NOMBRE Y APELLIDOS:

1. Marca la respuesta correcta: (1/2 cada una; las respuestas incorrectas restan 1/6)

a) La función de distribución de una variable aleatoria es siempre …

☐ Menor que 1 ☒ Menor ó igual que 1
☐ 1 ☐ Mayor ó igual que 1

b) Para una muestra $\{x_1, x_2, \cdots, x_n\}$, la expresión $\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2}$ se llama …

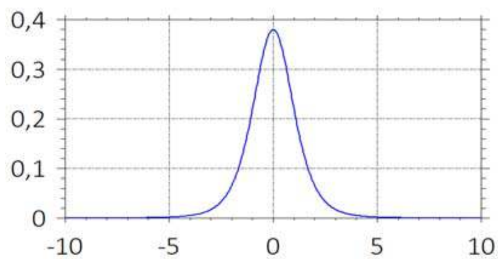
☐ Mediana ☐ Media
☐ Varianza muestral ☒ Desviación típica muestral

c) Para contrastar hipótesis sobre la varianza de una polación normal se usa la distribución …

☐ Normal estándar
☒ χ^2 de Pearson
☐ F de Fisher-Snedecor
☐ t de Student

d) El gráfico adjunto corresponde a la función de densidad …

☐ N(5,1)
☐ Exponencial de parámetro 5
☒ t de Student con 5 grados de libertad
☐ χ^2 con 5 grados de libertad



2. Admitiendo que en general la proporción de estudiantes que escriben con bolígrafo azul sea 0.2, y si se eligen 5 estudiantes al azar: (2)

(a) ¿Qué modelo de distribución sigue la variable aleatoria "*número de estudiantes (de esos 5) que escriben con bolígrafo azul*"?
(b) Calcula la probabilidad de que exactamente 1 de esos 5 escriba con bolígrafo azul.
(c) Calcula la probabilidad de que a lo sumo 2 de esos 5 escriban con bolígrafo azul.
(d) Si en lugar de 5 se eligen 50 estudiantes al azar, ¿cuántos podemos esperar que escriban con bolígrafo azul?

(a) Binomial de parámetros 5 y 0,2.

(b) Llamo X a la variable; $P(X=1)=\binom{5}{1}\cdot 0,2^1\cdot 0,8^4=\boxed{0,4096}$

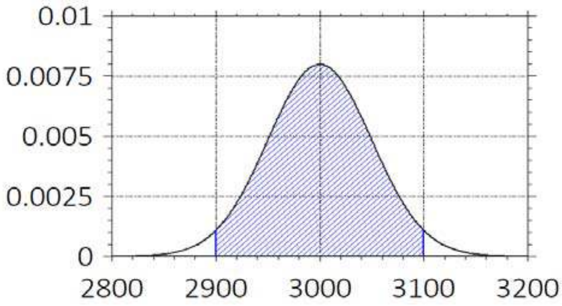
$$P(X\le 2)=P(X=0)+P(X=1)+P(X=2)=\binom{5}{0}\cdot 0,2^0\cdot 0,8^5+\binom{5}{1}\cdot 0,2^1\cdot 0,8^4+\binom{5}{2}\cdot 0,2^2\cdot 0,8^3=$$
$$=0,3277+0,4096+0,2048=\boxed{0,9421}$$

(c) La esperanza de una variable binomial de parámetros 50 y 0.2: $50\cdot 0,2=\boxed{10}$

(Sigue a la vuelta)

3. Se supone que la distribución teórica de la variable "*Peso de una malla de naranjas*" es normal de media μ_N = 3000 y desviación σ_N = 50 (en gramos). (2)

(a) Sobre la siguiente curva de densidad N(3000,50), representa la probabilidad de que una malla de naranjas tenga un peso entre 2900 y 3100 g. Calcula esta probabilidad con ayuda de la tabla.



Llamo X al peso de una malla de naranjas y Z a una variable N(0,1). Entonces:

$$P(2900 \le X \le 3100) =$$
$$= P\left(\frac{2900-3000}{50} \le \frac{X-3000}{50} \le \frac{3100-3000}{50}\right) =$$
$$= P(-2 \le Z \le 2) = P(Z \le 2) - P(Z \le -2) =$$
$$= 0{,}9772 - (1-0{,}9772) = 2 \cdot 0{,}9772 - 1 =$$
$$= 1{,}9544 - 1 = \boxed{0{,}9544}$$

(b) Si la variable "*Peso de una malla de limones*" es también normal, pero con media μ_L = 2000 y desviación σ_L = 30 (en gramos) y compramos una malla de cada, ¿cuál es la probabilidad de que el peso total supere 5070 g? (Se supone que son variables independientes).

$$\left.\begin{array}{ll} X = \text{"}\textit{Peso de una malla de naranjas}\text{"}; & X \to N(3000,50) \\ Y = \text{"}\textit{Peso de una malla de limones}\text{"}; & Y \to N(2000,30) \end{array}\right\} X+Y \to N\left(3000+2000, \sqrt{50^2+30^2}\right)$$

$$\left.\begin{array}{l} 3000+2000 = 5000 \\ \sqrt{50^2+30^2} = \sqrt{2500+900} = \sqrt{3400} = 58{,}31 \end{array}\right\} X+Y \to N(5000, 58{,}31)$$

$$P(X+Y > 5070) = P\left(\frac{X+Y-5000}{58{,}31} > \frac{5070-5000}{58{,}31}\right) = P(Z > 1{,}2) = 1 - 0{,}8849 = \boxed{0{,}1151}$$

NOMBRE Y APELLIDOS:

4. Una compañía embotella agua mineral. Se midió el contenido en sodio en 8 botellas elegidas al azar y se encontró una media muestral de 5,2 y una desviación muestral de 0,2. Construye un intervalo de confianza al 99% para el contenido medio en sodio. (1)

Intervalo de confianza para la media de una población normal: $\overline{X} \pm \frac{\hat{s}}{\sqrt{n}} \cdot t_{1-\alpha/2}^{(n-1)}$

$$1-\alpha = 0{,}99 \Rightarrow \alpha = 0{,}01 \Rightarrow \alpha/2 = 0{,}005 \Rightarrow 1-\alpha/2 = 0{,}995$$

$$t_{1-\alpha/2}^{(n-1)} = t_{0{,}995}^{(7)} = 3{,}499$$

$$\left(5{,}2 - \frac{0{,}2}{\sqrt{8}} \cdot 3{,}499 \;,\; 5{,}2 + \frac{0{,}2}{\sqrt{8}} \cdot 3{,}499 \right)$$

$$(5{,}2 - 0{,}247 \;,\; 5{,}2 + 0{,}247)$$

$$\boxed{(4{,}953 \;,\; 5{,}447)}$$

5. Un fabricante de electrodomésticos inició una campaña publicitaria para un nuevo producto. Al cabo de algún tiempo, un sondeo en la población mostró que 81 personas de 300 encuestadas conocían dicho producto. Determina mediante el contraste de hipótesis adecuado si se puede afirmar que más de la cuarta parte de la población conoce el producto (nivel: 0,05). (1)

$$\left.\begin{array}{l} H_0 : p = 0{,}25 \\ H_1 : p > 0{,}25 \end{array}\right\} \qquad \alpha = 0.05 \qquad \hat{p} = \frac{81}{300} = 0{,}27$$

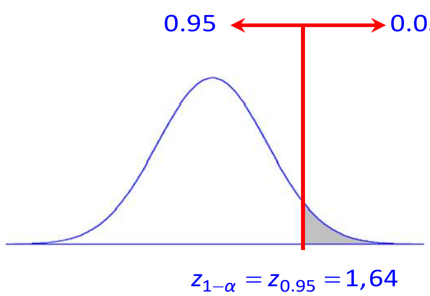
Estadístico de contraste: $Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0{,}27 - 0{,}25}{\sqrt{0{,}25 \cdot (1-0{,}25)/300}} = 0{,}8$

Distribución aproximada bajo la hipótesis nula: normal estándar.

Región crítica a la derecha de nivel 0,05:

$C = [1{,}64 \;, \infty)$

$Z_{obs} = 0{,}8 \notin C$, por tanto, no se rechaza la hipótesis nula.



$$\boxed{\text{A la vista de los datos observados, no se puede afirmar que más de la cuarta parte de la población conozca el producto.}}$$

(Sigue a la vuelta)

6. Con el fin de estudiar el rendimiento de distintas variedades de arándano se cultivaron cuatro variedades en tres viveros. La variable respuesta fue el diámetro medio en mm. Con los datos recogidos se obtuvieron las siguientes sumas de cuadrados: (2)

Entre variedades:	16.43
Entre viveros:	0.18
Total:	16.86

(a) Construye la tabla de análisis de la varianza.

(b) Razona a partir de la tabla si hay diferencia significativa en el tamaño medio según la variedad.

(c) Calcula e interpreta los coeficientes de determinación.

(a)

Fuente de variación	SC	g.l.	SCM	F	Percentil F(0,95;3;6)
Variedades	16.43	3	5.4767	131.44	4,757
Viveros	0.18	2	0.09	2.16	
Residual	0.25	6	0.0417		
Total	16.86	11			

(b) Efecto de la variadad sobre el tamaño medio:

H0: $\mu_i = \mu$ para todo i = 1, … , 4.

H1: Algún μ_i es distinto de los demás.

El contraste se resuelve mediante la F de variedades. Su distribución bajo la hipótesis nula es F con 3 y 6 grados de libertad. La región crítica de nivel 0.05 empieza en 4,757. El valor observado de F, 131,44, está dentro de la región crítica y por tanto rechazamos la hipótesis nula. Concluimos que hay efecto significativo al nivel 0.05 de la variedad sobre el tamaño medio.

(c) Variedades: **0.9745**

Viveros: **0.0107**

El 97,45% de la variabilidad en el tamaño medio se debe al efecto de la variedad.

El 1,07% de la variabilidad en el tamaño medio se debe al efecto del vivero.