	Autocorrelaciones Parciales de la serie Prov. Manabí con una diferenciación estacional.....	..154
Figura 3.55	Autocorrelaciones residuales de la serie Prov. Manabí con el modelo $SARIMA_{12}(0,0,2)x(2,1,0)$ .....	..156
Figura 3.56	Autocorrelaciones Parciales residuales de la serie Prov. Manabí con el modelo $SARIMA_{12}(0,0,2)x(2,1,0)$ .....	..157
Figura 3.57	Comparación de la serie real de la provincia Manabí con los pronósticos del modelo $SARIMA_{12}(0,0,2)x(2,1,0)$ .....	..158
Figura 3.58	Gráfico de las predicciones de la serie de la Prov. Manabí con el modelo $SARIMA_{12}(0,0,2)x(2,1,0)$ .....	..160
Figura 3.59	Gráfico de la serie de la Provincia Guayas.....	..161
Figura 3.60	Autocorrelaciones de la serie de la Prov. Guayas.....	..162
Figura 3.61	Autocorrelaciones de la serie de la Prov. Guayas con una diferenciación estacional.....	..163
Figura 3.62	Autocorrelaciones parciales de la serie de la Prov. Guayas con una diferenciación estacional.....	..163
Figura 3.63	Autocorrelaciones residuales de la serie Prov. Guayas con el modelo $SARIMA_{12}(1,0,0)x(2,1,0)$ .....	..168
Figura 3.64	Autocorrelaciones parciales residuales de la serie Prov. Guayas con el modelo $SARIMA_{12}(1,0,0)x(2,1,0)$ .....	..169
Figura 3.65	Comparación de la serie real de la provincia Guayas con los pronósticos del modelo $SARIMA(1,0,0)(2,1,0)_{12}$ .....	..169

Figura 3.66	Gráfico de las predicciones de la serie de la Prov. Guayas con el modelo SARIMA <sub>12</sub> (1,0,0)x(2,1,0).....	..171
Figura 3.67	Gráfico de la serie de la Provincia El Oro.....	..172
Figura 3.68	Autocorrelaciones de la serie de la Prov. El Oro con una diferencia estacional.....	..173
Figura 3.69	Autocorrelaciones parciales de la serie de la Prov. El Oro con una diferencia estacional.....	..174
Figura 3.70	Autocorrelaciones parciales residuales de la serie Prov. El Oro con el modelo SARIMA <sub>12</sub> (1,0,0)x(0,1,1).....	..177
Figura 3.71	Autocorrelaciones parciales residuales de la serie Prov. El Oro con el modelo SARIMA <sub>12</sub> (1,0,0)x(0,1,1).....	..177
Figura 3.72	Comparación de la serie real de la provincia El Oro con los pronósticos del modelo SARIMA(1,0,0)(0,1,1) <sub>12</sub> .....	..178
Figura 3.73	Gráfico de las predicciones de la serie de la Prov. El Oro con el modelo SARIMA <sub>12</sub> (1,0,0)x(0,1,1).....	..180
Figura 3.74	Gráfico de la serie de la Provincia Napo.....	..181
Figura 3.75	Autocorrelaciones y autocorrelaciones parciales de la serie de la Prov. Napo .....	..182
Figura 3.76	Autocorrelaciones estimadas de la serie de la Prov. Napo..	..183
Figura 3.77	Autocorrelaciones parciales estimadas de la serie de la Prov. Napo con una diferencia estacional.....	..183
Figura 3.78	Comparación de los pronósticos de los modelos propuestos con valores reales del año 1997, de la Prov. Napo.....	..188
Figura 3.79	Gráfico de las predicciones de la serie de la Prov. Napo con el modelo SARIMA <sub>12</sub> (0,1,1)x(1,0,1).....	..190
Figura 3.80	Gráfico de la serie de la Provincia Pastaza.....	..186
Figura 3.81	Autocorrelaciones de la serie de la Prov. Pastaza.....	..192
Figura 3.82	Autocorrelaciones parciales de la serie de la Prov. Pastaza .....	..188
Figura 3.83	Autocorrelaciones de los residuos del modelo SARIMA <sub>12</sub> (1,0,1)x(1,0,2) de la serie de la Prov. Pastaza.....	..197
Figura 3.84	Autocorrelaciones de los residuos del modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Pastaza.....	..198
Figura 3.85	Autocorrelaciones parciales de los residuos del modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Pastaza.....	..198
Figura 3.86	Comparación de los pronósticos de los modelos propuestos con valores reales del año 1997, de la Prov. Pastaza.....	..199
Figura 3.87	Gráfico de las predicciones de la serie de la Prov. Pastaza con el modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1).....	..202

## ÍNDICE DE TABLAS

		<b>Pág.</b>
Tabla I	Población de los Continentes.....	18
Tabla II	Población y tasa anual de crecimiento del Ecuador (1950-2000).....	20
Tabla III	Tasas de crecimiento de América Latina y centroamérica.	22
Tabla IV	Tasas de natalidad y fecundidad en el Ecuador .....	24
Tabla V	Población del Ecuador.....	30
Tabla VI	Tasa incremental del PIB.....	79
Tabla VII	Pruebas T_Student para los parámetros del modelo $SARIMA_{12}(0,1,0)x(1,1,1)$ .....	85
Tabla VIII	Pruebas T_Student para los parámetros del modelo $SARIMA_{12}(0,1,0)x(1,1,0)$ .....	85
Tabla IX	Pruebas para los Residuos del modelo $SARIMA_{12}$ $(0,1,0)(1,1,1)$ .....	87
Tabla X	Pruebas para los Residuos del modelo $SARIMA_{12}$ $(0,1,0)x(1,1,0)$ .....	88
Tabla XI	Valores pronosticados de la serie Total con el modelo $SARIMA_{12}(0,1,0)x(1,1,0)$ .....	91
Tabla XII	Pruebas para los parámetros del modelo $SARIMA(0,1,0)(1,1,1)$ de la serie del total de Hombres..	99
Tabla XIII	Pruebas para los parámetros del modelo $SARIMA_{12}$ $(0,1,0)x(0,1,1)$ de la serie del total de Hombres.....	99
Tabla XIV	Pruebas para los parámetros del modelo $SARIMA_{12}$ $(0,1,0)x(1,1,0)$ de la serie del total de Hombres.....	100
Tabla XV	Pruebas para los residuos del modelo $SARIMA_{12}$ $(0,1,0)x(1,1,1)$ de la serie del Total de Hombres.....	101
Tabla XVI	Pruebas para los Residuos del modelo $SARIMA_{12}$ $(0,1,0)x(0,1,1)$ de la serie Total de Hombres.....	101
Tabla XVII	Pruebas para los Residuos del modelo $SARIMA_{12}$ $(0,1,0)x(1,1,0)$ de la serie Total de Hombres.....	102

Tabla XVIII	Valores pronosticados de la serie del total de hombres con el modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,0).....	104
Tabla XIX	Pruebas para los parámetros del modelo SARIMA <sub>12</sub> (0,1,0)x(0,1,1) de la serie del total de Mujeres.....	..109
Tabla XX	Pruebas para los parámetros del modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,0) de la serie del total de Mujeres.....	..109
Tabla XXI	Pruebas para los residuos del modelo SARIMA <sub>12</sub> (0,1,0)x(0,1,1) de la serie del Total de Mujeres.....	..110
Tabla XXII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,0) de la serie Total de Mujeres.....	..110
Tabla XXIII	Valores pronosticados de la serie del total de Mujeres con el modelo SARIMA <sub>12</sub> (0,1,0)x(0,1,1).....	..112
Tabla XXIV	Pruebas T_Student para los parámetros del modelo SARIMA(0,1,0)(0,1,1) <sub>12</sub> de la serie de la Prov. Pichincha..	..116
Tabla XXV	Pruebas T_Student para los parámetros del modelo SARIMA(0,1,0)(1,1,0) <sub>12</sub> de la serie de la Prov. Pichincha..	..117
Tabla XXVI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,0)x(1,1,0) de la serie de la Prov. Pichincha.....	..117
Tabla XXVII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,0)x(0,1,1) de la serie de la Prov. Pichincha.....	..118
Tabla XXVIII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,0)x(1,1,0) de la serie de la Prov. Pichincha.....	..119
Tabla XXIX	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,0)x(1,1,0) de la serie de la Prov. Pichincha.....	..119
Tabla XXX	Predicción de valores de la serie de la Prov. Pichincha.....	..121
Tabla XXXI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,0,0)x(1,1,1) de la serie de la Prov. Chimborazo.....	..127
Tabla XXXII	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,0,3)x(1,1,1) de la serie de la Prov. Chimborazo.....	..127
Tabla XXXIII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,0,0)(1,1,1) de la serie de la Prov. Chimborazo.....	..128
Tabla XXXIV	Predicción de valores de la serie de la Prov. Chimborazo con el modelo SARIMA <sub>12</sub> (0,0,0)(1,1,1).....	..131
Tabla XXXV	Pruebas T_Student para los parámetros del modelo SARIMA(1,0,1)(1,1,0) <sub>12</sub> de la serie de la Prov. Loja.....	..135
Tabla XXXVI	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,1)(1,1,0) de la serie de la Prov. Loja.....	..136
Tabla XXXVII	Predicción de valores de la serie de la Prov. Loja con el modelo SARIMA <sub>12</sub> (1,0,1)(1,1,0).....	..137

Tabla XXXVIII	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,3)x(1,1,0) de la serie de la Prov. Azuay....	..143
Tabla XXXIX	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,0)x(1,1,0) de la serie de la Prov. Azuay....	..144
Tabla XL	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,0)(1,1,0) de la serie de la Prov. Azuay.....	..144
Tabla XLI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,1,1)x(1,0,2) de la serie de la Prov. Azuay....	..147
Tabla XLII	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,1,0)x(1,0,1) de la serie de la Prov. Azuay....	..148
Tabla XLIII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,0)x(1,0,1) de la serie de la Prov. Azuay.....	..148
Tabla XLIV	Predicción de valores de la serie de la Prov. Azuay con el modelo SARIMA <sub>12</sub> (0,1,0)(1,0,1).....	..150
Tabla XLV	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,2)x(2,1,0) de la serie de la Prov. Manabí...	..155
Tabla XLVI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,0,2)x(2,1,0) de la serie de la Prov. Manabí...	..156
Tabla XLVII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,0,2)(2,1,0) de la serie de la Prov. Manabí.....	..157
Tabla XLVIII	Predicción de valores de la serie de la Prov. Manabí.....	..159
Tabla XLIX	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,3)x(2,1,0) de la serie de la Prov. Guayas...	..165
Tabla L	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,0)x(2,1,0) de la serie de la Prov. Guayas...	..166
Tabla LI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,0)x(1,1,0) de la serie de la Prov. Guayas...	..166
Tabla LII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,0)x(2,1,0) de la serie de la Prov. Guayas.....	..167
Tabla LIII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,0)x(1,1,0) de la serie de la Prov. Guayas.....	..167
Tabla LIV	Predicción de valores de la serie de la Prov. Guayas.....	..170
Tabla LV	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,2)x(0,1,1) de la serie de la Prov. El Oro....	..175
Tabla LVI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,0)x(0,1,1) de la serie de la Prov. El Oro....	..175
Tabla LVII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,0)x(0,1,1) de la serie de la Prov. El Oro.....	..176
Tabla LVIII	Predicción de valores de la serie de la Prov. El Oro.....	..179
Tabla LIX	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Napo.....	..184
Tabla LX	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,1)x(1,0,1) de la serie de la Prov. Napo.....	..185



Tabla LXI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,1,1)x(1,0,1) de la serie de la Prov. Napo.....	..185
Tabla LXII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Napo.....	..186
Tabla LXIII	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,1)x(1,0,1) de la serie de la Prov. Napo.....	..186
Tabla LXIV	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,1)x(1,0,1) de la serie de la Prov. Napo.....	..187
Tabla LXV	Predicción de valores de la serie de la Prov. Napo.....	..189
Tabla LXVI	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,1)x(1,0,2) de la serie de la Prov. Pastaza.....	..194
Tabla LXVII	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (1,0,1)x(0,1,1) de la serie de la Prov. Pastaza.....	..194
Tabla LXVIII	Pruebas T_Student para los parámetros del modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Pastaza.....	..195
Tabla LXIX	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,1)x(1,0,2) de la serie de la Prov. Pastaza.....	..195
Tabla LXX	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (1,0,1)x(0,1,1) de la serie de la Prov. Pastaza.....	..196
Tabla LXXI	Pruebas para los Residuos del modelo SARIMA <sub>12</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Pastaza.....	..196
Tabla LXXII	Predicción de valores de la serie de la Prov. Pastaza.....	..201

## INTRODUCCIÓN

El número de los habitantes que ocupan un área (tal como un país o el mundo) continuamente sigue siendo modificado por aumentos (nacimientos e inmigraciones) y pérdidas (muertes y emigraciones).

El tamaño de la población es limitado por la cantidad de alimentos, enfermedades y otros factores ambientales. Las poblaciones humanas además son afectadas por las medidas gubernamentales de la reproducción y por los desarrollos tecnológicos, especialmente en medicina y salud pública que tienden a reducir la mortalidad y han prolongado la vida.

Algunos aspectos de las sociedades humanas son muy fundamentales para el desarrollo normal de la población, como el tamaño, composición, y el porcentaje de cambio de sus poblaciones.

Tales factores afectan la economía, salud educación, estructura familiar, patrones de crímenes, lenguaje, cultura, de hecho, virtualmente cada aspecto de sociedad humana es rozado por tendencias de la población.

El estudio de poblaciones humanas se llama demografía, una disciplina con orígenes intelectuales que yacen en el siglo 18, cuando por primera vez se reconoció que la mortalidad y natalidad humana se podría examinar como un

fenómeno con regularidades estadísticas. Su recorrido cronológico es muy largo limitando la evidencia demográfica por muchos siglos en el pasado. La actual comprensión de la demografía permite proyectar (con precaución) los cambios de la población hacia varias décadas en el futuro.

# Capítulo 1

## 1. LA NATALIDAD EN EL ECUADOR

Antes de examinar lo que sucede con la natalidad y la población en el Ecuador, se realiza un análisis de los componentes y del crecimiento de las poblaciones en el mundo.

### 1.1 Los componentes básicos en el cambio de la población

En su nivel más básico, los componentes del cambio de la población de hecho son pocos. Una población cerrada (es decir, una en la cual la inmigración y la emigración no ocurren) puede cambiar según la siguiente ecuación simple: la población (cerrada) al final de un período de tiempo va a ser igual a la población de inicio del período más nacimientos y menos muertes ocurridas durante ese período. Es decir solamente ocurrirán cambios de aumentos por nacimientos y disminución por muertes en una población cerrada. Entonces vemos que la población de un país puede cambiar por los

nacimientos, las defunciones, las emigraciones y las inmigraciones. La idea de una población cerrada no es una abstracción en el mundo; a menos que uno crea que ha habido migración substancial hacia y desde este planeta, la población del mundo en su totalidad es cerrada.

## **1.2 Fertilidad**

Se llama fertilidad al número de nacimientos que una mujer tiene en promedio en una región o país determinado.

En muchas partes del mundo, la fertilidad humana es considerablemente más baja que el potencial biológico. Esto se contrasta fuertemente con las regulaciones culturales, especialmente las referentes a la unión y sexualidad, y por esfuerzos conscientes de parte de las parejas casadas de limitar su maternidad.

Un grupo de alta fertilidad según estudios realizados es el Hutterites de Norteamérica, una secta religiosa que ve la regulación de la fertilidad como un incumplimiento y la alta fertilidad como una bendición. Se conocen que las mujeres de Hutterites que se casaron entre 1921 y 1930 tuvieron un promedio de 10 niños por mujer. Mientras tanto, mujeres en muchas partes de Europa y de Norteamérica durante los años 70 y 80 tuvieron un promedio de dos

niños por mujer un 80 por ciento menos que lo alcanzado por las mujeres de Hutterites. Incluso en las poblaciones altamente fértiles de países desarrollados en África, Asia y América Latina el índice de nacimiento de niños está muy lejos de el de los Hutterites.

A los inicios del decimoctavo siglo en Francia y Hungría, hubo una declinación dramática en la fertilidad en las sociedades más desarrolladas de Europa y de Norteamérica, y en los dos siglos siguientes la curva de la variable de la fertilidad tuvo declinaciones en aproximadamente 50 por ciento de todos estos países. Desde los años 60 la fertilidad se ha disminuido intencionalmente en muchos países desarrollados, y las reducciones notables más rápidas han ocurrido en la República Popular de China.

### **1.3 Distribución de la población**

Las características más importantes de una población -- además de su tamaño y de su variación-- son las maneras por las cuales distribuyen sus miembros según la categoría de edad, sexo, etc.

#### **1.3.1 Distribución por edad**

Quizás la más fundamental de las características de una población es la distribución por edad. Comúnmente los

demógrafos usan los gráficos de barras para describir distribuciones de la edad y del sexo de una población, en el cual la longitud de cada barra representa el número (o el porcentaje) de personas en un grupo de edad, en el caso del sexo el número (o proporción) de varones y de mujeres. El número de barras se grafica de acuerdo a los intervalos en que se tome la edad. En la mayoría de las poblaciones la proporción de personas más viejas es mucho más pequeña que la de las más jóvenes.

En tres naciones las poblaciones revelan marcadas diferencias con sus características: alta fertilidad y crecimiento rápido de la población (México), fertilidad baja y crecimiento lento (Estados Unidos), y fertilidad muy baja y crecimiento negativo (República Federal de Alemania).

Aunque se crea lo contrario, el factor principal que tiende a cambiar la distribución de edad de una población no es el índice de mortalidad, sino el índice de fertilidad.

Un cambio (incremento o decremento) en la mortalidad generalmente afecta en cierta medida a todas las categorías de edad, pero tiene efectos limitados sobre cada categoría de edad. Sin embargo, un cambio en la fertilidad afecta el número

de personas en una sola categoría, el grupo de la edad cero, los recién nacidos. Esto significa que las estructuras de edades jóvenes corresponden a las poblaciones altamente fértiles, típico de países subdesarrollados y las estructuras de edades más viejas son las de las poblaciones de baja fertilidad, tales como aquellas del mundo industrializado.

### **1.3.2 Distribución por sexo**

Un segundo aspecto importante de las estructuras de las poblaciones es el número de varones y mujeres que lo componen. Generalmente, nacen más varones que mujeres (una relación típica sería 105 o 106 varones por cada 100 mujeres). Por otra parte, es muy común para los varones experimentar una alta mortalidad después de cumplir años en todas las edades, prácticamente. Esta diferencia es aparentemente de origen biológico. Las excepciones ocurren en los países tales como India, donde la mortalidad de las mujeres puede ser más alta que la de varones en su niñez y en las edades de maternidad debido a la desigual asignación de recursos dentro de la familia y de la mala calidad del cuidado médico maternal.



La regla general expone que nacen más varones que mujeres, pero las mujeres experimentan un nivel más bajo de mortalidad, esto significa que, durante la niñez los hombres sobrepasan el número de mujeres en una cierta edad y la diferencia se va dando a medida que va incrementando la edad, hasta llegar a un cierto punto en la vida de adulto donde el número de varones y de mujeres llegan a ser iguales, y mientras que se alcanzan edades más altas el número de mujeres llega a ser más grande. Por ejemplo, en Europa y Norteamérica, en 1985 entre las personas mayores a 70 años de edad, existían sólo de 61 a 63 varones por cada 100 mujeres. (Según la División Poblacional de las Naciones Unidas, la proporción para la Unión Soviética era solamente 40% del total de la población mundial, lo cual puede ser atribuible a la alta mortalidad masculina durante la II Guerra mundial así como a los posibles aumentos en la mortalidad masculina durante los años 80.)

La distribución del sexo dentro de una población tiene implicaciones significativas para los modelos de la unión matrimonial. Una escasez de varones de una determinada edad crea una variación en los índices de la unión de mujeres en un mismo grupo de edad o generalmente en edades

menores, y esto probablemente reduce su fertilidad. En muchos países, la convención social dicta un modelo matrimonial en el cual los varones sean levemente mayores que sus esposas. Así si hay una subida dramática de la fertilidad, tal como el "baby boom" en el período que seguía la II Guerra mundial, un "ajuste de la unión matrimonial" puede resultar eventual; es decir, el número de varones de edad social correcta para la unión es escaso para el número de mujeres algo más jóvenes. Esto puede conducir al aplazamiento de la unión de estas mujeres. De una manera similar, una declinación dramática de la fertilidad en tal sociedad es probable que conduzca a una eventual escasez de mujeres elegibles para la unión, lo que puede conducir a la unión temprana de estas mujeres. Todos estos efectos se desarrollan lentamente; toma por lo menos de 20 a 25 años para mantener un cambio en la fertilidad de esta manera.

#### **1.4 Población en los países desde 1950**

Después de la segunda guerra mundial hubo una rápida disminución de la mortalidad en muchos de los países desarrollados. Esto resultó, en parte, por los esfuerzos en tiempos de guerra por

mantener la salud de las fuerzas armadas de los países industrializados que luchaban en áreas tropicales. Muchas personas y gobiernos recibieron técnicas para reducir la incidencia de enfermedades y muertes, estos esfuerzos fueron aceptados fácilmente en muchos países desarrollados del mundo, pero no fueron acompañados por los cambios de las clases sociales y culturales que habían ocurrido y tampoco consideraron el aumento de la fertilidad en países industrializados.

La reducción de la mortalidad acompañado por un aumento de la fertilidad, tenía un resultado simple y predecible: acelerado crecimiento de la población. En 1960 muchos países desarrollados tuvieron índices de crecimiento de hasta 3 por ciento al año, excediéndose el doble o triple de los índices más altos experimentados por las poblaciones europeas. Desde ese aumento de la población ese índice dobló en solamente 23 años, las poblaciones de tales países se ampliaron dramáticamente. En 25 años desde 1950 a 1975, la población de México aumentó de 27.000.000 a 60.000.000; Irán de 14.000.000 a 33.000.000; Brasil de 53.000.000 a 108.000.000; y China de 554.000.000 a 933.000.000.

Los índices más grandes de crecimiento de población fueron alcanzados en América Latina y en Asia a mediados y finales de los años 60.

Desde entonces, estas regiones han experimentado cambios, por lo general han sido disminuciones temporales de fertilidad a lo largo de periodos de disminuciones de mortalidad, dando como resultado declinaciones moderadas y ocasionalmente largas en el comportamiento del crecimiento de la población. Las declinaciones más dramáticas han sido las de la República Popular de China, en donde el índice de crecimiento fue estimado sobre el 2 por ciento al año en los años 60, y a mediados de los años 80, siguiendo la adopción oficial de una política concertada a retrasar el matrimonio y a limitar la maternidad dentro de la unión. El Este de Asia ha experimentado las declinaciones más dramáticas de la población que cualesquiera de las regiones desarrolladas.

En el Sur de Asia el índice ha declinado solamente de 2,4 a 2,0 por ciento; en América latina, aproximadamente de 2,7 a 2,3 por ciento. Mientras tanto, el crecimiento de la población de África ha acelerado a partir de 2,6 por ciento a más de 3 por ciento en un mismo período, siguiendo más tarde con declinaciones significativas en la mortalidad que no han sido acompañadas por reducciones similares en la fertilidad.

### **1.5 Población en países industrializados desde 1950**

Para muchos países industrializados, el período después de la II Guerra mundial fue nombrado como el "baby boom". La fertilidad incrementó como resultado de uniones y los nacimientos aumentaron en gran cantidad. Un grupo de cuatro países en particular –Estados Unidos, Canadá, Australia, y Nueva Zelanda– experimentaron substanciales y sostenibles incrementos en la fertilidad de los niveles que existían en el período de la preguerra. En los Estados Unidos, por ejemplo, la fertilidad aumentó en un dos tercios, alcanzando en los años 50 niveles no vistos desde 1910.

Un segundo grupo de países industrializados, incluyendo la mayoría de Europa occidental y algunos países europeos del este (notablemente Checoslovaquia y el este de Alemania), experimentaron también el "baby boom".

En muchos países europeos la fertilidad fue muy baja en los años 30; pero después de la II guerra con el baby boom la fertilidad incrementó sustancialmente en un período de 3 y 4 años, para luego mantenerse estable en las dos siguientes décadas. A mediados de los años sesenta, los niveles de fertilidad en estos países comenzaron a disminuir nuevamente, y en muchos casos, bajaron hasta los niveles comparables con los de la década del 30.

Un tercer grupo de países industrializados, que consiste en la mayoría del este de Europa los cuales junto con Japón, mostraron modelos diferentes de fertilidad. La mayoría no registraron fertilidad baja en los años 30 sino experimentaron declinaciones substanciales en los años 50 después de un corto período del baby boom. Por los años 80 los niveles de fertilidad en la mayoría de los países industrializados eran muy bajos, debido a dos razones ocurrió este fenómeno: el aplazamiento del matrimonio, la maternidad de mujeres jóvenes quienes se incorporaron a la fuerza laboral, y una reducción en el nacimiento de niños de mujeres casadas.

### **1.6 Proyecciones poblacionales**

El cambio demográfico está sujeto a un fenómeno a largo plazo. De forma contraria a las poblaciones de insectos, las poblaciones humanas han estado raramente sujetas a la "explosión" o "colapso" numérico. Por otra parte, el ímpetu de prolongar la vida hacen que los efectos de los cambios de la fertilidad lleguen a ser evidentes solamente en un futuro lejano. Por éstas y otras razones, ahora es práctico emplear la teoría de la proyección de las poblaciones como medios para entender mejor las tendencias.

Las proyecciones de la población representan simplemente las manipulaciones de un conjunto de hipótesis acerca de la fertilidad, la mortalidad y los índices de migración en un futuro. Esto no puede ser mantenido fuertemente por cuanto las proyecciones no son predicciones, aunque son malinterpretadas como tal con mucha frecuencia. Una proyección es un cálculo basado en las hipótesis explícitas que pueden o no pueden estar correctas. Mientras más grande es la aritmética de una proyección esta se la hace correctamente, su utilidad es determinada por la plausibilidad de sus hipótesis centrales.

Si las asunciones incorporan tendencias futuras plausibles, entonces las salidas de la proyección pueden ser plausibles y útiles. Porque el curso de tendencias demográficas es duro de anticipar en el futuro, la mayoría de los demógrafos calculan un conjunto de las proyecciones alternativas que, tomado juntas, se espera que definan un rango de futuros plausibles, mucho mejor que predecir o pronosticar basado en una sola. Debido a que las tendencias demográficas cambian a veces de maneras inesperadas, es importante que todas las proyecciones demográficas sean actualizadas sobre una base regular e incorporar nuevas tendencias y datos para nuevamente desarrollarlas.

La División Población de las Naciones Unidas prepara cada dos años un conjunto estándar de las proyecciones para el mundo y sus países. Estas proyecciones incluyen una variante baja, media, y alta para cada país y región.

### **1.7 Población actual del mundo y sus estimaciones**

La población del mundo en 1998 fue de 5,9 mil millones de personas y está creciendo actualmente en 1,33 por ciento por año, según el Departamento de asuntos sociales de las Naciones Unidas. Se espera que la población del mundo para mediados del siglo XXI esté en el rango de 7,3 a 10,7 mil millones.

Se proyecta que la tasa de crecimiento poblacional anual va a disminuir desde 1.33 % en 1995-2000 a 0.34% en 2045-2050. Desde 1804, cuando el mundo pasó de mil millones, tomó 123 años alcanzar 2 mil millones de personas en 1927, 33 años para alcanzar 3 mil millones en 1960, 14 años para alcanzar 4 mil millones en 1974, 13 años para lograr 5 mil millones en 1987 y 12 años para alcanzar 6 mil millones en 1999. Y tomará 14 años alcanzar 7 mil millones en 2013, 15 años alcanzar 8 mil millones en 2028, y para 2054 tomará 26 años en alcanzar los 9 mil millones de habitantes.



El nivel promedio global de fertilidad ahora está en 2,7 nacimientos por mujer; al inicio de los años 50 el número promedio era 5 nacimientos por mujer. La fertilidad ahora está declinando en todas las regiones del mundo. Por ejemplo, durante los últimos 25 años, el número de niños por pares ha disminuido de 6.6 a 5.1 en África, de 5.1 a 2.6 en Asia, y de 5,0 a 2,7 en América latina y el Caribe.

En una investigación realizada en 1998 se hizo notable la influencia del SIDA sobre la esperanza de vida promedio. Por ejemplo, en el continente africano 29 países fueron los más golpeados por el SIDA, haciendo que la esperanza de vida promedio se redujera en 7 años de estimado antes del SIDA. Actualmente en Botswana la influencia del SIDA es muy alto en la esperanza de vida (en donde son infectados uno de cada 4 adultos) puesto que tiende a bajar de 61 años (1990-1995) a 41 años (2000-2005). De acuerdo con las proyecciones de las Naciones Unidas, la población de Botswana para el 2025 puede ser 23 por ciento más pequeña de lo que habría sido en ausencia del SIDA. Sin embargo, debido que la fertilidad es alta, todavía se espera que la población de Botswana se duplique entre 1995 y 2050.

Como resultado de una baja en la tasa de fertilidad se estimará una declinación rápida en el índice anual de la población cambiando a un valor negativo de -0,23 por ciento por año a mediados del siglo XXI. La población en 2050 será como máximo 10,7 mil millones y como mínimo 7,3 mil millones según la variante baja.

El 97% del aumento de la población del mundo ocurre en las regiones menos desarrolladas. Cada año la población de Asia está aumentando en 50 millones, la población de África en 17 millones, y la de América latina y del Caribe en casi 8 millones. África tiene la tarifa de crecimiento más alta entre todas las áreas importantes (2,36 por ciento). África media, África del este y África occidental tienen índices de crecimiento de 2,5 por ciento. Europa, por otra parte, tiene la tarifa de crecimiento más baja (0,03 por ciento).

Tabla I

**Población de los Continentes**

(Población en millones)

	<b><u>1950</u></b>	<b><u>1998</u></b>
<b>Mundo</b>	<b>2,521</b>	<b>5,901</b>
Países desarrollados	813	1,182
Países no desarrollados	1,709	4,719
África	221	749
Asia	1,402	3,585
Europa	547	729
América Latina y el Caribe	167	504
Norte América	172	305
Oceanía	13	30

Fuente: Datos de las Naciones Unidas.

**1.8 Población del Ecuador**

Para todo el país cada día es más evidente la importancia que tiene para el desarrollo nacional los cambios en la población. El Ecuador es una de las naciones del mundo donde los cambios en la población

están ocurriendo a un paso acelerado y tales cambios tienen también un efecto directo e importante sobre el desarrollo económico y social.

Según los resultados del censo de 1982, la población del Ecuador era de 8.060.712 habitantes, menor a la que se había estimado para esa fecha; así por ejemplo, las Naciones Unidas habían pronosticado una población de 8.537.000 habitantes para 1982. Los resultados oficiales representan un aumento de 1.539.002 habitantes sobre el censo de 1974 cuyos resultados fueron de 6.521.710 habitantes; las Naciones Unidas habían estimado un aumento de 2.015.000 o casi una cuarta parte adicional de crecimiento. Según el censo, la tasa promedio de crecimiento entre 1974 y 1982 fue el 2,5 por ciento por año, en lugar del 3,1 por ciento que las Naciones Unidas habían calculado.

El Instituto Nacional de Estadísticas y Censos y CELADE (Centro latinoamericano de Demografía) han estudiado los censos pasados del Ecuador, los habían ajustado, para que fueran más consistentes entre sí, y estimado el crecimiento del Ecuador entre 1950 y 1982 conjuntamente con las tendencias de crecimiento más probables para el país hasta el año 2000, estos datos aparecen en la Tabla II, donde se presentan los publicados por las Naciones Unidas en 1983

en un informe sobre tendencias mundiales de población, y el estudio realizado en diciembre de 1984 por CONADE, INEC y CELADE.

**Tabla II**

**Población y tasa anual de crecimiento del Ecuador (1950-2000)**

<b>Año</b>	<b>Población</b>	<b>Año</b>	<b>Tasa anual</b>
1950	3.307.000	1950-55	2,84
1960	4.422.000	1960-65	2,99
1970	5.958.000	1965-70	2,98
1975	6.981.000	1970-75	2,91
Proyectado1			
1980	8.021.000	1975-80	3,04
1984	9.090.000	1980-85	3,13
1985	9.380.000	1985-90	3,09
1990	1.094.900	1995-2000	2,78
2000	14.596.000		
Proyectado2			
1985	9.377.980	1980-85	2,87
1990	10.781.613	1985-90	2,79
1995	12.314.210	1990-95	2,65
2000	13.939.400	1995-2000	2,48
<b>Proyectado1</b>	<i>Fuente Naciones Unidas, 1982</i>		
<b>Proyectado2</b>	<i>Fuente CONADE, CELADE, INEC, 1984</i>		

En las siguientes estimaciones de crecimiento de las Naciones Unidas indican que a partir de 1960, la población del Ecuador habría crecido más rápidamente en promedio que la de cualquier otra nación en América del Sur con la excepción de Bolivia, Venezuela y Paraguay. El cuadro permite analizar que mientras algunos países, en especial de América Central, están creciendo muy rápidamente (cerca del 3 por ciento anual) otros (Argentina, Chile, Uruguay) mantienen una limitada tasa de incremento de su población.

Tabla III

## Tasas de crecimiento de América Latina y centroamérica

<i>Países</i>	Tasas de crecimiento total (por mil)						
	<i>1970-1975</i>	<i>1975-1980</i>	<i>1980-1985</i>	<i>1985-1990</i>	<i>1990-1995</i>	<i>1995-2000</i>	<i>2000-2005</i>
<b>América Latina</b>	24.9	23.5	20.9	19.2	17.4	16.0	14.6
<b>Argentina</b>	16.7	15.1	15.2	14.1	13.3	12.6	11.9
<b>Bolivia</b>	24.5	23.6	19.3	21.8	24.1	23.3	21.5
<b>Brasil</b>	23.8	23.5	21.2	18.0	15.1	13.4	12.4
<b>Chile</b>	17.0	15.1	15.5	16.7	16.3	13.6	11.8
<b>Colombia</b>	23.6	22.8	21.4	19.9	19.5	18.7	16.8
<b>Costa Rica</b>	25.7	29.7	29.0	28.6	30.5	24.8	20.3
<b>Cuba</b>	17.7	8.5	8.2	9.9	6.3	4.3	3.0
<b>Ecuador</b>	29.1	28.4	26.7	24.1	22.0	19.7	17.4
<b>El Salvador</b>	27.1	21.5	8.1	13.9	20.7	20.4	18.2
<b>Guatemala</b>	27.6	25.1	25.4	24.6	26.3	26.4	25.8
<b>Paraguay</b>	24.7	31.5	29.5	31.2	27.0	25.9	24.6
<b>Uruguay</b>	1.5	5.9	6.4	6.4	7.1	7.3	7.0
<b>Venezuela</b>	34.3	33.8	25.4	25.8	22.7	20.2	18.2
<b>Nicaragua</b>	32.5	31.3	30.6	23.5	29.0	27.3	26.7
<b>Panamá</b>	26.9	24.7	21.2	20.2	18.6	16.4	14.3
<b>Perú</b>	27.8	26.7	23.6	20.3	17.4	17.3	16.0

<b>Haití</b>	17.0	20.6	23.6	24.7	18.7	18.4	18.2
<b>Rep.Dominic</b>	26.5	24.2	22.5	21.8	19.1	16.5	14.3
<b>Honduras</b>	30.3	33.5	31.9	30.6	29.4	27.4	24.9
<b>México</b>	31.0	26.8	22.1	19.6	18.2	16.3	14.2

Fuente: CELADE

### 1.9 Fecundidad en el Ecuador

A lo largo de la historia, el Ecuador ha tenido tasas muy altas de natalidad. Las estimaciones más recientes de las Naciones Unidas demuestran que ha habido una disminución gradual, pero que las tasas continúan siendo sumamente altas tanto en comparación con las otras naciones de Sudamérica como con naciones de Europa y Norte América.



Tabla IV

## Tasas de natalidad y fecundidad del Ecuador

<i>Años</i>	<i>Tasa bruta de natalidad(TBN)</i>	<i>Tasa global de Fecundidad(TGF)</i>
1950-55	46,8	6,90
1960-65	45,6	6,90
1965-70	44,5	6,70
1970-75	41,2	6,05
1975-80	38,2	5,40
1980-85	36,8	5,00

Datos proporcionados por la CONADE

La fecundidad ha experimentado un descenso suave pero sostenido a finales de la década de los 60's, se manifiesta con mayor intensidad en las dos últimas décadas, descensos que involucran cambios relevantes en aspectos socio-demográficos y económicos de la población.

Es importante advertir que, no obstante la disminución ocurrida, el nivel de la fecundidad del Ecuador, considerado como "moderado", es todavía superior al de América Latina (3.2 hijos por mujer para el período 1990-1995), e inferior a lo observado en países como Bolivia (6 hijos) y Guatemala (5.7 hijos).

La tasa global de fecundidad (TGF, promedio de hijos por mujer durante su período reproductivo), bajo a nivel nacional de 5.3 hijos en 1982 a 4.1 en 1990 (TGF, corregida por metodología indirecta). De mantenerse el nivel actual de la fecundidad, implicaría que cada mujer tendrá cerca de cuatro hijos durante su vida reproductiva, lo que permitiría que la población se duplique en una generación (alrededor de 28 años), determinando un significativo potencial de crecimiento.

El descenso de la fecundidad se ha dado de manera diferencial por áreas, como manifestación del desigual desarrollo socio-económico alcanzado, dando lugar a descensos con diferentes intensidades.

Las mujeres del área urbana en forma persistente presentan tasas de fecundidad inferiores a las del área rural, diferencias de 3 hijos en 1982 (4 y 7 hijos, respectivamente) y algo más de 2 hijos para 1990 (3.3 vs. 5.5 hijos). Entre éstas, el descenso fue mayor en el área rural (20%) que en las ciudades (17%).

A pesar del importante descenso de la fecundidad en ambas áreas, para 1990, la fecundidad de las mujeres rurales es más alta en un 65% que la de las urbanas ( más de dos hijos de diferencia).

Los descensos se dan en todas las provincias pero con diferente intensidad, más alta en Manabí y Zamora Chinchipe, entre las provincias serranas mayores descensos Cañar y Cotopaxi.

El descenso de la fecundidad observado en el país no se dio simultáneamente en todos los grupos de población. Se inició en grupos que poseían características económicas-sociales especiales, relacionadas inequívocamente con determinantes como: la educación, tipo de actividad económica, lugar de residencia, etc. Al relacionar la fecundidad con el nivel de instrucción de la mujer, el comportamiento reproductivo determina una relación inversa: a menor educación, mayor promedio de hijos por mujer, a mayor nivel de educación el promedio de hijos disminuye.

La relación educación y fecundidad se ve estrechamente vinculada al desarrollo histórico y económico-social alcanzado por nuestras sociedades. La jerarquía ocupacional de la mujer ejerce cierta influencia en la fecundidad por su asociación positiva con el grado de escolaridad y el ingreso monetario; las mujeres económicamente

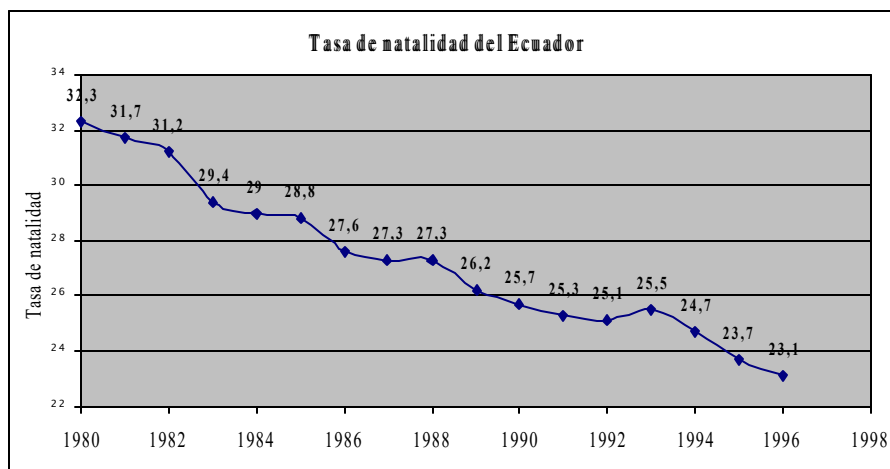
activas tienen en promedio 3 hijos, casi 2 hijos menos que aquellas que no realizan ninguna actividad económica (4.6 hijos).

En términos generales, la baja en la fecundidad se debe, en gran parte, al aumento en los niveles de educación, al proceso de urbanización, y a una mayor prevalencia de uso de anticonceptivos.

Para 1984 cada mujer en el Ecuador tenía un promedio de 5 niños, esta tasa era más de dos veces y media el tamaño de Europa (1,87niños) o de Norte América (2,04 niños).

**Figura 1.1**

**Tasas de natalidad del Ecuador**



Datos proporcionados por el INEC.

El Ecuador tiene una de las tasas más altas de natalidad en América del Sur. Sólo Bolivia tiene tasas de natalidad más altas que las del Ecuador. En América Central, Nicaragua y Honduras tienen tasas más altas de natalidad que las del Ecuador.

Cabe recalcar que según datos recolectados por la Encuesta Nacional de Fecundidad realizada por el INEC, la fecundidad en la región Costa era más alta en 1979 que la de la Sierra. Estos diferenciales también se encontraron a niveles de tamaño de las ciudades (grandes, intermedias y pequeñas) y de la población rural.

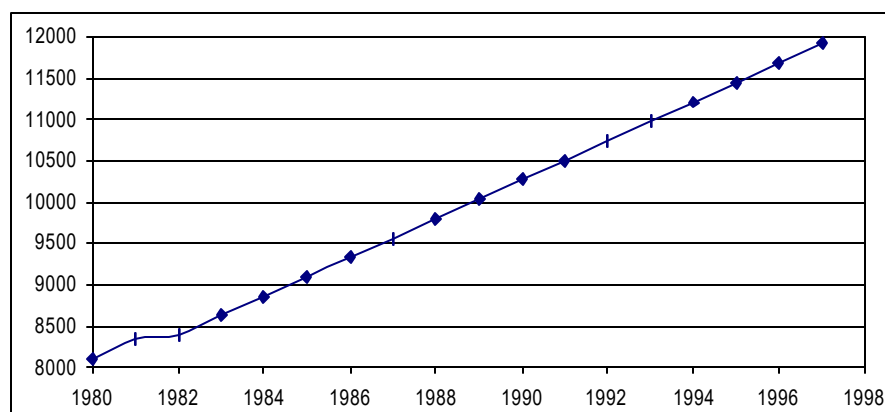
### **1.10 Crecimiento de la población del Ecuador**

El crecimiento de la población es el balance neto entre nacimientos, defunciones y migración. Debido a que el Ecuador durante la década del 60 y 70 ha tenido comparativamente poca inmigración y emigración internacional, su crecimiento se debió casi exclusivamente a los nacimientos y defunciones. La combinación de tasas de nacimiento elevadas y las tasas de mortalidad bastantes

reducidas son las causantes del rápido crecimiento, durante este período, de la población del país.

En 1984 Ecuador añadió 257.000 personas por un año a su población. El resto significa un incremento de 700 personas por día o una persona adicional cada dos minutos.

**Figura 1.2**  
**Población del Ecuador**



Datos proporcionados por el INEC.

**Tabla V**  
**Población del Ecuador**

<b>Año</b>	<b>Población (en miles)</b>	<b>Numero de nacidos</b>	<b>Tasa bruta de natalidad</b>
1980	8123.3	262778	32.3
1981	8361.3	264963	31.7
1982	8409	262102	31.2
1983	8637.8	253990	29.4
1984	8868.2	257044	29
1985	9098.8	262260	28.8
1986	9329.6	257234	27.6
1987	9561.4	261312	27.3
1988	9794.4	267652	27.3
1989	10028.6	262652	26.2
1990	10264.1	263612	25.7
1991	10501.5	265611	25.3
1992	10740.7	269903	25.1
1993	10980.9	279678	25.5
1994	11221	277625	24.7
1995	11460.1	271340	23.7
1996	11698.4	270578	23.1
1997	11936.8	169869	14.2

Datos proporcionados por el INEC.

*Ecuador puede escoger la rapidez con la que su población crecerá en el futuro.*

Es incorrecto suponer que la rapidez con que la población del Ecuador crecerá en el futuro es cosa del destino. Al contrario, el crecimiento de la población debe constituir un hecho social consciente, derivado del derecho consagrado en la Declaración de los derechos Humanos de las Naciones Unidas, según el cual la persona o la pareja puede decidir sobre el número de hijos a los que deberá alimentar y educar. Para que esta situación se cumpla es menester que se incrementen los ya importantes esfuerzos en educación, para que todas las personas conozcan su responsabilidad en el aumento de la población, social y familiarmente considerada.

Así comenzamos este estudio de natalidad del Ecuador comprendiendo que contamos con una población cuyo tamaño está aumentando rápidamente con cambios importantes en su distribución; además que tiene mucha tierra por explorar que se podría ocupar si de alguna, manera se logra que esta tierra proporcione un medio de vida adecuado para aquellos que deseen vivir en ella.



# Capítulo 2

## 2. MARCO TEORICO

### 2.1 Series de tiempo y Procesos Estocásticos

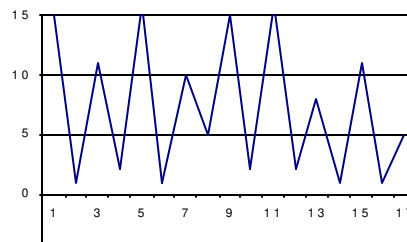
Una serie de tiempo es un conjunto de observaciones generadas secuencialmente en el tiempo. Si el conjunto de los instantes de observaciones es continuo, se llama a la serie de tiempo *continua*. Si el conjunto es discreto, la serie de tiempo se llama *discreta*. Así las observaciones de una serie de tiempo discreta hechas en los tiempos  $\tau_1, \tau_2, \dots, \tau_t, \dots, \tau_N$ , pueden ser denotadas por  $Y(\tau_1), Y(\tau_2), \dots, Y(\tau_t), \dots, Y(\tau_N)$ . Considérese sólo las series de tiempo discretas donde las observaciones son realizadas cada intervalo  $h$ . Si se tiene  $N$  valores sucesivos de una serie, se puede escribir  $Y_1, Y_2, \dots, Y_t, \dots, Y_N$  para denotar observaciones realizadas en intervalos de tiempo equidistantes  $\tau_0+h, \tau_0+2h, \dots, \tau_0+th, \dots, \tau_0+Nh$ . Para algunos propósitos los valores de  $\tau_0$  y  $h$  no son importantes, pero si las observaciones necesitan tiempos para ser definidas, estos dos

valores pueden ser especificados. Si se toma  $\tau_0$  como el origen y  $h$  como la unidad tiempo, se puede considerar  $Y_t$  como la observación en el tiempo  $t$ . Las series de tiempo discretas se las puede obtener:

1. Por muestreo.
2. Por la acumulación de valores de una variable sobre un periodo de tiempo. Como se observa en la figura 2.1, se encuentra el gráfico de valores que se han tomado de la lectura del número de artículos fabricados por una máquina.

**Figura 2.1**

**Ejemplo del gráfico de una serie de tiempo**



### 2.1.1 Series de tiempo Determinísticas y Estadísticas

Si los valores futuros de una serie de tiempo son exactamente determinados por alguna función matemática tal como:

$Y_t = \cos(2\pi f t)$ , dice que la serie de tiempo es *determinística*.

Si los valores futuros pueden ser descritos en términos de una distribución de probabilidad, se dice que la serie de tiempo es no determinística o *estadística*.

En el presente trabajo se considera sólo las series de tiempo estadísticas.

### 2.1.2 Procesos estocásticos

Un fenómeno estadístico que ocurre en el tiempo de acuerdo a leyes de probabilidad es llamado un *proceso estocástico*.

Las series de tiempo a ser analizadas entonces pueden ser vistas como una particular realización, producida por un mecanismo de probabilidad subyacente del sistema bajo estudio. En otras palabras, un análisis de una serie de tiempo se considerará como una realización de un proceso estocástico.

Se puede considerar la observación  $Y_t$  en el tiempo  $t$ , como una variable aleatoria  $Y_t$  con función de probabilidad  $p(Y_t)$ . Similarmente las observaciones en dos tiempos:  $t_1$  y  $t_2$  pueden considerarse como realizaciones de 2 variables aleatorias  $Y_{t_1}$  y  $Y_{t_2}$  con función de densidad conjunta  $P(Y_{t_1}, Y_{t_2})$ . En general, las observaciones hechas en una serie de tiempo pueden ser descritas por una variable aleatoria  $N$ -dimensional  $(Y_1, Y_2, \dots, Y_N)$  con distribución de probabilidad  $P(Y_1, Y_2, \dots, Y_N)$ .

En una serie temporal se dispone de una observación para cada período de tiempo, por lo que se la puede considerar como una muestra de tamaño 1 tomada en períodos sucesivos de tiempo en un proceso estacionario, es decir, es una realización de un proceso estacionario. A diferencia de un muestreo aleatorio simple donde cada extracción es independiente de las demás, en una serie de tiempo el dato “extraído” para un período concreto no será, en general, independiente de los datos de períodos anteriores.

Así, si se disponen de  $n$  datos de una serie de tiempo, con ellos hay que estimar  $n$  medias y  $n$  varianzas. Para poder, a partir de una sola realización efectuar inferencias sobre un proceso estacionario es preciso imponer restricciones al proceso estacionario, éstas son que sea estacionario y ergódico.

### **2.1.3 Procesos estocásticos estacionarios**

Una clase muy especial de los procesos estocásticos son los llamados procesos estacionarios que se basan en la suposición de que un proceso se encuentra en un estado particular de equilibrio estadístico.

Un **proceso estocástico** se dice que es **estrictamente estacionario** si sus propiedades no son afectadas por un cambio en el tiempo original, esto es, si la distribución de probabilidad conjunta asociada con  $m$  observaciones  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_m}$ , realizadas en un conjunto de tiempos  $t_1, t_2, \dots, t_m$ , es la misma que la asociada con  $m$  observaciones  $Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_m+k}$ , realizadas en los tiempos  $t_1+k, t_2+k, \dots, t_m+k$ .

$$P(Y_{t_1}, Y_{t_2}, \dots, Y_{t_m}) = P(Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_m+k})$$

Pero este enfoque es complicado, por lo que se recurre a los momentos. Se dice que un proceso estacionario es estacionario de primer orden o en media si:

$$\forall t: E(Y_t) = \mu$$

es decir la media permanece constante a lo largo del tiempo.

Se dice que un proceso estacionario es estacionario de segundo orden (o en sentido amplio) cuando además de ser estacionario de primer orden, se verifican las siguientes condiciones.

- 1) La varianza es finita y constante a lo largo del tiempo

$$\forall t: \text{Var}(Y_t) = E[(Y_t - \mu)]^2 = \sigma^2 < \infty$$

2) La autocovarianza entre dos períodos distintos de tiempo viene afectada únicamente por el lapso de tiempo transcurrido entre estos dos períodos:

$$\forall t: E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k$$

es decir

$$\text{Cov}(Y_t, Y_{t+k}) = \gamma_k$$

que es una autocovarianza de orden  $k$ , por ser este el lapso de tiempo que separa a  $Y_t$  de  $Y_{t+k}$ . El valor de  $\gamma_k$  es independiente de  $t$ , así la varianza sería  $\gamma_0$  (la autocovarianza de orden cero).

Y se define a la autocorrelación de orden  $K$  como:

$$\rho_k = \frac{E[(Y_t - \mu)(Y_{t+k} - \mu)]}{\sqrt{E[(Y_t - \mu)^2]E[(Y_{t+k} - \mu)^2]}}$$

$$\rho_k = \frac{\text{Cov}(Y_t, Y_{t+k})}{\sigma_Y^2}$$

Entonces para un proceso estacionario, la varianza  $\sigma_Y^2 = \gamma_0$  es la misma en el tiempo  $t+k$  y en el tiempo  $t$ . Así la autocorrelación de orden  $k$  (autocorrelación entre  $Y_t$  y  $Y_{t+k}$ ) es:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

## 2.2 Matrices definidas positivas y de autocovarianza

La matriz de covarianza asociada con un proceso estacionario para observaciones  $(Y_1, Y_2, \dots, Y_n)$  hechas en  $n$  tiempos sucesivos es:

$$\Gamma_n = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \dots & \gamma_0 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{n-3} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & 1 \end{bmatrix} = \sigma_Y^2 P_n$$

Una matriz de covarianza  $\Gamma_n$  de esta forma, la cual es simétrica con elementos constantes en la diagonal, será llamada *matriz de autocovarianza* y la matriz de correlación  $P_n$  correspondiente será llamada *matriz de autocorrelación*.

Ahora consideremos una función lineal de variables aleatorias

$$Y_t, Y_{t-1}, \dots, Y_{t-n+1}$$

$$L_t = l_1 Y_t + l_2 Y_{t-1} + \dots + l_n Y_{t-n+1} \quad (1)$$

Entonces  $\text{Cov}(Y_i, Y_j) = \gamma_{|j-i|}$  para un proceso estacionario, la varianza de  $L_t$  es:

$$\text{Var}[L_t] = \sum_{i=1}^n \sum_{j=1}^n l_i l_j \gamma_{|j-i|}$$

Lo cual es mayor que cero, si es que no todas las  $l$ 's son cero. Por lo tanto una matriz de autocovarianza y de autocorrelación son definidas positivas para un proceso estacionario.

En un proceso estacionario en sentido amplio las autocorrelaciones están dadas por:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad k \geq 0$$

además  $\gamma_k = \gamma_{-k}$ ,  $\therefore \rho_k = \rho_{-k}$

Cuando el proceso estocástico es estacionario en sentido amplio, pueden estimarse los parámetros  $\mu$ ,  $\gamma_0$ ,  $\gamma_1$ ,  $\gamma_2, \dots$  a partir de una sola realización. Si se dispone de una muestra  $Y_1, Y_2, \dots, Y_T$  de tal proceso estacionario se utilizan los siguientes estimadores:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T Y_t$$

$$\hat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^T (Y_t - \hat{\mu})^2$$

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (Y_{t+k} - \hat{\mu})(Y_t - \hat{\mu})$$



### 2.3 Tipos de Procesos Lineales

Un modelo estocástico lineal general es descrito por suposiciones de una serie de tiempo a ser generadas por una agregación lineal de impactos aleatorios. Para representaciones prácticas es adecuado emplear modelos que usen pocos parámetros. Estos modelos a veces pueden lograrse mediante representaciones de procesos lineales en términos de un número de parámetros autorregresivos y de medias móviles.

Los procesos estocásticos consisten en combinaciones lineales de variables aleatorias.

- 1) Proceso puramente aleatorio (Ruido blanco):

$$Y_t = \varepsilon_t$$

$$\text{donde } E(\varepsilon_t) = 0 \quad \forall t$$

$$E(\varepsilon_t^2) = \sigma^2 \quad \forall t$$

$$E(\varepsilon_t, \varepsilon_{t'}) = 0 \quad t \neq t'$$

- 2) Proceso autoregresivo de orden p: AR(p)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

- 3) Procesos de medias móviles de orden q: MA(q)

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

4) Procesos ARMA(p,q) es una combinación de los dos anteriores

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

### Operador de retardos L

$$LY_t = Y_{t-1}$$

$$\text{Entonces } L^2 Y_t = L(LY_t) = L(Y_{t-1}) = Y_{t-2}$$

$$\text{En general: } L^K Y_t = Y_{t-K}$$

$$\text{Además } \Delta = 1 - L$$

$$\text{Ej: } \Delta Y_t = Y_t - Y_{t-1} = Y_t - LY_t = (1 - L) Y_t$$

El operador L se puede usar para expresar un modelo con retardos:

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} = \varepsilon_t$$

Equivale a:

$$\underbrace{[1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p]}_{\text{Operador polinomial de retardos: } \phi(L)} Y_t = \varepsilon_t$$

$$\phi(L)Y_t = \varepsilon_t$$

Si  $\phi$  se iguala a cero se tiene la ecuación polinomial

### 2.3.1 Ecuaciones de Primer orden

Se tiene la ecuación homogénea de la forma:

$$\phi(L)Y_t = 0$$

Donde  $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$

Entonces en una ecuación de primer orden

$$Y_t - \phi_1 Y_{t-1} = 0$$

Asumimos que la solución es del tipo  $\lambda^t$ :

$$\lambda^t - \phi_1 \lambda^{t-1} = 0$$

$$\lambda - \phi_1 = 0 \quad \text{entonces } \lambda = \phi_1$$

Por tanto la solución de la ecuación homogénea es:

$$Y_t = \alpha \lambda^t = \alpha \phi_1^t \quad \text{con } \alpha \text{ constante}$$

$$\text{si } t = 0: \quad Y_0 = \alpha$$

y  $Y_t = Y_0 \phi_1^t$  describe la trayectoria temporal de  $Y_t$  a partir de su valor inicial dado.

Una condición suficiente y necesaria para que el proceso estocástico sea estacionario es que  $|\phi_1| < 1$ .

Puesto que

Si  $0 < \phi_1 < 1$  se tiene un decrecimiento exponencial.

Si  $-1 < \phi_1 < 0$  se tiene un decrecimiento exponencial alternado.

Alternativamente podemos usar la ecuación polinomial de retardos

$$1 - \phi_1 L = 0$$

entonces,  $L_1 = 1 / \phi_1$  es la raíz de dicho polinomio.

Y  $|L_1| > 1$  para que se cumpla la condición de estacionariedad.

Por lo tanto:  $|1 / \phi_1| > 1$ , y se tiene que  $|\phi_1| < 1$ .

### 2.3.2 Ecuaciones de Segundo orden

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} = 0$$

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = 0$$

La ecuación característica es:  $\lambda^2 - \phi_1 \lambda - \phi_2 = 0$

$$\lambda_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 - 4\phi_2}}{2}$$

En una ecuación homogénea de segundo orden, cuando las raíces son reales y diferentes, la condición de estacionariedad implica que

$|\lambda_1| < 1$  y  $|\lambda_2| < 1$ , si es raíz única debe verificarse que  $|\lambda_u| < 1$ .

Se puede demostrar que si las raíces caen fuera del círculo unidad se tiene que:

$$\phi_1 + \phi_2 < 1$$

$$-\phi_1 + \phi_2 < 1$$

$$\phi_2 > -1$$

Como alternativa a la ecuación característica y usando la ecuación polinomial de retardos, se tiene:

$$1 - \phi_1 L - \phi_2 L^2 = 0$$

$$L_1 = 1/\lambda_1$$

$$L_2 = 1/\lambda_2$$

En este caso las raíces  $L_1$  y  $L_2$  deben caer fuera del círculo unidad.

## 2.4 Modelos Lineales

### 2.4.1 Modelos Autorregresivos (AR)

Un modelo AR(p) es de la forma:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \dots - \phi_p Y_{t-p} = \varepsilon_t$$

$$\text{ó } \phi(L) Y_t = \varepsilon_t \quad \text{donde } \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

#### Modelo AR(1)

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

$$\text{ó } (1 - \phi_1 L) Y_t = \varepsilon_t$$

Para que el proceso sea estacionario la raíz del polinomio  $1 - \phi_1 L$  debe caer fuera del círculo unidad, es decir:

$$|L| = \left| \frac{1}{\phi_1} \right| > 1$$

$$|\phi_1| < 1$$

Supongamos que el proceso se inicia en el período  $-N$ :

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

$$Y_t = \phi_1^2 Y_{t-2} + \phi_1 \varepsilon_{t-1} + \varepsilon_t$$

...

$$Y_t = \sum_{j=0}^{t+N-1} (\phi_1^j \varepsilon_{t-j}) + \phi_1^{t+N} Y_{-N}$$

Tomando esperanzas:

$$E[Y_t] = \phi_1^{t+N} E[Y_{-N}]$$

Si consideramos que  $Y_{-N}$  (el primer valor) es dado:

$$E[Y_t] = \phi_1^{t+N} Y_{-N}$$

$$\text{Si } |\phi_1| > 1: E[Y_t] \xrightarrow{t \rightarrow \infty} \infty$$

$$\text{Si } |\phi_1| < 1: E[Y_t] \xrightarrow{t \rightarrow \infty} 0$$

Si  $t \rightarrow \infty$ , sino que  $t$  es el tiempo actual y  $-N \rightarrow \infty$

Si se considera que el proceso inicia en  $-\infty$  y  $|\phi_1| < 1$ , entonces

$E[Y_t] = 0$  independiente del valor inicial.

La *varianza* será:

$$\gamma_0 = E[(Y_t - E(Y_t))^2] = E[(Y_t - \phi_1^{t+N} Y_{-N})^2], \text{ resolviendo:}$$

$$\gamma_0 = \sigma_\varepsilon^2 \frac{1 - \phi_1^{2(t+N-1)}}{1 - \phi_1^2},$$

para  $\phi_1 = 1$  y  $|\phi_1| > 1$

Se hace  $\infty$  cuando el proceso arranca desde  $-\infty$ , es decir el proceso es de carácter explosivo.

Si  $|\phi_1| < 1$  y se inicia en  $-\infty$

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}$$

En lo que sigue y mientras no se diga lo contrario, se supondrá que el proceso arranca desde  $-\infty$  y que se cumple la condición de estacionariedad.

Las *autocovarianzas*:

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

Multiplicando por  $Y_{t-\tau}$  y tomando esperanzas se tiene:

$$E[Y_t Y_{t-\tau}] = \phi_1 E[Y_t Y_{t-\tau}] + E[\varepsilon_t Y_{t-\tau}]$$

Si  $\tau > 0$

$$\gamma_\tau = \gamma_0 \phi_1^\tau$$

$$\rho_\tau = \phi_1^\tau$$

**Modelo AR(2)**

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \quad (1)$$

$$\text{ó } (1 - \phi_1 L - \phi_2 L^2) Y_t = \varepsilon_t$$

Para que el proceso sea estacionario se requiere que las raíces de la ecuación  $1 - \phi_1 L - \phi_2 L^2$  estén fuera del círculo unidad.

Si se cumplen las condiciones de estacionariedad:  $E[Y_t] = 0$

Multiplicando a la ecuación (1) por  $Y_{t-\tau}$  y tomando esperanzas:

$$\text{Si } \tau > 0 \quad \gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2}$$

Dividiendo para  $\gamma_0$ :

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2}$$

Para determinar  $\phi_1$  y  $\phi_2$ :

Sea  $\tau=1$  y  $\tau=2$ , se obtiene:

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

$$\phi_1 = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2}$$

$$\phi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$



**Modelo AR(p)**

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (2)$$

$$\text{ó} \quad \phi(L)Y_t = \varepsilon_t$$

$$\text{con } \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

Para que el proceso sea estacionario se requiere que las raíces de la ecuación  $\phi(L) = 0$  estén fuera del círculo unidad.

Si se multiplica la ecuación (2) por  $Y_{t-\tau}$  y tomando esperanzas:

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \dots + \phi_p \gamma_{\tau-p} + E[\varepsilon_t Y_{t-\tau}]$$

para  $t = 0$

$$\gamma_0 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_\varepsilon^2$$

para  $\tau > 0$

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \dots + \phi_p \gamma_{\tau-p} \quad (2)$$

Si se divide por  $\gamma_0$ :

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \dots + \phi_p \rho_{\tau-p} \quad (3)$$

Tomando  $\rho_0, \rho_1, \dots, \rho_{p-1}$  como condiciones iniciales determinadas a partir de  $\phi_1, \phi_2, \dots, \phi_p$  (conocidos), la solución de (4) permite calcular  $\rho_\tau$  para todo  $\tau \geq p$ .

A la inversa, si se conoce  $\rho_1, \rho_2, \dots, \rho_p$ , utilizando (4) para  $\tau=1, 2, \dots, p$ , se puede calcular  $\phi_1, \phi_2, \dots, \phi_p$ :

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} \\ &\dots \\ \rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \dots + \phi_p\end{aligned}$$

Matricialmente:

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_p \end{bmatrix} \quad \left\langle \begin{array}{l} \text{Ecuaciones de} \\ \text{Yule - Walker} \end{array} \right\rangle$$

#### 2.4.2 Modelos de Medias Móviles MA

Un modelo de medias móviles de orden q se define por:

$$Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}$$

Sea el operador de retardos:

$$\theta(L) = 1 - \theta_1L - \theta_2L^2 - \dots - \theta_qL^q$$

este modelo se expresa:  $Y_t = \theta(L)\varepsilon_t$

Se tiene que :  $E[Y_t] = 0$ .

**Modelo MA(1)**

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

ó

$$Y_t = (1 - \theta_1 L) \varepsilon_t$$

Si se multiplica por  $Y_{t-\tau}$  y se toma esperanzas:

$$E[Y_t Y_{t-\tau}] = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-\tau} - \theta_1 \varepsilon_{t-\tau-1})]$$

Para  $\tau = 0$  y puesto que  $E[\varepsilon_t \varepsilon_t'] = \sigma_\varepsilon^2$   $t \neq t'$ :

$$\gamma_0 = E[Y_t^2] = E[\varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 - 2\theta_1 \varepsilon_t \varepsilon_{t-1}] = [1 + \theta_1^2] \sigma_\varepsilon^2 \quad (5)$$

Para  $\tau = 1$ :

$$\gamma_1 = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} - \theta_1 \varepsilon_{t-2})]$$

$$\gamma_1 = -\theta_1 \sigma_\varepsilon^2$$

Para  $\tau > 1$ :

$$\gamma_\tau = 0$$

Dividiendo para  $\gamma_0$ , se tiene:

$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2} \quad \tau = 1$$

$$\rho_\tau = 0 \quad \tau > 1$$

De la ecuación (5) se deduce que un modelo MA(1) es siempre estacionario con independencia del valor de  $\theta_1$ .

Si al modelo  $Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$  lo expresamos como:

$$\varepsilon_t - \theta_1 \varepsilon_{t-1} = Y_t$$

$$(1 - \theta_1 L) \varepsilon_t = Y_t$$

$$\varepsilon_t = \frac{1}{1 - \theta_1 L} Y_t$$

$$\varepsilon_t = \sum_{j=0}^{\infty} (\theta_1 L)^j Y_t$$

Vemos que es equivalente a un proceso  $AR(\infty)$ , la condición que permite pasar de un modelo a otro, (es decir  $|\theta_1| < 1$ ) se denomina **condición de invertibilidad**.

### Modelo MA(2):

Un modelo MA (2) se define por:

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

Para

$$\tau = 0 \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2) \sigma_\varepsilon^2$$

$$\rho_1 = \frac{-\theta_1 + \theta_1 \theta_2}{1 + \theta_1^2 + \theta_2^2}$$

$$\tau = 1 \quad \gamma_1 = (-\theta_1 + \theta_1 \theta_2) \sigma_\varepsilon^2$$

$$\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

$$\rho_\tau = 0, \tau > 2$$

$$\tau = 2 \quad \gamma_2 = (-\theta_2) \sigma_\varepsilon^2$$

$$\tau > 2 \quad \gamma_\tau = 0$$

**Modelo MA (q)**

$$Y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

Si se multiplica ambos lados por  $Y_{t-\tau}$  y se toma esperanzas:

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_\varepsilon^2$$

$$\gamma_\tau = \begin{cases} (-\theta_\tau + \theta_1 \theta_{\tau+1} + \dots + \theta_{q-\tau} \theta_q) \sigma_\varepsilon^2 & \text{Si } \tau = 1, 2, \dots, q \\ 0 & \text{Si } \tau > q \end{cases}$$

De aquí que:

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1 \theta_{\tau+1} + \dots + \theta_{q-\tau} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & \text{Si } \tau = 1, 2, \dots, q \\ 0 & \text{Si } \tau > q \end{cases}$$

**2.4.3 Modelos Mixtos Autorregresivos Medias Móviles (ARMA)**

Un modelo ARMA(p,q) se define de la siguiente forma:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

ó

$$\phi(L) Y_t = \theta(L) \varepsilon_t$$

$$\text{donde } \begin{cases} \phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \\ \theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q \end{cases}$$

Para que el modelo sea estacionario se requiere que las raíces de  $\phi(L) = 0$  caigan fuera del círculo unidad.

Si se cumplen las condiciones de estacionariedad, el modelo ARMA(p,q) se puede expresar como un MA( $\infty$ ):

$$Y_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t = \Psi(L) \varepsilon_t$$

$$\phi(L) \Psi(L) \equiv \theta(L)$$

Es decir, en el modelo ARMA(p,q) la media es cero, si se añade al segundo miembro un término constante  $\delta$ , la media del proceso está dada por:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \delta + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

Es decir:

$$\phi(L) Y_t = \delta + \theta(L) \varepsilon_t$$

$$E[\phi(L) Y_t] = E[\delta + \theta(L) \varepsilon_t]$$

$$\phi(L) E[Y_t] = \delta + \theta(L) E[\varepsilon_t]$$

$$E[Y_t] = \left( \frac{1}{\phi(L)} \right)^\delta$$

Cabe mencionar las siguientes observaciones:

a.- Puesto que es estacionario se tiene que:  $\forall t : E[Y_t] = \mu$

b.- El operador  $L$  al ser aplicado a una constante ( $\delta$ ) nos da el valor de la constante, es decir  $L(\delta) = \delta$ . Entonces  $L^2(\delta) = \delta$

$$(1 - \phi_1 L - \phi_2 L^2) \delta = 1 - \phi_1 \delta - \phi_2 \delta,$$

el operador  $\left( \frac{1}{\phi(L)} \right)$  aplicado a  $\delta$  es igual a  $\left( \frac{1}{1 - \phi_1 - \dots - \phi_p} \right)$  por  $\delta$

$$\text{Entonces } E[Y_t] = \mu = \frac{1}{1 - \phi_1 - \dots - \phi_p}$$

### **Modelo ARMA(1,1)**

Un modelo ARMA(1,1) es de la forma:

$$Y_t - \phi_1 Y_{t-1} = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Es decir:

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Si se multiplica por  $Y_{t-\tau}$  y se toma esperanzas:

$$E[Y_t Y_{t-\tau}] = E[\phi_1 Y_{t-1} Y_{t-\tau} + \varepsilon_t Y_{t-\tau} - \theta_1 \varepsilon_{t-1} Y_{t-\tau}]$$

$$\gamma_\tau = \phi_1 E[Y_t Y_{t-\tau}] + E[\varepsilon_t Y_{t-\tau}] - \theta_1 E[\varepsilon_{t-1} Y_{t-\tau}]$$

Puesto que:

$$E[\varepsilon_t Y_t] = \sigma_\varepsilon^2, \text{ y que } E[\varepsilon_{t-1} Y_t] = (\phi_1 - \theta_1) \sigma_\varepsilon^2$$

Se tiene:

$$\tau = 0$$

$$\gamma_0 = \phi_1 \gamma_1 + E[\varepsilon_t Y_t] - \theta_1 E[\varepsilon_{t-1} Y_t]$$

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_\varepsilon^2 - \theta_1(\phi_1 - \theta_1) \sigma_\varepsilon^2$$

$$\tau = 1$$

$$\gamma_1 = \phi_1 \gamma_0 + E[\varepsilon_t Y_{t-1}] - \theta_1 E[\varepsilon_{t-1} Y_{t-1}]$$

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2$$

$$\tau > 1$$

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + E[\varepsilon_t Y_{t-\tau}] - \theta_1 E[\varepsilon_{t-1} Y_{t-\tau}]$$

$$\gamma_\tau = \phi_1 \gamma_{\tau-1}$$

Reemplazando  $\gamma_1$  en  $\gamma_0$ :

$$\gamma_0 = \phi_1 [\phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2] + \sigma_\varepsilon^2 - \theta_1 (\phi_1 - \theta_1) \sigma_\varepsilon^2$$

despejando

$$\gamma_0 = \frac{1 - 2\theta_1\phi_1 + \theta_1^2}{1 - \phi_1^2} \sigma_\varepsilon^2$$

a su vez reemplazando éste valor en  $\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2$ , tenemos:

$$\gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_\varepsilon^2$$

De aquí:

$$\rho_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - 2\theta_1\phi_1 + \theta_1^2} \sigma_\varepsilon^2$$

$$\rho_\tau = \phi_1 \rho_{\tau-1} \quad \tau > 1$$



**Modelo ARMA(p,q)**

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

Multiplicando por  $Y_{t-\tau}$  y tomando esperanzas:

$$\gamma_\tau - \phi_1 \gamma_{\tau-1} - \dots - \phi_p \gamma_{\tau-p} = E[\varepsilon_t Y_{t-\tau}] - \theta_1 E[\varepsilon_{t-1} Y_{t-\tau}] - \dots - \theta_q E[\varepsilon_{t-q} Y_{t-\tau}]$$

puesto que  $E[\varepsilon_t Y_{t'}] = 0$  para  $t' < t$  tenemos:

$$\gamma_\tau - \phi_1 \gamma_{\tau-1} - \dots - \phi_p \gamma_{\tau-p} = 0 \quad \text{para } \tau > q$$

de esta ecuación en diferencias homogéneas se obtiene la solución que permite calcular las autocovarianzas a partir de  $\tau > q$ .

Si se divide para  $\gamma_0$ :

$$\rho_\tau - \phi_1 \rho_{\tau-1} - \dots - \phi_p \rho_{\tau-p} = 0 \quad \text{para } \tau > q$$

Para calcular los  $q$  primeros valores de  $\rho_\tau$  se utiliza la parte de medias móviles del modelo.

En los modelos ARMA (p,q) es conveniente factorizar la parte autorregresiva y de medias móviles, para ver si existen raíces comunes.

Sean  $\lambda_1, \dots, \lambda_p$  las raíces de la ecuación:

$$\lambda^p - \phi_1 \lambda^{p-1} - \dots - \phi_p = 0$$

y sean  $\delta_1, \dots, \delta_q$  las raíces de la ecuación:

$$\delta^q - \theta_1 \delta^{q-1} - \dots - \theta_q = 0$$

Entonces el modelo se puede expresar:

$$(1 - \lambda_1 L)(1 - \lambda_2 L) \dots (1 - \lambda_p L) Y_t = (1 - \delta_1 L) \dots (1 - \delta_q L) \varepsilon_t$$

En caso de existir alguna raíz repetida (por ejemplo:  $\lambda_1 = \lambda_q$ ) el modelo estaría sobreparametrizado innecesariamente, pues un modelo ARMA(p-1, q-1) tendría las mismas propiedades.

## 2.5 Procesos no estacionarios

Muchas de las series económicas son no estacionarias y para su tratamiento es necesario aplicar transformaciones que las hagan estacionarias.

Consideremos el modelo:

$$Y_t = Y_{t-1} + \varepsilon_t$$

No es estacionario, pues  $|\phi_1| = 1$ .

A este modelo se lo conoce como ***caminata aleatoria***.

Si el proceso comienza en un pasado remoto, mediante sustituciones sucesivas se tiene:

$$Y_t = Y_{t-1} + \varepsilon_t$$

$$Y_t = (Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$Y_t = Y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

...

$$Y_t = \sum_{j=0}^{\infty} \varepsilon_{t-j}$$

Su varianza es infinita, por lo que el proceso no es estacionario.

$$\text{Var}(Y_t) = \text{Var}(\varepsilon_t) + \text{Var}(\varepsilon_{t-1}) + \dots$$

Si suponemos que el proceso comienza en  $-N$ , la varianza para el período  $t$  viene dada por:

$$\gamma_{0,t} = (t+N) \sigma_\varepsilon^2$$

como se ve, el proceso tampoco es estacionario pues la varianza toma un valor distinto para cada valor de  $t$ .

Idénticamente las autocovarianzas están dadas por:

$$\gamma_{\tau,t} = (t+N-\tau) \sigma_\varepsilon^2$$

El coeficiente de autocorrelación:

$$\rho_{\tau,t} = \frac{\gamma_{\tau,t}}{\sqrt{\gamma_{0,t}} \sqrt{\gamma_{0,t-\tau}}} = \frac{t+N-\tau}{\sqrt{t+N} \sqrt{t+N-\tau}} = \sqrt{\frac{t+N-\tau}{t+N}}$$

$\gamma_{\tau,t}$  y  $\rho_{\tau,t}$  dependen de  $t$ , cuando su valor sólo debe depender de  $\tau$ .

Sin embargo el proceso  $Y_t = Y_{t-1} + \varepsilon_t$  puede transformarse en estacionario:

Sea

$$W_t = Y_t - Y_{t-1} = \Delta Y_t$$

Es decir:

$$W_t = \varepsilon_t$$

El proceso transformado  $W_t$ , al tomar primeras diferencias de  $Y$  es un ruido blanco y por lo tanto es estacionario.

Para obtener  $Y_t$  a partir de  $W_t$ :

$$Y_t = W_t + Y_{t-1}$$

$$Y_t = W_t + W_{t-1} + Y_{t-2}$$

...

$$Y_t = W_t + W_{t-1} + W_{t-2} + \dots$$

Es decir el proceso  $Y_t$  se obtiene sumando (o integrando) el proceso estacionario  $W_t$ .

Por esta razón se dice que el proceso  $Y_t$  (llamado caminata aleatoria), pertenece a la clase de modelos integrados. Esta clase está conformada por todos los procesos no estacionarios que se pueden transformar en estacionarios mediante la forma de diferencias de un determinado orden.

Los **modelos integrados** son aquellos que se obtienen por medio de suma o integración de un proceso estacionario.

### 2.5.1 Modelos ARIMA

A un proceso integrado  $Y_t$  se le denomina **proceso ARIMA**  $(p,d,q)$  si tomando diferencias de orden  $d$  se obtiene un proceso estacionario  $W_t$  del tipo ARMA(p,q).

Sea  $Y_t$  el proceso original (no estacionario), tomando diferencias de orden  $d$  se tiene el modelo  $W_t$  estacionario ARMA(p,q):

$$W_t = \Delta^d Y_t$$

$$(1 - \phi_1 L - \dots - \phi_p L^p) W_t = (1 - \phi_1 L - \dots - \phi_p L^p) \varepsilon_t$$

Reemplazando:

$$(1 - \phi_1 L - \dots - \phi_p L^p) (1 - L)^d Y_t = (1 - \phi_1 L - \dots - \phi_p L^p) \varepsilon_t$$

Entonces un modelo ARIMA(p,d,q) puede representarse por:

$$\phi(L) (1 - L)^d Y_t = \theta(L) \varepsilon_t$$

Una clase considerable amplia de series se puede modelizar por medio de los procesos ARIMA, gracias a que aplicando

transformaciones de tipo no lineal muchas series pasan a ser representables por medio de estos modelos, por ejemplo hay series en las que a lo largo de un período extenso de tiempo la varianza puede estar afectada por una tendencia, la cual no desaparece al tomar diferencias. Si se presenta esta característica la transformación adecuada puede consistir en la toma de logaritmos.

## 2.6 Modelos estacionarios multiplicativos (SARIMA)

En algunas series se observa que ocurren dos tipos de relaciones; una entre las observaciones para meses sucesivos de un año en particular y otra entre las observaciones para el mismo mes en años sucesivos. Por ejemplo, puede suceder que los datos relativos al mismo mes de diferentes años tiendan a situarse en forma análoga con respecto a la media anual. Esto conduce a pensar que sería adecuado hacer intervenir retardos múltiples de doce en un modelo ARIMA.

Teóricamente se tendrían que poner valores muy grandes de  $p$  y  $q$  para que estos retardos sean tomados en cuenta; pero esto nos llevaría a estimar un número demasiado grande de parámetros. Para evitar esto, Box y Jenkins proponen un tipo particular de modelos **ARIMA estacionarios**, los modelos multiplicativos de tipo

$$\nabla^d \phi_p(B) (\nabla^S)^D \phi_p(B^S) Y_t = \Theta_q(B) \Theta_Q(B^S) \varepsilon_t \quad (6)$$

donde S es el periodo de la estacionalidad (S = 12 para series anuales, S = 4 para series trimestrales, etc);

$$\nabla = 1 - B, \quad \nabla^S = 1 - B^S$$

$\phi_p, \phi_P, \Theta_q, \Theta_Q$  son polinomios de grado p, P, q, Q con raíces de módulo superior a uno;  $\varepsilon_t$  es un ruido blanco.

Un proceso  $Y_t$  que satisface (6) se llama:

$$\text{SARIMA}_S [(p,d,q), (P,D,Q)]$$

El razonamiento que conduce a obtener un modelo SARIMA, consiste en aplicar en s series, obtenidas a partir de  $Y_t$ , poniendo para cada serie los meses idénticos, la misma transformación:

$$\frac{(\nabla^S)^D \phi_p(B^S)}{\Theta_p(B^S)}$$

y suponer que la serie obtenida: 
$$\alpha_t = \frac{(\nabla^S)^D \phi_p(B^S)}{\Theta_p(B^S)} \quad (7)$$

no tiene estacionalidades y es entonces modelizable por un ARIMA (p,d,q):

$$\nabla^d \phi_p(B) \alpha_t = \Theta_q(B) \varepsilon_t \quad (8)$$

De la combinación de las ecuaciones (7) y (8) se obtiene (6).

Las series estacionales pueden detectarse analizando las funciones de autocorrelación y de autocorrelación parcial estimadas, pues ellas presentan grandes valores en módulo para los índices múltiplos de  $s$ . La identificación de los parámetros  $P$ ,  $D$ ,  $Q$  de los factores estacionales se hace de manera análoga al procedimiento descrito en el capítulo 3 para los parámetros  $p, d, y q$ , de los procesos ARIMA. Luego se procede a la fase de verificación de los coeficientes y de los ruidos blancos, lo que lleva aceptar el modelo propuesto o, alternativamente, sugiere la manera de mejorarlo.

Una vez que se tiene las bases teóricas de los modelos a utilizar, se procede al desarrollo del análisis de la serie en estudio.



# Capítulo 3

## 3. ANALISIS DE SERIES DE TIEMPO

### 3.1 Método de Box y Jenkins

La búsqueda de un modelo ARIMA compatible con los datos, se hace de una manera secuencial que se puede describir así:

1. Se comienza por buscar, a partir de las observaciones, los valores plausibles para  $(p,d,q)$ . Esta etapa constituye la fase de identificación a priori del modelo. Se requiere de muchos cálculos, y en general no se obtiene una sola tripleta de valores  $(p,d,q)$ . Esto se explica porque un proceso estacionario puede aproximarse, tanto como se quiera, por un proceso ARMA( $p,q$ ), donde el valor de  $p$  puede ser cualquiera.
2. Para cada tripleta obtenida de valores de  $(p,d,q)$ , se procede a la fase de estimación de los coeficientes de los polinomios autoregresivo y media móvil.

Las siguientes etapas del modelo constituyen la fase de identificación a posteriori, es decir después de la estimación.

3. Al final de la etapa 2., se dispone de varios modelos estimados. Estos deben de someterse a varias pruebas estadísticas (pruebas de significación para los coeficientes, pruebas concernientes a las no correlaciones de los ruidos blancos, etc), para así verificar los resultados obtenidos con las hipótesis hechas. Esta fase se la llama fase de verificación y aquí se puede llegar a la conclusión que todos los modelos estimados sean rechazados.

4. Puede darse el caso que varios modelos pasen la fase de verificación. En este caso se escoge el método que tenga el mayor poder predictivo o la mayor cantidad de información.

### 3.1.1 Elección de $D$

Si el proceso  $(Y_t; t \in Z)$ , generado por las observaciones, es un ARIMA con  $d > 0$ , es evidente que este proceso no es un proceso estacionario, luego su función de autocorrelación depende de  $h$  y  $t$ , así:

$$p(h, t) = \frac{Cov(Y_t, Y_{t+h})}{\sqrt{Var(Y_t)}\sqrt{Var(Y_{t+h})}}$$

Se puede demostrar además que la función  $p(h, t)$ , tiende a 1 cuando  $t$  tiende a  $+\infty$ .

Si la función  $p(h)$  queda relativamente próxima a uno para un gran número de valores de  $h$ , se puede entonces pensar que es necesario diferenciar la serie para volverla estacionaria. En la práctica, sin embargo, el criterio de proximidad a 1 de los primeros valores de la función  $p(h)$  se reemplaza por el de la proximidad entre sí de los primeros valores de esta función, aún cuando el valor  $p(h)$  sea muy diferente de uno.

A la serie diferenciada una vez se le aplica nuevamente este criterio para ver si es conveniente diferenciar una segunda vez, y así sucesivamente. Para las series económicas los valores de  $d$  que se obtienen en general son 0, 1 y 2; los valores superiores son muy raros.

### 3.1.2 Elección de $P Y Q$

Para la elección de  $p$  y  $q$  se analiza esencialmente la forma de la función de autocorrelación estimada  $\hat{p}(h)$  y la función de autocorrelación parcial estimada  $\hat{r}(h)$  de la serie diferenciada  $d$  veces.

Si la función  $\hat{p}(h)$  tiende hacia cero para  $h > q$ , se puede pensar que se trata de una MA(q); y si la función  $\hat{r}(h)$  tiende a cero para valores de  $h > p$  se puede pensar que se trata de un AR(p).

### 3.1.3 Fase de verificación

Una vez realizadas las estimaciones de los modelos procedemos a efectuar las fases de verificación y de elección del mejor modelo.

Las pruebas estadísticas utilizadas en esta fase son de dos tipos: las pruebas concernientes a los parámetros  $\phi_j$  y  $\theta_j$  de los procesos autorregresivo y media móvil, respectivamente, del modelo propuesto ARMA(p,q), de la serie diferenciada  $d$  veces, y las pruebas concernientes a las hipótesis del ruido blanco  $\varepsilon_t$ .

#### Pruebas concernientes a los parámetros

Nos interesa saber si es posible disminuir el número de parámetros; es decir, si disponemos de un modelo ARMA (p, q), deseamos saber si es posible formular un modelo ARMA(p-1, q) o un modelo ARMA (p, q-1). Esto quiere decir, que queremos examinar el nivel de significación del coeficiente  $\phi_p$  o de  $\theta_q$ , para

lo cual utilizamos una prueba T –Student (bajo hipótesis de normalidad del ruido blanco).

Sea  $\hat{\phi}_p$  el estimador de  $\phi_p$  en la formalización ARMA (p,q) y  $Var(\hat{\phi}_p)$ , su varianza estimada. Se aceptará con riesgo ( $\alpha$ ) del 5%, la hipótesis de que el modelo es ARMA (p-1, q) si  $|\hat{\phi}_p| / (Var(\hat{\phi}_p))^{1/2} < 1.96$ ; de lo contrario se acepta la modelización ARMA(p, q).

De manera similar se procede para examinar el nivel de significación del parámetro  $\theta_q$ . También se puede efectuar esta prueba para examinar el nivel de significación de un parámetro  $\phi_{p+1}$  o  $\theta_{q+1}$ ; no es conveniente probar la hipótesis de que un modelo sea ARMA (p+1, q+1), o un ARMA(p-1, q-1).

### **Pruebas concernientes al ruido blanco**

Se pretende verificar si la serie de residuos:

$$\varepsilon_t = \frac{\hat{\Phi}(B)}{\hat{\Theta}(B)} Y_t$$

es coherente con la hipótesis de que los  $\varepsilon_t$  forman un ruido blanco; esto es, verificar si los residuos son independientes, centrados, de varianza  $\sigma^2$  y gaussianos. Estas pruebas son ampliamente conocidas, salvo la de independencia que requiere ciertas precisiones que se verán a continuación.

➤ **Prueba de Box-Pierce (“Overcoat”)**

La prueba se basa en que el estadístico:

$$Q = N \sum_{h=1}^k \hat{\rho}_h[\varepsilon]$$

sigue asintóticamente una ley  $\chi^2$  con  $k-p-q$  grados de libertad, bajo la hipótesis de independencia de los  $\varepsilon_t$  ( $\hat{\rho}_h[\varepsilon]$  representa la correlación estimada de orden  $h$ , de los residuos). Luego, se rechaza la hipótesis de independencia de los ruidos blancos al nivel  $\alpha$ , si  $Q > \chi^2_{1-\alpha}[k-p-q]$ . El número  $k$  debe ser suficientemente grande; habitualmente se lo toma entre 15 y 20. Cuando la hipótesis se declara inadecuada, se puede utilizar la función de autocorrelación estimada de los residuos, para ver en que forma se modifica el modelo.

➤ **Periodograma acumulativo**

El periodograma está diseñado para detectar patrones periódicos no aleatorios en un ruido blanco.

El espectro potencia  $p(f)$  para el ruido blanco tiene un valor constante de  $2\sigma_a^2$  sobre un dominio de frecuencias de  $0 - 0.5$  ciclos. En consecuencia, el espectro acumulativo del ruido blanco:

$$P(f) = \int_0^f p(g)dg$$

graficando para los valores de  $f$ , dan una Línea recta que va de  $(0.0)$  a  $(0.5, \sigma_a^2)$ . Esto es,  $P(f)/\sigma_a^2$  es una línea recta que va desde  $(0.0)$  a  $(0.5, 1)$ .

Si el modelo es adecuado, el gráfico del periodograma acumulativo muestra puntos alrededor de la recta, en caso contrario, el periodograma acumulativo mostraría desviaciones sistemáticas de esa línea. En particular, periodicidades en los ruidos blancos tenderían a producir ordenadas grandes que forman una especie de perturbaciones alrededor de la recta.

La relación de probabilidad entre el periodograma acumulativo y el espectro integrado es la misma que entre la función de frecuencia acumulativa empírica y la función de distribución

acumulativa. Por esta razón, se puede utilizar una prueba de Kolmogorov-Smirnov para localizar líneas límites alrededor de la línea teórica. Si los  $\varepsilon_t$  son ruido blanco, entonces el periodograma acumulativo se desviaría de la recta dentro de la región marcada por estas líneas límites, con un cierto margen de probabilidad.

### 3.1.4 Uso de los ruidos blancos

Cuando la función de autocorrelación de los ruidos blancos de un modelo ajustado indica un modelo inadecuado, es necesario considerar en que manera el modelo podría ser modificado.

Supongamos que los ruidos blancos  $\varepsilon_t$  del modelo:

$$\phi_0(L) \nabla^{do} Y_t = \theta_0(L) \varepsilon_t$$

parecen no verificar las hipótesis de independencia y normalidad. Utilizando las funciones de autocorrelación parcial estimadas de los  $\varepsilon_t$  y los métodos explicados anteriormente, podemos tratar de ajustar un modelo:

$$H(L) \nabla^{dr} \varepsilon_t = \phi(L) \beta_t$$

para la serie  $\beta_t$ . Eliminando  $\beta_t$  de las ecuaciones anteriores, llegamos a plantear un modelo:

$$\phi_0(L) H(L) \nabla^{do} \nabla^{dr} Y_t = \theta_0(L) \phi(L) \beta_t$$

que puede ser ajustado nuevamente y verificado.



### 3.1.5 Elección del modelo

Puede suceder que varios modelos pasen la fase de verificación. En este caso existen criterios que no poseen una base teórica sólida, pero son útiles a la hora de realizar una elección.

#### a) Criterio del mayor poder predictivo

En un modelo ARMA, el error de previsión al horizonte 1, es tal que  $V[\varepsilon_t(1)] = \sigma^2$ . Se puede entonces proponer escoger el modelo que conduzca a un error de predicción pequeño.

Algunos criterios son:

1. La estimación  $\sigma$  de la varianza
2. El coeficiente de determinación

$$R^2 = 1 - \frac{\sigma^2}{V}$$

donde  $V$ , es la varianza de la serie inicial diferenciada  $d$  veces.

3. El coeficiente de determinación modificado:

$$R^2 = 1 - \frac{\sigma^2 / (N - p - q)}{V / (N - 1)}$$

este coeficiente permite tomar en cuenta los grados de los polinomios autorregresivos y media móvil.

4. El estadístico de Fisher:

$$F = \frac{(V - \sigma^2) / pq}{\sigma^2 / (N - p - q)}$$

En el criterio 1) se trata de minimizar, mientras que los otros criterios se trata de maximizar.

#### b) Criterio de información

Otra aproximación, introducida por Akaike (1969), consiste en suponer que los modelos ARMA(p,q) dan aproximaciones de la realidad y que la verdadera ley desconocida de las observaciones  $\nabla^d Y_t$  (observaciones  $Y_t$  diferenciadas d veces) no satisface forzosamente tal modelo. Podemos entonces fundamentar la elección del modelo sobre una medida de la desviación entre la verdadera ley desconocida y el modelo propuesto. La medida habitual es la cantidad de información de Kullback. Sea  $f_0(X)$  la densidad desconocida de las observaciones,  $\{f(x), f \in F_{p,q}\}$  la familia de densidades correspondientes al modelo ARMA (p,q); la desviación entre la verdadera ley y el modelo se mide por:

$$I(f_0, F_{p,q}) = \min_{f \in F_{p,q}} \int \log \frac{f_0(x)}{f(x)} f_0(x) dx$$

Esta cantidad es siempre positiva y solo se anula si la verdadera ley  $f_0$  pertenece a  $F_{p,q}$ .  $I(f_0, F_{p,q})$ , el valor de la cantidad de información es evidentemente desconocido; pero si se dispone de un buen estimador  $\hat{I}(f_0, F_{p,q})$  de esta cantidad, será posible utilizarlo como criterio.

Se escoge entonces el modelo  $F_{p,q}$  que conduce al más pequeño valor de la estimación  $\hat{I}(f_0, F_{p,q})$ .

Los estimadores de la cantidad de información que han sido propuestos son:

1.  $AIC(p,q) = \text{Log} \sigma^2 + \frac{2(p+q)}{N}$
2.  $BIC(p,q) = \text{Log} \sigma^2 + (p+q) \frac{\text{Log} N}{N}$
3.  $AIC(p,q) = \text{Log} \sigma^2 + (p+q)c \frac{\text{Log} N}{N}$ , con  $c > 2$

El primero de estos estimadores es el más utilizado; sin embargo, solo los dos restantes son convergentes y conducen a una selección asintótica correcta del modelo.

Para hallar el modelo apropiado de las series utilizaremos el paquete estadístico STATGRAPHICS, para todos los cálculos matemáticos de series de tiempo necesarios.

### 3.2 Clasificación de las series utilizadas para el análisis

Para un análisis exhaustivo del índice de natalidad del Ecuador, se propone estudiar a la variable clasificada en los siguientes grupos:

- **Total:** Esta serie corresponde al número total de nacidos vivos en el Ecuador cada mes.
- **Hombres:** Esta serie contiene el número total de hombres nacidos vivos en el Ecuador mensualmente.
- **Mujeres:** Contiene el número de mujeres nacidas vivas en el Ecuador cada mes.
- **Por Provincias:** Corresponde al total (hombres y mujeres) de nacidos vivos de cada provincia, mensualmente.

Cada grupo va a ser analizado independientemente, es decir que se va a particularizar el estudio de tal modo que se establecerá un modelo para cada conjunto anteriormente descrito.

A estos grupos se los ha considerado como variables, las cuales representan datos mensuales que obtienen información sobre el número de nacidos vivos registrados en el país desde el año 1991 a 1997, cuyo período corresponde a este estudio. Esta información incluye algunos hechos ocurridos en años anteriores pero inscritos en el año de la investigación, estos datos constituyen las inscripciones tardías. Las investigaciones de las cuales se hace referencia son

realizadas por el Instituto Nacional de Estadísticas y Censos (INEC), es esta fuente de donde se ha obtenido todos los datos necesarios para este estudio.

Esta información es recolectada en forma directa, a través de las oficinas del Registro Civil, Identificación y Cedulación, y del Ministerio de Salud, entidades que funcionan en las cabeceras provinciales, cantonales y parroquiales de la República que se encargan de registrar y legalizar al hecho vital.

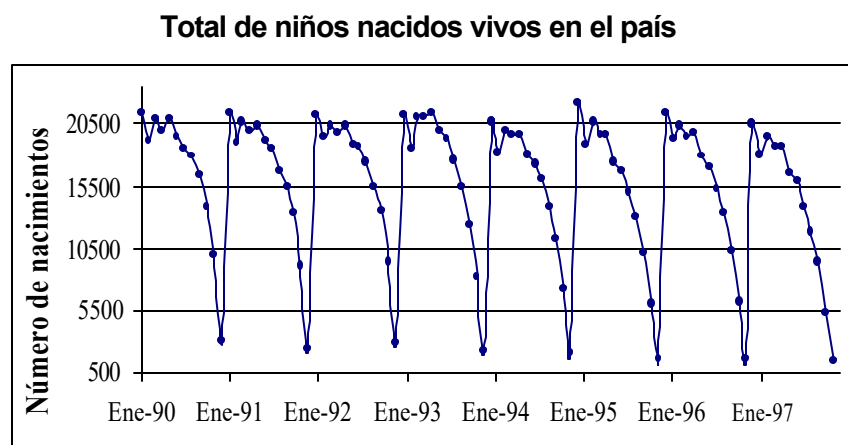
La investigación se efectúa en el momento que se realiza la inscripción de los hechos y la recolección se la hace mensualmente para luego ser publicado en forma anual.

En el momento en que fue realizada la recolección de datos para este estudio, estuvieron a disposición datos hasta el año 1997, es por este motivo que sólo hasta esta fecha se ha considerado el período de estudio.

### 3.3 Análisis de la Serie del Total de nacidos vivos en el país

Siguiendo los pasos del método de Box y Jenkins, el cual fue explicado anteriormente, se procede en primer lugar a la fase de identificación de los parámetros, por lo que en la figura 3.1 se muestra el gráfico de la serie original del Total.

**Figura 3.1**



En este gráfico se puede apreciar claramente que el comportamiento de la serie se repite cada 12 datos, es nuestro caso particular esto quiere decir cada 12 meses. Entonces se puede afirmar que existe una estacionalidad de 12, motivo por el cual se trabajará con los modelos SARIMA  $_{12}(p,d,q) \times (P,D,Q)$ .

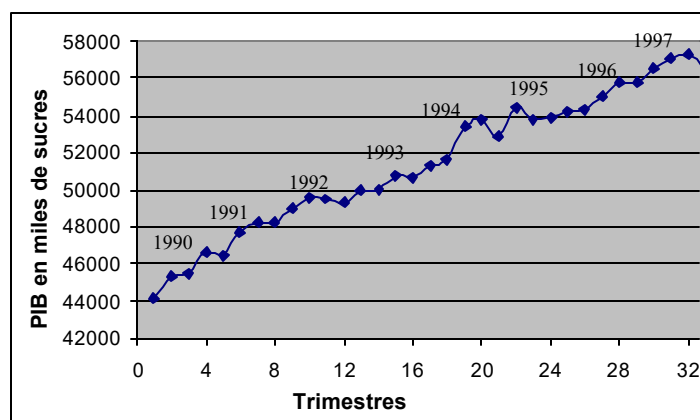
Como se puede observar, en enero del año 1995 está el valor mas alto de la serie, lo mismo ocurre en los demás meses de enero de todos los años, esto podría llevarnos a pensar que en este mes ocurre

mas nacimientos que en otros, pero cabe recalcar que los últimos nacimientos del mes de diciembre se los ingresa en forma tardía, pues en la mayoría de los casos éstos son registrados a principios del mes siguiente, es decir en enero.

Pero no sólo en el mes de enero de 1995 está el mayor número de nacimientos, sino que también este fenómeno se lo encuentra en sus primeros meses. Cabe indicar que unos meses atrás el país tuvo una recuperación de tipo económica, puesto que en el año 1994 el PIB tuvo su máximo aumento en el periodo 1990 a 1995, entonces existía una buena situación económica, deduciendo así que las familias pudieron haber planificado el nacimiento de sus miembros en esta buena época económica, motivo por el cual pudo haber existido este gran número de nacimientos meses después, es decir, ya en 1995, como se había dicho anteriormente.

**Figura 3.2**

**Producto Interno Bruto del Ecuador. Período 1990-1997**



Si se observa el periodo 1990-1995 vemos que el PIB tuvo un incremento muy evidente en el año 1994, en incluso si se compara el porcentaje de incremento entre un año y otro, este año es el que tuvo un incremento mas drástico en comparación con los otros.

**Tabla VI**

**Tasa incremental del PIB**

<b><i>Incrementos del PIB</i></b>	<b><i>%</i></b>
Del 91 al 92	3,57
Del 92 al 93	2,03
Del 93 al 94	4,32
Del 94 al 95	2,34

Realizando un análisis mas detallado de los años del período de estudio, se tiene que en Diciembre y Enero de todos los años, se registran la mayor y menor cantidad de nacimientos en nuestro país, pero este hecho es justificado por las inscripciones tardías que se nombraron anteriormente. Fuera de estos dos meses, en general, en Marzo es el mes donde mas nacimientos ocurren, seguido por el mes de Mayo y Febrero, dentro del período 1990-1997.



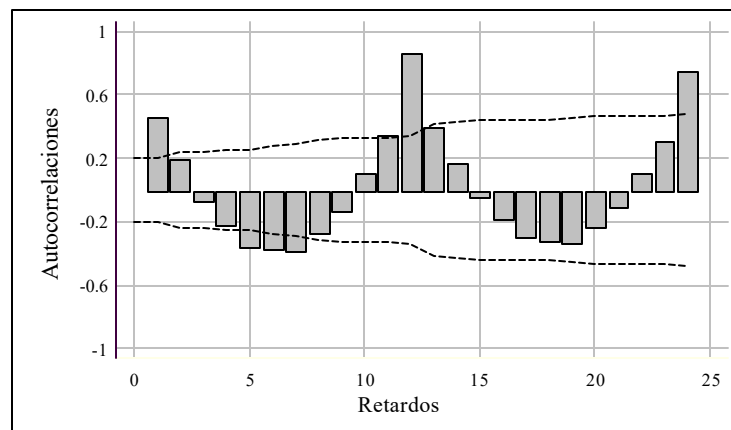
Por otro lado el peor año, fue 1997, es decir nació la menor cantidad de niños en este año, cabe indicar que a mediados de 1997 comenzó el fenómeno del niño siendo mas grave aún este hecho a finales del mismo año.

Tomando los datos anualmente, dentro de este período de estudio, el año en el que más nacimientos se registraron fue en 1992.

Continuando con el análisis para hallar un modelo adecuado para la serie, se considera necesario graficar las autocorrelaciones y las autocorrelaciones parciales para identificar si la serie necesita ser diferenciada.

**Figura 3.3**

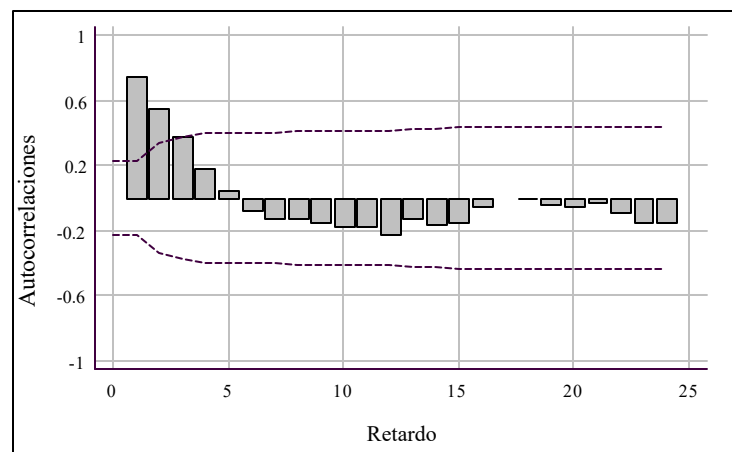
**Autocorrelaciones de la serie Total**



En la figura 3.3 se observa que existen barras continuas de las autocorrelaciones que sobresalen de las bandas de 95 % de confianza, en el primer período (primeras doce barras), por lo que es necesario una diferenciación de tipo estacional de la serie.

**Figura 3.4**

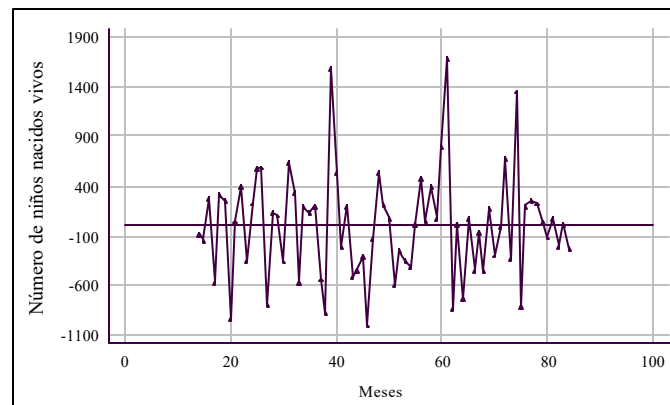
**Autocorrelaciones estimadas de la serie Total con una diferenciación estacional**



En la figura 3.4 se muestra el gráfico de las autocorrelaciones estimadas de la serie Total diferenciada una vez estacionalmente, de acuerdo a éste se puede decir que es necesario diferenciar la serie estacionariamente también, ya que las 3 primeras barras de las autocorrelaciones salen de las bandas de confianza.

Figura 3.5

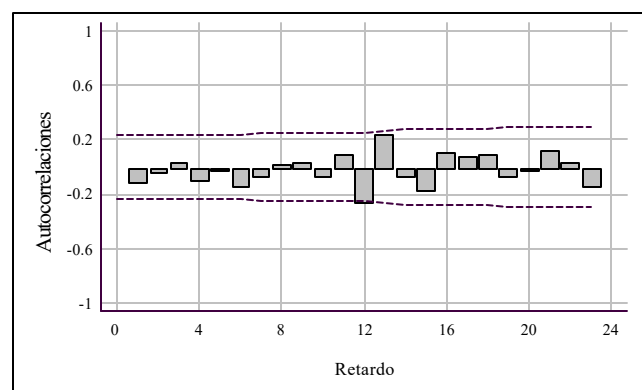
**Gráfico de la Serie Total con una diferenciación estacionaria y una diferenciación estacional**



En la figura 3.5 se muestra el gráfico de la serie Total diferenciada una vez estacionariamente y una vez estacionalmente, obsérvese que los valores de la serie fluctúan alrededor de la recta  $y = 0$ , debido a que han sido centrados.

Figura 3.6

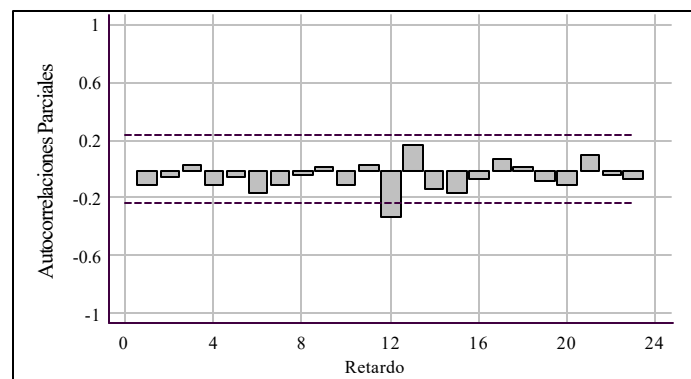
**Autocorrelaciones estimadas de la serie Total con una diferenciación estacionaria y una diferenciación estacional**



Para la parte media móvil del modelo se puede proponer un parámetro SMA(1), debido a que como muestra la figura 3.6 la décima segunda barra que representa las autocorrelaciones sobresale de manera significativa respecto a las otras barras, entonces se hace necesario un parámetro de tipo estacional.

**Figura 3.7**

**Autocorrelaciones parciales estimadas de la serie Total con una diferenciación estacional y una diferenciación estacionaria**



Mediante la figura 3.7 de las autocorrelaciones parciales se puede observar que la décima segunda barra es la que sobresale de las bandas de confianza, por lo que se puede proponer un parámetro SAR(1), es decir de tipo estacional, ya que la barra se sale en el primer periodo entonces su comportamiento es estacional mas no estacionario.

Por los motivos expresados anteriormente en los gráficos de las figuras de las autocorrelaciones y de las autocorrelaciones parciales se proponen los siguientes modelos:

SARIMA(0,1,0)(1,1,1)<sub>12</sub>

SARIMA(0,1,0)(1,1,0)<sub>12</sub>

Una vez propuestos los modelos se procede a la fase del desarrollo de las pruebas para los coeficientes de los polinomios autorregresivo y media móvil.

La prueba que se le aplica a cada parámetro, tiene el siguiente contraste de hipótesis:

$$H_0: \theta_i = 0$$

vs

$$H_1: \theta_i \neq 0$$

Rechace  $H_0$  en favor de  $H_1$ , si el valor  $p < 0.05$ .

El valor  $p$  de la prueba para cada uno de los coeficientes del modelo SARIMA(0,1,0)(1,1,1)<sub>12</sub> se muestra en la Tabla VI. Los valores  $p$  menores a 0.05 tienen evidencia estadística para decir que son significativamente diferentes de cero, con un 95% de nivel de confianza.

Tabla VII

## Pruebas T\_Student para los parámetros del modelo

SARIMA(0,1,0)(1,1,1)<sub>12</sub>

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	-0.3560	0.1166	-3.0538	0.0032
<b>MCE = 253364</b>				

El valor p del único parámetro para el modelo es menor que 0.05, entonces se dice que es significativamente diferente de cero, lo que indica que el valor estimado del parámetro se lo puede utilizar en el modelo.

Tabla VIII

## Pruebas T\_Student para los parámetros del modelo

SARIMA(0,1,0)(1,1,0)<sub>12</sub>

Parámetro	Estimación	Error estándar	Valor T	Valor P
SMA(1)	0.7905	0.0437	16.3106	0
<b>MCE = 412836</b>				

En este modelo el parámetro también pasa la prueba y se dice que es significativamente diferente de cero.

Se tiene hasta ahora que los dos modelos propuestos han pasado las pruebas de los parámetros, por lo que todavía no se puede decidir cual de ellos escoger.

A continuación se trabajará con estos modelos en las pruebas para verificar si los residuos son independientes, para la cual se tiene el siguiente contraste de hipótesis:

$H_0$ : Los residuos del modelo son independientes

vs

$H_1$ :  $\neg H_0$

Se rechaza la hipótesis nula si el valor p de la prueba es menor que 0.10 con un 90% de nivel de confianza.

Se considerará las siguientes pruebas, para los residuos de los modelos:

- Corrida arriba y abajo de la mediana
- Corrida arriba y abajo
- Box-Pierce

En las siguientes tablas se muestra el valor del estadístico Z y el valor P de las pruebas de corrida arriba y abajo, en la prueba de Box-Pierce se presenta el valor de su estadístico de prueba y el valor p, concernientes a los ruidos blancos.

Tabla IX

## Pruebas para los Residuos de la serie Total con el modelo

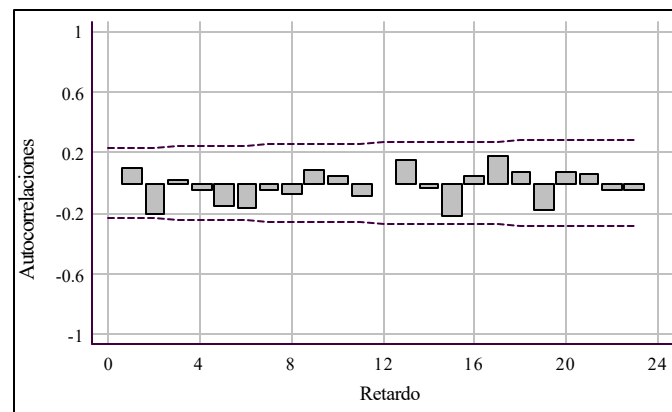
SARIMA(0,1,0)(1,1,1)<sub>12</sub>

Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	-0.1204	1.0958
Corrida arriba y abajo	0.1426	0.8866
Box-Pierce	19.5973	0.5469

Para este modelo el valor P de todas las pruebas es mayor que 0.10, entonces no se puede rechazar la hipótesis que los residuos son independientes.

Figura 3.8

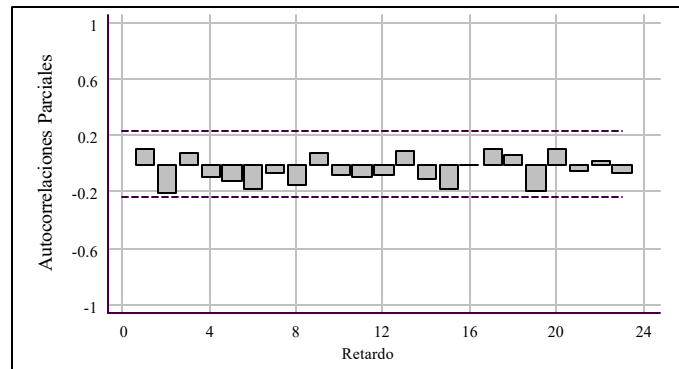
## Autocorrelaciones de los residuos de la serie Total con el modelo

SARIMA(0,1,0)(1,1,1)<sub>12</sub>

Como se puede observar ninguna de las barras de las autocorrelaciones sobresale de las bandas de confianza.



**Figura 3.9**  
**Autocorrelaciones parciales de los residuos de la serie Total con**  
**el modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,1)**



Como se puede observar en el gráfico de las autocorrelaciones parciales de los residuos de la serie mediante el modelo SARIMA(0,1,0)x(1,1,1)<sub>12</sub>, ninguna de las barras quedan fuera de las bandas de confianza.

**Tabla X**

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,0)**

Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	-0.3612	0.7180
Corrida arriba y abajo	0.4277	0.6689
Box-Pierce	20.500	0.5518

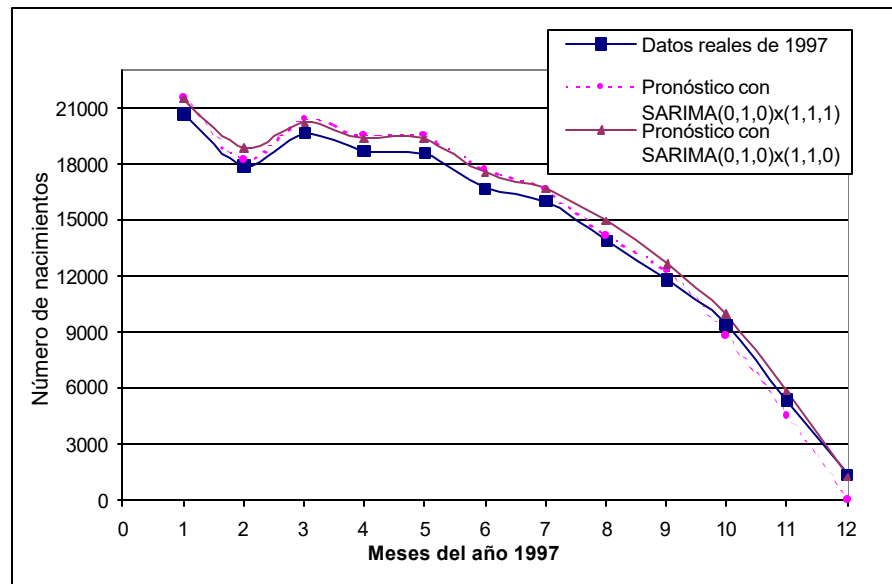
En la tabla X se muestra que los valores p de las pruebas para el otro modelo propuesto son mayores que 0.10, así que el modelo también pasa la prueba y se puede afirmar que los residuos son independientes.

Puesto que los dos modelos han pasado la prueba de los residuos, se procede con ellos a la fase de elección del modelo adecuado, para esto se utilizará el criterio del mayor poder predictivo seleccionando el de menor MCE.

Revisando las tablas anteriores (VII, VIII) de los modelos, se ve que el modelo de menor varianza es el SARIMA(0,1,0)(1,1,1)<sub>12</sub>, por lo tanto se escoge este modelo por el momento, hasta realizar la comparación con los datos reales y el otro modelo.

Para realizar una comparación de los datos reales con los valores que pronostican los modelos propuestos se ha dejado un año (1997) para poder verificar que tan bien se ajusta el modelo a los datos.

**Figura 3.10**  
**Comparación de la serie real del Total con los pronósticos de los**  
**modelos propuestos**



Se aprecia en la Figura 3.10 que los dos modelos son buenos, pero los datos pronosticados del modelo  $SARIMA(0,1,0) \times (1,1,0)_{12}$  (línea con triángulos) son los que más se ajustan a los datos reales, además la línea de rayas y puntos que corresponde al otro modelo, se aleja de los datos reales en los últimos meses; por lo tanto se concluye que el mejor modelo que estima la serie del Total de niños nacidos vivos en el Ecuador es:  $SARIMA(0,1,0) \times (1,1,0)_{12}$ , cuyos valores del pronóstico, los reales y sus intervalos de confianza se muestran en la Tabla XI;

además se indica el error relativo entre los reales y los pronósticos de cada valor. Su respectivo gráfico se muestra en la Figura 3.11.

**Tabla XI**

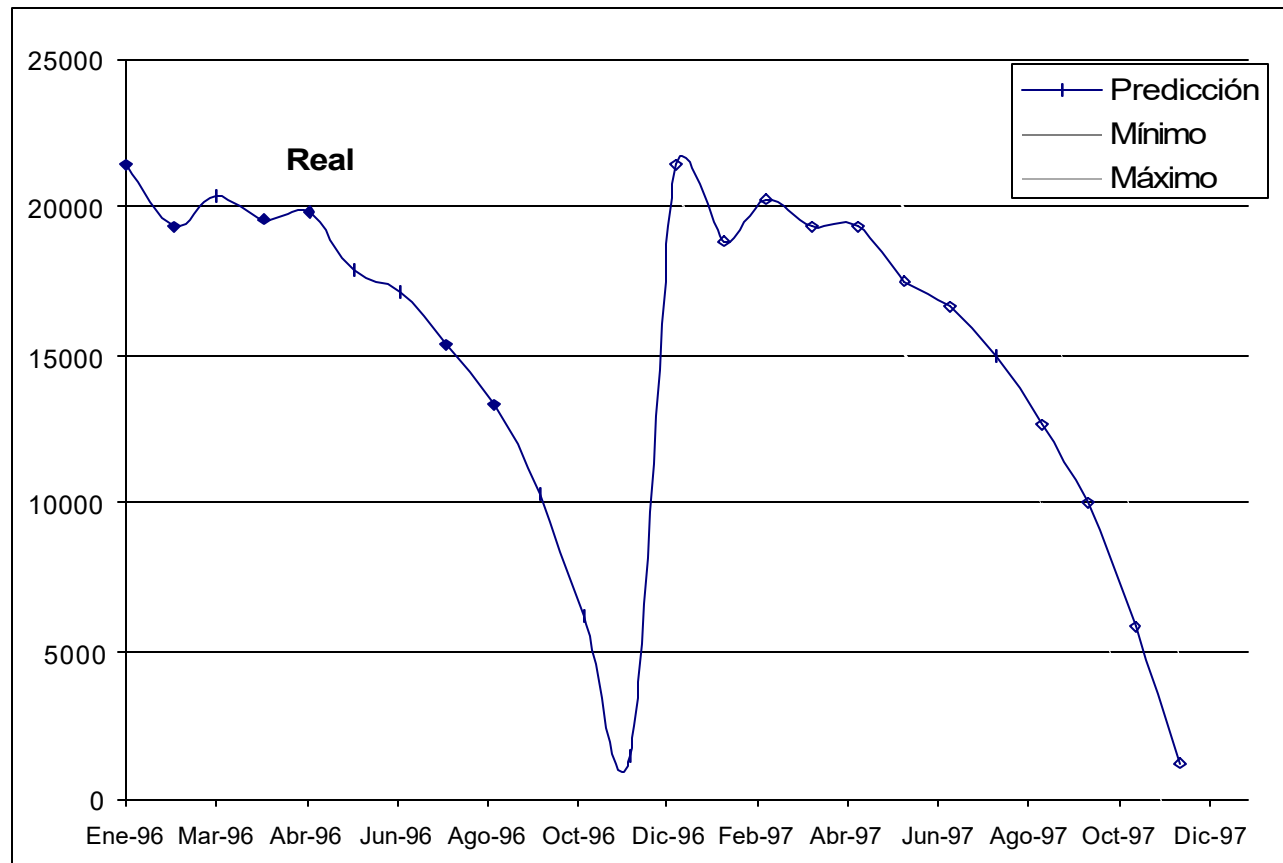
**Valores pronosticados de la serie Total con el modelo**

**SARIMA<sub>2</sub> (0,1,0)x(1,1,0)**

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
Ene-97	20659	21429	20459	22494	0.04
Feb-97	17962	18869	17444	20322	0.05
Mar-97	19621	20258	18498	22022	0.03
Abr-97	18652	19333	17307	21376	0.04
May-97	18609	19332	17244	21793	0.04
Jun-97	16663	17491	14982	19966	0.05
Jul-97	15958	16635	14024	19406	0.04
Ago-97	13884	14987	12087	17842	0.08
Sep-97	11782	12683	9906	16009	0.08
Oct-97	9369	9987	6742	13175	0.07
Nov-97	5390	5864	2458	9206	0.09
Dic-97	1320	1234	-2324	4724	0.07

Figura 3.11

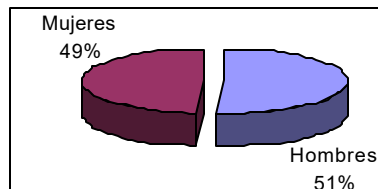
Gráfico de las predicciones de la serie del Total de nacimientos en el país



Siguiendo con el análisis del número de hombres y mujeres nacidos en el país, cabe indicar primero que en el Ecuador, durante el período 1990-1997, nacieron más hombres que mujeres, pero es con una diferencia mínima de 2%, que en cifras es aproximadamente 30.000 hombres de diferencia durante este periodo, lo que corresponde a 3.700 nacimientos de varones mas que mujeres cada año.

**Figura 3.12**

**Total de hombres y mujeres nacidos en el Ecuador**

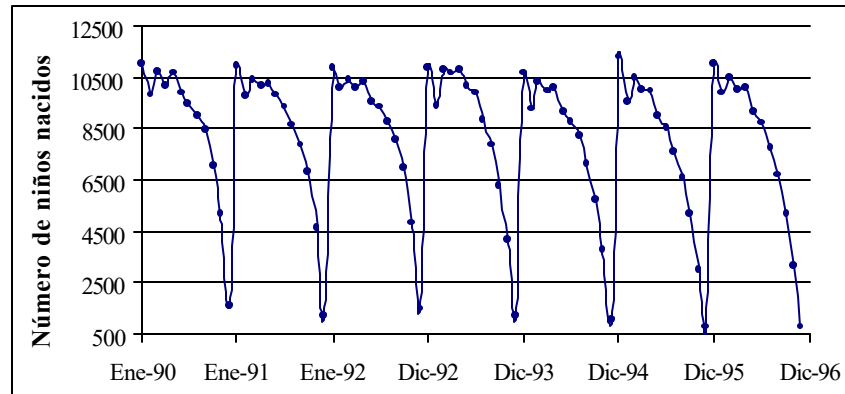


**3.4 Análisis de la serie del total de hombres nacidos en el país**

Como se procedió anteriormente, primero analizaremos el gráfico de la serie para notificar si existe alguna estacionalidad de los datos y así poder diferenciar, si es necesario.

Figura 3.13

## Total de Hombres nacidos en el país



Se puede observar en la Figura 3.13 mediante el gráfico de la serie original del Total de hombres nacidos en el Ecuador, que el comportamiento de los datos se repite cada 12 meses, igual que el en Total. Entonces es necesario diferenciar la serie ya que existe una estacionalidad de tamaño 12, motivo por el cual también aquí se trabajará con los modelos multiplicativos SARIMA  $(p,d,q) \times (P,D,Q)_{12}$ .

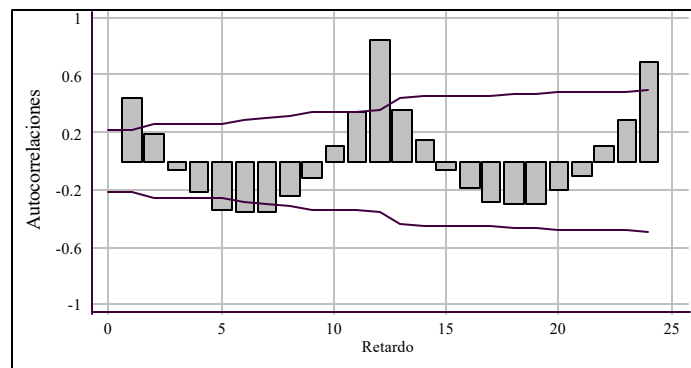
También en esta serie en enero de 1995 es donde mas nacimientos se han registrado, seguido por el año 1990 lo cual indica que, debido a los registros tardíos, a finales de 1994 y 1989 es donde mas nacimientos han ocurrido realmente. Por otro lado la fecha donde menos nacimientos se han registrado es en diciembre de 1996, seguido por diciembre de 1995. Pues se ve que son los dos últimos años, lo que lleva a pensar que el siguiente año será aún mas bajo

que éstos, pero solo se comprobará con los pronósticos que el número de nacimientos va decreciendo cada año.

A continuación se muestra el gráfico de las autocorrelaciones y las autocorrelaciones parciales de la serie original del total de hombres nacidos vivos, para verificar su estacionalidad.

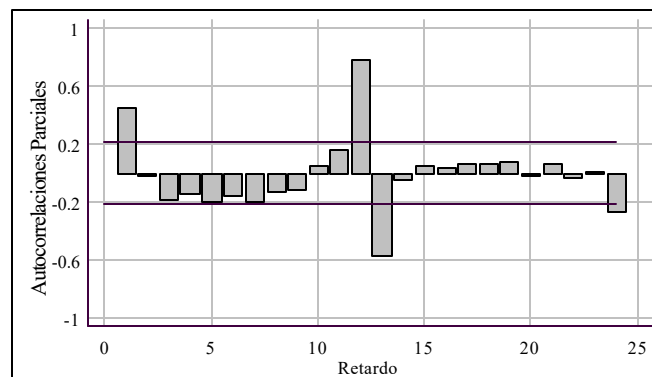
**Figura 3.14**

**Autocorrelaciones del Total de Hombres nacidos en el Ecuador**



**Figura 3.15**

**Autocorrelaciones parciales del Total de Hombres nacidos**

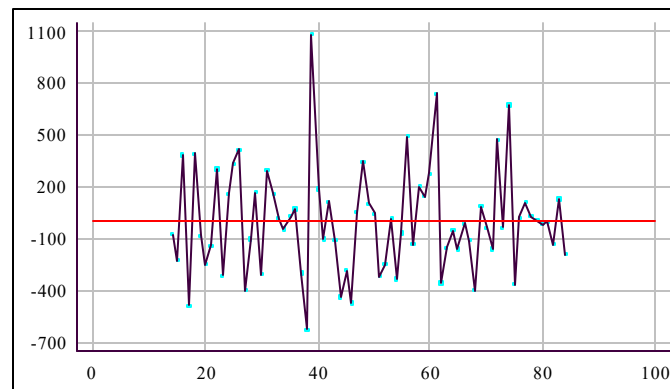




Se puede apreciar mediante la Figura 3.14 y 3.15 que es preciso diferenciar la serie estacionalmente y estacionariamente, puesto que el gráfico muestra barras seguidas que se salen de las bandas de confianza tanto al inicio de cada periodo (12 meses) como en el primer periodo, entonces se procede a diferenciar la serie de las 2 formas.

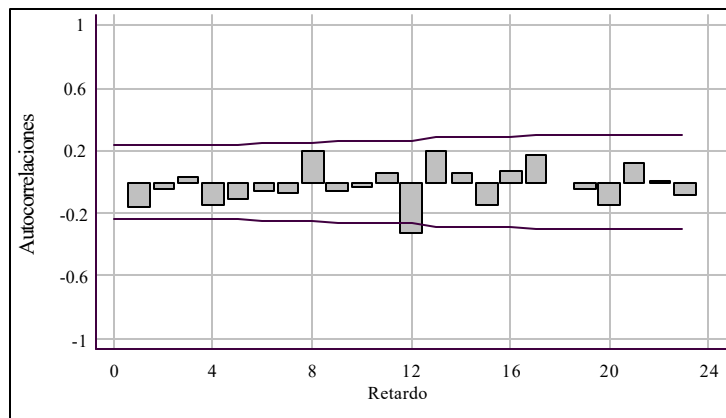
**Figura 3.16**

**Gráfico de la serie del total de Hombres con una diferenciación estacional y una estacionaria.**



Se puede observar que a través de una diferenciación estacional y estacionaria, la serie queda centrada, entonces se procede a analizar las autocorrelaciones y autocorrelaciones parciales de la nueva serie diferenciada para identificar el número de parámetros del modelo.

**Figura 3.17**  
**Autocorrelaciones del Total de Hombres con una diferenciación**  
**estacional y una estacionaria**

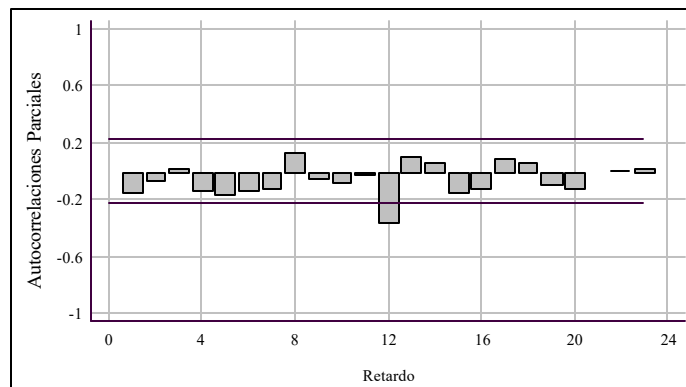


Como se muestra en la figura 3.17 la mayoría de las autocorrelaciones estimadas no sobresalen de las bandas de confianza a excepción de la doceava lo que implica adicionar al modelo un parámetro SMA(1), es decir un parámetro estacional en la parte media móvil del modelo.

En la siguiente figura se ve que también sobresale la barra número 12, así que se hace necesario un parámetro estacional, SAR(1) en la parte autorregresiva del modelo.

Figura 3.18

**Autocorrelaciones parciales del Total de Hombres con una  
diferenciación estacional y una estacionaria**



De acuerdo a lo analizado anteriormente se propone los siguientes modelos para la serie del total de hombres nacidos en el país:

$$\text{SARIMA}(0,1,0)(1,1,1)_{12}$$

$$\text{SARIMA}(0,1,0)(0,1,1)_{12}$$

$$\text{SARIMA}(0,1,0)(1,1,0)_{12}$$

A continuación se van a realizar las pruebas para identificar estadísticamente si los parámetros son significativamente diferentes de cero.

Entonces se tiene el siguiente contraste de hipótesis:

$$H_0: \theta_i = 0$$

vs

$$H_1: \theta_i \neq 0$$

Rechace  $H_0$  en favor de  $H_1$ , si el valor p es menor a 0.05.

**Tabla XII**

**Pruebas para los parámetros del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,1)**  
de la serie del total de Hombres

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	-0.2119	0.3054	-0.6940	0.49
SMA(1)	-0.2671	0.2989	0.8938	0.3745
<b>MCE = 78092.6</b>				

El valor p de la prueba para los dos parámetros del modelo son mayores a 0.05, entonces se dice que no son significativamente diferentes de cero.

**Tabla XIII**

**Pruebas para los parámetros del modelo SARIMA<sub>2</sub> (0,1,0)x(0,1,1)**  
de la serie del total de Hombres

Parámetro	Estimación	Error estándar	Valor T	Valor P
SMA(1)	0.4555	0.1083	4.2055	0
<b>MCE = 77230.7</b>				

En este modelo el parámetro de la parte media móvil del modelo pasa la prueba por lo que se dice que es significativamente diferente de cero.

**Tabla XIV**

**Pruebas para los parámetros del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,0)  
de la serie del total de Hombres**

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	-0.4659	0.1074	-4.3377	0
<b>MCE = 76973.3</b>				

De acuerdo a lo mostrado en la tabla se puede afirmar que el parámetro es significativamente diferente de cero, por su valor p que es menor a 0.05.

Entonces, se tiene que dos modelos tienen sus parámetros significativamente diferente de cero, SARIMA(0,1,0)(0,1,1)<sub>12</sub> y SARIMA(0,1,0)(1,1,0)<sub>12</sub>.

Ahora se pasará a la prueba de los ruidos blancos, cuyo contraste de hipótesis es el siguiente:

$H_0$ : Los residuos del modelo son independientes

vs

$H_1$ :  $\neg H_0$

Se rechaza la hipótesis nula si el valor p de la prueba es menor que 0.10 con un 90% de nivel de confianza.

Tabla XV

**Pruebas para los residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,1) de la serie del Total de Hombres**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	0.3612	0.7180
Corrida arriba y abajo	0.1426	0.8866
Box-Pierce	16.9292	0.7154

Para este modelo el valor p de todas las pruebas es mayor que 0.10, entonces no se puede rechazar la hipótesis que los residuos son aleatorios, pero de acuerdo a las otras pruebas nos indican que los parámetros de este modelo no son significativamente diferentes de cero, entonces se elimina la posibilidad de escoger este modelo.

Tabla XVI

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(0,1,1) de la serie Total de Hombres**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.1204	1.0958
Corrida arriba y abajo	0.1426	0.66
Box-Pierce	16.1803	0.8068

Tabla XVII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,0) de  
la serie Total de Hombres**

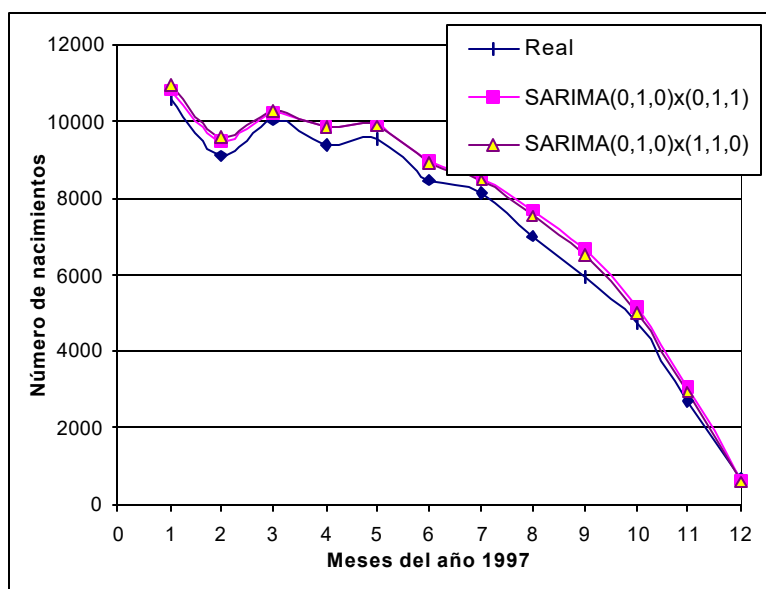
<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.1204	1.0958
Corrida arriba y abajo	0.4277	0.6689
Box-Pierce	17.132	0.756

En las tablas anteriores se muestra que los valores P de las pruebas para los modelos propuestos son mayores que 0.10, así que estos modelos también pasan la prueba y se puede afirmar que sus residuos son independientes.

Utilizando el criterio de menor varianza, el modelo de MCE menor es SARIMA(0,1,0)x(1,1,0)<sub>12</sub>, con una varianza de 76973. Además se puede apreciar en el gráfico comparativo de la Figura 3.19 que los valores que predice el modelo SARIMA(0,1,0)x(1,1,0)<sub>12</sub>, se ajustan mas a los valores reales, que el otro modelo propuesto; por lo tanto se tiene que el mejor modelo es éste.

Figura 3.19

Comparación de la serie real del Total de hombres con los pronósticos de los modelos propuestos



A continuación se muestra el gráfico de los valores que predice el modelo  $SARIMA(0,1,0) \times (1,1,0)_{12}$  en la Figura 3.20 y sus valores en la Tabla XVIII.



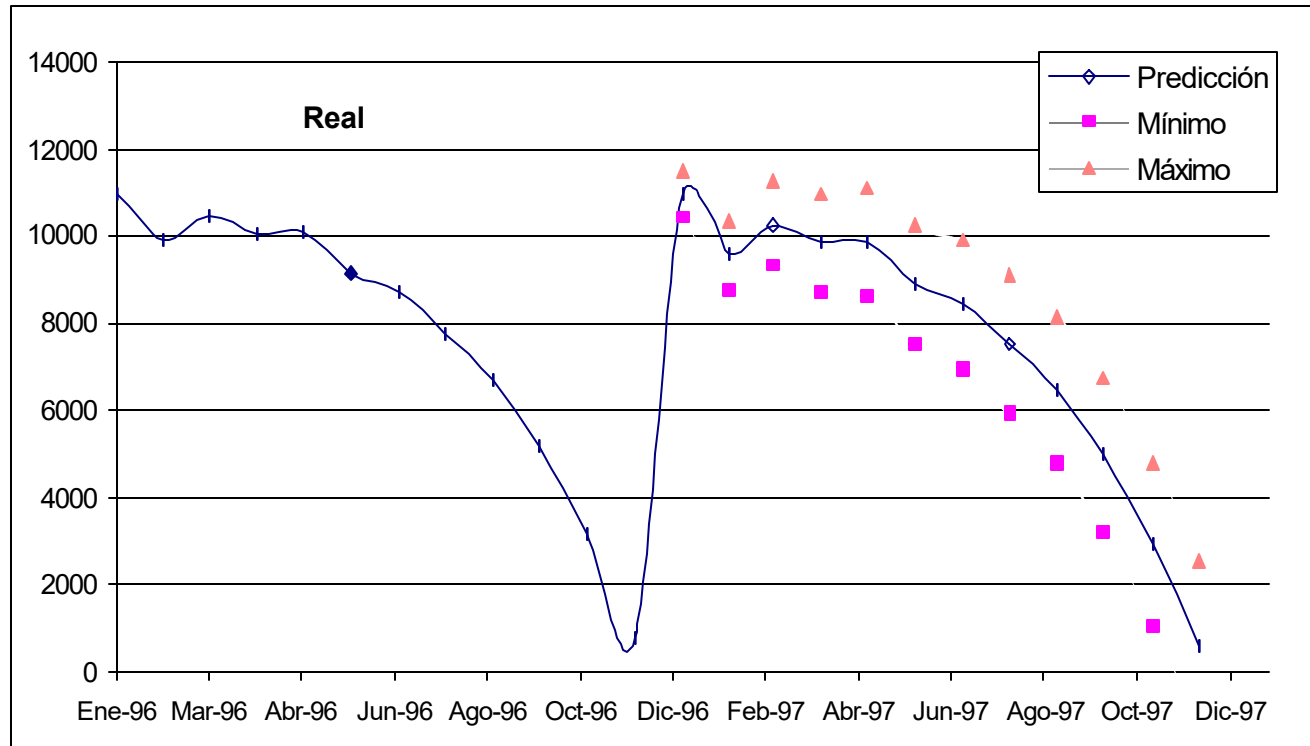
Tabla XVIII

Valores pronosticados de la serie del total de hombres con el  
 modelo SARIMA<sub>12</sub> (0,1,0)x(1,1,0)

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
Ene-97	10586	10970	10410	11531	0.04
Feb-97	9122	9586	8793	10378	0.05
Mar-97	10039	10282	9311	11252	0.02
Abr-97	9410	9867	8747	10988	0.05
May-97	9527	9881	8628	11134	0.04
Jun-97	8484	8908	7536	10281	0.05
Jul-97	8143	8454	6972	9937	0.04
Ago-97	7007	7539	5954	9124	0.08
Sep-97	5919	6491	4810	8172	0.10
Oct-97	4733	4999	3227	6771	0.06
Nov-97	2694	2922	1063	4780	0.08
Dic-97	668	596	-1345	2537	0.11

Figura 3.20

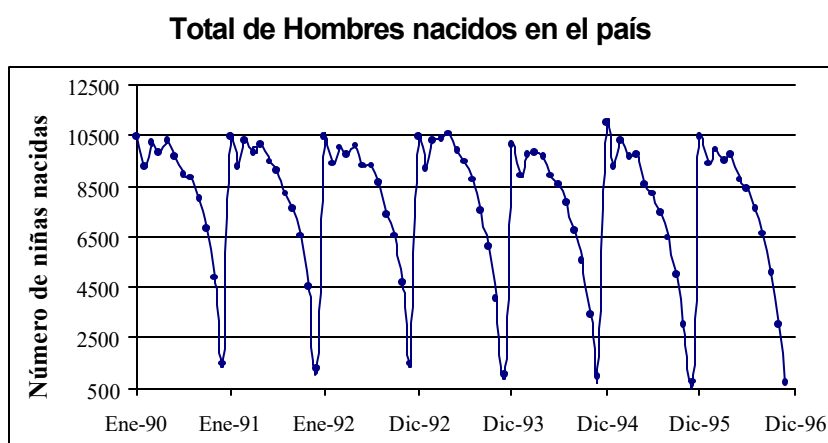
Gráfico de las predicciones de la serie Total de hombres con el modelo  $SARIMA_{12}(0,1,0)x(1,1,0)$



### 3.5 Análisis de la serie del total de mujeres nacidas en el país

Por medio del gráfico de la serie del número total de mujeres se puede indicar que en el Ecuador nacen más niños que niñas y al igual que en el gráfico de los datos de los Hombres y del Total en esta serie también se observa un comportamiento periódico.

**Figura 3.21**



Observando la serie por meses, como ya se había mencionado en enero es cuando más nacimientos ocurren y en diciembre es cuando menos nacimientos se registran, pero para esta serie, en enero de 1995 fue el mes que más nacimientos se registraron y en diciembre de 1996 se registró el número más bajo de nacimientos.

La cantidad de nacimientos de enero del 95 está seguida por la cantidad correspondiente al mes de mayo de 1993.

Revisando los datos de este periodo y tomando las cantidades de nacimientos por año, en 1995 es donde menos nacimientos ocurrieron, y en 1990 nacieron mas mujeres que en los otros años.

Ahora se procede a identificar el valor de  $d$  y  $D$  para los modelos de la serie con los gráficos de la autocorrelaciones y autocorrelaciones parciales.

El gráfico de las autocorrelaciones y las autocorrelaciones parciales es similar al de la serie del Total motivo por el cual se hace necesario una diferencia estacional y estacionaria de la serie del total de mujeres.

Una vez diferenciada la serie se tiene los siguientes gráficos:

**Figura 3.22**

**Autocorrelaciones del Total de Mujeres con una diferencia estacional y una diferencia estacionaria**

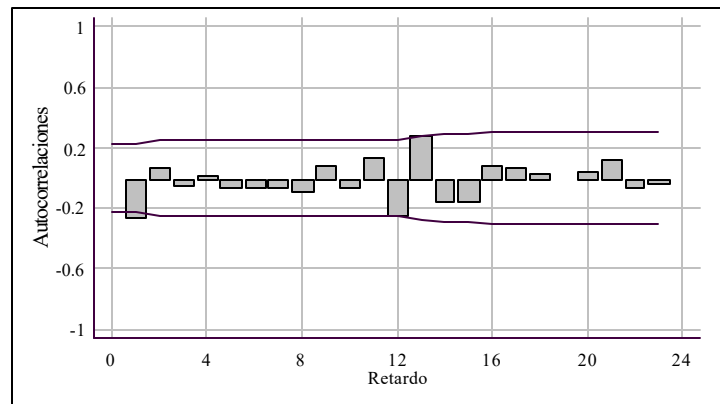
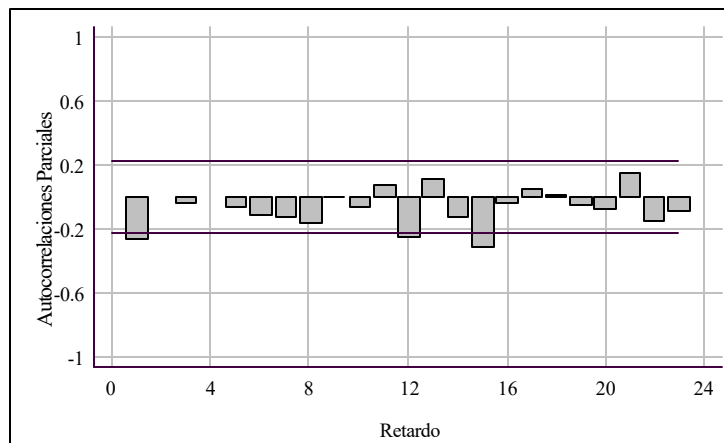


Figura 3.23

**Autocorrelaciones parciales del Total de Mujeres con una  
diferencia estacional y estacionaria**



Una vez realizados estos dos tipos de diferenciación la serie queda centrada.

Debido a que en el gráfico de las autocorrelaciones se observa que la primera y la décima segunda barra sobresalen de las bandas, entonces se propone incluir un parámetro estacional SMA; lo mismo ocurre en el gráfico de las autocorrelaciones parciales por lo tanto se propone un modelo que incluya un parámetro estacional SAR.

SARIMA(0,1,0)(0,1,1)12

SARIMA(0,1,0)(1,1,0)12

Ahora se pasará a la fase de verificación para probar si los parámetros son significativamente diferentes de cero.

Tabla XIX

Pruebas para los parámetros del modelo SARIMA(0,1,0)(0,1,1)<sub>12</sub>  
de la serie del total de Mujeres

Parámetro	Estimación	Error estándar	Valor T	Valor P
SMA(1)	0.3932	0.1107	3.551	0
<b>MCE = 75901.6</b>				

Tabla XX

Pruebas para los parámetros del modelo SARIMA(0,1,0)(1,1,0)<sub>12</sub>  
de la serie del total de Mujeres

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	-0.3451	0.1163	-2.9668	0.0041
<b>MCE = 77492.6</b>				

En los dos modelos se tiene que el valor p de las pruebas es menor a 0.05 por lo que se puede afirmar que sus parámetros son significativamente diferentes de cero, entonces son válidos para incluirlos en el modelo.

Tabla XXI

**Pruebas para los residuos del modelo SARIMA<sub>12</sub>(0,1,0)x(0,1,1) de la serie del Total de Mujeres**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y debajo de la mediana	1.0836	0.2785
Corrida arriba y abajo	1.5682	0.1168
Box-Pierce	19.1996	0.6330

Para este modelo el valor P de todas las pruebas es mayor que 0.10, entonces no se puede rechazar la hipótesis que los residuos son aleatorios.

Tabla XXII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,0) de la serie Total de Mujeres**

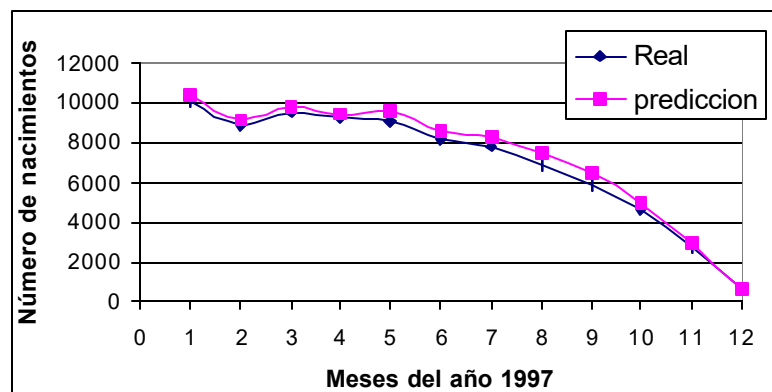
<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	1.805	0.0709
Corrida arriba y abajo	0.9980	0.3183
Box-Pierce	19.0215	0.644

Podemos observar que la prueba de corrida arriba y abajo de la mediana tiene valor p menor a 0.10, entonces se dice que los residuos de este modelo no son aleatorios.

Por lo tanto se escoge el primer modelo propuesto (SARIMA<sub>12</sub> (0,1,0)x(0,1,1)), incluso si se guía por el criterio de menor varianza este modelo tiene una MCE menor que es 75901.6. También se muestra el gráfico comparativo de los valores reales y los pronosticados con el modelo seleccionado.

**Figura 3.24**

**Comparación de la serie real del Total de Mujeres con los pronósticos del modelo SARIMA<sub>2</sub>(0,1,0)x(0,1,1)**





A continuación se muestra el gráfico más amplio en la Figura 3.25 y sus valores en la Tabla XXIII.

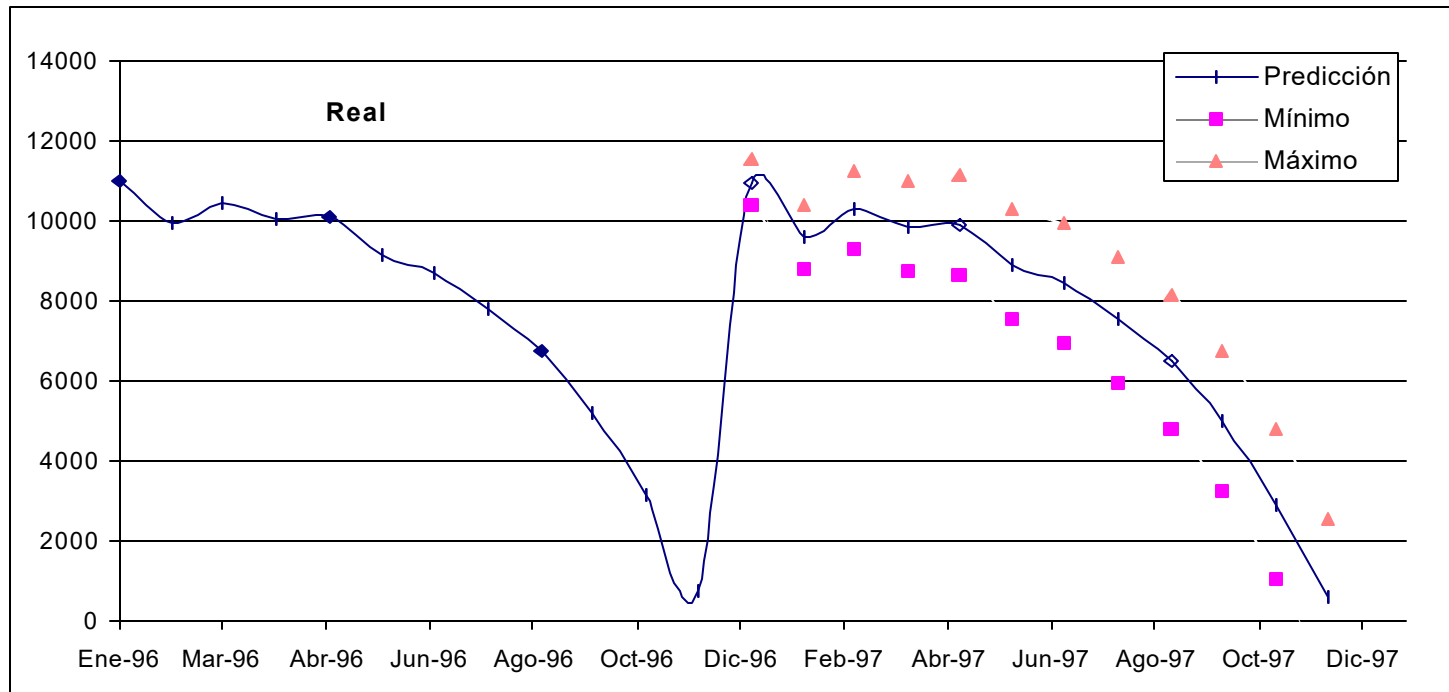
**Tabla XXIII**

**Valores pronosticados de la serie del total de Mujeres con el modelo SARIMA<sub>12</sub> (0,1,0)x(0,1,1)**

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	10073	10376	9821	10930	0.03
<i>Feb-97</i>	8840	9146	8361	9931	0.03
<i>Mar-97</i>	9582	9870	8909	10831	0.03
<i>Abr-97</i>	9242	9457	8347	10567	0.02
<i>May-97</i>	9082	9598	8357	10839	0.06
<i>Jun-97</i>	8179	8590	7231	9949	0.05
<i>Jul-97</i>	7815	8281	6813	9749	0.06
<i>Ago-97</i>	6877	7464	5895	9034	0.09
<i>Sep-97</i>	5863	6487	4823	8152	0.11
<i>Oct-97</i>	4636	5041	3286	6796	0.09
<i>Nov-97</i>	2696	2969	1129	4809	0.10
<i>Dic-97</i>	652	621	-1301	2543	0.05

Figura 3.25

Gráfico de las predicciones de la serie Total de Mujeres con el modelo SARIMA<sub>12</sub> (0,1,0)x(0,1,1)



Como se dijo anteriormente, se realizará un análisis por provincia, para el cual se han considerado las siguientes:

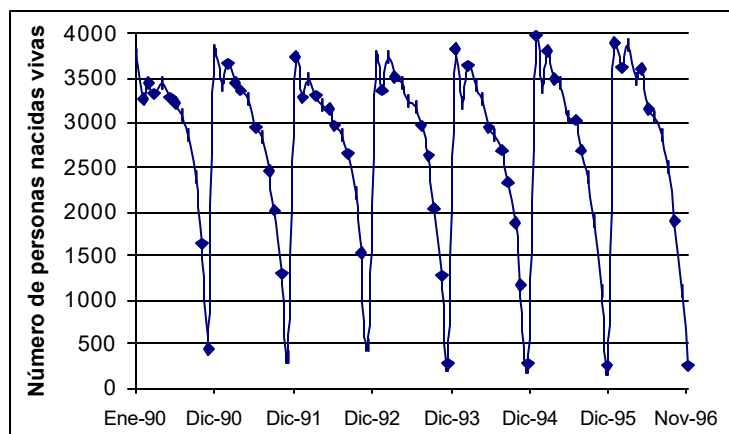
- Pichincha
- Chimborazo
- Loja
- Azuay
- Guayas
- Manabí
- El Oro
- Napo
- Pastaza

### **3.6 Análisis de serie de la provincia Pichincha**

Esta provincia es una de las principales del país, puesto que en ella se encuentra la capital del Ecuador motivo por el cual fue seleccionada para el presente análisis.

Figura 3.26

## Gráfico de la serie de la Provincia del Pichincha



Se encuentra que esta serie sigue el mismo comportamiento de las otras series ya analizadas, pero no se puede asegurar que tenga la misma serie para modelar sus datos, así que se procede a su respectivo análisis empezando con la observación de sus autocorrelaciones y autocorrelaciones parciales.

Puesto que los gráficos de sus autocorrelaciones son idénticos a los de las series anteriores, se procedió a diferenciar la serie una vez en forma estacional y una vez en forma estacionaria, entonces se propuso los siguientes modelos:

$$\text{SARIMA}(0,1,0) \times (0,1,1)_{12}$$

$$\text{SARIMA}(0,1,0) \times (1,1,0)_{12}$$

Además se va a incluir el siguiente modelo, que a diferencia de los otros no se diferencia la serie estacionariamente, sólo se considera una diferenciación de tipo estacional y con dos parámetros en la parte autorregresiva estacional y estacionaria.

SARIMA (1,0,0)x(1,1,0)<sub>12</sub>

Las pruebas para los parámetros se describen en las tablas siguientes.

**Tabla XXIV**

**Pruebas T\_Student para los parámetros del modelo**

**SARIMA(0,1,0)(0,1,1)<sub>12</sub> de la serie de la Prov. Pichincha**

Parámetro	Estimación	Error estándar	Valor T	Valor P
SMA(1)	0.3975	0.1159	3.43	0.0010
<b>MCE = 11850.8</b>				

El valor p del parámetro del modelo es menor que 0.05, entonces se dice que es significativamente diferente de cero.

Tabla XXV

**Pruebas T\_Student para los parámetros del modelo**  
**SARIMA(0,1,0)(1,1,0)<sub>12</sub> de la serie de la Prov. Pichincha**

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	-0.3265	0.1130	-2.8899	0.0051
<b>MCE = 12187.6</b>				

En este modelo el parámetro también pasa la prueba y se dice que es significativamente diferente de cero.

Se tiene hasta ahora que los dos modelos propuestos han pasado las pruebas de los parámetros, pero todavía no se puede decidir cual de ellos escoger, hasta verificar sus residuos.

Tabla XXVI

**Pruebas T\_Student para los parámetros del modelo**  
**SARIMA<sub>12</sub> (1,0,0)x(1,1,0) de la serie de la Prov. Pichincha**

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.7333	0.0812	9.0363	0
SAR(1)	-0.3987	0.1091	-3.6547	0.0005
<b>MCE = 10439.3</b>				

En el modelo que se propuso sólo con una diferenciación estacional y con dos parámetros autorregresivos estacional y no estacional respectivamente, también se puede afirmar que son significativamente diferentes de cero.

A continuación se trabajará con estos modelos en las pruebas para verificar si los residuos son independientes, para las cuales se utiliza el contraste de hipótesis mencionado anteriormente.

**Tabla XXVII**

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(0,1,1)  
de la serie de la Prov. Pichincha**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	1.5652	0.1175
Corrida arriba y abajo	1.2831	0.1995
Box-Pierce	13.9352	0.9038

Para este modelo el valor P de todas las pruebas es mayor que 0.10, entonces no se puede rechazar la hipótesis que los residuos son aleatorios.

Tabla XXVIII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(1,1,0)**  
**de la serie de la Prov. Pichincha**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	1.8060	0.0709
Corrida arriba y abajo	1.2831	0.1995
Box-Pierce	13.156	0.9287

En la tabla XXVIII se muestra el valor p de cada una de las pruebas para el modelo propuesto y se ve que el valor de la primera prueba no es mayor a 0.10, así que se dice que los residuos de este modelo no son independientes.

Tabla XXIX

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (1,0,0)x(1,1,0)**  
**de la serie de la Prov. Pichincha**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.3561	0.7218
Corrida arriba y abajo	1.3683	0.1712
Box-Pierce	13.3425	0.9232

Para el otro modelo propuesto, se comprueba que sus residuos son independientes.

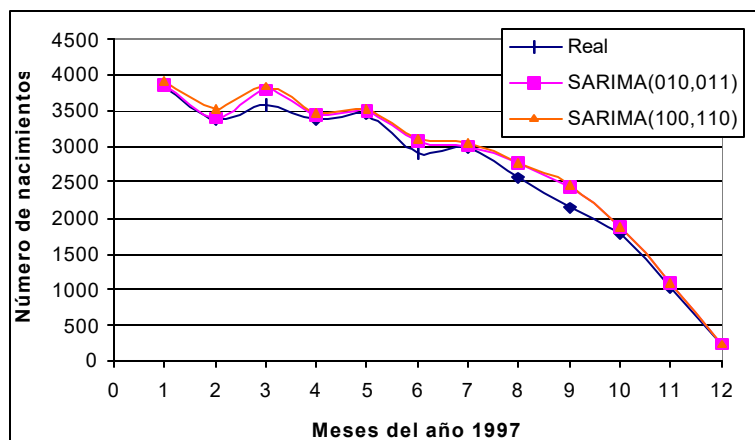


Estas pruebas han ayudado a eliminar un modelo, quedando así solo dos para pasar a la fase de elección del modelo mediante el criterio del mayor poder predictivo.

Revisando las tablas anteriores (XXIV, XXVI) de los modelos, se ve que el modelo de menor varianza es el **SARIMA<sub>12</sub>(1,0,0)x(1,1,0)**, con una varianza de 10.439. Además se muestra un gráfico comparativo de los valores reales con los pronósticos de los modelos propuestos, y como se aprecia las curvas de los dos modelos se aproximan en igual forma a la curva de los valores reales, por lo tanto se escoge el modelo de menor varianza.

**Figura 3.27**

**Comparación de la serie real de la provincia Pichincha con los pronósticos de los modelos propuestos**



La predicción de los valores para el modelo seleccionado se muestran en la Tabla XXX y sus gráfico en la figura 3.28.

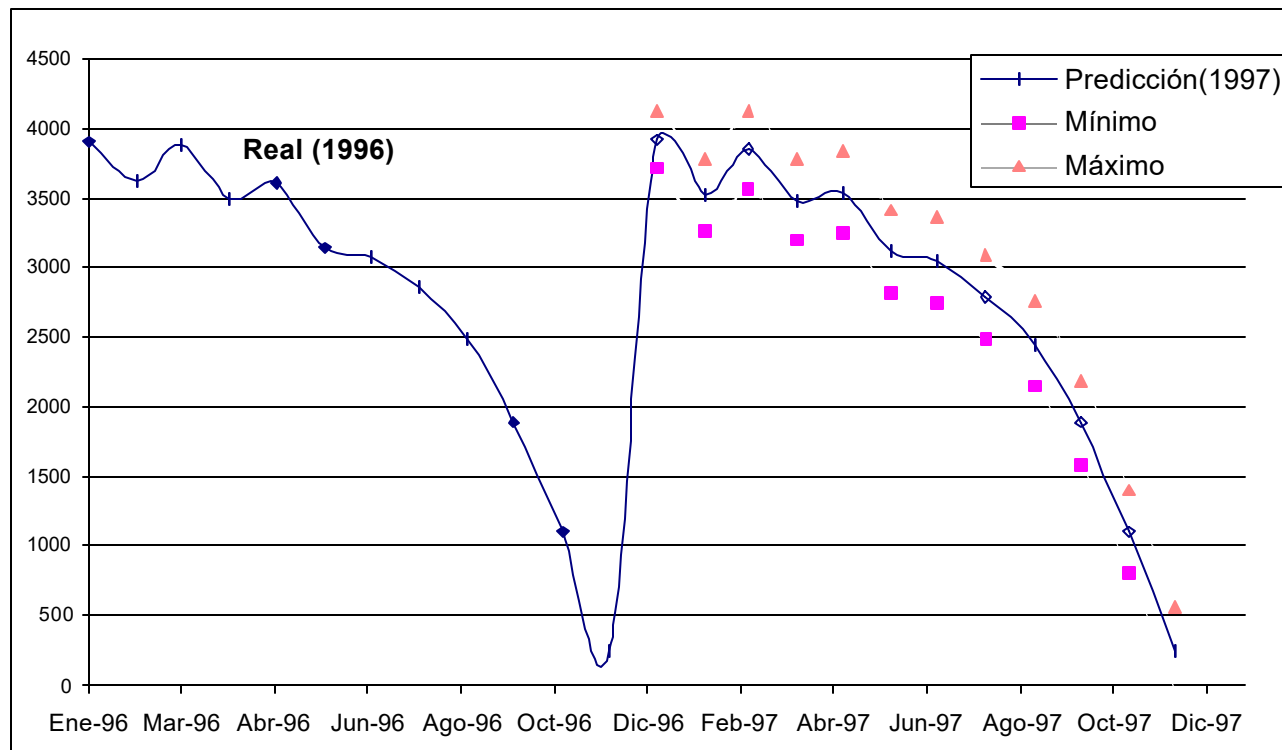
**Tabla XXX**

**Predicción de valores de la serie de la Prov. Pichincha**

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
Ene-97	3852	3917	3710	4125	0.02
Feb-97	3367	3526	3269	3783	0.05
Mar-97	3592	3845	3565	4125	0.07
Abr-97	3378	3485	3193	3777	0.03
May-97	3443	3541	3243	3839	0.03
Jun-97	2910	3112	2810	3413	0.07
Jul-97	2993	3050	2747	3353	0.02
Ago-97	2569	2786	2482	3090	0.08
Sep-97	2155	2450	2146	2755	0.14
Oct-97	1779	1884	1579	2188	0.06
Nov-97	1017	1100	795	1405	0.08
Dic-97	225	249	-56	554	0.11

Figura 3.28

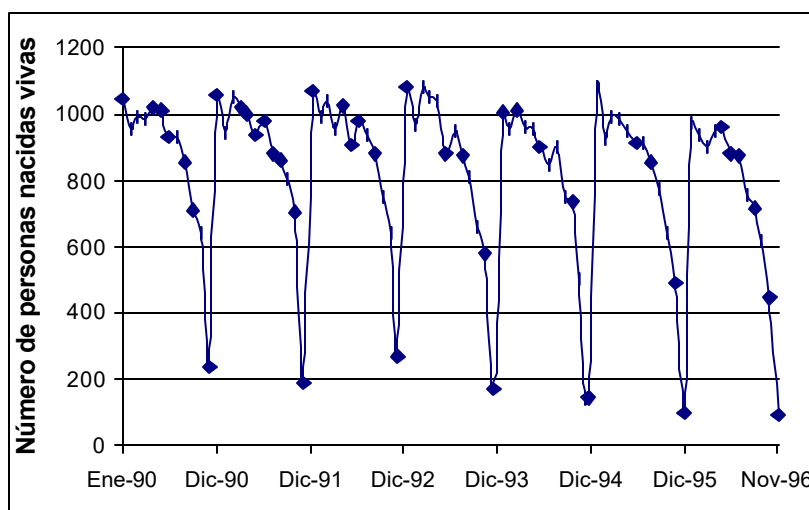
Gráfico de las predicciones de la serie de la Prov. Pichincha



### 3.7 Análisis de la provincia Chimborazo

Figura 3.29

Gráfico de la serie de la Provincia Chimborazo



En las siguientes figuras se muestran los gráficos de las autocorrelaciones de la serie original, es decir, que la serie todavía no ha sufrido ninguna alteración, entendiéndose por alteración, cuando se realiza algún tipo de diferenciación a la serie.

Figura 3.30

## Autocorrelaciones de la serie Prov. Chimborazo

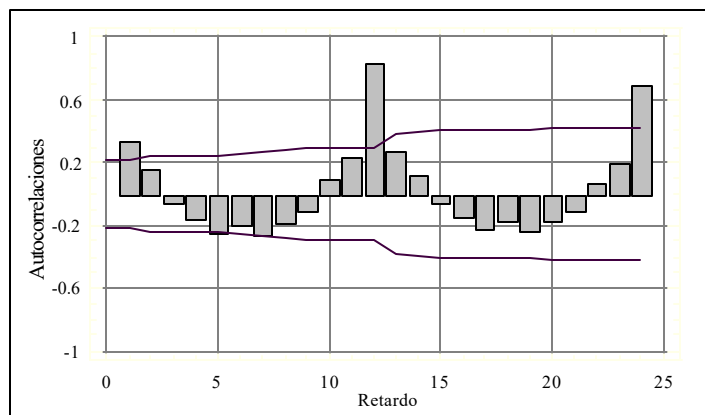
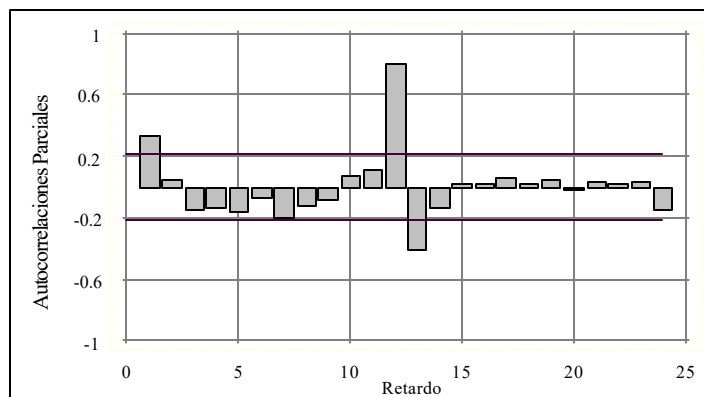


Figura 3.31

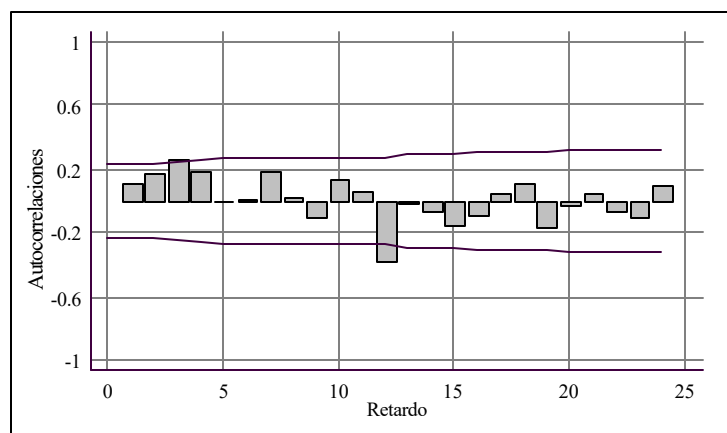
## Autocorrelaciones parciales de la serie Prov. Chimborazo



Mediante el gráfico de la serie (Figura 3.29) se puede observar que la serie también es estacional, ya que tiene un comportamiento similar cada cierto período de tiempo, también vemos que esa periodicidad es de 12. Se hace necesario también una diferenciación estacional, puesto que las barras sobresalen cada periodo de 12 barras.

Figura 3.32

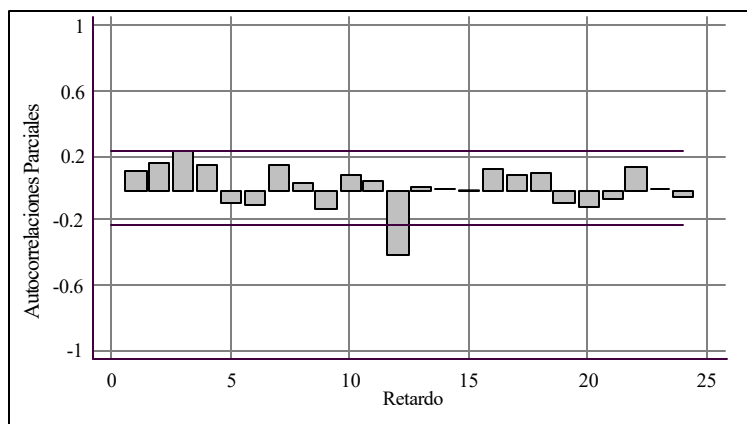
**Autocorrelaciones estimadas de la serie de la Prov. Chimborazo  
con una diferenciación estacional**



Como se aprecia en el gráfico anterior, la serie diferenciada es diferente al gráfico de las anteriores series diferenciadas, motivo por el cual se debe de tratar de un modo distinto; es decir con otros parámetros. Entonces se incluye un parámetro MA(1) estacional, puesto que sobresale de décima segunda barra en este gráfico de autocorrelaciones, además se ve que sobresale la tercera barra por lo que también se tomará en cuenta un parámetro no estacional, MA(3).

Figura 3.33

**Autocorrelaciones parciales estimadas de la serie de la Prov.  
Chimborazo con una diferenciación estacional**



De acuerdo al gráfico de las autocorrelaciones parciales se debe adicionar un parámetro AR(1) estacional, ya que la décima segunda barra es la que sobresale de las bandas de confianza.

Como se ve, en esta serie no es necesario incluir parámetros no estacionales al modelo.

Por las razones descritas en los dos gráficos se propone los siguientes modelos:

$$\text{SARIMA}_{12}(0,0,0)\times(1,1,1)$$

$$\text{SARIMA}_{12}(0,0,3)\times(1,1,1)$$

En las siguientes tablas se muestran las pruebas para los dos modelos.

**Tabla XXXI**

**Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>**

**(0,0,0)x(1,1,1) de la serie de la Prov. Chimborazo**

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	-1.2588	0.0939	-13.4097	0
SMA(1)	-1.1883	0.1194	-9.9545	0
<b>MCE = 2355.9</b>				

**Tabla XXXII**

**Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>**

**(0,0,3)x(1,1,1) de la serie de la Prov. Chimborazo**

Parámetro	Estimación	Error estándar	Valor T	Valor P
MA(1)	-0.1372	0.1195	-1.1484	0.2549
MA(2)	-0.2778	0.1181	-2.3529	0.0216
MA(3)	-0.2099	0.1192	-1.7606	0.0829
SAR(1)	-0.8840	0.0980	-9.0229	0
SMA(1)	-0.4243	0.1791	-2.3690	0.0207
<b>MCE = 2204.6</b>				



El valor p de las pruebas para los parámetros MA son mayores que 0.05, entonces se dice que no son significativamente diferentes de cero, por lo tanto no es necesario adicionar estos parámetros al modelo.

Con esta prueba de los parámetros se ha eliminado la posibilidad de seleccionar el segundo modelo, así que la siguiente prueba de los residuos se la realizará sólo para el primer modelo propuesto.

**Tabla XXXIII**

**Pruebas para los Residuos del modelo SARIMA<sub>12</sub>(0,0,0)(1,1,1) de**

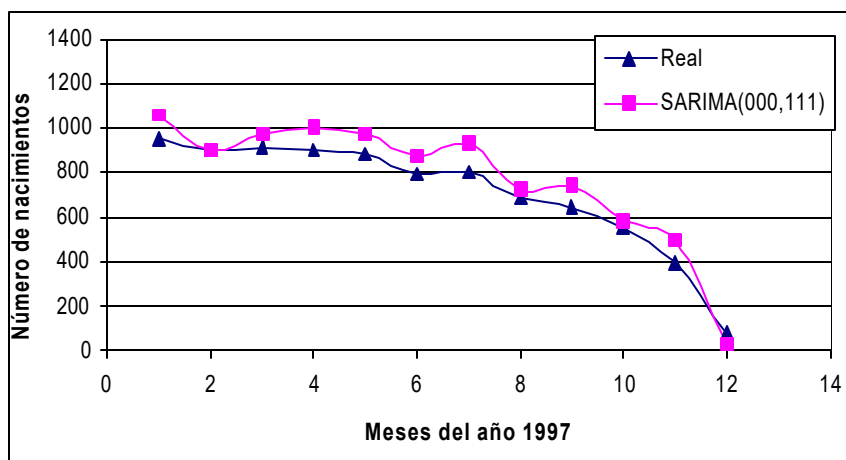
**la serie de la Prov. Chimborazo**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.1187	0.9055
Corrida arriba y abajo	-0.8965	0.37
Box-Pierce	28.3213	0.1653

Todos los valores p de las pruebas son mayores que 0.10, así se puede afirmar que los residuos de este modelo son independientes.

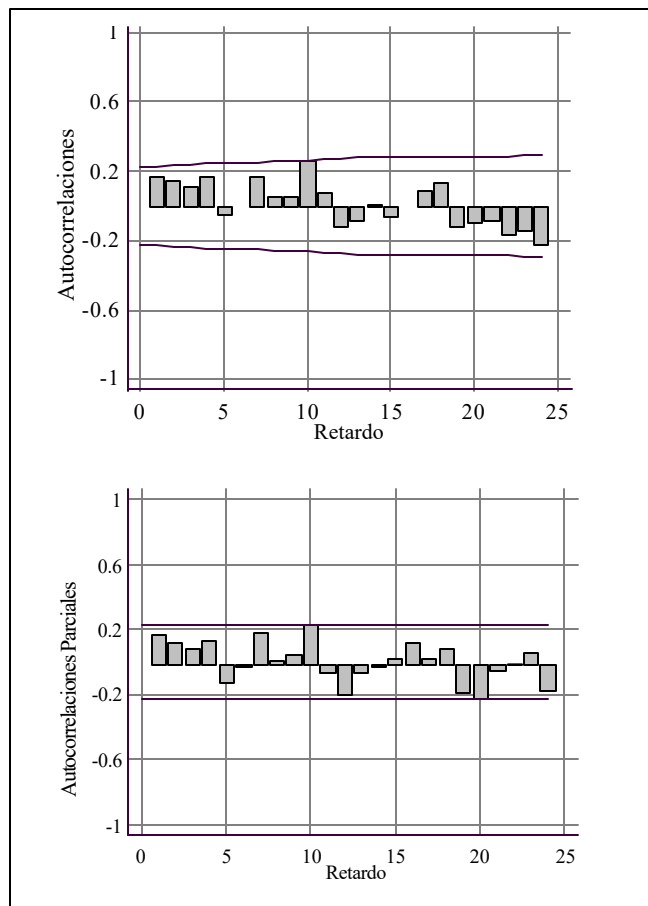
**Figura 3.34**

**Comparación de la serie real de la provincia Chimborazo con los pronósticos del modelo SARIMA<sub>2</sub> (0,0,0)(1,1,1)**



En la figura 3.35 se muestra el gráfico de los residuos del modelo seleccionado.

**Figura 3.35**  
**Autocorrelaciones y Autocorrelaciones parciales de los Residuos**  
**del modelo SARIMA<sub>2</sub>(0,0,0)(1,1,1) de la serie de la Prov.**  
**Chimborazo**



Como se muestra en el gráfico de las autocorrelaciones, todos los residuos quedan dentro de las bandas de confianza, entonces se asegura que el modelo seleccionado es adecuado, en la siguiente

tabla se encuentran los valores de sus predicciones, y su respectivo gráfico en la figura 3.36.

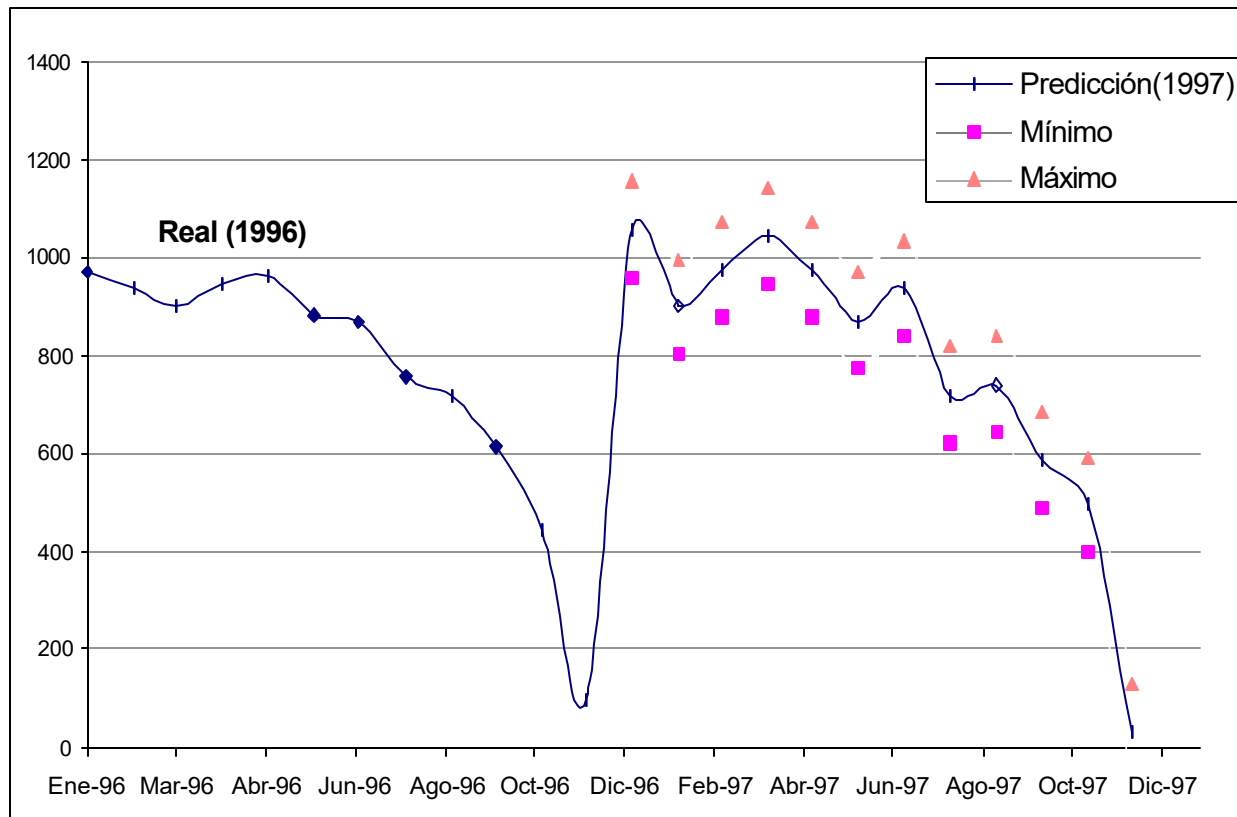
**Tabla XXXIV**

**Predicción de valores de la serie de la Prov. Chimborazo con el modelo SARIMA<sub>12</sub>(0,0,0)(1,1,1)**

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	949	1057	960	1155	0.11
<i>Feb-97</i>	903	900	803	998	0.00
<i>Mar-97</i>	916	975	878	1073	0.06
<i>Abr-97</i>	908	1044	947	1142	0.15
<i>May-97</i>	883	976	879	1074	0.11
<i>Jun-97</i>	796	872	775	970	0.10
<i>Jul-97</i>	800	938	841	1036	0.17
<i>Ago-97</i>	683	720	623	818	0.05
<i>Sep-97</i>	646	742	644	839	0.15
<i>Oct-97</i>	558	587	489	684	0.05
<i>Nov-97</i>	398	479	399	594	0.20
<i>Dic-97</i>	76	42	-65	129	0.45

Figura 3.36

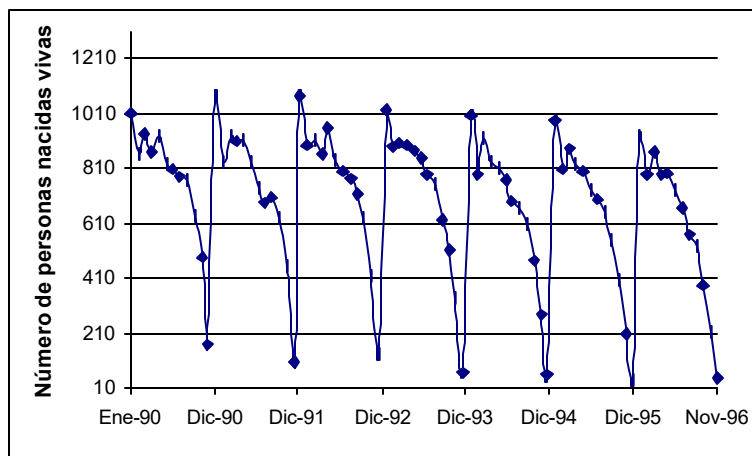
Gráfico de las predicciones de la serie de la Prov. Chimborazo con el modelo SARIMA<sub>12</sub>(0,0,0)x(1,1,1)



### 3.7 Análisis de serie de la provincia Loja

Figura 3.37

Gráfico de la serie de la Provincia Loja



En esta provincia se nota claramente un decrecimiento marcado de la serie, así cada año transcurrido de este período han ido naciendo menos niños en la provincia de Loja.

Esta serie al igual que las anteriores tiene sus gráficos de autocorrelaciones y autocorrelaciones parciales con las barras cada doce meses sobresalidas de las bandas de confianza motivo por el cual se considera necesario diferenciarla estacionalmente.

Figura 3.38

**Autocorrelaciones estimadas de la serie de la Prov. Loja con una diferencia estacional**

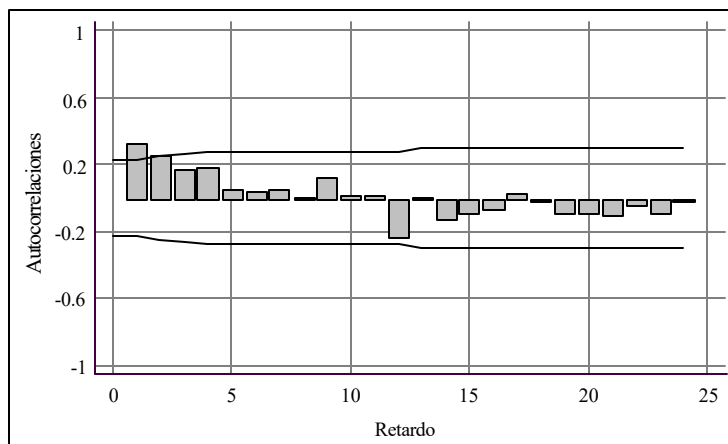
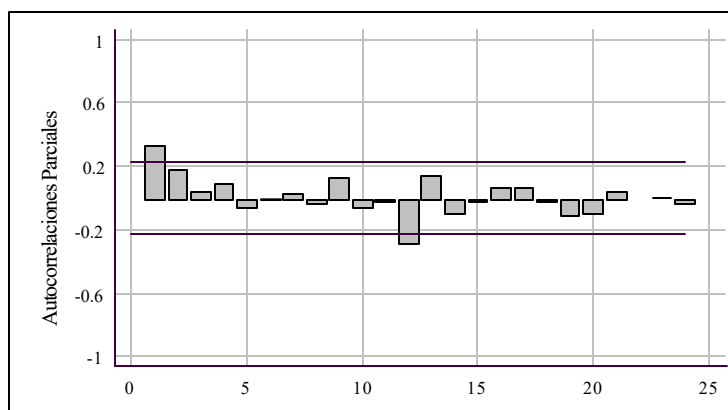


Figura 3.39

**Autocorrelaciones parciales estimadas de la serie de la Prov. Loja con una diferencia estacional**



Observando la figura 3.38 nos indica que debemos adicionar un parámetro MA(1), ya que la primera barra de las autocorrelaciones sobresale de las bandas, mientras que en el gráfico de las autocorrelaciones parciales (Fig. 3.39) indica adicionar dos parámetros para la parte autorregresiva del modelo, uno estacional y otro no estacional, debido a que la primera y la décima segunda barra sobresalen de las bandas.

Por los motivos expuestos, se tiene el siguiente modelo:

$$\text{SARIMA}_{12}(1,0,1)(1,1,0)$$

**Tabla XXXV**

**Pruebas T\_Student para los parámetros del modelo**

**SARIMA(1,0,1)(1,1,0)<sub>12</sub> de la serie de la Prov. Loja**

<b>Parámetro</b>	<b>Estimación</b>	<b>Error estándar</b>	<b>Valor T</b>	<b>Valor P</b>
AR(1)	0.945	0.053	17.714	0
MA(1)	0.579	0.126	4.586	0
SAR(1)	-0.504	0.103	-4.911	0
<b>MCE = 1372.56</b>				



Tabla XXXVI

**Pruebas para los Residuos del modelo SARIMA<sub>12</sub> (1,0,1)(1,1,0) de  
la serie de la Prov. Loja**

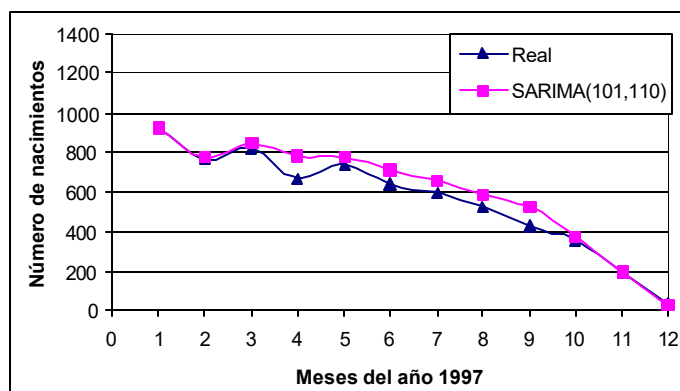
Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	0.356	0.721
Corrida arriba y abajo	-0.897	0.370
Box-Pierce	11.011	0.963

Como se observa en los cuadros anteriores los parámetros del modelo propuesto son significativamente diferentes de cero y sus residuos son independientes, entonces el modelo propuesto es el apropiado para ajustar esta serie.

A continuación en la figura 3.40 se muestra la comparación de los datos reales con los pronósticos del modelo SARIMA<sub>12</sub> (1,0,1)(1,1,0).

Figura 3.40

**Comparación de la serie real de la provincia Loja con los  
pronósticos del modelo SARIMA(1,0,1)x(1,1,0)<sub>12</sub>**



Al observar este gráfico vemos que el pronóstico está muy cerca de los valores reales, sólo en el cuarto mes aparece una pequeña desviación y en el resto de los meses la línea que representa los valores reales se ajusta muy bien a la línea de los pronósticos.

**Tabla XXXVII**

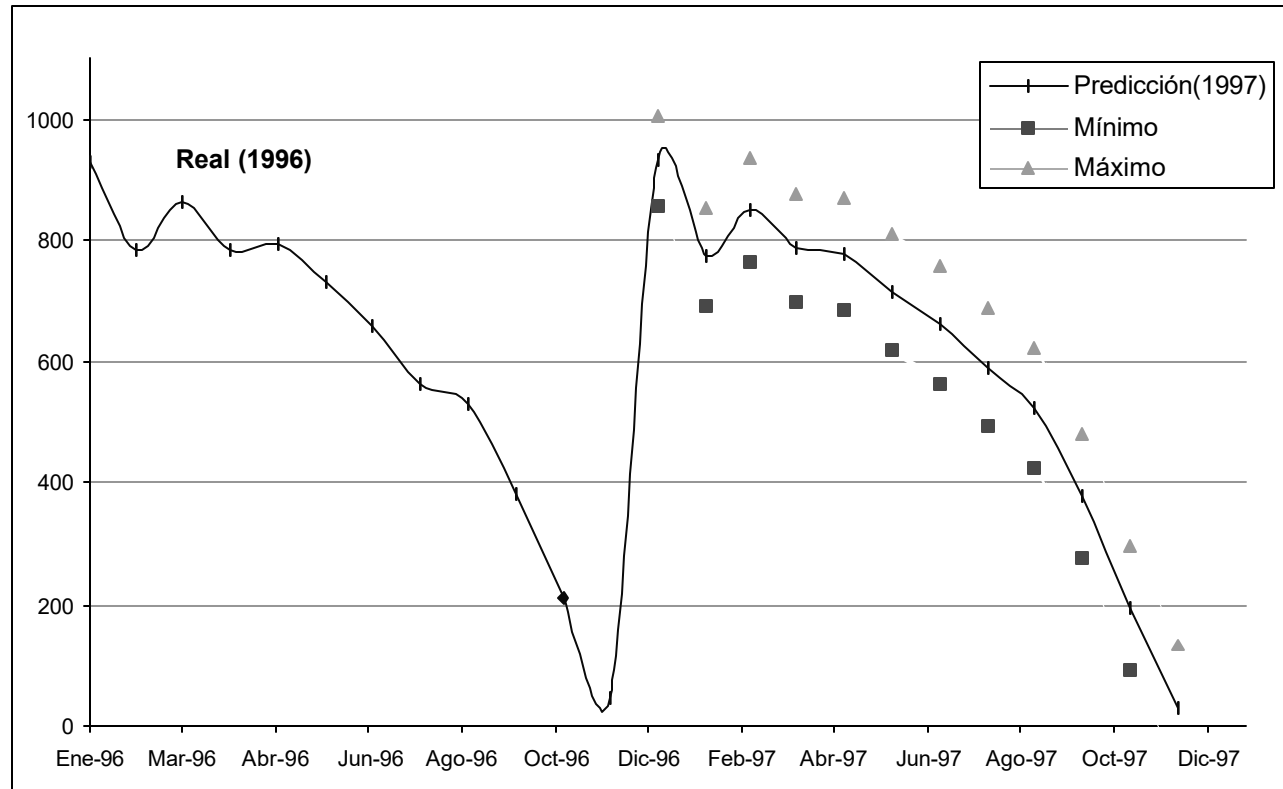
**Predicción de valores de la serie de la Prov. Loja con el modelo**

**SARIMA<sub>12</sub>(1,0,1)(1,1,0)**

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	927	931	855	1006	0.00
<i>Feb-97</i>	765	772	692	853	0.01
<i>Mar-97</i>	821	849	764	933	0.03
<i>Abr-97</i>	666	786	698	874	0.18
<i>May-97</i>	742	776	685	867	0.05
<i>Jun-97</i>	645	715	621	808	0.11
<i>Jul-97</i>	597	661	565	757	0.11
<i>Ago-97</i>	524	591	493	689	0.13
<i>Sep-97</i>	426	523	423	623	0.23
<i>Oct-97</i>	357	377	276	479	0.06
<i>Nov-97</i>	194	195	92	298	0.01
<i>Dic-97</i>	37	30	-73	134	0.18

Figura 3.41

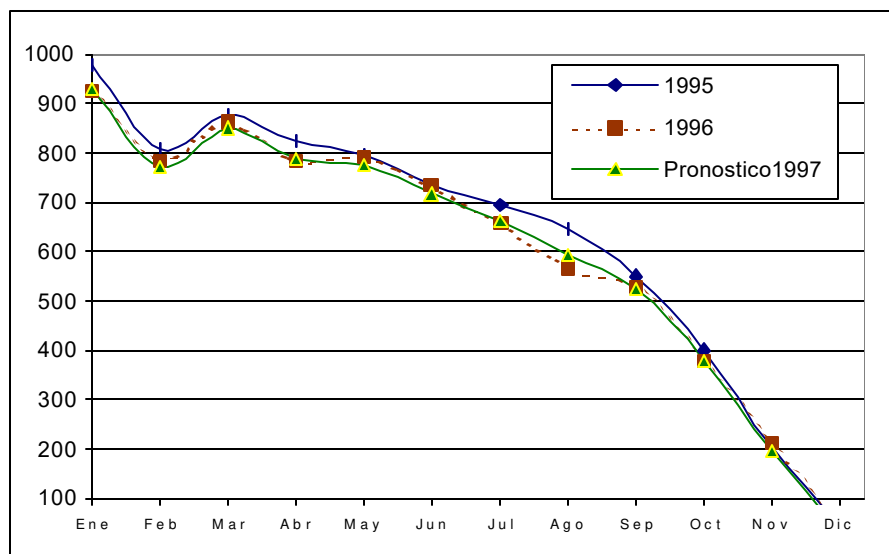
Gráfico de las predicciones de la serie de la Prov. Loja con el modelo SARIMA<sub>12</sub>(1,0,1)x(1,1,0)



En el siguiente gráfico se observa una comparación entre los años 1995, 1996 y los pronósticos de 1997, para así poder mostrar que en el año 1997 sigue disminuyendo el número de nacimientos como en los años anteriores para esta provincia.

**Figura 3.42**

**Comparación de los años 95 y 96 con los pronósticos del 97 de la serie de la Prov. Loja**

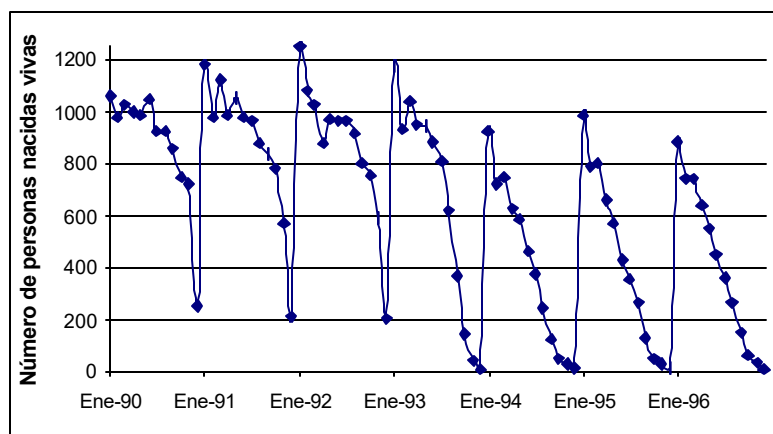


Entonces para el año 1997 nacerán menos niños que en el año 1996 en casi todos los meses, por lo que se confirma que esta serie sigue su comportamiento decreciente.

### 3.8 Análisis de la serie de la provincia Azuay

Figura 3.43

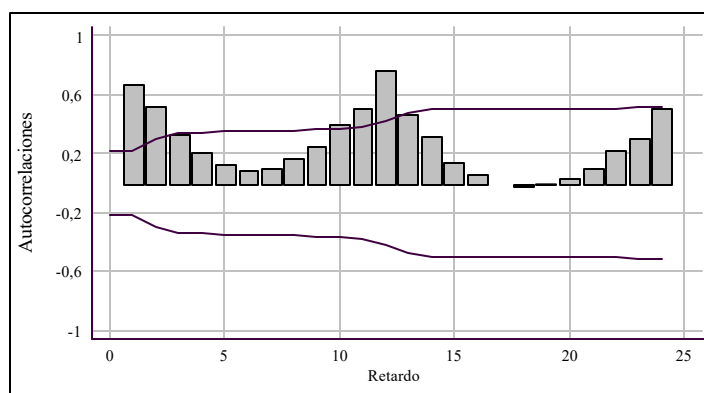
Gráfico de la serie de la Provincia Azuay



En la provincia del Azuay se observa que el número de nacimientos va disminuyendo cada año, siendo más evidente en los años 1994, 1995 y 1996.

Figura 3.44

Autocorrelaciones de la serie Prov. Azuay



Mediante este gráfico se puede suponer en primer instante que es necesario diferenciar la serie una vez estacionalmente y posiblemente estacionariamente, pero se probará dos modelos: uno con diferenciación estacional y otro con diferenciación estacionaria. La razón de evitar diferenciar la serie es que perderíamos muchos datos al realizar la diferenciación estacional, ya que se tiene una estacionalidad de 12, entonces se pierden 12 datos.

**Figura 3.45**

**Autocorrelaciones parciales de la serie Prov. Azuay**

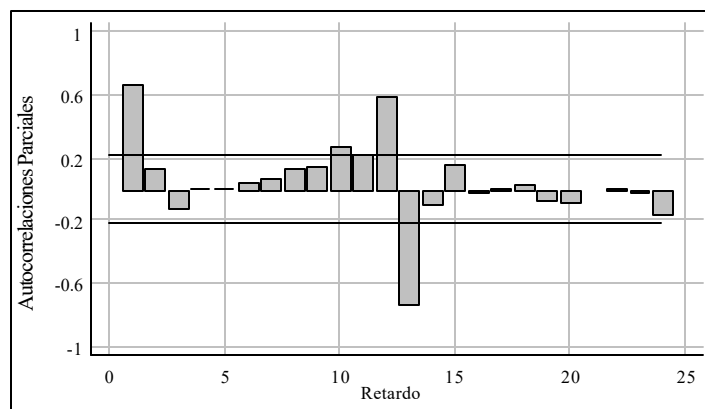
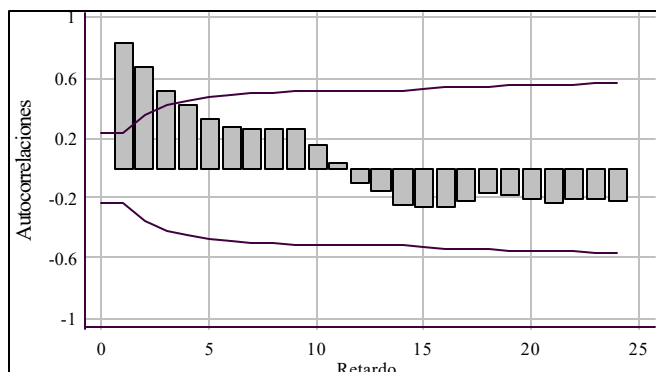


Figura 3.46

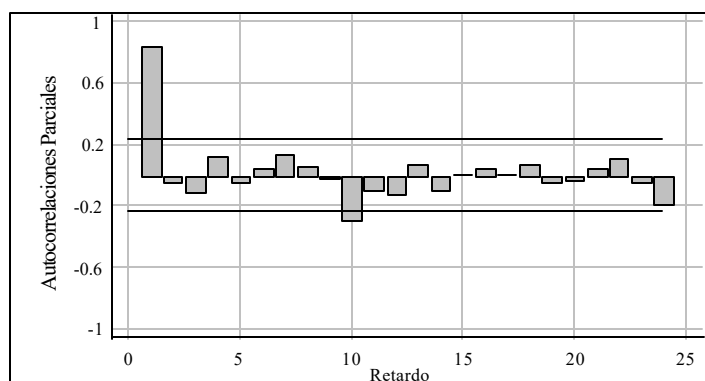
**Autocorrelaciones estimadas de la serie Prov. Azuay con una diferencia estacional**



Para la parte media móvil del modelo, el gráfico sugiere adicionar al modelo 3 parámetros estacionarios.

Figura 3.47

**Autocorrelaciones parciales estimadas de la serie Prov. Azuay con una diferencia estacional**



Y para la parte autorregresiva el gráfico sugiere poner al modelo un parámetro estacionario y uno no estacionario, debido a que la primera y la décima barra sobresalen de las bandas de confianza.

Entonces el modelo propuesto con una diferenciación estacional es:

$$\text{SARIMA}_{12}(1,0,3)(1,1,0)$$

Ahora vemos las pruebas para los coeficientes de este modelo:

**Tabla XXXVIII**

**Pruebas T\_Student para los parámetros del modelo SARIMA<sub>12</sub>  
(1,0,3)x(1,1,0) de la serie de la Prov. Azuay**

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.925	0.067	13.914	0
MA(1)	-0.092	0.141	-0.651	0.517
MA(2)	-0.011	0.142	-0.078	0.938
MA(3)	0.171	0.137	1.249	0.216
SAR(1)	-0.402	0.108	-3.708	0.0004
<b>MCE = 6430.54</b>				

Como vemos los tres parámetros son rechazados del modelo, puesto que el valor p de la prueba son mayores a 0.05 para los parámetros MA, motivo por el cual se va a eliminar estos parámetros y a continuación se verá la tabla de las pruebas para este nuevo modelo:

$$\text{SARIMA}_{12}(1,0,0)(1,1,0)$$



Tabla XXXIX

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,0)x(1,1,0) de la serie de la Prov. Azuay

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.916	0.048	18.919	0
SAR(1)	-0.386	0.106	-3.646	0.0005
<b>MCE = 6425.86</b>				

Los dos parámetros del modelo tienen sus valores p de la prueba mayores a 0.05, por lo que se dice que son significativamente diferentes de cero, entonces se acepta el modelo sin los parámetros MA.

Tabla XL

Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (1,0,0)(1,1,0) de

la serie de la Prov. Azuay

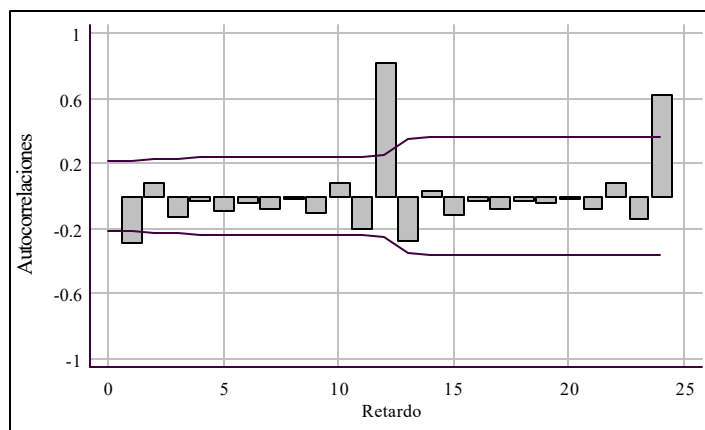
Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	-0.356	0.721
Corrida arriba y abajo	-0.047	0.962
Box-Pierce	21.806	0.472

El resultado de las pruebas de los residuos nos indican que éstos son independientes. Entonces por ahora se tiene que este modelo puede ser el adecuado para modelar la serie.

Ahora observemos que sucede al diferenciar la serie original, pero estacionariamente.

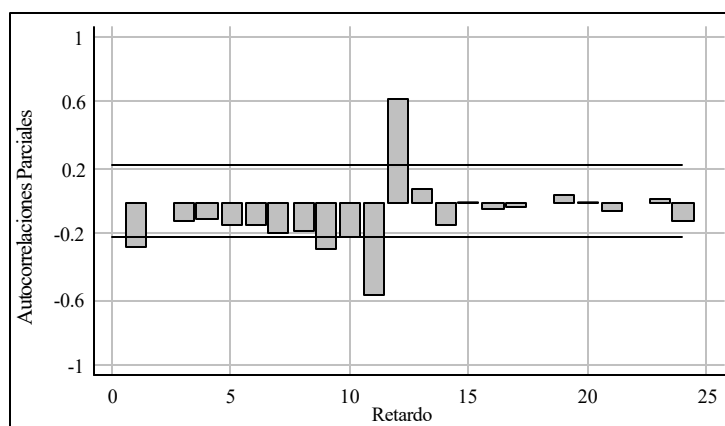
**Figura 3.48**

**Autocorrelaciones estimadas de la serie Prov. Azuay con una diferencia estacionaria**



Para la parte estacionaria media móvil(MA) del modelo el gráfico sugiere incluir un parámetro, pero para la parte estacional media móvil (SMA), nos sugiere incluir dos parámetros, debido a que sobresale la primera, la décima segunda y la vigésima cuarta barra.

**Figura 3.49**  
**Autocorrelaciones parciales estimadas de la serie Prov. Azuay**  
**con una diferencia estacionaria**



En el gráfico de las autocorrelaciones parciales, nos sugiere incluir dos parámetros uno estacional y otro no estacional, por la primera y la décima segunda barra que sobresalen de las bandas de confianza. Se ha ignorado las otras dos barras que sobresalen, pero se probará con este modelo, y si las prueba indican unas modificaciones a los parámetros se las realizará en ese momento.

Por los motivos explicados se tiene el siguiente modelo:

$$\text{SARIMA}_{12} (1,1,1) \times (1,0,2)$$

Tabla XLI

Pruebas T\_Student para los parámetros del modelo

SARIMA<sub>2</sub> (1,1,1)x(1,0,2) de la serie de la Prov. Azuay

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.192	1.812	0.106	0.916
MA(1)	0.108	1.798	0.060	0.952
SAR(1)	0.986	0.012	80.508	0
SMA(1)	0.402	0.104	3.876	0
SMA(2)	-0.166	0.095	-1.746	0.085
<b>MCE = 6041.37</b>				

La tabla anterior nos indica que tres parámetros no son significativamente diferentes de cero, ya que el valor p de las pruebas son mayores a 0.05, por lo tanto se va a eliminar los dos parámetros primeros AR. MA y se reducirá a un solo parámetro SMA. A continuación se presentará la tabla correspondiente a este modelo nuevo.

Tabla XLII

**Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>  
(0,1,0)x(1,0,1) de la serie de la Prov. Azuay**

Parámetro	Estimación	Error estándar	Valor T	Valor P
SAR(1)	0.988	0.015	64.237	0
SMA(1)	0.296	0.101	2.937	0.004
<b>MCE = 6038.29</b>				

Tabla XLIII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,0)x(1,0,1)  
de la serie de la Prov. Azuay**

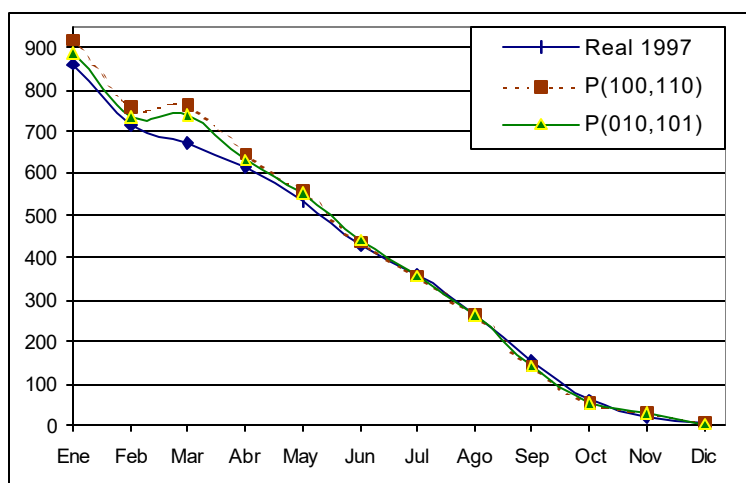
Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	0.778	0.437
Corrida arriba y abajo	-0.132	1.105
Box-Pierce	23.98	0.348

La tabla muestra que los valores p de las pruebas son mayores a 0.10, por lo que se afirma que los residuos del modelo propuesto son independientes.

Ahora se tiene dos modelos con características favorables, pero para poder seleccionar uno, es necesario realizar una comparación de los valores predichos con los reales.

**Figura 3.50**

**Comparación de los datos reales con los pronósticos de los modelos de la serie de la Prov. Azuay**



Como se observa en el gráfico, los valores del modelo que mas se aproxima a la serie real es  $SARIMA_{12}(0,1,0) \times (1,0,1)$ , además es el modelo que menor media cuadrática tiene entre los dos modelos propuestos. Por lo tanto se escoge este modelo para la serie de la provincia del Azuay.

A continuación se mostrarán los valores y gráficos de los pronósticos, así como los de los intervalos.

Tabla XLIV

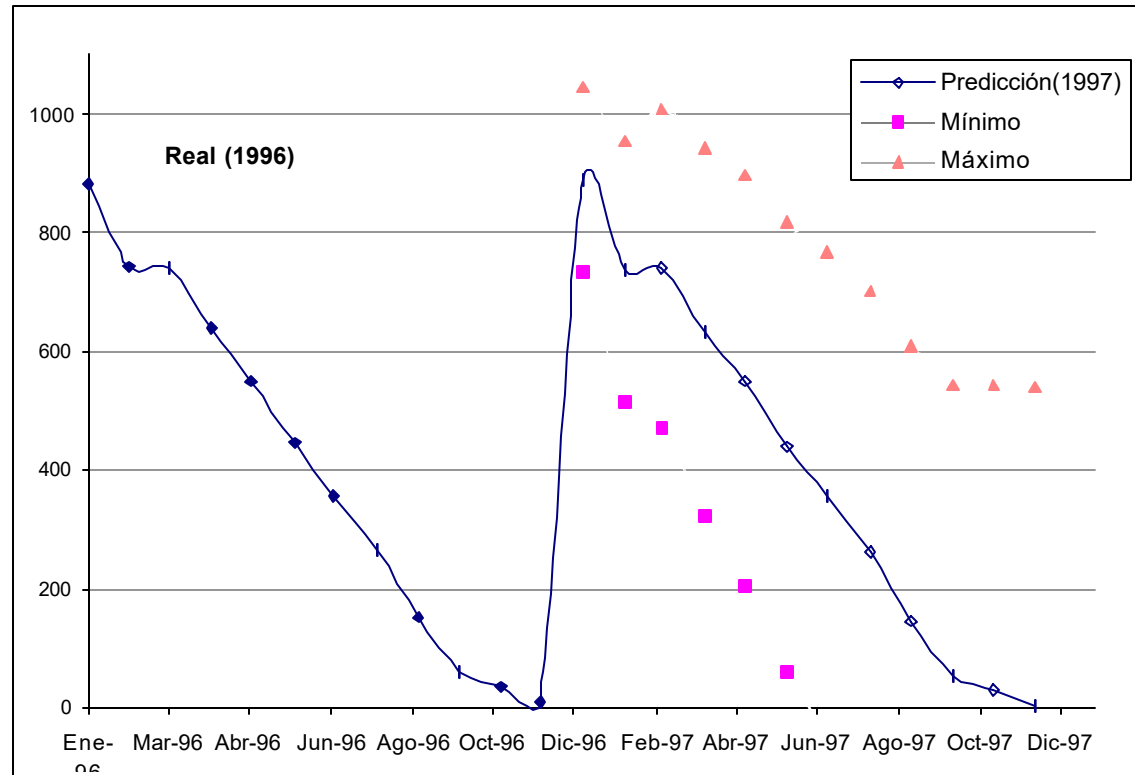
Predicción de valores de la serie de la Prov. Azuay con el modelo

$SARIMA_{12}(0,1,0)(1,0,1)$

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	859	890	735	1045	0.04
<i>Feb-97</i>	712	736	517	956	0.03
<i>Mar-97</i>	674	740	471	1009	0.10
<i>Abr-97</i>	616	633	322	943	0.03
<i>May-97</i>	533	551	204	898	0.03
<i>Jun-97</i>	430	439	59	819	0.02
<i>Jul-97</i>	362	357	-53	768	0.01
<i>Ago-97</i>	263	262	-177	701	0.01
<i>Sep-97</i>	154	144	-321	610	0.06
<i>Oct-97</i>	61	54	-436	545	0.11
<i>Nov-97</i>	22	29	-485	544	0.34
<i>Dic-97</i>	8	4	-535	540	0.50

Figura 3.51

Gráfico de las predicciones de la serie de la Prov. Azuay con el modelo SARIMA<sub>12(0,1,0)x(1,0,1)</sub>

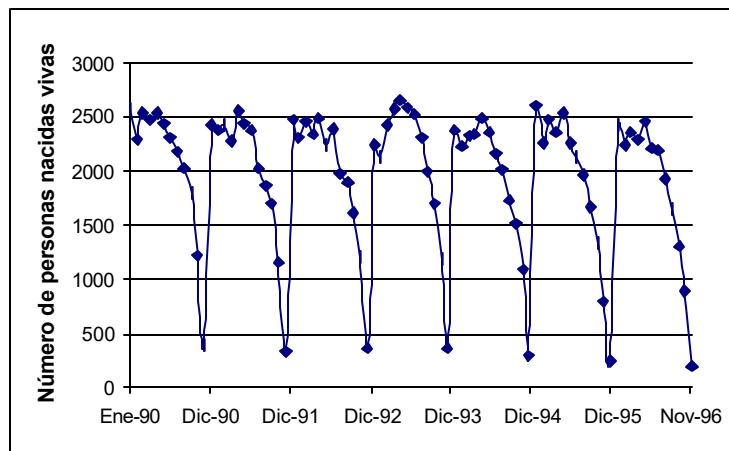




### 3.9 Análisis de la serie de la provincia Manabí

Figura 3.52

Gráfico de la serie de la Provincia Manabí



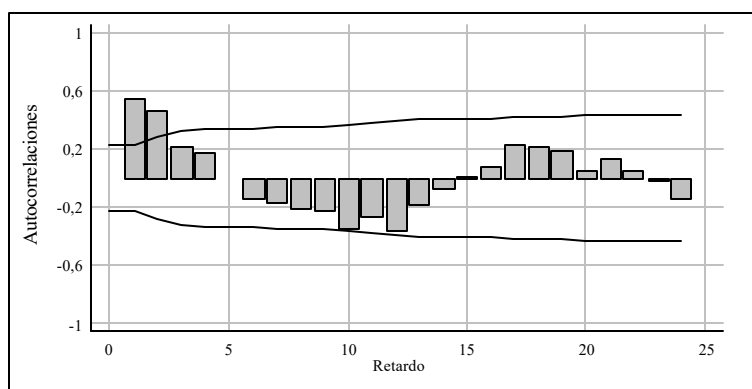
Manabí es la segunda provincia después del Guayas que mayor número de nacimientos tiene, en Mayo de 1993 es cuando ocurrió la mayor cantidad de nacimientos, en esta provincia, incluso superando las cifras de enero que es por lo general cuando más nacimientos se registran.

Con los gráficos de las autocorrelaciones obtendremos el modelo adecuado.

Se ha omitido el gráfico de las autocorrelaciones de la serie original, debido a que es el mismo que el de la serie total, y se procede a

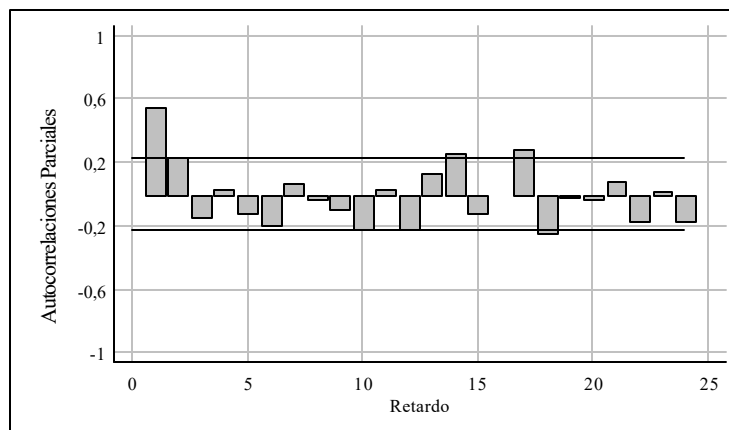
mostrar el gráfico de las autocorrelaciones estimadas de la serie diferenciada estacionalmente.

**Figura 3.53**  
**Autocorrelaciones de la serie Prov. Manabí con una**  
**diferenciación estacional**



Mediante el gráfico de las autocorrelaciones de la serie se puede decidir incluir dos parámetros al modelo en la parte media móvil, puesto que sobresalen las dos primeras barras de las bandas de confianza.

**Figura 3.54**  
**Autocorrelaciones Parciales de la serie Prov. Manabí con una**  
**diferenciación estacional**



Para el lado autorregresivo del modelo, se hace necesario un parámetro no estacional AR(1) por la primera barra, y también dos parámetros estacionales, SAR(2), ya que dos autocorrelaciones sobresalen de las bandas de confianza luego de 12 barras.

Entonces se tiene que considerar los siguientes modelos para la serie de la provincia de Manabí:

$$\text{SARIMA}_{12}(1,0,2)\times(2,1,0)$$

$$\text{SARIMA}_{12}(0,0,2)\times(2,1,0)$$

Cabe recalcar que al realizar los gráficos de las autocorrelaciones, con un solo parámetro estacional SAR(1), los residuos se salen

todavía de las bandas, por que se consideró necesario adicionar el otro parámetro como lo indicaba el gráfico de las autocorrelaciones parciales.

**Tabla XLV**

**Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>**

**(1,0,2)x(2,1,0) de la serie de la Prov. Manabí**

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.198	0.149	1.330	0.1881
MA(1)	-0.682	0.105	-6.471	0
MA(2)	-0.729	0.092	-7.950	0
SAR(1)	-0.586	0.1040	-5.636	0
SAR(2)	-0.831	0.065	-12.719	0
<b>MCE = 7631.9</b>				

Por los valores p expuestos en esta tabla se ve que no es necesario incluir el parámetro AR(1), ya que es el único valor p que no es menor a 0.05, por lo que se dice que no es diferente de cero.

Por otro lado, en el siguiente modelo todos los parámetros son significativamente diferentes de cero (Tabla XLVI), razón por la cual se elimina la posibilidad de escoger el modelo anterior y se selecciona el segundo, para el mismo que se realizará la prueba de sus residuos.

Tabla XLVI

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>  
(0,0,2)x(2,1,0) de la serie de la Prov. Manabí

Parámetro	Estimación	Error estándar	Valor T	Valor P
MA(1)	-0.772	0.077	-10.008	0
MA(2)	-0.768	0.079	-9.602	0
SAR(1)	-0.594	0.106	-5.629	0
SAR(2)	-0.789	0.073	-10.873	0
<b>MCE = 7641.1</b>				

Figura 3.55

Autocorrelaciones residuales de la serie Prov. Manabí con el  
modelo SARIMA<sub>2</sub>(0,0,2)x(2,1,0)

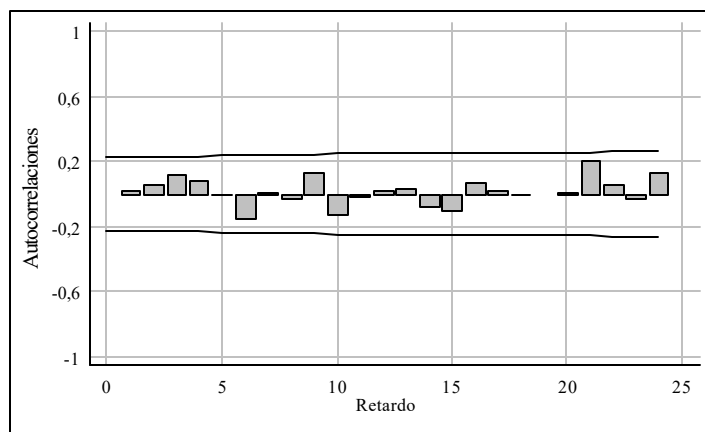


Figura 3.56

**Autocorrelaciones Parciales residuales de la serie Prov. Manabí  
con el modelo SARIMA<sub>2</sub>(0,0,2)x(2,1,0)**

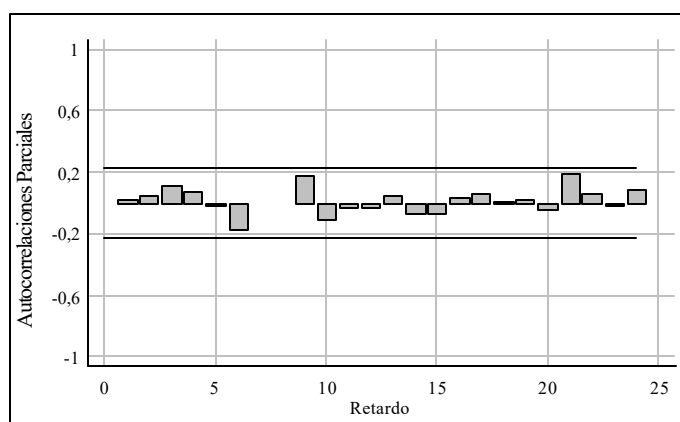


Tabla XLVII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,0,2)(2,1,0) de  
la serie de la Prov. Manabí**

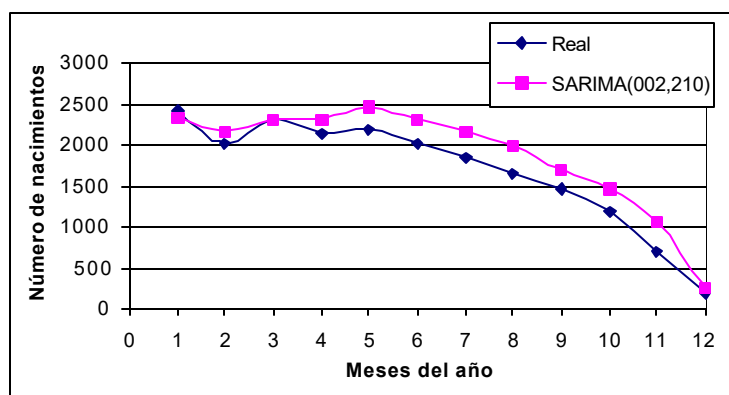
Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	-0.119	1.094
Corrida arriba y abajo	0.519	0.604
Box-Pierce	12.024	0.915

Este cuadro muestra que el valor p de cada una de las pruebas son mayores que 0.10, por lo que se afirma que los residuos del modelo son independientes.

En la siguiente figura se muestra una comparación de los valores reales con los valores que nos proporciona el modelo escogido.

**Figura 3.57**

**Comparación de la serie real de la provincia Manabí con los pronósticos del modelo SARIMA<sub>12</sub> (0,0,2)x(2,1,0)**



Una vez que hemos observado que el modelo se ajusta bien a los valores reales, y el modelo ha pasado las pruebas se procede a graficar sus pronósticos y los intervalos (dentro de los cuales se hayan los pronósticos) en la Figura 3.58 y se muestran sus valores en la Tabla XLVIII.

Tabla XLVIII

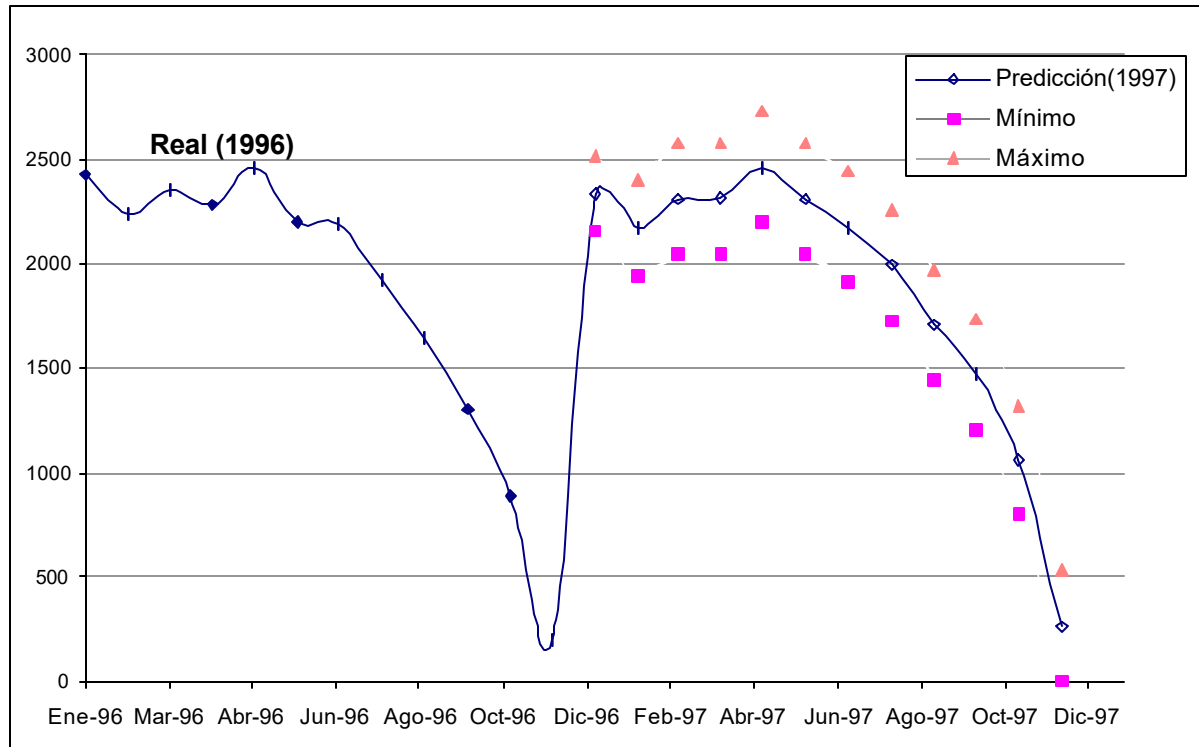
## Predicción de valores de la serie de la Prov. Manabí

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	2424	2337	2158	2516	0.04
<i>Feb-97</i>	2029	2171	1945	2397	0.07
<i>Mar-97</i>	2314	2311	2046	2576	0.00
<i>Abr-97</i>	2153	2314	2049	2579	0.07
<i>May-97</i>	2182	2465	2201	2730	0.13
<i>Jun-97</i>	2012	2313	2048	2577	0.15
<i>Jul-97</i>	1854	2177	1912	2442	0.17
<i>Ago-97</i>	1654	1991	1726	2256	0.20
<i>Sep-97</i>	1461	1706	1442	1971	0.17
<i>Oct-97</i>	1193	1469	1205	1734	0.23
<i>Nov-97</i>	703	1060	795	1324	0.51
<i>Dic-97</i>	195	262	-3	526	0.34



Figura 3.58

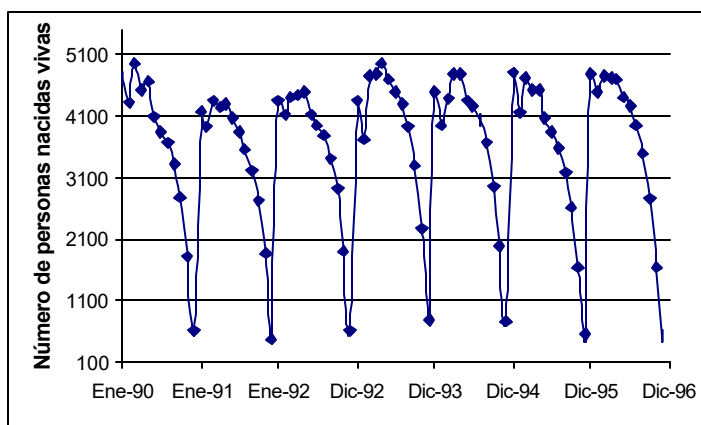
Gráfico de las predicciones de la serie de la Prov. Manabí con el modelo SARIMA<sub>12(0,0,2)x(2,1,0)</sub>



### 3.10 Análisis de la serie de la provincia Guayas

Figura 3.59

Gráfico de la serie de la Provincia Guayas

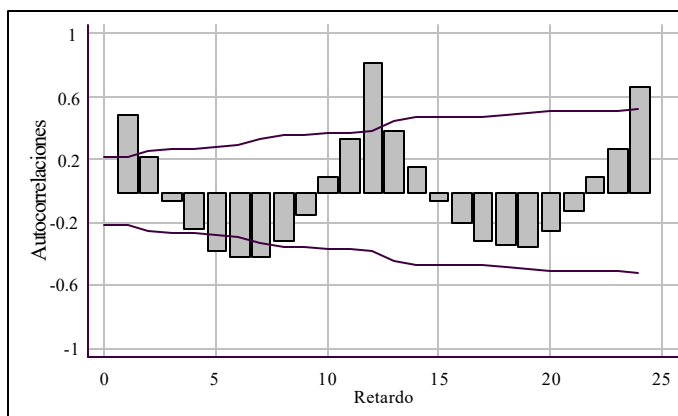


Considerando la suma de todos los meses de cada año, en 1996 fue el año en el que ocurrieron la mayor cantidad de nacimientos de todo este período.

Pero si se toma un análisis por meses se tiene que en Marzo de 1990 nació la mayor cantidad de niños en la provincia del Guayas, aproximadamente 5000 personas nacieron. Por el contrario en Diciembre de 1991, fue la fecha donde menos personas nacieron en esta provincia.

Figura 3.60

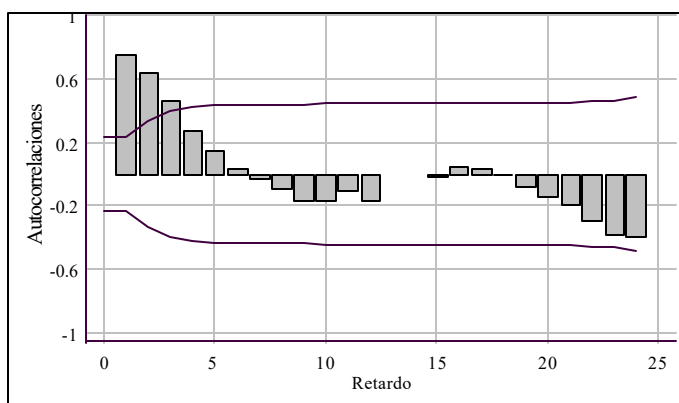
## Autocorrelaciones de la serie de la Prov. Guayas



Se puede verificar mediante el gráfico de las autocorrelaciones de la serie de la provincia del Guayas que es igual al de la serie del Total de nacimientos en el Ecuador, lo mismo ocurre con el gráfico de las autocorrelaciones parciales, así que se continúa el análisis con la observación de las autocorrelaciones de la serie diferenciada estacionalmente debido al comportamiento que tiene la serie cada 12 meses, esto es que hay barras seguidas que se salen de las bandas dentro de cada cierto periodo.

Figura 3.61

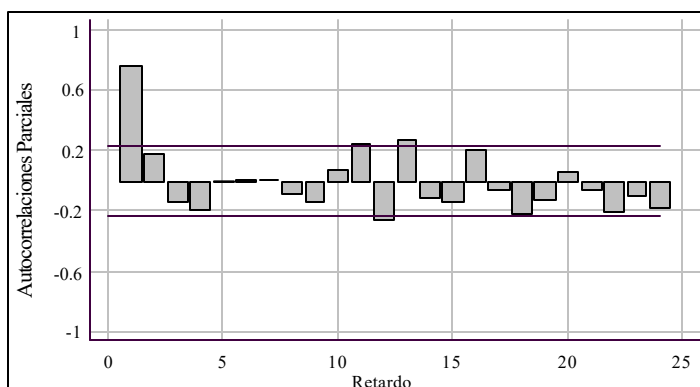
**Autocorrelaciones de la serie de la Prov. Guayas con una  
diferenciación estacional**



Como se observa, las tres primeras barras sobresalen de las bandas, así se propone incluir tres parámetros estacionarios para el lado de medias móviles del modelo, MA.

Figura 3.62

**Autocorrelaciones parciales de la serie de la Prov. Guayas con  
una diferenciación estacional**



Como se observa en el gráfico de las autocorrelaciones parciales, se hace necesario adicionar un parámetro no estacional autorregresivo en el modelo, y también dos parámetros estacionales, debido a que sobresalen la primera, la décima segunda y la décima tercera barra de las autocorrelaciones parciales.

Con estas justificaciones para proponer los parámetros del modelo, se tiene por escoger de tres posibilidades:

$SARIMA_{12}(1,0,3) \times (2,1,0)$

$SARIMA_{12}(1,0,0) \times (2,1,0)$

$SARIMA_{12}(1,0,0) \times (1,1,0)$

Se ha propuesto el primer modelo, tratando de cubrir los requerimientos que indican los gráficos de las autocorrelaciones, el segundo modelo es parecido al primero, pero sin los parámetros MA, y el tercero se le ha disminuido un parámetro AR, para ir eliminando las posibilidades para escoger un modelo adecuado se les aplicará la prueba de los parámetros.

Tabla XLIX

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,3)x(2,1,0) de la serie de la Prov. Guayas

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.721	0.161	4.473	0
MA(1)	-0.204	0.204	-0.999	0.321
MA(2)	-0.212	0.181	-1.172	0.246
MA(3)	-0.018	0.164	-0.112	0.911
SAR(1)	-0.717	0.123	-5.833	0
SAR(2)	-0.165	0.133	-1.241	0.219
<b>MCE = 27656.8</b>				

Como se muestra en esta tabla, para este modelo los parámetros del lado media móvil no son diferentes de cero, entonces desde ya se considera eliminar la posibilidad de escoger este modelo.

Ahora veamos si los parámetros de los otros modelos pasan la prueba de ser diferentes de cero.

Tabla L

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,0)x(2,1,0) de la serie de la Prov. Guayas

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.832	0.065	12.890	0
SAR(1)	-0.674	0.117	-5.780	0
SAR(2)	-0.274	0.131	-2.089	0.040
<b>MCE = 26945.2</b>				

Los parámetros de este modelo se dice que son significativamente diferentes de cero, puesto que su valor p es menor a 0.05, entonces no se lo puede descartar aún a este modelo.

Tabla LI

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,0)x(1,1,0) de la serie de la Prov. Guayas

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.845	0.062	13.715	0
SAR(1)	-0.589	0.101	-5.844	0
<b>MCE = 27952.1</b>				

Para este modelo la prueba de los parámetros también indica que son diferentes de cero.

Ahora se probará que los residuos son independientes.

**Tabla LII**

**Pruebas para los Residuos del modelo SARIMA<sub>42</sub> (1,0,0)x(2,1,0)**

**de la serie de la Prov. Guayas**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.593	0.553
Corrida arriba y abajo	-0.613	0.540
Box-Pierce	16.443	0.744

**Tabla LIII**

**Pruebas para los Residuos del modelo SARIMA<sub>42</sub> (1,0,0)x(1,1,0)**

**de la serie de la Prov. Guayas**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.119	0.906
Corrida arriba y abajo	-1.180	0.238
Box-Pierce	16.531	0.789

Los dos modelos muestran en las tres pruebas un valor p mayor a 0.10, por lo que se demuestra que los residuos son independientes. Hasta ahora no se ha eliminado mediante esta prueba ninguno de estos dos modelos, por lo que se pasará a escoger el modelo por el



criterio del mayor poder predictivo, seleccionando el modelo que tenga menor varianza.

Revisando las tablas L y LI de los modelos que no se han eliminado, se tiene que la menor MCE la tiene el modelo SARIMA<sub>12</sub>(1,0,0)x(2,1,0).

**Figura 3.63**

**Autocorrelaciones residuales de la serie Prov. Guayas con el modelo SARIMA<sub>12</sub>(1,0,0)x(2,1,0)**

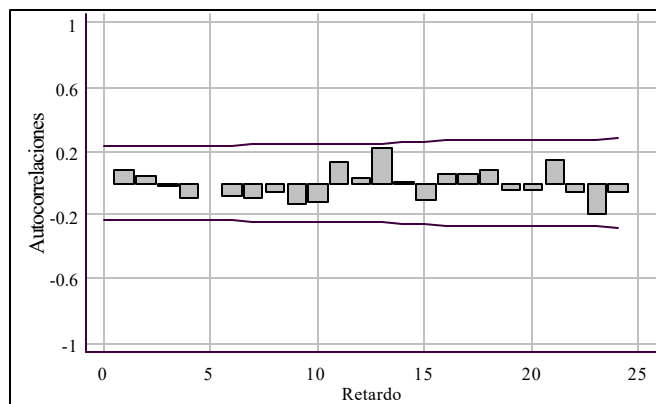
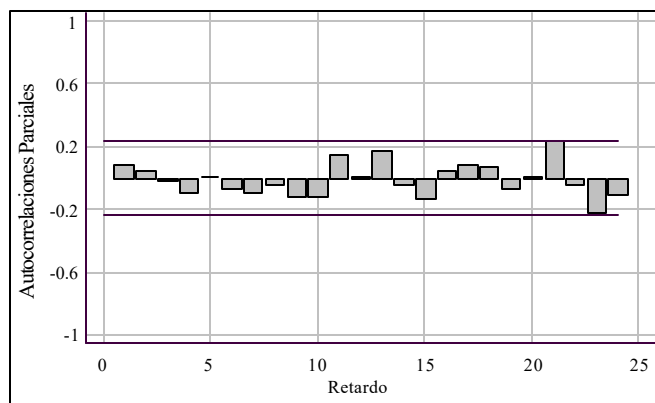


Figura 3.64

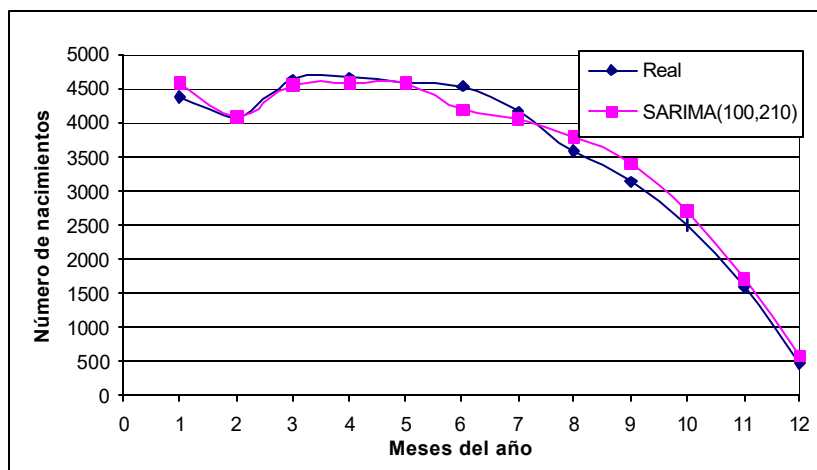
**Autocorrelaciones parciales residuales de la serie Prov. Guayas  
con el modelo SARIMA<sub>2</sub>(1,0,0)x(2,1,0)**



Se muestra en la siguiente figura que los valores de nos proporciona el modelo seleccionado se ajusta bien a los valores reales de la serie de la provincia del Guayas.

Figura 3.65

**Comparación de la serie real de la provincia Guayas con los  
pronósticos del modelo SARIMA(1,0,0)(2,1,0)<sub>12</sub>**



Luego de revisar la similitud que existe entre los valores reales y los pronósticos, se mostrará el gráfico de las predicciones con los intervalos de confianza, y sus respectivos valores en siguiente tabla.

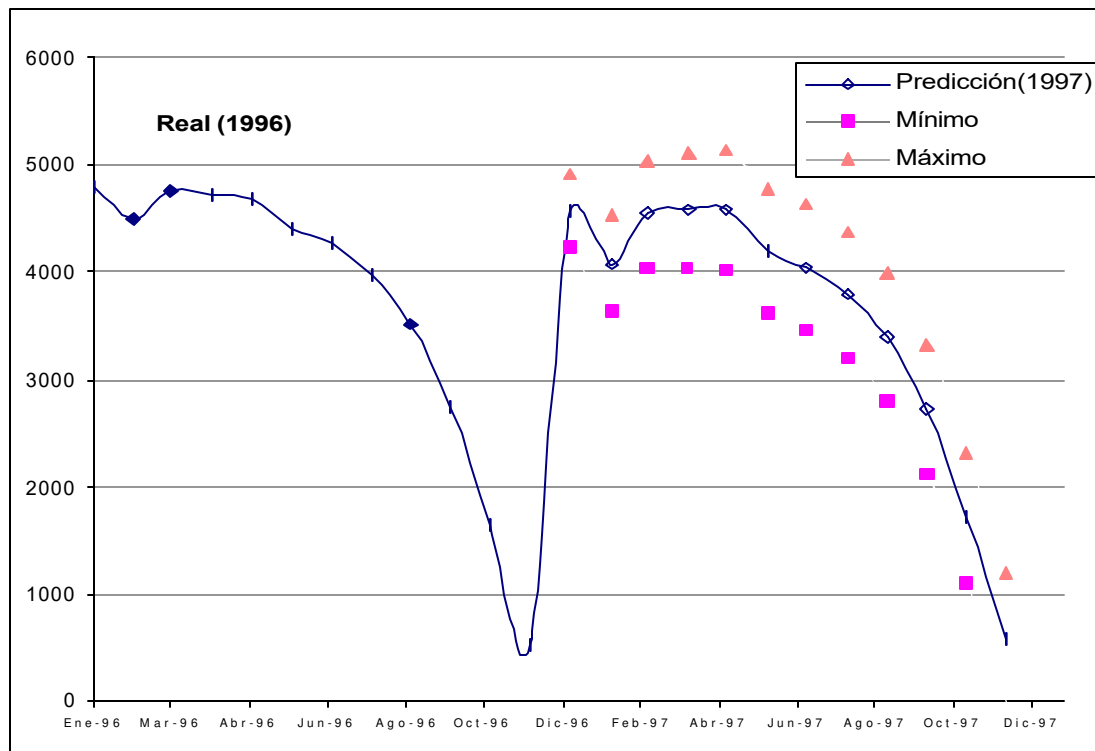
**Tabla LIV**

**Predicción de valores de la serie de la Prov. Guayas**

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	4366	4574	4236	4912	0.05
<i>Feb-97</i>	4092	4082	3643	4522	0.00
<i>Mar-97</i>	4634	4545	4047	5043	0.02
<i>Abr-97</i>	4670	4582	4047	5116	0.02
<i>May-97</i>	4575	4579	4020	5138	0.00
<i>Jun-97</i>	4516	4198	3623	4772	0.07
<i>Jul-97</i>	4164	4048	3463	4634	0.03
<i>Ago-97</i>	3587	3788	3195	4381	0.06
<i>Sep-97</i>	3130	3392	2794	3990	0.08
<i>Oct-97</i>	2512	2722	2120	3324	0.08
<i>Nov-97</i>	1570	1712	1108	2317	0.09
<i>Dic-97</i>	476	582	-24	1187	0.22

Figura 3.66

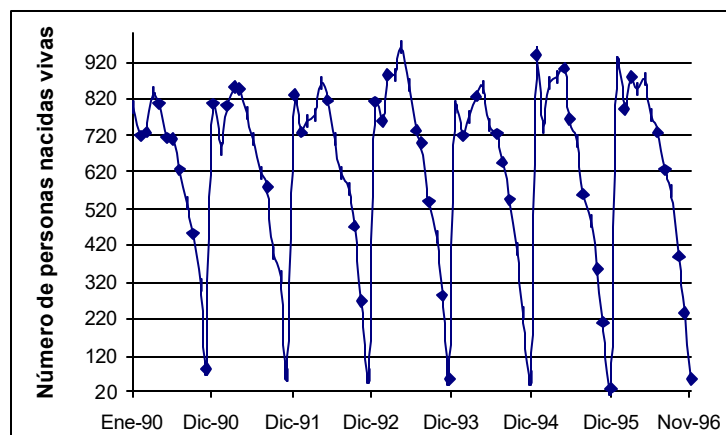
Gráfico de las predicciones de la serie de la Prov. Guayas con el modelo SARIMA<sub>12</sub>(1,0,0)x(2,1,0)



### 3.11 Análisis de la serie de la provincia El Oro

Figura 3.67

Gráfico de la serie de la Provincia El Oro



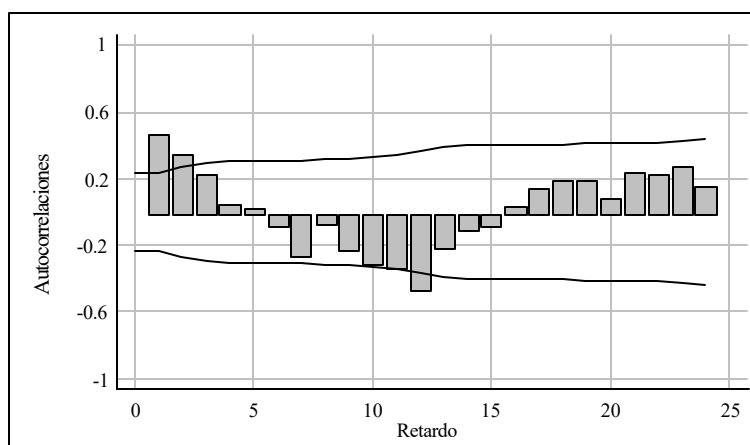
En la provincia de El Oro se obtuvo la mayor cantidad de nacimientos en Mayo de 1993, seguido por Enero de 1995, mientras que en Diciembre de 1995 es la fecha en la que se registraron menos nacimientos dentro del período que se realiza este análisis (1990-1996).

Si se observa el gráfico de esta provincia se puede verificar que en enero de 1995 fue donde se registró una de las mayores cantidades de nacimientos, y en diciembre de ese mismo año se registró la menos cantidad de nacimientos de este período.

El gráfico de las autocorrelaciones y las autocorrelaciones parciales de la serie original, sin ninguna alteración, es el mismo que para la serie del Total de nacimientos que ya se analizó, por lo que se ha decidido también diferenciar estacionalmente a esta serie.

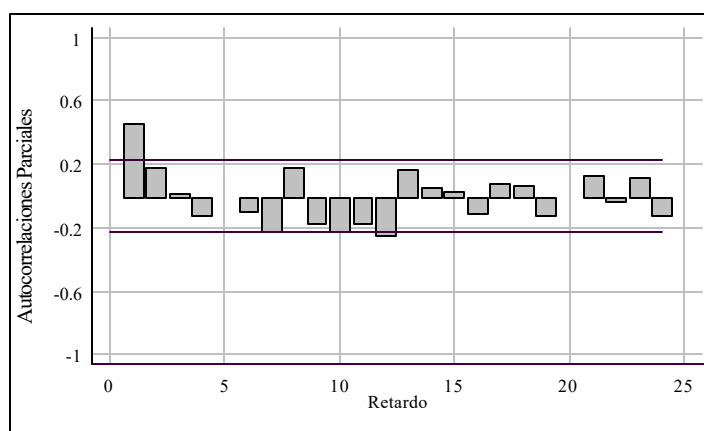
**Figura 3.68**

**Autocorrelaciones de la serie de la Prov. El Oro con una  
diferencia estacional**



Observando la Figura 3.68 se puede decidir la cantidad de parámetros necesarios para el modelo, entonces se considera necesario adicionar en la parte media móvil del modelo dos parámetros no estacionales y uno estacional, puesto que la primera, segunda y décima segunda barra de las autocorrelaciones sobresalen de las bandas.

**Figura 3.69**  
**Autocorrelaciones parciales de la serie de la Prov. El Oro con una**  
**diferencia estacional**



Como lo muestra la figura, sólo la primera barra sale de las bandas de confianza, así que se adiciona un parámetro AR no estacional; mientras que para la parte estacional no es necesario algún parámetro.

Por los motivos expuestos, se propone los siguientes modelos:

$$\text{SARIMA}_{12}(1,0,2)\times(0,1,1)$$

$$\text{SARIMA}_{12}(1,0,0)\times(0,1,1)$$

Tabla LV

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,2)x(0,1,1) de la serie de la Prov. El Oro

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.423	0.320	1.324	0.189
MA(1)	0.023	0.327	0.069	0.945
MA(2)	-0.174	0.171	-1.019	0.312
SMA(1)	0.826	0.047	17.593	0
<b>MCE = 1414.9</b>				

Para este modelo propuesto, la prueba T\_Student nos indica que los parámetros estacionarios MA y el AR no son significativamente diferentes de cero, así que probaremos con el otro modelo.

Tabla LVI

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,0)x(0,1,1) de la serie de la Prov. El Oro

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.467	0.104	4.486	0
SMA(1)	0.838	0.043	19.387	0
<b>MCE = 1380.5</b>				



Como se ve en esta tabla, que quitándole los parámetros MA, las pruebas nos indican que se puede aceptar los parámetros propuestos, ya que son significativamente diferentes de cero, pero todavía queda por realizar la prueba de los residuos para este modelo para asegurarnos que es el adecuado para modelar esta serie de la provincia de El Oro.

**Tabla LVII**

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (1,0,0)x(0,1,1)  
de la serie de la Prov. El Oro**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	0.593	0.553
Corrida arriba y abajo	0.802	0.422
Box-Pierce	14.667	0.876

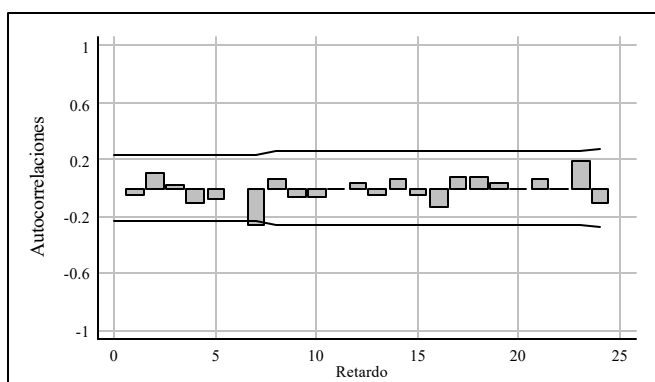
Consiguientemente, todas las pruebas nos indican que los residuos del modelo son independientes, ya que el valor p de la prueba es mayor a 0.10. A continuación se muestran los gráficos de las autocorrelaciones de sus residuos.

Cabe indicar que si a la serie se le aplica el modelo seleccionado de la serie Total, SARIMA<sub>12</sub>(0,1,0)x(1,1,0), que sería diferenciar una vez la serie estacionariamente, las autocorrelaciones de los residuos no se mantienen dentro de las bandas de confianza, entonces se descarta

este modelo, además la varianza que muestra es mayor que la de el modelo seleccionado para la serie de la provincia de El Oro.

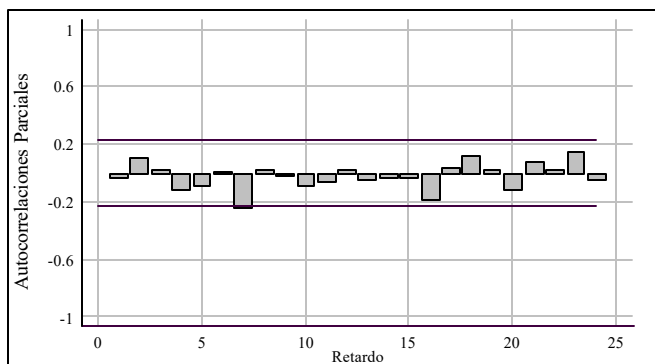
**Figura 3.70**

**Autocorrelaciones parciales residuales de la serie Prov. El Oro  
con el modelo SARIMA<sub>2</sub>(1,0,0)x(0,1,1)**



**Figura 3.71**

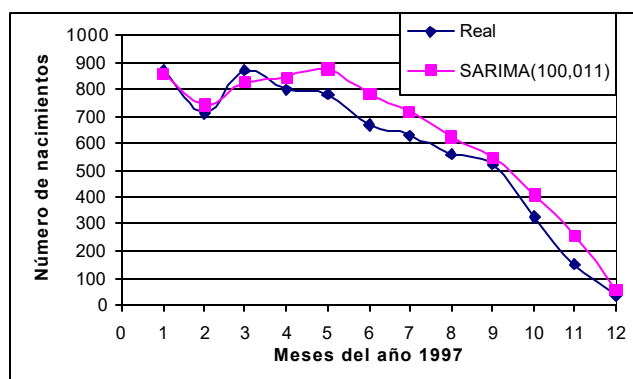
**Autocorrelaciones parciales residuales de la serie Prov. El Oro  
con el modelo SARIMA<sub>2</sub>(1,0,0)x(0,1,1)**



Tanto el gráfico de las autocorrelaciones como el de las autocorrelaciones parciales muestran que todas las barras se encuentran dentro de las bandas de confianza, por lo tanto se ha seleccionado este modelo: SARIMA<sub>12</sub> (1,0,0)x(0,1,1).

**Figura 3.72**

**Comparación de la serie real de la provincia El Oro con los pronósticos del modelo SARIMA(1,0,0)(0,1,1)<sub>12</sub>**



Los valores que da el modelo aunque no estén tan ligados a los reales, se ha seleccionado este modelo porque es el que tiene menor varianza, y sus residuos quedan dentro de las bandas como ya se indico anteriormente.

Enseguida se muestra el gráfico de las predicciones con los valores de los intervalos de confianza y sus respectivos valores en la siguiente tabla.

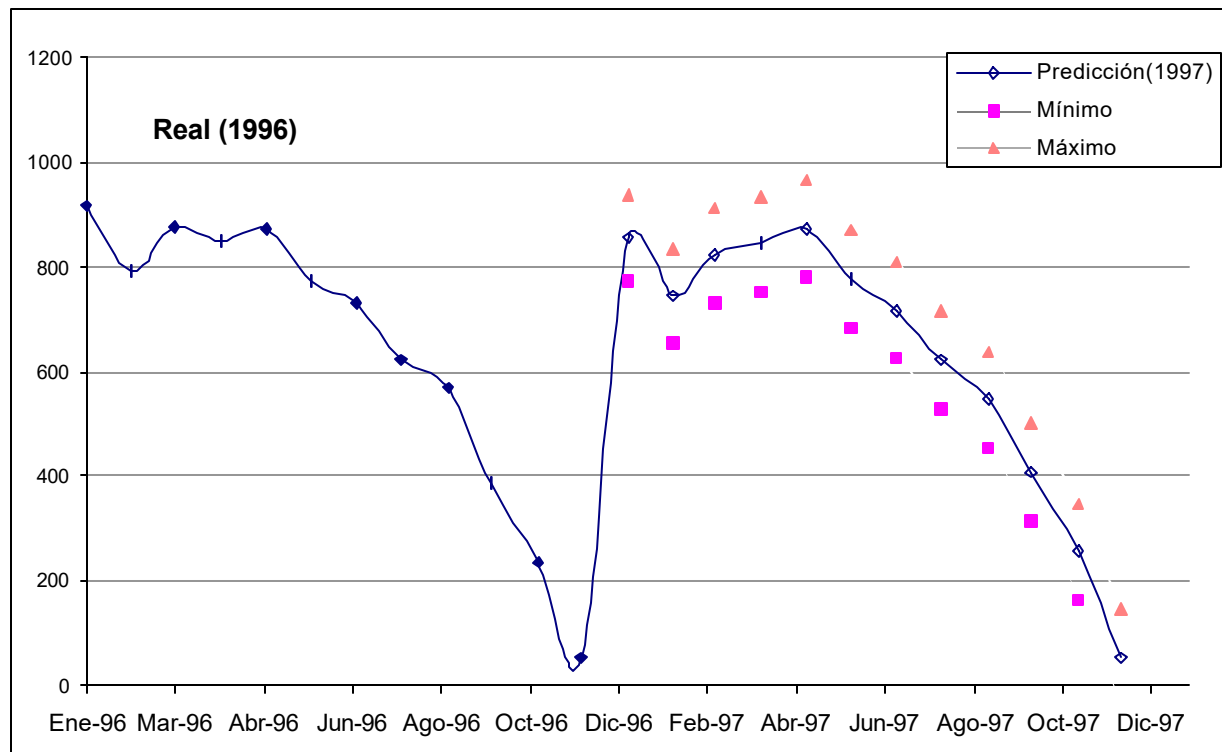
Tabla LVIII

## Predicción de valores de la serie de la Prov. El Oro

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	866	856	775	937	0.01
<i>Feb-97</i>	711	744	655	834	0.05
<i>Mar-97</i>	865	823	732	915	0.05
<i>Abr-97</i>	797	844	752	936	0.06
<i>May-97</i>	779	874	782	966	0.12
<i>Jun-97</i>	666	777	686	869	0.17
<i>Jul-97</i>	626	718	626	809	0.15
<i>Ago-97</i>	562	623	532	715	0.11
<i>Sep-97</i>	523	547	455	639	0.05
<i>Oct-97</i>	325	408	317	500	0.26
<i>Nov-97</i>	149	257	166	349	0.73
<i>Dic-97</i>	37	55	-37	146	0.48

Figura 3.73

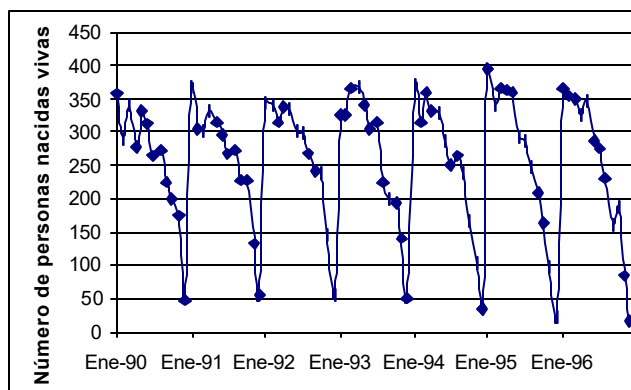
Gráfico de las predicciones de la serie de la Prov. El Oro con el modelo SARIMA<sub>12</sub>(1,0,0)x(0,1,1)



### 3.12 Análisis de la serie de la provincia Napo

Figura 3.74

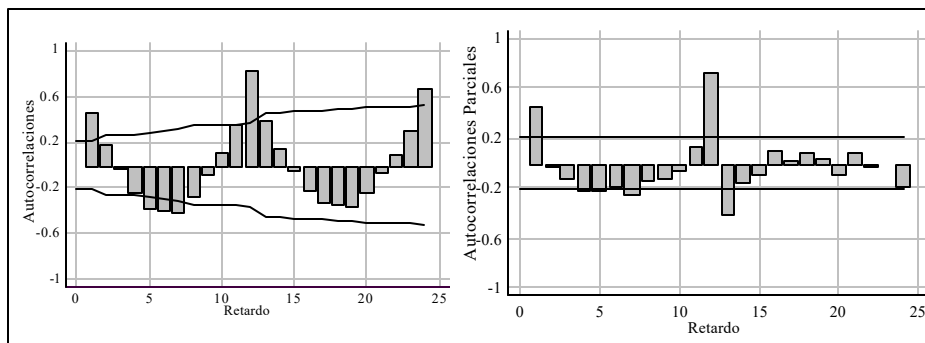
Gráfico de la serie de la Provincia Napo



En la provincia Napo del Ecuador, dentro del período de 1990 a 1996, la mayor cantidad de personas nacidas fue en Enero de 1995, esta cantidad fue aproximadamente de 400 personas; por otro lado en Diciembre de 1996 fue la fecha en la que se registraron la menor cantidad de nacimientos dentro de este periodo. También se puede acotar que del año 1994 a 1995 en el mes de Diciembre hay una marcada diferencia, puesto que desde ese año la curva va descendiendo mas rápido que los años anteriores.

Enseguida se muestran los gráficos correspondientes a las autocorrelaciones de la serie original.

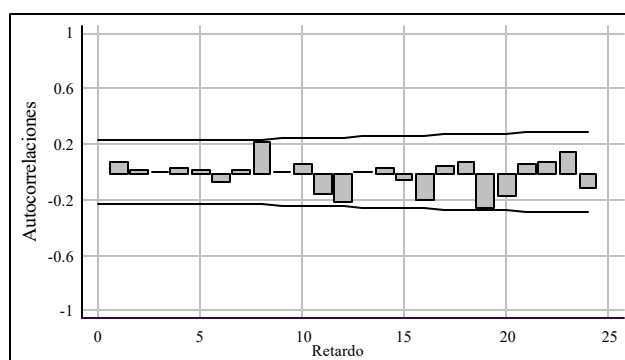
**Figura 3.75**  
**Autocorrelaciones y autocorrelaciones parciales de la serie de la**  
**Prov. Napo**



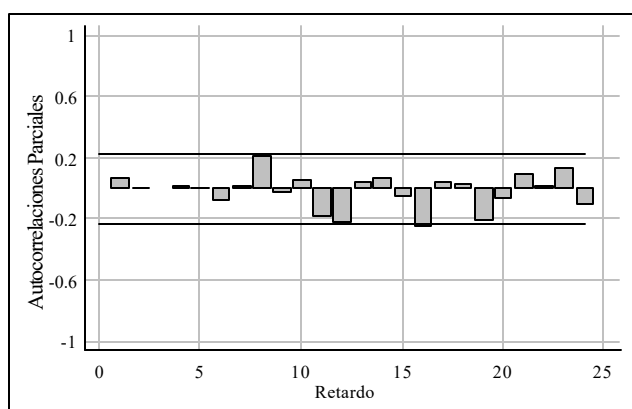
Se ve que la serie de esta provincia del Oriente, también tiene el mismo comportamiento que la del Total, así que se procede a diferenciarla estacionalmente por sus barras seguidas que sobresalen de las bandas cada cierto periodo. A continuación se muestra el gráfico de las autocorrelaciones de la serie diferenciada estacionalmente.

**Figura 3.76**

**Autocorrelaciones estimadas de la serie de la Prov. Napo con una diferencia estacional**

**Figura 3.77**

**Autocorrelaciones parciales estimadas de la serie de la Prov. Napo con una diferencia estacional**



El comportamiento de las autocorrelaciones en la serie de la provincia Napo, es diferente, como se observa en los gráficos anteriores,



ninguna barra sobresale de las bandas de confianza, motivo por el cual se va proponer los siguientes:

Aplicando los dos tipos de diferenciación

$$\text{SARIMA}_{12}(0,1,1)\times(1,1,1)$$

Sin diferenciación se propone un modelo con estos parámetros, mediante la observación de sus autocorrelaciones y sus autocorrelaciones parciales de la serie original.

$$\text{SARIMA}_{12}(1,0,1)\times(1,0,1)$$

Con sólo una diferenciación no estacional

$$\text{SARIMA}_{12}(0,1,1)\times(1,0,1)$$

A continuación se realizan las pruebas de los parámetros de los modelos propuestos.

**Tabla LIX**

**Pruebas T\_Student para los parámetros del modelo**

**SARIMA<sub>2</sub> (0,1,1)x(1,1,1) de la serie de la Prov. Napo**

Parámetro	Estimación	Error estándar	Valor T	Valor P
MA(1)	0.976	0.019	50.166	0
SAR(1)	0.565	0.118	4.776	0
SMA(1)	0.952	0.044	21.487	0
<b>MCE = 502.7</b>				

Tabla LX

Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>

(1,0,1)x(1,0,1) de la serie de la Prov. Napo

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.908	0.076	11.940	0
MA(1)	0.857	0.122	7.027	0
SAR(1)	1.006	0.008	127.629	0
SMA(1)	0.396	0.091	4.369	0
<b>MCE = 537.9</b>				

Tabla LXI

Pruebas T\_Student para los parámetros del modelo

SARIMA<sub>2</sub>(0,1,1)x(1,0,1) de la serie de la Prov. Napo

Parámetro	Estimación	Error estándar	Valor T	Valor P
MA(1)	0.833	0.042	19.808	0
SAR(1)	1.037	0.011	96.980	0
SMA(1)	0.684	0.063	10.924	0
<b>MCE = 498.9</b>				

Cada uno de los parámetros de los tres modelos propuestos son significativamente diferentes de cero, pero veamos que sucede con las pruebas de los residuos de cada modelo.

Tabla LXII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub>(0,1,1)x(1,1,1) de la serie de la Prov. Napo**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-3.009	0.003
Corrida arriba y abajo	0.143	0.887
Box-Pierce	19.211	0.508

Tabla LXIII

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (1,0,1)x(1,0,1) de la serie de la Prov. Napo**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	-0.110	1.087
Corrida arriba y abajo	-0.305	0.760
Box-Pierce	25.779	0.173

Tabla LXIV

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,1)x(1,0,1)**  
**de la serie de la Prov. Napo**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	2.111	0.035
Corrida arriba y abajo	0.395	0.693
Box-Pierce	28.873	0.117

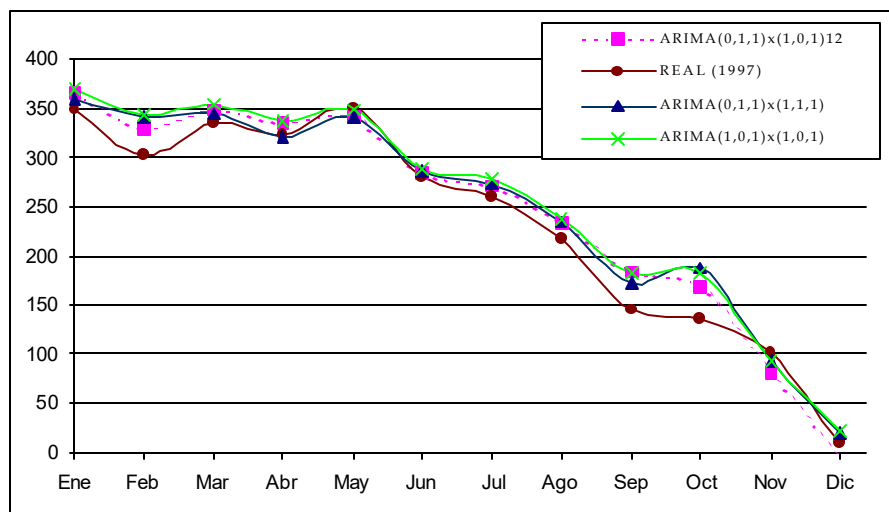
Como se puede observar, las pruebas indican que sólo en un modelo ( SARIMA<sub>12</sub> (1,0,1)x(1,0,1) ) se puede afirmar que sus residuos son independientes, en los otros dos modelos falla la prueba Corrida arriba y abajo de la mediana. Pero es necesario todavía verificar que sucede con los valores que nos dan sus pronósticos, es decir observar el modelo que ajusta mejor los datos.

Si se revisa los cuadros anteriores, podemos verificar que el modelo de menor MCE es SARIMA<sub>12</sub> (0,1,1)x(1,0,1).

En la figura 3.56, se realiza una comparación de los pronósticos de cada modelo con los valores reales del año 1997.

Figura 3.78

Comparación de los pronósticos de los modelos propuestos con valores reales del año 1997, de la Prov. Napo



Como se observa la línea punteada se aproxima más que las otras a la real, la cual representa al modelo **SARIMA<sub>12</sub> (0,1,1)x(1,0,1)**, las otras líneas que representan los demás modelos se alejan mucho de los reales, en especial el modelo que incluye los dos tipos de diferenciaciones, así que se confirma que no es necesario diferenciar la serie dos veces, sólo con una diferenciación estacionaria, se ha podido hallar el modelo adecuado. Por lo tanto se escoge este modelo para la serie del número de nacimientos mensuales de la provincia Napo, sus valores y gráficos se muestran a continuación.

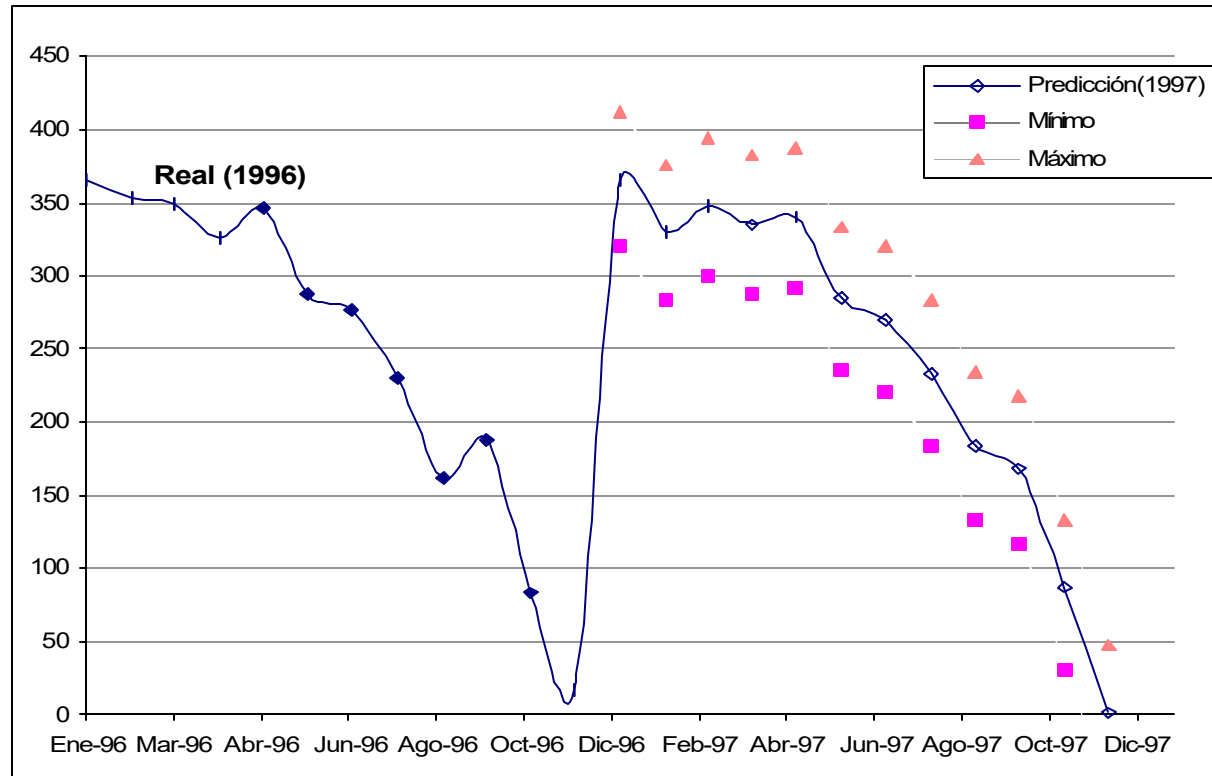
Tabla LXV

## Predicción de valores de la serie de la Prov. Napo

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	347	365	320	411	0.05
<i>Feb-97</i>	302	329	283	376	0.09
<i>Mar-97</i>	335	348	301	395	0.04
<i>Abr-97</i>	322	335	287	382	0.04
<i>May-97</i>	349	340	292	388	0.03
<i>Jun-97</i>	281	285	236	333	0.01
<i>Jul-97</i>	259	270	221	320	0.04
<i>Ago-97</i>	217	233	183	283	0.08
<i>Sep-97</i>	144	184	134	235	0.28
<i>Oct-97</i>	136	168	117	219	0.23
<i>Nov-97</i>	100	87	30	133	0.13
<i>Dic-97</i>	10	4	-49	48	0.60

Figura 3.79

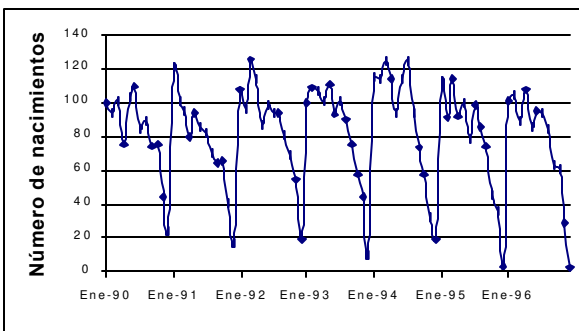
Gráfico de las predicciones de la serie de la Prov. Napo con el modelo  $SARIMA_{12}(0,1,1) \times (1,0,1)$



### 3.13 Análisis de la serie de la provincia Pastaza

**Figura 3.80**

**Gráfico de la serie de la Provincia Pastaza**



En la provincia Pastaza, dentro del período de estudio, la mayor cantidad de nacimientos ocurrieron en Marzo de 1992, y por el contrario la menor cantidad ocurrió en Diciembre y Noviembre de 1996. Es importante señalar que la mayoría de los meses del año 1994, son en los que se registran las mas altas cantidades de nacimientos, en comparación con los otros años, esto se verifica en la figura, puesto que la curva en el año 1994 se eleva mucho mas que la de los otros años.

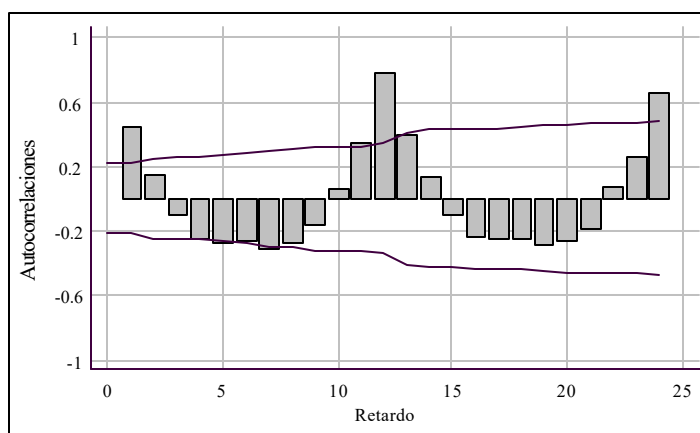
Como las series anteriores, se puede asegurar mediante el gráfico que esta serie tiene una estacionalidad de 12, motivo por el cual se trabaja con modelos SARIMA.



Seguindo con el análisis, a continuación se observa el gráfico de las autocorrelaciones para decidir cuantos parámetros son necesarios en esta serie.

**Figura 3.81**

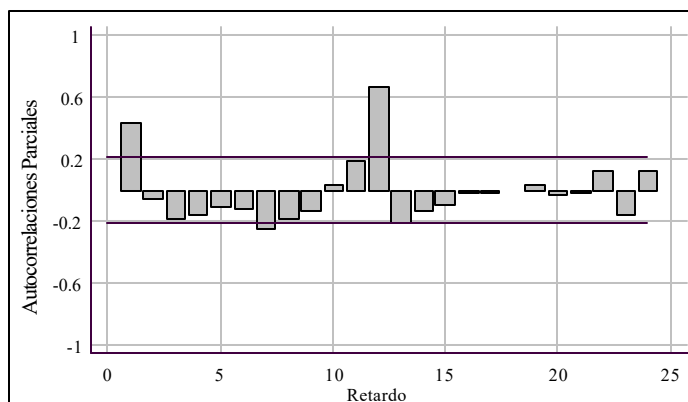
**Autocorrelaciones de la serie de la Prov. Pastaza**



Mediante este gráfico se puede afirmar que es necesario incluir en el modelo un parámetro no estacional, y dos estacionales, esto se justifica porque la primera (no estacional), décima segunda y la vigésima cuarta (estacional) barra sobresalen de las bandas de confianza.

Figura 3.82

## Autocorrelaciones parciales de la serie de la Prov. Pastaza



Como se observa en este gráfico la primera barra sobresale, entonces se puede adicionar un parámetro AR no estacional, pero también es necesario un parámetro estacional AR, ya que la décima segunda barra también sale de las bandas de confianza.

Entonces se pueden proponer los siguientes modelos:

$$\text{SARIMA}_{12}(1,0,1)\times(1,0,2)$$

Si diferenciamos la serie estacionalmente, el gráfico de las autocorrelaciones de la serie diferenciada, propone el siguiente modelo:

$$\text{SARIMA}_{12}(1,0,1)\times(0,1,1)$$

Pero si se realiza una diferenciación estacionaria y una diferenciación estacional, sus gráficos requieren que se proponga el siguiente modelo:  $\text{SARIMA}_{12}(0,1,1)\times(1,1,1)$

Ahora se pasará a la fase de verificación de los parámetros de cada uno de los modelos propuestos.

**Tabla LXVI**

**Pruebas T\_Student para los parámetros del modelo SARIMA<sub>2</sub>  
(1,0,1)x(1,0,2) de la serie de la Prov. Pastaza**

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.856	0.097	8.789	0
MA(1)	0.606	0.157	3.853	0
SAR(1)	0.998	0.009	115.731	0
SMA(1)	0.870	0.089	9.813	0
SMA(2)	-0.268	0.081	-3.306	0.001
<b>MCE = 107.4</b>				

**Tabla LXVII**

**Pruebas T\_Student para los parámetros del modelo  
SARIMA<sub>2</sub>(1,0,1)x(0,1,1) de la serie de la Prov. Pastaza**

Parámetro	Estimación	Error estándar	Valor T	Valor P
AR(1)	0.830	0.182	4.554	0
MA(1)	0.624	0.246	2.533	0.014
SMA(1)	0.845	0.045	18.569	0
<b>MCE = 102.5</b>				

Tabla LXVIII

Pruebas T\_Student para los parámetros del modelo

SARIMA<sub>12</sub>(0,1,1)x(1,1,1) de la serie de la Prov. Pastaza

Parámetro	Estimación	Error estándar	Valor T	Valor P
MA(1)	0.793	0.072	11.085	0
SAR(1)	-0.339	0.129	-2.619	0.011
SMA(1)	0.783	0.060	13.134	0
<b>MCE = 106.8</b>				

Como lo muestran las tablas anteriores, los parámetros de todos los modelos son significativamente diferentes de cero, por lo tanto no se puede hasta ahora descartar algún modelo propuesto. A continuación se calcula las pruebas para asegurar que los residuos de los tres modelos son independientes.

Tabla LXIX

Pruebas para los Residuos del modelo SARIMA<sub>12</sub> (1,0,1)(1,0,2) de

la serie de la Prov. Pastaza

Pruebas	Z	Valor P
Corrida arriba y abajo de la mediana	-0.109	1.087
Corrida arriba y abajo	1.003	0.316
Box-Pierce	17.351	0566

Tabla LXX

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (1,0,1)(0,1,1) de la serie de la Prov. Pastaza**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y debajo de la mediana	0.593	0.553
Corrida arriba y abajo	1.085	0.278
Box-Pierce	18.186	0.637

Tabla LXXI

**Pruebas para los Residuos del modelo SARIMA<sub>2</sub> (0,1,1)(1,1,1) de la serie de la Prov. Pastaza**

<b>Pruebas</b>	<b>Z</b>	<b>Valor P</b>
Corrida arriba y abajo de la mediana	0.120	0.904
Corrida arriba y abajo	-0.428	0.669
Box-Pierce	14.401	0.809

Los valores p de las pruebas para los tres modelos son mayores que 0.10, entonces podemos afirmar que los residuos de estos modelos son independientes.

En la figura 3.83 se muestra el gráfico de las autocorrelaciones de los residuos del modelo que no incluye diferenciación de ningún tipo, y en la Figura 3.84 y 3.85 se encuentran los gráficos de las

autocorrelaciones de los residuos del modelo  $SARIMA_{12}(0,1,1)\times(1,1,1)$ . En los gráficos se verifica que las autocorrelaciones de los residuos de los modelos quedan dentro de las bandas de confianza.

**Figura 3.83**

**Autocorrelaciones de los residuos del modelo  
 $SARIMA_{12}(1,0,1)\times(1,0,2)$  de la serie de la Prov. Pastaza**

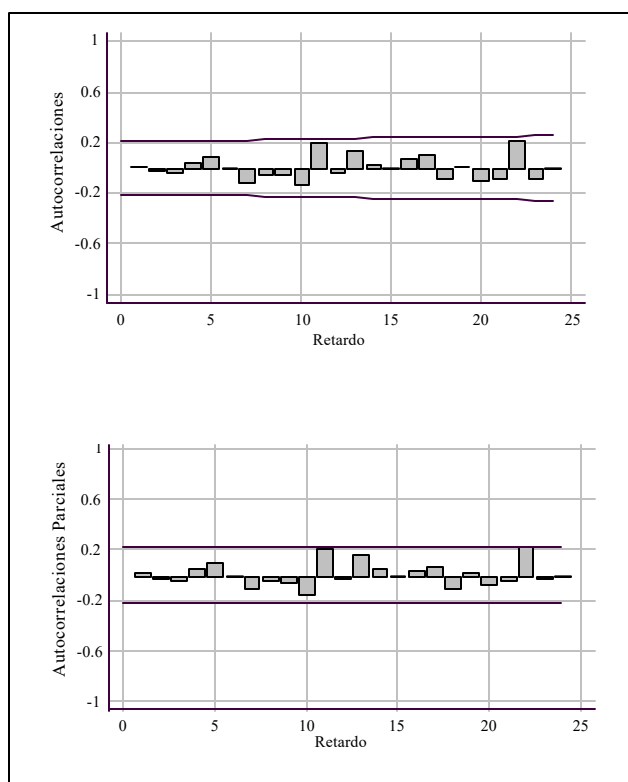


Figura 3.84

**Autocorrelaciones de los residuos del modelo**  
**SARIMA<sub>12</sub>(0,1,1)x(1,1,1) de la serie de la Prov. Pastaza**

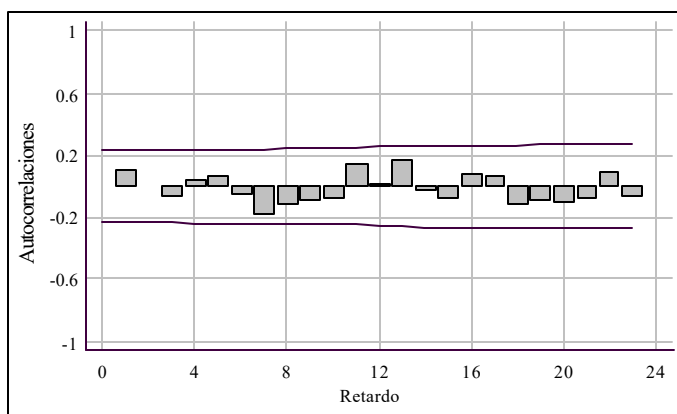
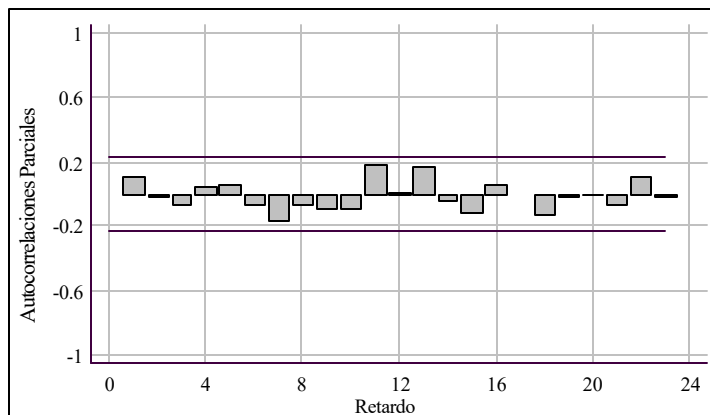


Figura 3.85

**Autocorrelaciones parciales de los residuos del modelo**  
**SARIMA<sub>2</sub>(0,1,1)x(1,1,1) de la serie de la Prov. Pastaza**



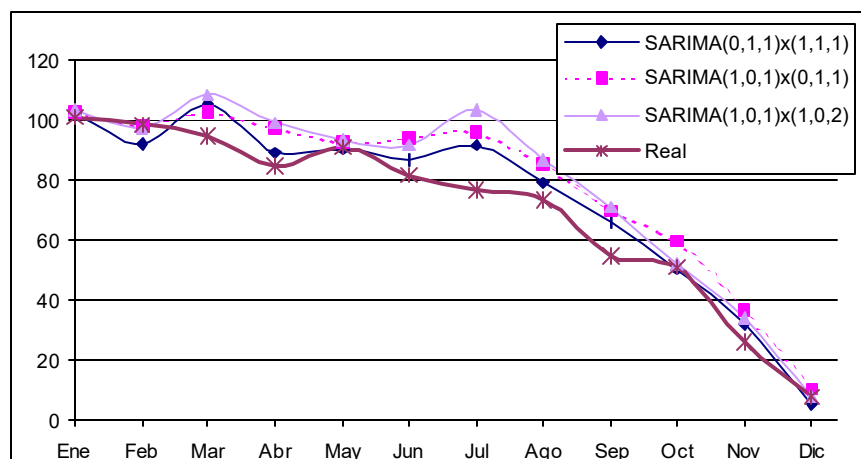
Como los dos tipos de pruebas para seleccionar el mejor modelo han resultado a favor de cada uno, todavía no se puede descartar uno de

ellos, pero siguiendo con el procedimiento de Box y Jenkins se escogerá el modelo de acuerdo al criterio del mayor poder predictivo. Para este criterio se considerará el modelo que tenga la menor media cuadrática, y el que mejor se ajuste a los datos reales.

Si se revisa los valores de las tablas LXVI, LXVII, LXVIII, la MCE menor la tiene el modelo  $SARIMA_{12}(1,0,1) \times (0,1,1)$ . Pero también es importante observar los valores que predicen cada modelo comparado con los reales, de esta forma se verá cual se ajusta mejor a los datos reales, cuyo objetivo general de este análisis es ese, el gráfico comparativo se muestra en la figura 3.86.

**Figura 3.86**

**Comparación de los pronósticos de los modelos propuestos con valores reales del año 1997, de la Prov. Pastaza**





Como se puede observar, la curva que mejor ajusta los datos es la que contiene los valores pronosticados por el modelo  **$SARIMA_{12}(0,1,1)x(1,1,1)$** , por lo tanto se concluye que este es el mejor modelo que predice los valores de la serie de la provincia Pastaza. Además se debe indicar que los valores de los otros modelos se van alejando de la curva real, en lugar de irse aproximando.

En la tabla LXXII se pueden observar los valores de los pronósticos del modelo seleccionado y su gráfico en la figura 3.87.

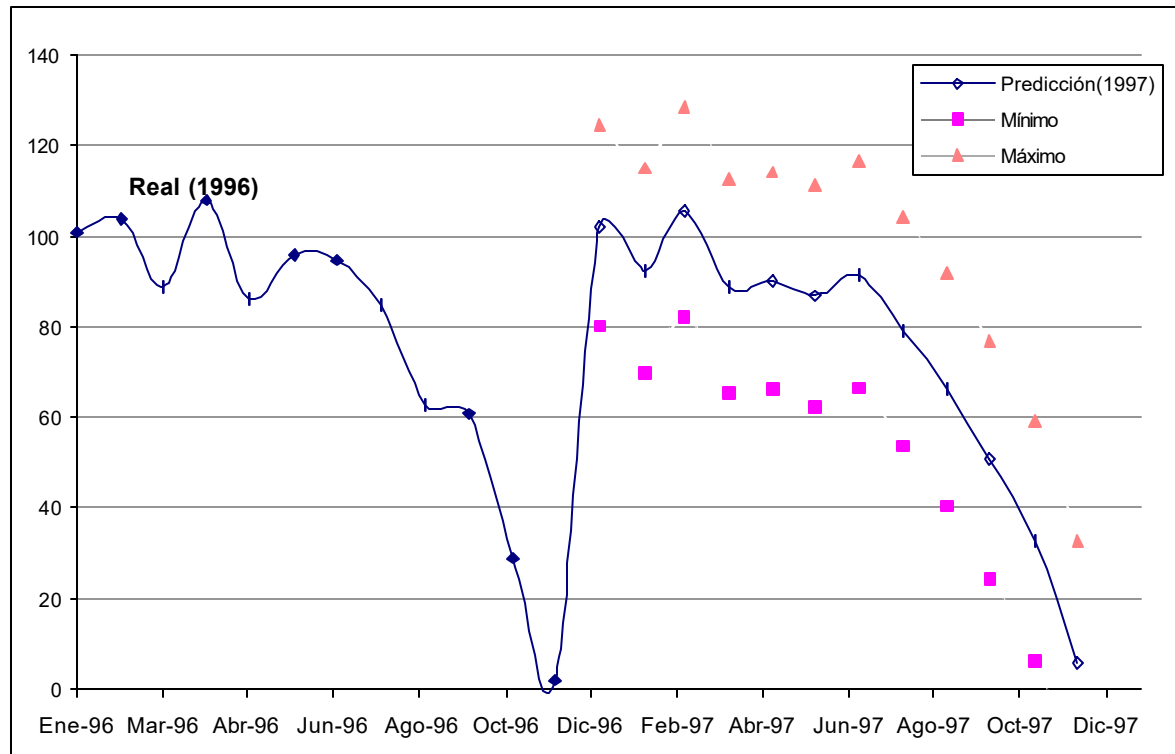
Tabla LXXII

## Predicción de valores de la serie de la Prov. Pastaza

<i>Mes</i>	<i>Real</i>	<i>Predicción</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Error</i>
<i>Ene-97</i>	101	102	80	125	0.01
<i>Feb-97</i>	99	92	70	115	0.07
<i>Mar-97</i>	95	106	83	129	0.11
<i>Abr-97</i>	85	89	66	113	0.05
<i>May-97</i>	91	90	66	114	0.01
<i>Jun-97</i>	82	87	62	111	0.06
<i>Jul-97</i>	77	92	67	117	0.19
<i>Ago-97</i>	74	79	54	105	0.07
<i>Sep-97</i>	55	66	41	92	0.21
<i>Oct-97</i>	51	51	25	77	0.00
<i>Nov-97</i>	26	33	6	59	0.26
<i>Dic-97</i>	8	6	-21	33	0.28

Figura 3.87

Gráfico de las predicciones de la serie de la Prov. Pastaza con el modelo  $SARIMA_{12}(0,1,1)x(1,1,1)$



## CONCLUSIONES

1. La fecha en la que más nacimientos se han registrado en el Ecuador, durante el periodo 1990 - 1997, es enero de 1995 con la cantidad de 22.276 personas, por el contrario la fecha en la que menos cantidad de nacimientos se registraron en el país fue en diciembre de 1997.
2. Todos los años en Ecuador nacen mas varones que mujeres, la diferencia en promedio es de un 2% lo que en cantidades corresponde a aproximadamente 3.700 nacimientos de varones mas que mujeres cada año.
3. El modelo que mejor ajusta la serie del total de nacimientos en el Ecuador es: SARIMA<sub>12</sub> (0,1,0)x(1,1,0), la misma que muestra una tendencia a disminuir cada año.

En promedio en el Ecuador nacen aproximadamente 525 personas diarias.

Las series del Total de nacimientos de hombres y mujeres muestran sus autocorrelaciones similares a las del Total, motivo por el cual tienen modelos muy parecidos.

4. A partir del año 1994 en el Ecuador se han registrado más nacimientos en la región Costa que en la región Sierra con una diferencia en promedio de 7.500 personas más, cada año. Pero dentro del período 1990-1993 ocurrió todo lo contrario, en la Sierra se registraron más nacimientos que en la Costa.
5. En el modelo adecuado para la serie de la provincia de Chimborazo no fue necesario adicionar parámetros estacionarios, sólo se incluyó parámetros estacionales:  $SARIMA_{12}(0,0,0)X(1,1,1)$ .
6. En la provincia Azuay se tiene un marcado decrecimiento cada año, siendo más evidente en los años 1994, 1995 y 1996.
7. Guayas, Pichincha y Manabí son las tres provincias del Ecuador que tienen la mayor cantidad de nacimientos, las mismas que representan en promedio un 23%, 17% y 12% respectivamente del total de nacimientos del país. A pesar de que la provincia de Pichincha tiene un porcentaje muy representativo del total de nacimientos del país, su modelo varía en

relación al del total, pues sólo se necesita diferenciar la serie estacionalmente: SARIMA<sub>12</sub> (1,0,0)X(1,1,0), ya que si se le aplica el mismo modelo del Total muestra una mayor varianza.

8. En la provincia del Guayas en 1996 fue el año en el que ocurrieron la mayor cantidad de nacimientos. Pero tomando un análisis por meses se tiene que en Marzo de 1990 se registraron la mayor cantidad de niños en esta provincia, con una cantidad de 5.000 personas aproximadamente.
9. En El Oro a excepción de las demás provincias, en Mayo de 1993 fue cuando se obtuvo la mayor cantidad de nacimientos. Además en esta provincia no varía mucho el número de nacimientos registrados, se mantiene casi constante.
10. En la provincia Napo existe una marcada diferencia en los meses de Diciembre de los años 1994 y 1995, puesto que hasta 1994 las cantidades en este mes son casi constantes mientras que a partir de 1995 la cantidad de nacimientos va descendiendo más rápido que en los años anteriores.
11. En la provincia Pastaza, la mayor cantidad de nacimientos ocurrieron en Marzo de 1992, y por el contrario la menor cantidad ocurrió en Diciembre

y Noviembre de 1996. Además, la mayoría de los meses del año 1994, son en los que se registran las más altas cantidades de nacimientos, en comparación con los otros años.

## RECOMENDACIONES

1. Realizar un estudio actualizando los datos, es decir considerando los años siguientes a 1997, puesto que en este año el país sufrió cambios que pudieron afectar a la población, tales como el fenómeno del Niño.  
Además analizar las principales causas de los aumentos y disminuciones del número de nacimientos.
2. Incluir en este análisis un estudio sobre variables que pudieron afectar a cada cambio que se ha sufrido el número de nacimientos indicados en esta tesis.
3. Considerar los resultados de este trabajo para proyectos que requieran proyecciones futuras así como el comportamiento de este tipo de series.
4. Difundir en una forma completa tratando de abarcar a toda la población ecuatoriana el tema de la planificación familiar, evitando así el aumento acelerado del índice de natalidad y por ende el aumento de la población.



5. Enfatizar a las instituciones correspondientes en ofrecer información actualizada, con la finalidad de que los futuros estudios se realicen de una manera eficiente en el momento requerido.

## BIBLIOGRAFÍA

1. BOX GEORGE, JENKINS GWILYM y REINSEL GREGORY, Time Series Analysis Forecasting and Control, Prentice Hall, New Jersey 07632, 1994.
  2. DENNIS H. WRONG, Population and Society, 1977.
  3. <http://www.britannica.com/bcom/eb/article/3.html>
  4. Folleto Correo Poblacional, 1995
  5. Anuario de Estadísticas vitales y defunciones, año 1990, INEC.
  6. Anuario de Estadísticas vitales y defunciones, año 1991, INEC.
  7. Anuario de Estadísticas vitales y defunciones, año 1992, INEC.
  8. Anuario de Estadísticas vitales y defunciones, año 1993, INEC.
  9. Anuario de Estadísticas vitales y defunciones, año 1994, INEC.
  10. Anuario de Estadísticas vitales y defunciones, año 1995, INEC.
  11. Anuario de Estadísticas vitales y defunciones, año 1996, INEC.
  12. Anuario de Estadísticas vitales y defunciones, año 1997, INEC.
-