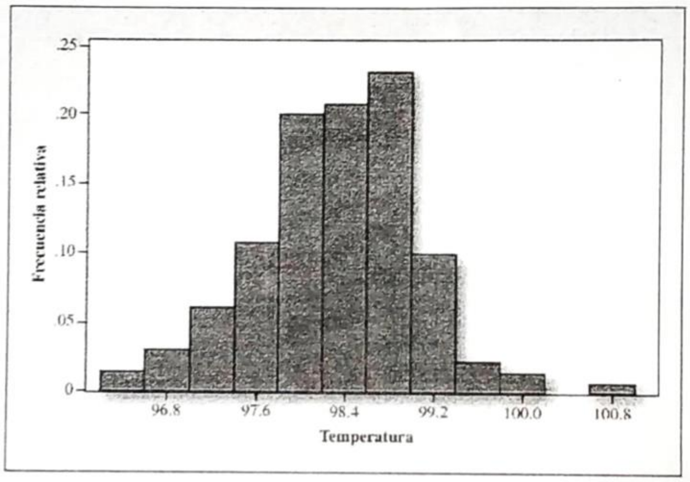
ESCUELA POLITÉCNICA NACIONAL
DEPARTAMENTO DE FORMACIÓN BÁSICA
PROBABILIDAD Y ESTADÍSTICA
2023-B
PRUEBA 1-GR11

Indicaciones importantes

- Se permite el uso de calculadora elemental y formulario.
- El tiempo asignado es de 60 minutos para la resolución.
- Respetar el formato de presentación solicitado.
- Recuerde que las interpretaciones y /o conclusiones de sus cálculos y gráficos son fundamentales.

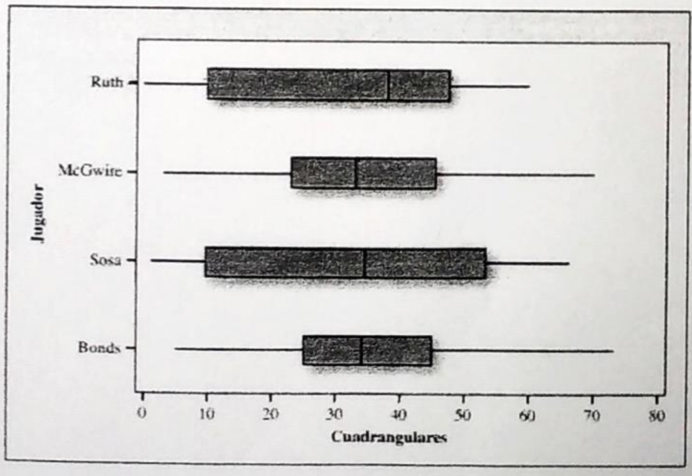
Problemas para resolver

1. {3,5 puntos} La temperatura corporal de 98,6 grados Farenheit como estándar en los seres humanos fue obtenida por un médico alemán en 1868. En un intento por verificar esta afirmación, Mackowiak, Wasserman y Levine, tomaron las temperaturas de 148 personas sanas en un periodo de tres días. Un conjunto de datos, que estrechamente se compara con el del artículo de Mackowiak, Las temperaturas corporales para estas 130 personas se muestran en el histograma de frecuencia relativa siguiente.



   a) ¿Cuál y qué tipo de variable muestra este gráfico?
   b) Describa e interprete el histograma. ¿Hay algunas observaciones poco comunes? ¿Escriba alguna explicación para éstas?
   c) Localice los 98.6 grados normales en el eje horizontal de la gráfica. ¿Parecen estar cerca del centro de distribución?

2. {3,5 puntos} En el verano de 2001, Barry Bonds empezó su búsqueda de romper el récord de Mark McGwire de 70 cuadrangulares conectados en una sola temporada. Al terminar la temporada de béisbol de 2003 de las ligas mayores, se registró el número de cuadrangulares conectados por temporada por cada uno de cuatro superestrellas de ligas mayores en su carrera y a continuación se presentan en las gráficas de caja:



...iba un párrafo corto que compare los patrones de bateo de cuadrangulares de estos cuatro jugadores.

(3,5 puntos) Un fabricante de jeans (pantalones vaqueros) tiene plantas en California, Arizona y Texas. Un grupo de 25 pares de jeans se selecciona al azar de la base de datos computarizada, registrándose el estado en el que se produjo cada uno:

CA AZ AZ TX CA CA CA TX TX TX
AZ AZ CA AZ TX CA AZ TX TX TX
CA AZ AZ CA CA

a) ¿Cuál es la variable que se mide? ¿Es cualitativa o cuantitativa?
b) Elabore la gráfica más adecuada para describir los datos.
c) ¿Qué proporción de los jeans se hizo en Texas?
d) ¿Cuál estado produjo más jeans?
e) Si desea averiguar si las tres plantas produjeron igual número de jeans, o si una produjo más que las otras, ¿cómo se usaría la gráfica de la parte *b* para ayudarse? ¿Qué conclusiones obtiene de estos datos?

ESCUELA POLITÉCNICA NACIONAL
DEPARTAMENTO DE FORMACIÓN BÁSICA
PROBABILIDAD Y ESTADÍSTICA • PRUEBA 1A
FECHA: 30 DE ABRIL 2019

DEPARTAMENTO DE FORMACIÓN BÁSICA
EPN

Nombres:	CI:	Firma:
Profesor CD:		Grupo CD:

Instrucciones:

1. No se permite el uso de DISPOSITIVOS ELECTRÓNICOS.
2. La duración de PRUEBA es de 120 minutos.
3. Se permite el uso de calculadora científica y del formulario plastificado.
4. Cualquier INTENTO DE COPIA se sancionará con una nota de CERO puntos.
5. JUSTIFICAR CADA RESPUESTA.

EJERCICIOS:

1. En una vía de la ciudad se ha recopilado la información del número de pasajeros que transportan 18 autobuses escolares y se han obtenido los siguientes datos:

$$13, 15, 17, 15, 17, 16, 20, 18, 19, 18, 10, 19, 14, 16, 17, 16, 17, 14$$

a) Calcule las siguientes medidas: media aritmética, moda, mediana, desviación estándar y rango intercuartil. (1 punto)

Media aritmética:

$$\begin{aligned}\bar{x} &= \sum_{i=1}^{18} x_i \\ &= \frac{13 + 15 + \ldots + 17 + 14}{18} \\ &= 16{,}1667\end{aligned}$$

Moda:
Para su cálculo se identifica la observación con más repeticiones:

$$\text{Moda} = 17$$

Mediana:
Se ordenan los datos en forma ascendente:

$$10, 13, 14, 14, 15, 15, 16, 16, 16, 17, 17, 17, 17, 18, 18, 19, 19, 20$$

para su cálculo debe tomarse en cuenta que existe un número par de observaciones, por tanto:

$$\text{Mediana} = \frac{x_9 + x_{10}}{2} = \frac{16 + 17}{2} = 16{,}5$$

Desviación estándar:

$$\begin{aligned}s &= \sqrt{\frac{1}{18-1}\sum_{i=1}^{18}(x_i - \bar{x})^2} \\ &= \sqrt{\frac{(10 - 16{,}1667)^2 + (13 - 16{,}1667)^2 + \ldots + (19 - 16{,}1667)^2 + (20 - 16{,}1667)^2}{17}} \\ &= 2{,}4314\end{aligned}$$

Rango Intercuartil:
Se inician calculando los percentiles 25 y 75:

$$\frac{nk}{100} = \frac{18(25)}{100} = 4{,}5 \approx 5, \qquad \frac{nk}{100} = \frac{18(75)}{100} = 13{,}5 \approx 14$$

de donde:

$$P_{25} = x_5 = 15, \qquad P_{75} = x_{14} = 18$$

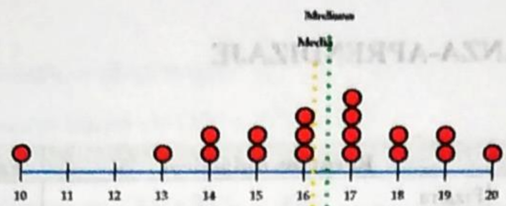
además:

$$\text{Rango Intercuartil} = 18 - 15 = 3$$

Si el coeficiente de asimetría de la distribución es $-0{,}7299$ y el coeficiente de curtosis es 3,4997, interprete las características de la distribución de datos. (0.5 puntos)

Dado que el coeficiente de asimetría es menor a cero, la distribución de los datos será simétrica a la izquierda, esto implica que la cola izquierda de la distribución es más larga que la cola derecha. Por otro lado, un valor positivo de curtosis implica que la distribución de datos es más apuntada que la distribución normal, en tal virtud, las observaciones se encuentran muy concentradas en torno a su media y existe poca presencia de valores extremos.

c) Represente los datos mediante un diagrama de puntos, en él ubique la media y la mediana, a continuación comente sobre la concentración de los datos. (0.5 puntos)



Los datos se encuentran concentrados alrededor de la media y mediana, pues ambas medidas de localización se encuentra cercanas.

2. En una empresa de selección de personal aplicó un examen a 60 aspirantes a ocupar varios cargos en una agencia bancaria. Se anotó el tiempo (en minutos) que cada uno de los aspirantes se demoró en contestar el examen. Todos los tiempos se resumieron en la siguiente tabla de frecuencias. (Valor: 3 puntos)

Tiempo	Punto medio (x_i)	Frecuencia absoluta (n_i)	Frecuencia relativa (f_i)	Frecuencia absoluta acumulada (N_i)	Frecuencia relativa acumulada (F_i)
-	75		0.1		
-		14			
-					
-			0.2	52	
-	115				

a) Complete la tabla de frecuencias teniendo en cuenta que la amplitud de cada clase W es constante. (1 punto)

Datos

n: Número total de personas

$f_1 = 0{,}1;\ \ f_4 = 0{,}2;\ \ n_2 = 14;\ \ N_4 = 52;\ \ F_1 = f_1 = 0{,}1;\ \ n = N_5 = 60;\ \ x_1 = 75;\ \ x_5 = 115;$

Planteamiento

$$f_4 = \frac{n_4}{n} \rightarrow n_4 = f_4 * n$$
$$n_4 = 0{,}2 * 60 = 12 \tag{1}$$

$$f_1 = \frac{n_1}{n} \rightarrow n_1 = f_1 * n$$
$$n_1 = 0{,}1 * 60 = 6 = N_2 \tag{2}$$

$$N_2 = n_1 + n_2 = 6 + 14$$
$$N_2 = 20 \tag{3}$$

$$f_2 = \frac{n_2}{n} = \frac{14}{60} = 0{,}23 \tag{4}$$

$$N_3 = N_4 - n_4 = 52 - 12 = 40 \tag{5}$$

$$N_2 = N_3 - n_3 \rightarrow n_3 = N_3 - N_2$$
$$n_3 = 40 - 20 = 20 \tag{6}$$

$$N_4 = N_5 - n_5 \rightarrow n_5 = N_5 - N_4$$
$$n_5 = 60 - 52 = 8 \tag{7}$$

$$f_3 = \frac{n_3}{n} = \frac{20}{60} = 0{,}33 \tag{8}$$

parte derecha de la caja es ligeramente mayor que la de la izquierda; ello quiere decir que los tiempos comprendidos entre el 50 % y el 75 % de la muestra están más dispersos que entre el 25 % y el 50 %. Además, ambos bigotes no presentan una gran diferencia entre ellos, esto implicaría la no existencia de datos atípicos.

3. Una institución educativa de la ciudad cuenta con 1200 estudiantes: un 20 % pertenece a formación inicial, otro 45 % pertenece a formación básica y el resto a bachillerato. Además, se conoce que los estudiantes provienen de tres sectores: 400 estudiantes viven en el sector norte, 600 estudiantes viven en el sector centro y el resto viven en el sector sur. Por último, 210 estudiantes de formación básica viven en el sector norte, 320 estudiantes de bachillerato viven en el sector centro, 110 estudiantes de formación inicial viven en el sector sur y 70 estudiantes del bachillerato viven en el sector norte.

Solución

A partir de la información proporcionada se obtiene que:

- El número de estudiantes de formación inicial es: 1200 * 20 % = 240.
- El número de estudiantes de formación básica es: 1200 * 45 % = 540.
- El número de estudiantes de bachillerato es: 1200 * 35 % = 420.

con lo cual se procede a completar la tabla:

	Formación Inicial (I)	Formación Básica (B)	Bachillerato (A)	Total
Sector Norte (SN)	120	210	70	400
Sector Centro (SC)	10	270	320	600
Sector Sur (SS)	110	60	30	200
Total	240	540	420	1200

Se escoge al azar a un estudiante de la institución educativa. Calcule la probabilidad de que dicho estudiante:

a) Sea de formación inicial o viva en el sector norte. (0.5 puntos)

$$\begin{aligned} Pr(I \cup SN) &= Pr(I) + Pr(SN) - Pr(I \cap SN) \\ &= \frac{240}{1200} + \frac{400}{1200} - \frac{120}{1200} \\ &= \frac{520}{1200} \approx 0{,}4333 \end{aligned}$$

b) Sea de bachillerato y no viva en el sector sur. (0.5 puntos)

$$\begin{aligned} Pr(A \cap SS^C) &= Pr(A) - Pr(A \cap SS) \\ &= \frac{420}{1200} - \frac{30}{1200} \\ &= \frac{390}{1200} \approx 0{,}3250 \end{aligned}$$

c) No viva en el sector centro o sea de formación básica. (0.5 puntos)

$$\begin{aligned} Pr(SC^C \cup B) &= Pr(SC^C) + Pr(B) - Pr(SC^C \cap B) \\ &= 1 - Pr(SC) + Pr(B) - Pr(B) + Pr(SC \cap B) \\ &= 1 - Pr(SC) + Pr(SC \cap B) \\ &= 1 - \frac{600}{1200} + \frac{270}{1200} \\ &= \frac{870}{1200} \approx 0{,}7250 \end{aligned}$$

d) Viva en el sector sur y, sea de formación básica o bachillerato. (0.5 puntos)

$$\begin{aligned} Pr(SS \cap (B \cup A)) &= Pr((SS \cap B) \cup (SS \cap A)) \\ &= Pr(SS \cap B) + Pr(SS \cap A) - Pr((SS \cap B) \cap (SS \cap A)) \\ &= \frac{60}{1200} + \frac{30}{1200} - \frac{0}{1200} \\ &= \frac{90}{1200} \approx 0{,}0750 \end{aligned}$$

$$f_5 = \frac{n_5}{n} = \frac{8}{60} = 0{,}133 \quad (9)$$

Planteamiento de Ecuaciones para encontrar el ancho del intervalo

$$\begin{aligned} 115 - 75 &= 4W \\ W &= 10 \end{aligned} \quad (10)$$

$$\begin{aligned} x_2 &= x_1 + W \\ x_2 &= 75 + 10 = 85 \end{aligned} \quad (11)$$

$$\begin{aligned} L_1 &= x_1 - \frac{W}{2} \\ L_1 &= 75 - 5 = 70 \end{aligned} \quad (12)$$

$$\begin{aligned} S_1 &= x_1 + \frac{W}{2} \\ S_1 &= 75 + 5 = 80 \end{aligned} \quad (13)$$

TABLA DE FRECUENCIAS COMPLETA

Tiempo	Punto medio (x_i)	Frecuencia absoluta (n_i)	Frecuencia relativa (f_i)	Frecuencia absoluta acumulada (N_i)	Frecuencia relativa acumulada (F_i)
70-80	75	6	0.1	6	0.1
80-90	85	14	0.23	20	0.33
90-100	95	20	0.33	40	0.66
100-110	105	12	0.2	52	0.86
110-120	115	8	0.13	60	1

b) Indique el porcentaje de aspirantes que rindieron el examen en menos de 80 minutos. (0.5 puntos) **Solución**

$$Porcentaje = \frac{6 * 100\,\%}{60} = 10\,\% \quad (14)$$

c) Halle el promedio y la desviación estándar de la variable de Tiempo. (0.5 puntos)
Media

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^{5}(n_i)(x_i)}{n} \\ \bar{x} &= \frac{(6)(75) + (14)(85) + (20)(95) + (12)(105) + (8)(115)}{60} \\ \bar{x} &= 95{,}33 \end{aligned} \quad (15)$$

Desviación estándar

$$\begin{aligned} \sigma &= \sqrt{\frac{\sum_{i=1}^{5}(n_i)(x_i)^2 - n\bar{x}^2}{n-1}} \\ \sigma &= \sqrt{\frac{553500 - (60)(95{,}33)^2}{59}} \\ \sigma &= 11{,}81 \end{aligned} \quad (16)$$

d) Dibuje el diagrama de caja, a partir de este gráfico responda lo siguiente: *i*) ¿En qué cuartil están más dispersos los datos?; *ii*) ¿Existen valores atípicos? (1 punto)

Solución
Método Gráfico (Ojiva)

b) $A^c \cap B^c = (A \cup B)^c = B^c = \{(1,4),(1,5),(1,6),\cdots,(6,6)\}$, $|A^c \cap B^c| = 1$.

3. $A = \{(1,4),(4,1),(2,4),(4,2),(4,5),(4,6)\}$, $P(A) = \frac{10}{36}$.

4. $A = \{(3,4),(4,3),(4,6),(6,4)\}$, $P(A) = \frac{4}{36}$.

$\square$

Escuela Politécnica Nacional

Estadística • Prueba 1

Semestre 2019-A | Departamento de Formación Básica

1. Se lanzan al aire 3 monedas distintas (Azul, Amarilla y Roja).

   a) Determinar el espacio muestral del experimento.

   b) Sean los sucesos A: "Salir al menos una cara", B: "Salir máximo un sello", determinar dichos sucesos como subconjuntos del espacio muestral y además encontrar:

      1) $A \cup B$,

      2) $A^c \cap B^c$.

   c) Encontrar la probabilidad de que exactamente una moneda haya salido cara.

   d) Encontrar la probabilidad de que exactamente una moneda haya salido cara y al menos dos monedas hayan salido sello.

Solución.

1. $S = \{ccc, ccs, csc, scc, css, scs, ssc, sss\}$.

2. $A = \{ccc, ccs, csc, scc, css, scs, ssc\}$, $B = \{ccc, ccs, csc, scc\}$.

   a) $A \cup B = \{ccc, ccs, csc, scc, css, scs, ssc\}$,

   b) $A^c \cap B^c = (A \cup B)^c = \{sss\}$.

3. $A = \{css, scs, ssc\}$, $P(A) = \dfrac{3}{8}$.

4. $A = \{css, scs, ssc\}$, $B = \{css, scs, ssc, sss\}$, $A \cap B = \{css, scs, ssc\}$, $P(A \cap B) = \dfrac{3}{8}$.

□

1. Se lanzan al aire dos dados distinguibles y se anotan los resultados de los dados.

   a) Determinar el espacio muestral del experimento.

   b) Sean los sucesos A: "Salir al menos un 2", B: "Salir máximo un tres", determinar dichos sucesos como subconjuntos del espacio muestral y además encontrar:

      1) $A \cup B$,

      2) $A^c \cap B^c$.

   c) Encontrar la probabilidad de que exactamente un dado haya salido 4.

   d) Encontrar la probabilidad de que haya salido un 4 y un múltiplo de 3.

Solución.

1. $S = \{(1,1), (1,2), (1,3), (1,4), \cdots, (5,6), (6,6)\}$, $|S| = 36$.

2. $A = \{(1,2), (2,1), (2,2), (2,3), (3,2) \cdots, (2,6)\}$, $|A| = 11$, $B^c = \{(3,3)\}$, $|B| = 35$

   a) $A \cup B = B$, $|A \cup B| = 35$

4. Se lanzan al aire dos dados indistinguibles y se anotan los resultados de los dados.

   a) Determinar el espacio muestral del experimento. (0.5 puntos)

   b) Sean los sucesos A: “Salir al menos un 2”, B: “Salir máximo un tres”, determinar dichos sucesos como subconjuntos del espacio muestral y además encontrar:

      1) $A \cup B$, (0.75 puntos)

      2) $A^c \cap B^c$. (0.75 puntos)

   c) Encontrar la probabilidad de que exactamente un dado haya salido 4. (0.5 puntos)

   d) Encontrar la probabilidad de que haya salido un 4 y un múltiplo de 3. (0.5 puntos)

EJERCICIOS:

1. En una fábrica se ha medido la longitud (en cm) de 20 piezas de las mismas características y se han obtenido los siguientes datos:

$$93, 80, 81, 92, 87, 77, 79, 73, 81, 87, 55, 85, 68, 90, 85, 80, 90, 91, 90, 88$$

   a) Calcule las siguientes medidas: media aritmética, moda, mediana, desviación estándar y rango intercuartil. (1 punto)
   b) Dado el coeficiente de asimetría $-1{,}4243$ y el coeficiente de curtosis 4,9818, interprete la distribución de datos. (0.5 puntos)
   c) Represente los datos mediante un diagrama de puntos, en el ubique la media y la mediana, a continuación comente sobre la concentración de los datos. (0.5 puntos)

2. La siguiente tabla de distribución de frecuencias representa la distribución del tiempo (en minutos) requerida por las personas atendidas en un dispensario médico.

Tiempo	Punto medio (x_i)	Frecuencia absoluta (n_i)	Frecuencia relativa (f_i)	Frecuencia absoluta acumulada (N_i)	Frecuencia relativa acumulada (F_i)
3-		2			
-				8	
-	18				0.6
-		10			

   a) Complete la tabla de frecuencias sabiendo que los intervalos tienen la misma longitud W.(1 punto)
   b) Halle el promedio y la desviación estándar de la variable de Tiempo.(1 punto)
   c) Dibuje el diagrama de caja, a partir de este gráfico responda lo siguiente: i) ¿En qué cuartil están más dispersos los datos?; ii) ¿Existen valores atípicos? (1 punto)

3. Se conoce que una pequeña ciudad cuenta con 900 habitantes: el 25 % son niños, el 30 % son jóvenes y el resto son adultos. Además, se conoce que 340 habitantes viven en la zona norte, 240 habitantes viven en la zona centro y el resto viven en la zona sur. Por último, 110 jóvenes viven en la zona sur, 150 adultos viven en la zona centro, 80 niños viven en la zona sur y 130 jóvenes viven en la zona norte.
   Se escoge al azar a un habitante de la pequeña ciudad. Calcule la probabilidad de que dicho habitante:

   a) Sea niño o viva en la zona norte. (0.5 puntos)
   b) Sea adulto y no viva en la zona sur. (0.5 puntos)
   c) No viva en la zona centro o sea joven. (0.5 puntos)
   d) Viva en la zona sur y, sea joven o adulto. (0.5 puntos)

	Niño (N)	Joven (J)	Adulto (A)	Total
Zona Norte (ZN)				
Zona Centro (ZC)				
Zona Sur (ZS)				
Total				

4. Se lanzan al aire 3 monedas distintas (Azul, Amarilla y Roja).

   a) Determinar el espacio muestral del experimento. (0.5 puntos)
   b) Sean los sucesos A: "Salir al menos una cara", B: "Salir máximo un sello", determinar dichos sucesos como subconjuntos del espacio muestral y además encontrar:
      1) $A \cup B$, (0.75 puntos)
      2) $A^c \cap B^c$. (0.75 puntos)
   c) Encontrar la probabilidad de que exactamente una moneda haya salido cara. (0.5 puntos)
   d) Encontrar la probabilidad de que exactamente una moneda haya salido cara y al menos dos monedas hayan salido sello. (0.5 puntos)