

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL
Facultad de Ciencias Naturales y Matemáticas
Departamento de Matemáticas

PERIODO:	2025 PAO 2	PROFESORES:	Crow P./ Gullqui V./ Pivaque G./ Roa H./Sandoya F.
MATERIA:	ESTADÍSTICA	FECHA:	Jueves, 22 de enero de 2026
EVALUACIÓN:	FINAL		

COMPROMISO DE HONOR

Yo, _____ al firmar este compromiso, reconozco que el presente examen está diseñado para ser resuelto de manera individual, que puedo usar una calculadora *sencilla, ordinaria* para cálculos aritméticos, un lápiz o esferográfico; que solo puedo comunicarme con la persona responsable de la recepción del examen; y, cualquier instrumento de comunicación que hubiere traído, debo apagarlo y depositarlo en la parte anterior del aula, junto con algún otro material que se encuentre acompañándolo. No debo además, consultar libros, notas, ni apuntes adicionales a las que se entreguen en esta evaluación. Los temas debo desarrollarlos de manera ordenada. *Firmo al pie del presente compromiso, como constancia de haber leído y aceptar la declaración anterior.* “Como estudiante de ESPOL me comprometo a combatir la mediocridad y actuar con honestidad, por eso no copio ni dejo copiar.”

Firma: _____ NÚMERO DE MATRÍCULA: _____ PARALELO: _____

PREGUNTA 1. (10 puntos)

Considere que una compañía de teléfonos celulares decide realizar un estudio de mercado para inferir el nivel de satisfacción de los consumidores con un nuevo modelo de celular.

- (5 puntos) ¿Cuál debería ser el tamaño de la muestra aleatoria que se debe considerar en el estudio para estimar la proporción de consumidores satisfechos, si se desea un error máximo del 2 % y un intervalo de confianza del 90 %?
- (5 puntos) La empresa encargada del estudio de mercado entrevistó a 1000 consumidores que adquirieron el nuevo modelo de celular y 400 de ellos manifestaron estar satisfechos con su elección. Calcule un intervalo de confianza al 95 % para la proporción poblacional p .

PREGUNTA 2. (20 puntos)

Un equipo de ingenieros civiles evalúa el tiempo promedio de fraguado (en horas) de un concreto estructural utilizado en dos obras distintas de un proyecto vial. La Obra A emplea un aditivo acelerante tradicional, mientras que la **Obra B** utiliza un nuevo aditivo que, según el proveedor, reduce el tiempo de fraguado.

Debido a restricciones operativas, se toman muestras independientes y de tamaño reducido del tiempo de fraguado del concreto en cada obra, obteniéndose los siguientes resultados:

Tiempos de fraguado (horas)	
Obra A (aditivo tradicional)	9.8, 10.1, 9.6, 10.3, 9.9, 10.0
Obra B (nuevo aditivo)	9.15, 9.25, 9.20, 9.30, 9.10

Previo al análisis inferencial, se establece que los tiempos de fraguado en ambas obras siguen una distribución normal. La empresa constructora debe decidir si adopta el nuevo aditivo en futuras obras viales. Dado que su implementación implica un mayor costo, solo se justificará su uso si existe evidencia estadística suficiente de que permite reducir el tiempo promedio de fraguado del concreto respecto al aditivo tradicional.

Con un nivel de significancia del $\alpha = 0.05$, y utilizando el procedimiento estadístico adecuado:

- (5 puntos) Formule las hipótesis estadísticas correspondientes de acuerdo con el contexto del problema.
- (5 puntos) Evalúe si la variabilidad del tiempo de fraguado puede considerarse comparable entre ambas obras. Con base en esa verificación, justifique qué formulación del estadístico de prueba utilizaría en el literal c).
- (5 puntos) Con base en el análisis previo, realice el contraste de hipótesis apropiado para comparar los tiempos promedio de fraguado entre ambas obras.
- (5 puntos) Interprete el resultado en el contexto del problema.

PREGUNTA 3. (20 puntos)

Una empresa constructora está ejecutando la rehabilitación de un tramo vial urbano. Para garantizar la calidad del pavimento, el laboratorio de control realiza mediciones sobre la resistencia a la compresión del material medidas en megapascal (MPa), evaluada mediante núcleos extraídos del tramo. Se desea realizar inferencias sobre el tramo intervenido mediante intervalos de confianza y planificación muestral, utilizando un nivel de confianza del 95 %. En un control preliminar se seleccionó una muestra aleatoria de 36 núcleos y se obtuvo una media muestral de 42.8 MPa y una desviación estándar muestral de 5.4 MPa.

- a) **(10 puntos)** Construya e interprete un intervalo de confianza para la resistencia media del pavimento.
- b) **(10 puntos)** Para futuros controles, la empresa desea que la estimación de la resistencia media tenga un error máximo de 1.2 MPa. Determine el tamaño mínimo de muestra n requerido para cumplir esta condición.

PREGUNTA 4. (20 puntos)

Una empresa de consultoría en marketing desea evaluar si el diseño de la portada de una encuesta enviada por correo influye en la tasa de respuesta. Para ello, se enviaron encuestas con contenido idéntico pero con dos tipos de portada: una minimalista (solo texto) y otra visualmente atractiva (a color con gráficos). Se plantea que la tasa de respuesta para la portada minimalista es menor que para la portada visualmente atractiva. Los resultados obtenidos se muestran en la siguiente tabla:

Tipo de Portada	Cantidad enviada	Cantidad de Respuestas
Minimalista	207	104
Visual Atractiva	213	109

- a) **(7 puntos)** Defina los parámetros de interés y establezca las hipótesis estadísticas relevantes.
- b) **(7 puntos)** Pruebe las hipótesis utilizando un nivel de significancia de 0.10.
- c) **(6 puntos)** Interprete su resultado en el contexto del problema.

PREGUNTA 5. (30 puntos)

En una planta industrial, el analista de producción busca explicar el consumo de energía Y (kWh por hora) a partir de tres variables operativas: X_1 : Temperatura ambiente ($^{\circ}\text{C}$), X_2 : Carga de producción (ton/h), X_3 : Horas efectivas de operación del turno (h). El analista

implementó en su computadora un modelo de regresión lineal múltiple:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Con el código en R:

```
lm(Consumo ~Temp + Carga + Horas)
summary(modelo):
Se obtuvo la siguiente salida:
Call:
lm(formula=Consumo~ Temp + Carga + Horas,data=planta)
Residuals:
    Min     1Q   Median     3Q      Max
-9.30  -2.45   0.08   2.60   8.95

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   15.200      3.800    4.000  0.00018 ***
Temp           0.280      0.090    3.111  0.00290 **
Carga          1.650      0.250    6.600  4.0e-08 ***
Horas          0.520      0.210    2.476  0.01620 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.10 on 56 degrees of freedom
Multiple R-squared:  0.698, Adjusted R-squared:  0.682
F-statistic: 43.1 on 3 and 56 DF,  p-value: < 2.2e-16
```

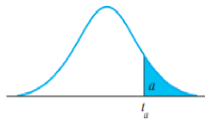
- a) **(4 puntos)** Escriba la ecuación del modelo estimado $\hat{Y}$ obtenido.
- b) **(4 puntos)** Interprete el significado práctico de cada coeficiente: intercepto y pendientes (β_1 , β_2 , β_3). Incluya unidades e indique qué representa cada coeficiente.
- c) **(4 puntos)** A partir de las pruebas t (columna $\text{Pr}(> |t|)$), determine qué variables son significativas al 5 % y justifique.
- d) **(3 puntos)** Analice el resumen de los residuos (mínimo, cuartiles y máximo) e indique si el comportamiento observado es consistente con los supuestos del modelo de regresión lineal, justifique.
- e) **(4 puntos)** Explique qué mide el coeficiente de determinación R^2 reportado. Interprete ese valor para este caso.
- f) **(4 puntos)** Interprete la prueba F del modelo: plantee H_0 y H_1 en palabras y concluya usando el valor- p .
- g) **(3 puntos)** Use el modelo para predecir el consumo de energía cuando: $Temp = 24^{\circ}\text{C}$, $Carga = 12 \text{ ton/h}$ y $Horas = 7.5 \text{ h}$.
- h) **(4 puntos)** En un párrafo corto: ¿para qué le sirve a un ingeniero este modelo al tomar decisiones operativas?

Probabilidades acumuladas de la normal estándar $P(Z \leq z)$, $Z \sim N(0, 1)$.



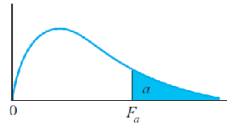
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974

$P(X \geq t_{\alpha,n}) = \alpha$



df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	df
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10

Puntos porcentuales de la distribución F



df ₂	a	df ₁								
		1	2	3	4	5	6	7	8	9
4	.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	.005	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14
5	.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	.005	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77

Correlación	Covarianza
$\rho_{x,y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$	$Cov(X,Y) = \sigma_{xy} = E(XY) - E(X)E(Y)$

Estimaciones (Intervalos de confianza)			
Media (n > 30)	Media (n < 30)	Proporción	Varianza
$\bar{X} \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}$	$\bar{X} \pm \frac{s}{\sqrt{n}} t_{\frac{\alpha}{2}}$	$\hat{p} \pm \sqrt{\frac{\hat{p}\hat{q}}{n}} Z_{\frac{\alpha}{2}}$	$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$

Prueba de hipótesis			
Una sola media		Dos medias	
$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}}$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sigma\sqrt{\left(1/n_1\right) + \left(1/n_2\right)}}$
Pruebas pareadas			
$t = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}$			
Una proporción		$t = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p\sqrt{\left(1/n_1\right) + \left(1/n_2\right)}}$	$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$
$Z = \frac{\hat{p} - p_0}{\sqrt{p_0q_0/n}} = \frac{X - np_0}{\sqrt{np_0q_0}}$			
Dos proporciones		$t' = \frac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{\left(s_1^2/n_1\right) + \left(s_2^2/n_2\right)}}$	$v = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$
$Z = \frac{(\hat{p}_1 - \hat{p}_2) - d_0}{\sqrt{\left(p_1q_1/n_1\right) + \left(p_2q_2/n_2\right)}}$			
Igualdad de varianzas			
$F = s_1^2/s_2^2$			

Pruebas Ji-Cuadrado			
Independencia	$e_{ij} = \frac{r_i * c_j}{n}$	$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$	$v = (r - 1)(c - 1)$
Bondad de ajuste	$e_i = n * p_i$	$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$	$v = (k - 1 - p)$
Kolmogorov	Smirnov $D = \max F_n(x) - F_0(x) $		

ANOVA	Regresión Lineal Simple
$SCT = \sum_{i=1}^n (y_i - \bar{y})^2$	$b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$
$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ MCR = SCR/1	$b_0 = \frac{\sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i}{n} = \bar{y} - b_1 \bar{x}$
$SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ MCE = SCE/(n - 2)	

ANOVA	Regresión Lineal Múltiple
$SCT = \sum_{i=1}^n (y_i - \bar{y})^2$	$\hat{\beta} = (X^T X)^{-1} X^T Y$
$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ MCR = SCR/p	
$SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ MCE = SCE/(n - p - 1)	$R^2 = \frac{SCR}{SCT}$