

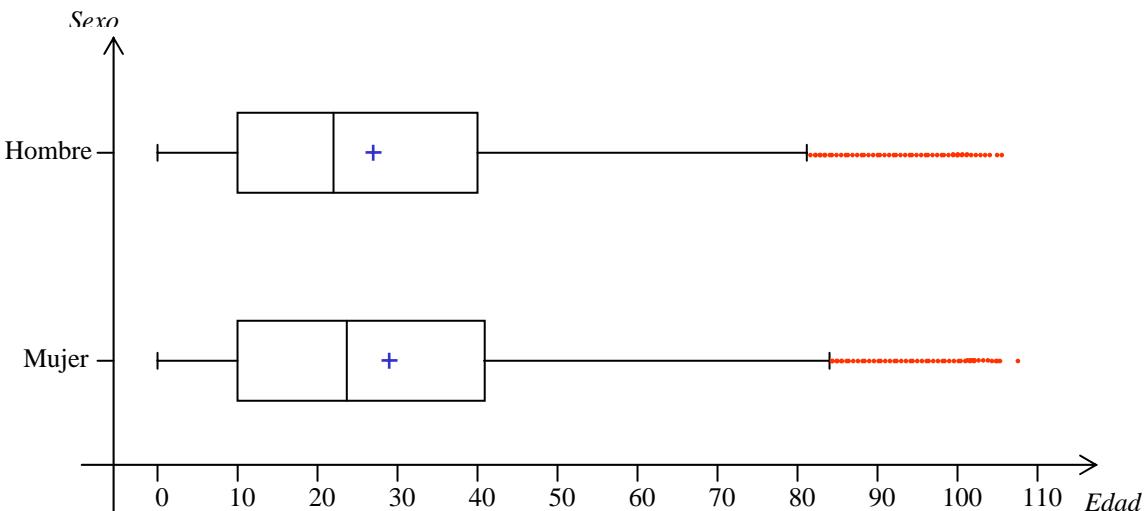
## SOLUCIÓN EXAMEN I

### MÉTODOS ESTADÍSTICOS I – EXAMEN I

Prof. Gudberto León

#### PARTE I: Marque con un círculo la respuesta correcta (0,5 puntos. c/u)

En los siguientes diagramas de caja se representan las distribuciones de las edades de los habitantes del Estado Mérida clasificados por sexo, según el Censo 2001:



Fuente: <http://www.sisov.mpd.gov.ve/>

Las siguientes preguntas se refieren 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 mujer es mayor que en la de hombre
    - b. En la distribución de mujer es menor que en la de hombre
    - c. En la distribución de mujer es aproximadamente igual a la de hombre
    - d. Los datos no presentan dispersión
    - e. No se puede identificar la dispersión
  2. El recorrido intercuartil en la distribución de los hombres es aproximadamente igual a:
    - a. 40
    - b. 30
    - c. 10
    - d. 85
    - e. 25
  3. En cuanto a la forma de la distribución, las edades de los hombres 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. Al comparar las distribuciones de edad de hombres y mujeres se concluye claramente que:
  - a. Hay aproximadamente 7 mujeres por cada hombre
  - b. En general los hombres son más viejos que las mujeres
  - c. Ambas distribuciones de edades son muy parecidas
  - d. La media de las mujeres es aproximadamente igual a la mediana
  - e. En Mérida hay más viejos que jóvenes
5. En la distribución de edades de las mujeres Ud. no recomendaría 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

#### PARTE II: Marque con un círculo la respuesta correcta o llene los espacios en blanco. (0,5 puntos. c/u)

1. En estadística, cuando se analiza una muestra aleatoria para generalizar las conclusiones obtenidas a la población, se conoce con el nombre de:
    - a. Muestreo
    - b. Estadística Descriptiva
    - c. Inferencia Estadística
    - d. Censo
    - e. Parámetro
  2. ¿Por qué es necesario elevar al cuadrado las diferencias con respecto a la media cuando calculamos la varianza?
    - a. Para que los valores extremos no afecten el cálculo.
    - b. Porque es posible que  $n$  sea muy pequeña.
    - c. Porque algunas de las diferencias serán positivas y otras negativas.
    - d. Para que la varianza sea representativa
    - e. Para facilitar los cálculos
3. Una organización de investigación realiza un estudio del precio de venta de computadoras personales. Se tienen 45 computadoras en el estudio. ¿Cuántas clases recomendaría?
  - a. 15
  - b. 20
  - c. 3
  - d. 13
  - e. 6
4. ¿Cuál de los siguientes es un ejemplo de parámetro?
  - a.  $\bar{X}$
  - b.  $N$
  - c.  $\mu$
  - d. La población
  - e. La muestra
5. Una gráfica de una distribución de frecuencias acumuladas se conoce como **OJIVA**

## SOLUCIÓN EXAMEN I

### PARTE III:

**Nota:** 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 de encontrarlas en su desarrollo. Solamente se responderán aquellas preguntas tendientes a aclarar enunciados de los problemas.

1. En Mérida Electronics están pensando en adoptar uno de dos programas de entrenamiento. Dos grupos distintos fueron entrenados para realizar el mismo trabajo. El grupo 1, conformado por 10 empleados, fue entrenado con el programa A; el grupo 2 de 13 empleados, con el B . En el grupo 1 se ocupó un tiempo promedio de 32.11 horas para entrenar a cada empleado, con una varianza de 68.09. Para el segundo, el tiempo promedio fue de 10.75 horas de entrenamiento para cada empleado, con una varianza de 71.14.
  - a. Calcule el tiempo promedio ocupado por los dos programas para entrenar a los empleados.
  - b. ¿Qué programa de entrenamiento presenta mayor homogeneidad en el tiempo de entrenamiento de cada empleado? Justifique. **(3 puntos)**
2. Según el Censo 2001, la situación conyugal declarada por los habitantes del Estado Mérida es la siguiente:

Situación Conyugal	%
Unido(a)	16,7
Casado(a)	26,52
Soltero(a)	48,44
Separado(a) de unión o matrimonio	2,61
Divorciado(a)	2,22
Viudo(a) de unión o matrimonio	3,51
<b>Total</b>	<b>100</b>

En relación a la tabla de al lado responda los siguientes:

- a. ¿De qué tipo son los datos?
- b. Realice y comente un gráfico conveniente para representar estos datos.
- c. Calcule e interprete la(s) medida(s) de posición adecuada(s) en este caso.

**(4 puntos)**

Fuente: <http://www.sisov.mpd.gov.ve/>

3. Un conjunto de 40 préstamos personales realizados por un banco se organizaron de acuerdo a su magnitud, en la siguiente distribución de frecuencias:

Monto (Millones de bolívares)	No. de préstamos	fr <sub>i</sub>	F <sub>i</sub>	Fr <sub>i</sub>
[1 – 2)	18	0,450	18	0,450
[2 – 3)	9	0,225	27	0,675
[3 – 4)	5	0,125	32	0,800
[4 – 5)	5	0,125	37	0,925
[5 – 6)	3	0,075	40	1

- a. Interprete en términos del problema las frecuencias f<sub>3</sub> y Fr<sub>4</sub>
- b. Grafique la información presentada en la distribución de frecuencias y comente acerca de la forma en que se distribuyen los datos.
- c. Calcule e interprete:
  - I. Media II. Mediana III. Desviación Estándar IV. Primer Cuartil
- d. ¿El 5% de los préstamos más altos son superiores a que monto (en Bs.)? Argumente su respuesta.
- e. ¿Qué tan confiable es la media aritmética como medida de tendencia central de estos datos? Argumente su respuesta.

**(8 puntos)**

### **FORMULARIO:**

$$\bar{x} = \frac{\sum_{i=1}^k m_i * f_i}{n}; \quad \bar{x} = \left( \sum_{i=1}^n x_i * w_i \right) / \sum_{i=1}^n w_i; \quad S^2 = \sum_{i=1}^k \frac{m_i^2 * f_i}{n} - \bar{x}^2; \quad S^2 = \sum_{i=1}^k \frac{(m_i - \bar{x})^2 * f_i}{n}; \quad CV = \frac{S}{\bar{x}} * 100;$$

$$Q = (Q_3 - Q_1)/2; \quad RQ = Q_3 - Q_1; \quad Md = LI_m + \left[ \left( \frac{n}{2} - Fa_m \right) / fm \right] * C_m; \quad P_h = LI_p + \left[ n * \left( \frac{h}{100} \right) - F_{ap} \right] / f_p * C_p;$$

$$C_i = R / K; \quad K = 1 + 3,3 * \log(n); \quad \overline{AR} / \overline{AB} = \overline{RS} / \overline{BC}; \quad ASP = 3(\bar{x} - Md) / S_*; \quad \gamma_1 = \left( \sum_{i=1}^k (m_i - \bar{x})^3 f_i / n \right) / S_*^3;$$

$$\beta_2 = \left( \sum_{i=1}^k (x_i - \bar{x})^4 / n \right) / S_*^4; \quad \gamma_1 = \left( \sum_{i=1}^k (x_i - \bar{x})^3 / n \right) / S_*^3 \quad \beta_2 = \left( \sum_{i=1}^k (m_i - \bar{x})^4 f_i / n \right) / S_*^4$$

## SOLUCIÓN EXAMEN I

### Solución Examen I

(04/11/05)

#### Punto III:

Grupo 1 (Programa A)	Grupo 2 (Programa B)
$n_1 = 10$	$n_2 = 13$
$\bar{x}_1 = 32,11$ horas	$\bar{x}_2 = 10,75$ horas
$s_1^2 = 68,09$	$s_2^2 = 71,14$

a. 
$$\bar{x} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2}{n_1 + n_2} = \frac{10 \cdot 32,11 + 13 \cdot 10,75}{10 + 13} = 20,04$$
  

$$\therefore \boxed{\bar{x} = 20,04 \text{ horas}}$$

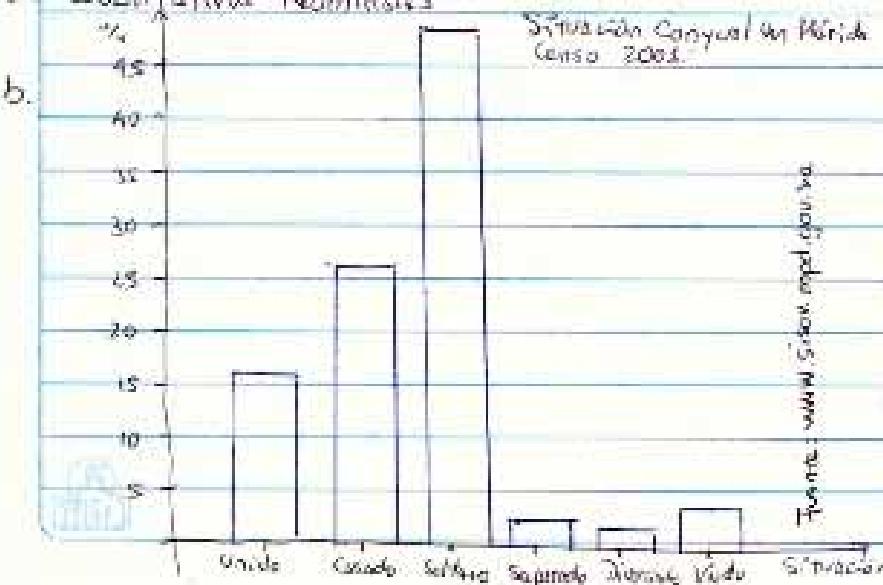
b. 
$$CV_1 = \frac{s_1}{\bar{x}_1} \times 100 = \frac{\sqrt{68,09}}{32,11} \times 100 \Rightarrow \boxed{CV_1 = 23,7\%}$$

$$CV_2 = \frac{s_2}{\bar{x}_2} \times 100 = \frac{\sqrt{71,14}}{10,75} \times 100 \Rightarrow \boxed{CV_2 = 70,46\%}$$

Así, el programa A presenta mayor homogeneidad en el tiempo de entrenamiento de los atletas, dado que  $CV_1 < CV_2$ .

2.

a. Cualitativas Nominales



c. Como los datos son cualitativa nominales sólo se puede calcular la moda:

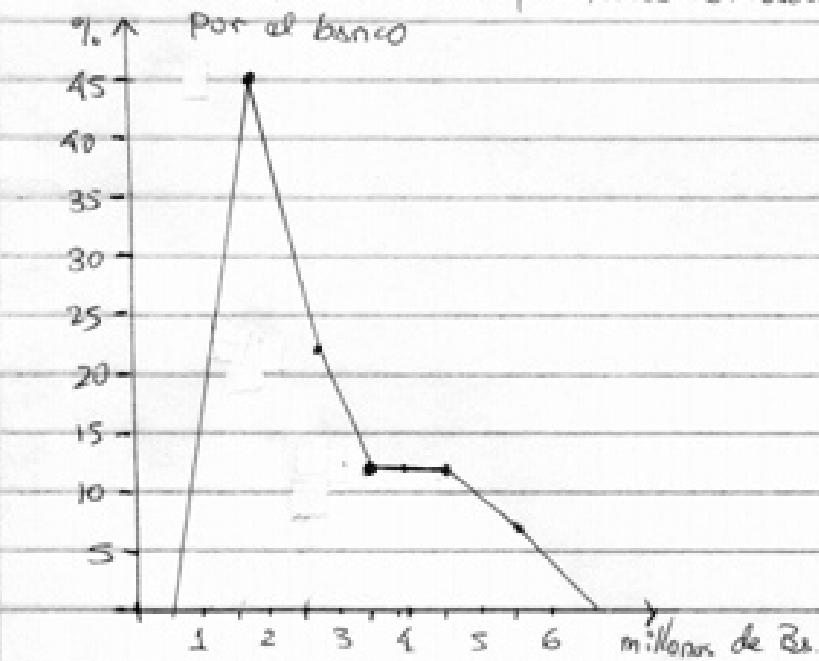
Moda: Soltero

## SOLUCIÓN EXAMEN I

3.a.  $f_3 = 5$ : Cinco de los préstamos realizados por el banco se encuentran entre 3 y 4 millones de bolívares

$F_{0,925} = 0,925$ : El 92,5% de los préstamos realizados por el banco son por montos entre 1 y 5 millones de bolívares

b. Distribución de los préstamos realizados



Fuente: Examen I Métodos Estadísticos J

c.

$m_i \cdot f_i$	$m_i^2 \cdot f_i$
27	40,5
22,5	56,25
17,5	61,25
22,5	101,25
16,5	90,75
106	350



**SOLUCIÓN EXAMEN I**

$$\text{I. } \bar{x} = \frac{\sum m_i \cdot f_i}{n} = \frac{106}{40} \Rightarrow \boxed{\bar{x} = 2,65 \text{ millones de Bs}}$$

$$\text{II. } M_d = L_I m + \left[ \frac{n - F_{ap}}{f_m} \right] \cdot C_m = 2 + \left[ \frac{20 - 18}{9} \right] \cdot 1$$

$$\Rightarrow \boxed{M_d = 2,22 \text{ millones de bolívares}}$$

$$\text{III. } S_x^2 = \sum_{i=1}^s \frac{m_i^2 f_i}{n} - \bar{x}^2 = \frac{350}{40} - 2,65^2$$

$$\Rightarrow \boxed{S_x^2 = 1,7275}$$

$$\text{Así } S_x = \sqrt{1,7275} \Rightarrow \boxed{S_x = 1,3143 \text{ millones de Bs}}$$

$$\text{IV. } Q_1 = P_{25} = L_I p + \left[ \frac{n + \left( \frac{h}{100} \right) - F_{ap}}{f_p} \right] + C_p$$

$$\Rightarrow P_{25} = 1 + \left[ \frac{40 + 25 - 0}{18} \right] + 1.$$

$$\therefore \boxed{P_{25} = 5,56 \text{ millones de bolívares}}$$

$$\text{d. } P_{95} = 5 + \left[ \frac{40 + 95 - 37}{3} \right] + 1. \Rightarrow \boxed{P_{95} = 5,33 \text{ millones de bolívares}}$$

2. Debido a que estos datos presentan poco dispersión ( $S_x = 1,3143$ ) y no existen valores atípicos, podemos concluir que la media de estos datos ( $\bar{x} = 2,65 \text{ millones de bolívares}$ ) es confiable como medida de tendencia central de la distribución de préstamos realizada por el banco.