

Solución del Examen

Nombre:
Apellido:

C.I.:

Fecha: 22/07/2004
Firma:

Prof. Gudberto León

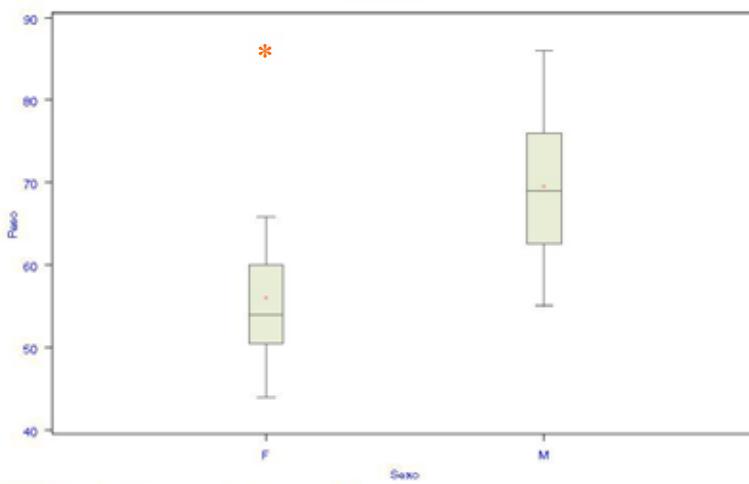
MÉTODOS ESTADÍSTICOS I – EXAMEN I

PARTE I: (Cada respuesta correcta tiene un valor de 0,5 puntos)

En los siguientes gráficos se representan las distribuciones de los pesos (Kg) de los estudiantes de Métodos Estadísticos I clasificados por sexo

Diagrama de Caja para la variable Peso

Datos de Estudiantes de Métodos Estadísticos I
Semestre B2004



Las siguientes preguntas son con respecto a las distribuciones representadas por los gráficos anteriores. Marque con un círculo la respuesta correcta:

1. En cuanto a la dispersión:
 - a. En la distribución de femenino es mayor que en la de masculino
 - b. En la distribución de femenino es menor que en la de masculino
 - c. En la distribución de femenino es igual a la de masculino
 - d. Los datos no presentan dispersión
 - e. No se puede identificar la dispersión
2. En cuanto a la dispersión de la parte central:
 - a. En la distribución de femenino es mayor que en la de masculino
 - b. En la distribución de femenino es menor que en la de masculino
 - c. En la distribución de femenino es igual a la de masculino
 - d. Los datos no presentan dispersión
 - e. No se puede identificar la dispersión
3. En cuanto a la forma de la distribución, los pesos de los masculinos presentan:
 - a. Asimetría positiva
 - b. Asimetría negativa
 - c. Simetría (aproximadamente)
 - d. No se puede determinar la forma
 - e. El gráfico no muestra esta información

4. En cuanto a la forma de la distribución, los pesos de los femeninos presentan:

- a. Asimetría positiva
- b. Asimetría negativa
- c. Simetría (aproximadamente)
- d. No se puede determinar la forma
- e. El gráfico no muestra esta información

5. En la distribución de pesos de las féminas no es recomendable usar como medida de tendencia central:

- a. La moda
- b. La mediana
- c. La media
- d. Cualquiera de las tres anteriores
- e. Ni media, ni mediana, ni moda

→ ¡Por la presencia de un valor atípico!

PARTE II:

Nota:

- i. Recuerde que este es un examen de desarrollo por lo que **debe** incluir **todos** los pasos necesarios que justifiquen los resultados. Favor encerrar sus respuestas de forma tal que sea fácil encontrarlas en su desarrollo. Solamente se responderán aquellas preguntas tendientes a aclarar enunciados de los problemas.
- ii. Las respuestas a las preguntas de la Parte II de este examen debe escribirlas en **papel tipo examen** (papel ministro) No son válidas las respuestas escritas en esta hoja de preguntas.

Solución del Examen

1. Sea x_1, x_2, \dots, x_n una colección de datos cuya media es \bar{x} . Demuestre que si a cada uno de los datos anteriores se les multiplica por la misma constante k , la media de estos “nuevos datos” viene dada por $\bar{x} * k$. (2 puntos)

SOLUCIÓN:

Por tanto los nuevos datos que se obtienen al modificar los datos originales los podemos escribir como: $x_1*k, x_2*k, \dots, x_n*k$.

Entonces, la media de estos “nuevos datos” viene dada por:

$$\bar{x} = \frac{\sum_{i=1}^n (x_i * k)}{n} = \frac{k \sum_{i=1}^n x_i}{n} = k * \bar{x}$$

2. En relación a los datos presentados en el siguiente cuadro, responda:

- ¿Qué nombre recibe este tipo de datos?
- Grafiqüe los datos.
- Interprete el gráfico obtenido en b.

(3,5 puntos)

**ÍNDICE GENERAL DE PRECIOS AL CONSUMIDOR
ÁREA METROPOLITANA DE CARACAS
Ene. 2002 – Sept. 2002 (en Porcentajes)
(BASE 1997 = 100)**

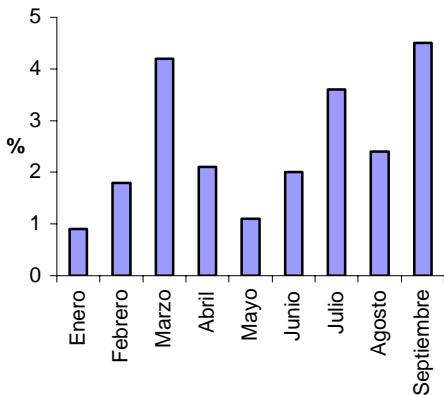
Mes	IPC
Enero	0,9
Febrero	1,8
Marzo	4,2
Abri	2,1
Mayo	1,1
Junio	2,0
Julio	3,6
Agosto	2,4
Septiembre	4,5

Fuente: BCV

SOLUCIÓN:

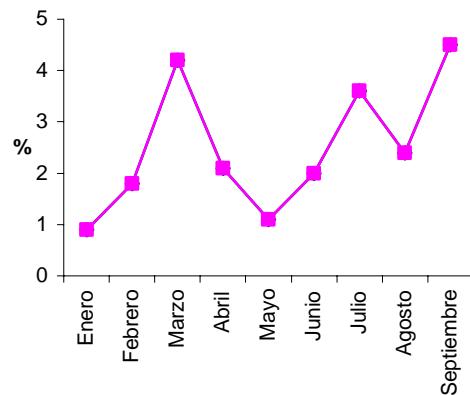
- Serie Cronológica o Serie de tiempo.
- Para representar gráficamente una serie de tiempo, se puede utilizar el gráfico de barras o el de curvas:

**ÍNDICE GENERAL DE PRECIOS AL CONSUMIDOR
ÁREA METROPOLITANA DE CARACAS
Ene. 2002 – Sept. 2002 (en Porcentajes)
(BASE 1997 = 100)**



Fuente: BCV

**ÍNDICE GENERAL DE PRECIOS AL CONSUMIDOR
ÁREA METROPOLITANA DE CARACAS
Ene. 2002 – Sept. 2002 (en Porcentajes)
(BASE 1997 = 100)**



Fuente: BCV

- Se puede observar un comportamiento inestable, de aumento y disminución, en el IPC en este período de tiempo. En los meses de Marzo, Julio y Septiembre se presentó un IPC bastante alto. Los meses con menor “inflación” (IPC) en el período Enero-Septiembre 2002 son Enero y Mayo.

Solución del Examen

3. Suponga que un inversionista piensa adquirir acciones en una de dos compañías, A o B, listadas en la Bolsa de Valores de Nueva York. Suponga también, que cada acción de la compañía A ha promediado 50 dólares en los últimos meses, con una desviación estándar de 10 dólares. Y durante el mismo período el precio promedio de las acciones en la compañía B fue de 12 dólares con una desviación estándar de 4 dólares. Responda lo siguiente:

- ¿Cómo puede determinar el inversionista cuáles acciones son más variables? Justifique su respuesta.
- Si las acciones de la compañía A tienen un incremento de 2 dólares que se mantiene en los próximos 6 meses
 - ¿Cuál será el precio promedio (media) de las acciones en esos 6 meses? Justifique su respuesta.
 - ¿y la desviación estándar? Justifique su respuesta.

(Basado en Berenson, M, Levine, D. y Krehbiel, T. (2000). *Estadística para Administración*)

(4 puntos)

SOLUCIÓN:

Compañía A	Compañía B
$\bar{X}_A = 50$	$\bar{X}_B = 12$
$S_A = 10$	$S_B = 4$

a.
$$CV_A = \frac{10}{50} * 100\% = 20\%$$

$$CV_B = \frac{4}{12} * 100\% = 33\%$$

La compañía B presenta la mayor variabilidad en sus acciones.

- b. Se tiene que los precios de las acciones de la compañía A han aumentado 2\$, es decir, $x_1+2, x_2+2, \dots, x_n+2$.
- Entonces, por propiedades de la media aritmética, la nueva media viene dada por $\bar{X}_A + 2 = 50 + 2 = 52\$$
 - Por propiedades de la desviación estándar se tiene que la nueva desviación estándar se mantiene igual en este caso, es decir, $S = 10$

4. Los sueldos (en decenas de miles de bolívares)^a de los empleados de una compañía se han agrupado en la siguiente distribución de frecuencias: **(8 puntos)**

Sueldos	f _i	fr _i	F _i	Fr _i
[35 - 45)			3	
[45 - 55)			16	0,2000
[55 - 65) 15				0,3875
[65 - 75) 0,2500	51			
[75 - 85) 0,2000				
[85 - 95) 11				0,9750
[95 - 105) 2				

Nota:

^a $35 = 35 * 10000 = \text{Bs.} 350000$

- Complete la distribución de frecuencias
- Construya el gráfico que usted considere conveniente para estos datos y analice la forma en que se distribuyen los mismos.
- Encuentre el valor del sueldo por arriba del cual se encuentra aproximadamente el 25% de los sueldos de los empleados. Justifique su respuesta.
- Calcule e interprete en términos del problema:
 - La Media
 - La Desviación Estándar
- ¿Qué porcentaje (aproximadamente) de empleados gana menos de Bs. 500000? Justifique su respuesta.
- ¿La media calculada en la parte (d) es representativa de estos datos? Justifique su respuesta.

SOLUCIÓN:

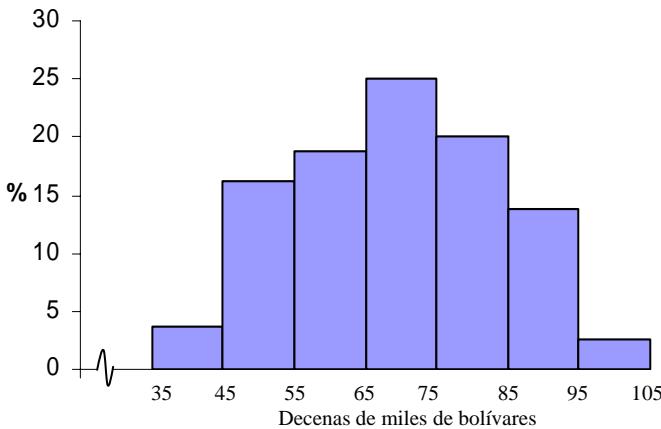
(a)

Sueldos	f _i	fr _i	F _i	Fr _i
[35 - 45) 3	0,0375	3	0,0375	
[45 - 55) 13	0,1625	16	0,2000	
[55 - 65) 15	0,1875	31	0,3875	
[65 - 75) 20	0,2500	51	0,6375	
[75 - 85) 16	0,2000	67	0,8375	
[85 - 95) 11	0,1375	78	0,9750	
[95 - 105) 2	0,0250	80	1	

Solución del Examen

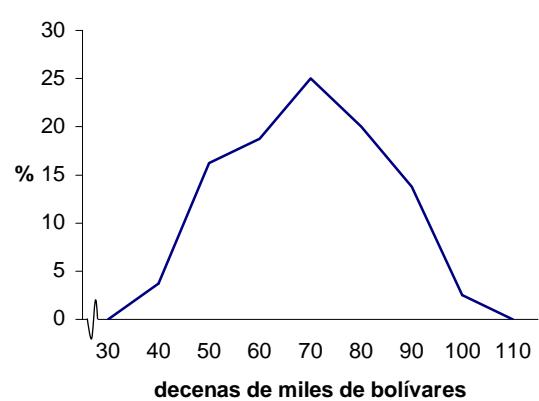
- (b) Se puede construir el histograma o el polígono.

Sueldos de los empleados de una compañía



Fuente: Datos de la Compañía

Sueldos de los empleados de una compañía



Fuente: Datos de la Compañía

- Claramente se observa que la distribución de los sueldos de esa compañía es **APROXIMADAMENTE SIMÉTRICA**. Los sueldos se concentran mayormente entre 550.000 y 850.000 bolívares. Muy pocos empleados tienen sueldos inferiores a 450.000 y superiores a 950.000 bolívares.
- (c) El valor del sueldo por arriba del cual se encuentra aproximadamente el 25% de los sueldos de los empleados, corresponde al cuartil tres o percentil 75. Así, usando la fórmula para percentiles y ubicando la clase del percentil 75 en la distribución de frecuencias, que en este caso es la quinta clase, se tiene que:

$$Q_3 = P_{75} = 75 + \left[\frac{80 * \left(\frac{75}{100} \right) - 51}{16} \right] * 10 = 80,625$$

Es decir, el sueldo por arriba del cual se encuentra aproximadamente el 25% de los sueldos es:

$$80,625 * 10000 = \boxed{806.250 \text{ bolívares}}$$

Y por debajo de ese valor se encuentra aproximadamente el 75% de los sueldos de los empleados de esta compañía.

Nota: Este valor también se puede obtener construyendo la Ojiva y luego usando el procedimiento gráfico de interpolación. Sin embargo, el resultado se obtiene de manera mucho más rápida usando la fórmula de los percentiles.

- (d) Para obtener la media y la desviación estándar se necesitan los siguientes cálculos:

m_i	$m_i * f_i$	$m_i^2 * f_i$
40	120	4800
50	650	32500
60	900	54000
70	1400	98000
80	1280	102400
90	990	89100
100	200	20000
	5540	400800

i. $\bar{X} = \frac{5540}{80} = 69,25$

Por tanto, el valor de la media es: $69,25 * 10000 = \boxed{692.500 \text{ bolívares}}$. Así, el sueldo promedio que paga la compañía a sus empleados es 692500 bolívares y representa el punto de equilibrio o centro de gravedad de los sueldos de los empleados.

- ii. En primer lugar se obtiene la varianza:

$$S^2 = \sum_{i=1}^n \frac{m_i^2 * f_i}{n} - \bar{x}^2$$

Solución del Examen

$$S^2 = \frac{400800}{80} - (69,25)^2 = 214,4375$$

Luego, se obtiene la desviación estándar:

$$S = \sqrt{S^2} = \sqrt{214,4375} = 14,643684$$

De tal manera que la desviación estándar es $14,643684 * 10000 = \boxed{146,436,85 \text{ bolívares}}$. Este valor representa el grado de dispersión de los sueldos con respecto a su media aritmética. El valor de la desviación estándar obtenida no es muy grande comparada con los valores de los datos en estudio. Eso significa que estos datos no presentan mucha dispersión alrededor de su media aritmética. Esto se puede corroborar observando el histograma o el polígono de la parte (b).

- (e) Para encontrar el porcentaje de empleados gana menos de Bs. 500.000 se despeja h de la fórmula de percentiles, obteniendo que:

$$h = \left[\frac{(P_h - LI_p) f_p}{C_p} + Fap \right] \frac{100}{n}$$

Es necesario ubicar la clase percentil en la distribución de frecuencias. En este caso es la clase que contiene a 500000 ó 55 en la notación simplificada. Así, la segunda clase es la *clase percentil*. Entonces, sustituyendo los valores en la fórmula anterior:

$$h = \left[\frac{(50 - 45) * 13}{10} + 3 \right] \frac{100}{80} = 11,9\%$$

Por tanto, **11,9%** es el porcentaje (aproximado) de empleados que gana menos de Bs.500.000.

Nota: Este valor también se puede obtener construyendo la Ojiva y luego usando el procedimiento gráfico de interpolación. Sin embargo, el resultado se obtiene de manera mucho más rápida usando la fórmula anterior.

- (f) La media aritmética es representativa del conjunto de datos cuando estos no presentan mucha dispersión. En este caso la desviación estándar calculada en la parte (d) ii. indica que los sueldos presentan poca dispersión, lo cual se corrobora visualmente con los gráficos de la parte (b). También se puede calcular el coeficiente de variación de los sueldos:

$$CV = \frac{S}{\bar{x}} * 100 = \frac{14,64}{69,25} * 100 = 21,14\%$$

El *CV* indica que la desviación estándar representa el 21,14% del valor de la media aritmética. Es decir, según el coeficiente de variación los sueldos de los empleados de la compañía, presentan poca dispersión alrededor de su media aritmética.

En conclusión, dado que estos datos presentan poca dispersión alrededor de la media, $\bar{X} = \boxed{692.500 \text{ bolívares}}$ es representativa de los sueldos de los empleados de la compañía

FORMULARIO:

$C_i = R / K$	$K = 1 + 3,3 * \log(n)$	$RQ = Q_3 - Q_1$
$\bar{x} = \sum_{i=1}^k \frac{m_i * f_i}{n}$	$\bar{x} = \left(\sum_{i=1}^n x_i * w_i \right) / \sum_{i=1}^n w_i$	$S^2 = \sum_{i=1}^n \frac{m_i^2 * f_i}{n} - \bar{x}^2 ; S^2 = \sum_{i=1}^n \frac{(m_i - \bar{x})^2 * f_i}{n}$
$CV = \frac{S}{\bar{x}} * 100$	$Q = \frac{Q_3 - Q_1}{2}$	$\frac{\overline{AR}}{\overline{AB}} = \frac{\overline{RS}}{\overline{BC}}$
$Md = LI_m + \left(\frac{n/2 - Fam}{fm} \right) * C_m$	$P_h = LI_p + \left[\frac{n * \left(\frac{h}{100} \right) - F_{ap}}{f_p} \right] * C_p$	$ASP = \frac{3(\bar{x} - Md)}{S_*}$
$\gamma_1 = \frac{\left(\sum_{i=1}^k (m_i - \bar{x})^3 f_i \right)}{S_*^3}$	$\beta_2 = \frac{\left(\sum_{i=1}^k (x_i - \bar{x})^4 \right)}{S_*^4}$	$\gamma_1 = \frac{\left(\sum_{i=1}^k (x_i - \bar{x})^3 \right)}{S_*^3} \quad \beta_2 = \frac{\left(\sum_{i=1}^k (m_i - \bar{x})^4 f_i \right)}{S_*^4}$