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

PERIODO:	2025 PAO 2		
MATERIA:	ESTADÍSTICA	**PROFESORES**:	Crow P./ Gullqui V./ Pivaque G./ Roa H./Sandoya F.
EVALUACIÓN:	TERCERA	**FECHA**:	Jueves, 05 de febrero de 2026

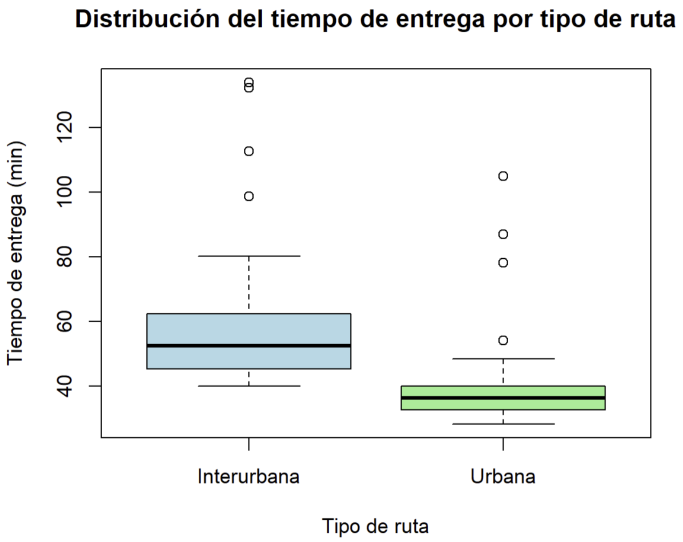
COMPROMISO DE HONOR

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. (15 puntos)

Una empresa de courrier analiza el tiempo de entrega (min) de paquetes según el tipo de ruta: Rutas Urbanas y Rutas Interurbanas. Se presentan dos boxplots y el siguiente resumen:



Ruta	Mediana	Q1	Q3	Máx
Urbana	35	28	45	120
Interurbana	50	40	70	140

a) **(5 puntos)** ¿Cuál tipo de ruta presenta mayor variabilidad operativa? Justique.

b) **(5 puntos)** ¿En qué ruta es más probable la presencia de valores atípicos altos? Justifique

c) **(5 puntos)** La empresa desea priorizar mejoras donde el 75 % de entregas supere los 60 min. ¿Dónde debe intervenir primero?

PREGUNTA 2. (15 puntos)

En una red de estaciones de recarga eléctrica se dan eventos de falla con una tasa $\lambda = 1.8$ fallas/día por cada estación. Se supervisan 5 estaciones. Calcule:

a) **(5 puntos)** Probabilidad de que al menos una estación presente 4 o más fallas en un día.

b) **(5 puntos)** Probabilidad de que el tiempo entre fallas sea menor a 6 horas.

c) **(5 puntos)** Con base en (a) y (b), ¿es razonable asignar un técnico permanente?

PREGUNTA 3. (15 puntos)
En un curso de estadística se selecciona una muestra aleatoria de 50 estudiantes. A cada estudiante se le pide que estime la cantidad de tiempo (en minutos) que dedicó en prepararse para su examen de mejoramiento. Si el tiempo dedicado por un estudiante sigue una distribución normal con desviación estándar de 30 minutos. Determine:

a) **(5 puntos)** La probabilidad de que un estudiante elegido al azar haya dedicado al menos 40 minutos por encima de la media poblacional.

b) **(5 puntos)** La probabilidad de que el promedio de tiempo de preparación en la muestra sea al menos 5 minutos mayor que el promedio poblacional

c) **(5 puntos)** La probabilidad de que el promedio de tiempo de preparación en la muestra difiera del promedio poblacional en más de 10 minutos

PREGUNTA 4. (15 puntos)
Dos plataformas de aprendizaje virtual presentan las siguientes tasas de finalización de cursos:

Plataforma	**Inscritos**	**Finalizan**
P1	600	330
P2	580	350

La institución considera migrar a la plataforma P2 únicamente si existe evidencia estadística suficiente de que su tasa de finalización es mayor que la de la plataforma P1. Migrar implica un costo adicional relevante.
A partir del análisis estadístico se obtiene el siguiente valor para el estadístico de prueba: $z = -1.48$

a) **(5 puntos)** Defina los parámetros de interés y plantee el contraste de hipótesis correspondiente al contexto descrito.

b) **(5 puntos)** Utilizando el valor de z proporcionado, concluya el contraste para $\alpha = 0.10$ y para $\alpha = 0.05$.

c) **(5 puntos)** Con base en el resultado estadístico y considerando el costo de migración, ¿qué decisión recomendaría a la institución? Justifique.

PREGUNTA 5. (15 puntos)
Una agencia ambiental utiliza un sistema automatizado para clasificar el nivel de contaminación del aire registrado en distintos puntos de una ciudad. Cada medición es clasificada en uno de los siguientes niveles: Alto, Medio, Bajo. Las mediciones provienen de tres tipos de zonas de la ciudad, las cuales contribuyen al total de observaciones de acuerdo con las siguientes proporciones: zona industrial, 0.30; zona residencial, 0.45; y zona verde, 0.25.
A partir de registros históricos, se conoce la probabilidad de que una medición sea clasificada como nivel Alto dado el tipo de zona donde se realizó la medición:

$$P(\text{Alto} \mid \text{Industrial}) = 0.55$$
$$P(\text{Alto} \mid \text{Residencial}) = 0.25$$
$$P(\text{Alto} \mid \text{Verde}) = 0.10$$

Con base en la información proporcionada, responda lo siguiente:

a) **(5 puntos)** Calcule la probabilidad total de que una medición seleccionada al azar sea clasificada como nivel Alto, es decir, $P(\text{Alto})$.

b) **(5 puntos)** Calcule la probabilidad de que una medición con nivel de contaminación Alto provenga de una zona Industrial, es decir, $P(\text{Industrial} \mid \text{Alto})$.

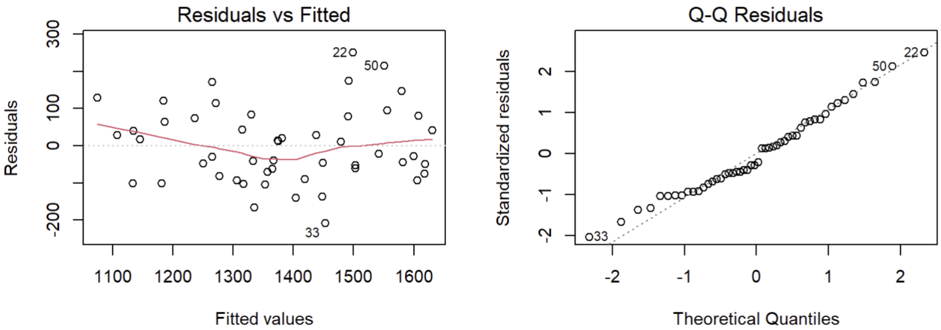
c) **(5 puntos)** Desde el punto de vista de la gestión ambiental, ¿en qué tipo de zona sería más razonable enfocar una auditoría ambiental? Justifique su respuesta utilizando los resultados obtenidos.

PREGUNTA 6. (25 puntos)
La administración de un campus universitario desea modelar el consumo promedio de agua (Y, en litros por hora) con el fin de mejorar la planificación de recursos y reducir costos operativos. Para ello, se ajustó un modelo de regresión lineal múltiple utilizando los siguientes predictores: Temperatura promedio del día (°C), Número de estudiantes presentes en el campus, Horas-clases programadas al día.
A continuación se muestra un resumen parcial de la salida obtenida con el código en R:

```
lm(Consumo_agua ~Temp + Num_estud + Horas_clases)
summary(modelo):
Se obtuvo la siguiente salida:
Call:
lm(formula=Consumo_agua ~Temp + Num_estud + Horas_clases,data=datos_campus)
Residuals:
     Min        1Q    Median       3Q       Max
-228.603  -102.854    -4.361   92.025   301.975

Coefficients:
             Estimate Std. Error t value   Pr(>|t|)
(Intercept)   599.82      138.5   4.331   4.48e-05  ***
Temp          3.5260       2.55   1.380   0.171760
Num_estud     0.0739     0.0098   7.524   8.93e-11  ***
Horas_clases  34.240     12.070   2.837   0.005830  **
---
Signif.codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 128.2 on 76 degrees of freedom
Multiple R-squared: 0.4823,  Adjusted R-squared: 0.4619
F-statistic: 23.6 on 3 and 76 DF,  p-value: < 6.736e-11
```



a) **(3 puntos)** Identifique cuáles variables explicativas muestran evidencia estadística significativa y cuáles consideraría para el modelo de predicción? Justifique su respuesta.

b) **(6 puntos)** Interprete el significado práctico de cada coeficiente: intercepto y pendientes significativas. Incluya unidades e indique qué representa cada coeficiente.

c) **(4 puntos)** Interprete el valor de R^2 en el contexto del problema. ¿Es razonable considerar que el modelo tiene una capacidad explicativa adecuada para fines operativos? Explique.

d) **(4 puntos)** Explique qué evalúa la prueba F global del modelo y cómo debe interpretarse su resultado en este caso particular.

e) **(4 puntos)** A partir del diagnóstico de residuos proporcionado, identifique e indique qué supuestos del modelo podrían estar siendo vulnerados.

f) **(4 puntos)** Utilizando el modelo ajustado, se desea estimar el consumo de agua en un día con las siguientes características: Temperatura promedio: 26 °C, Número de estudiantes: 8 000, Horas clases programadas al día: 10.

Formulario I Parcial

DATOS AGRUPADOS	COVARIANZA
$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i m_i$ $S^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i (m_i - \bar{x})^2$ $\tilde{x} = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} a_i$ $Mo = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} a_i$ $AC = \frac{x_{(n)} - x_{(1)}}{k}$	$S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ $r = \frac{S_{xy}}{S_x S_y}$ **PROBABILIDAD** $P(E) = \frac{n(E)}{n(\Omega)}$ $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ $P(B) = \sum_{i=1}^{k} P(A_i)\, P(B \mid A_i)$ $P(A_i \mid B) = \frac{P(A_i) P(B \mid A_i)}{\sum_{j=1}^{k} P(A_j) P(B \mid A_j)}$

Momento con respecto al origen	Momento con respecto a la media
Discreta: $E(X^r) = \sum_x x^r f(x)$ **Continua:** $E(X^r) = \int_{-\infty}^{\infty} x^r f(x)\, dx$	**Discreta:** $\mu_r = \sum_x (x - \mu)^r f(x)$ **Continua:** $\sigma^2 = E(X^2) - [E(X)]^2$ $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$

MODELOS ALEATORIOS DISCRETOS	MODELOS ALEATORIOS CONTINUOS
Binomial: $X \sim \mathrm{B}(n, p)$ $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ $E(X) = np, \quad V(X) = np(1-p)$ **Binomial Negativa:**$X \sim \mathrm{BN}(k, p)$ $P(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$ $E(X) = \frac{k}{p}, \quad V(X) = \frac{k(1-p)}{p^2}$ **Poisson:**$X \sim \mathrm{P}(\lambda)$ $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$ $E(X) = \lambda, \quad V(X) = \lambda$	**Uniforme:**$X \sim \mathrm{U}(a, b)$ $f(x) = \frac{1}{b-a}, \quad a \le x \le b$ $E(X) = \frac{a+b}{2}, \quad V(X) = \frac{(b-a)^2}{12}$ **Normal:**$X \sim \mathrm{N}(\mu, \sigma^2)$ $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$ **Normal estándar:**$Z \sim \mathrm{N}(0, 1)$ $Z = \frac{x - \mu}{\sigma}, \quad f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ **Exponencial:**$X \sim \mathrm{E}(\beta)$ $f(x) = \frac{1}{\beta} e^{-x/\beta}, \quad \beta > 0 \quad F(X) = 1 - e^{-\frac{x}{\beta}}$ $E(X) = \beta, \quad V(X) = \beta^2$

TEOREMA DEL LÍMITE CENTRAL
$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \longrightarrow N(0, 1)$

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

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

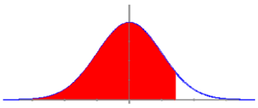
Prueba de hipótesis	
Una sola media	Dos medias
$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ $\quad t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$	$Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{\left(\sigma_1^2/n_1\right) + \left(\sigma_2^2/n_2\right)}}$ $\quad Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sigma\sqrt{\left(1/n_1\right) + \left(1/n_2\right)}}$
Pruebas pareadas	
$t = \dfrac{\bar{d} - d_0}{s_d/\sqrt{n}}$	$t = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{S_p\sqrt{\left(1/n_1\right) + \left(1/n_2\right)}}$ $\quad S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$
Una proporción	
$Z = \dfrac{\hat{p} - p_0}{\sqrt{p_0 q_0/n}} = \dfrac{X - np_0}{\sqrt{np_0 q_0}}$	
Dos proporciones	
$Z = \dfrac{(\hat{p}_1 - \hat{p}_2) - d_0}{\sqrt{\left(p_1 q_1/n_1\right) + \left(p_2 q_2/n_2\right)}}$	$t' = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{\left(s_1^2/n_1\right) + \left(s_2^2/n_2\right)}}$ $\quad v = \dfrac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\dfrac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \dfrac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$
Igualdad de varianzas	
$F = s_1^2/s_2^2$	

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

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

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

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