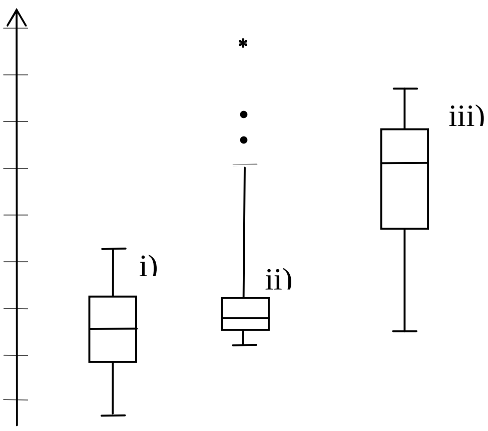
SOLUCIÓN EXAMEN I

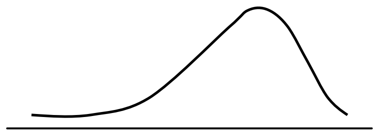
Nombre: Apellido: C.I.: Fecha: Firma:

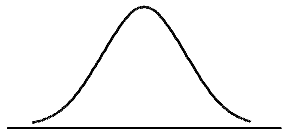
Prof. Gudberto León

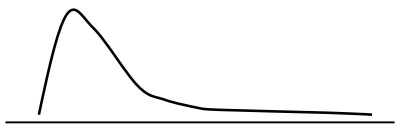
MÉTODOS ESTADÍSTICOS I – EXAMEN I

PARTE I: (***Cada respuesta correcta tiene un valor de 1 punto)***
En los siguientes gráficos se representan distintas distribuciones de frecuencias correspondientes a grupos de datos del mismo tipo (suponga que las escalas son iguales)




(a)


(b)


(c)

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

1. El DC que indica la menor dispersión en la parte central de la distribución es:
   a. DC i)
   b. DC ii)
   c. DC iii)
   d. Todos muestran la misma dispersión
   e. No se puede identificar la dispersión de la parte central
2. El DC que señala la presencia de valores atípicos es:
   a. DC i)
   b. DC ii)
   c. DC iii)
   d. Todos señalan valores atípicos
   e. No se puede identificar valores atípicos con los DC
3. Suponga que estos diagramas de caja representan las distribuciones del número de delitos diarios de tres ciudades. ¿En cuál ciudad existe menor criminalidad?
   a. i)
   b. ii)
   c. iii)
   d. Falta más información
   e. Las tres ciudades presentan aproximadamente la misma cantidad de delitos
4. Una de las siguientes proposiciones es cierta:
   a. En la curva (a) se cumple que $\bar{x}$ = Md pero distintas a la Mo.
   b. En la curva (c) se cumple que Md > $\bar{x}$ > Mo
   c. En la curva (a) se cumple que $\bar{x}$ < Md < Mo
   d. En la curva (b) se cumple que $\bar{x}$ > Md > Mo
   e. En la curva (c) se cumple que $\bar{x}$ = Md = Mo
5. Los tres polígonos de frecuencias (a), (b) y (c) representan las mismas distribuciones que los tres diagramas de caja i), ii) y iii). Escriba los **tres pares correctos**, polígono – diagrama de caja:

   ___a___-___iii_
   ___b _-_____i_
   ___c___-____ii_

SOLUCIÓN EXAMEN I

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.*

1. Una cadena de almacenes de ropa desea conocer la distribución de edad de los clientes que visitan su tienda más nueva. Con este fin, el departamento de estadística de esta cadena, selecciona una muestra aleatoria de 30 clientes. Se obtienen los siguientes resultados en cuanto a la edad de los clientes y en base a estos se construye una distribución de frecuencias: **(6 puntos)**

Datos no agrupados

43	33	18	23	19	16	51	54	18	26
25	21	17	30	28	27	27	17	32	21
35	40	39	36	48	47	38	41	50	19

Distribución de frecuencias

Clases	mi	fi	fri	Fi	Fri
[15 22)	19	9	0,3000	9	0,3000
[22 29)	26	6	0,2000	15	0,5000
[29 36)	33	4	0,1333	19	0,6333
[36 43)	40	5	0,1667	24	0,8000
[43 50)	47	4	0,1333	28	0,9333
[50 57)	54	2	0,0667	30	1
		30	1,0000		

a. Construya un diagrama de tallo y hoja
b. Comente acerca de la forma en que se distribuyen los datos
c. Calcule e **<u>interprete</u>** (en términos de este problema):
   i. La media
   ii. La desviación estándar
d. ¿Por encima de qué edad se encuentra el 75% de las edades de los clientes? Justifique su respuesta
e. ¿La media calculada en la parte c. es representativa de estos datos muestrales? Justifique su respuesta

2.

Medios de comunicación a través de los que se informan los merideños (Febrero 2006)

Medios	%
Internet	9,19
Periódico	23,01
Radio	24,40
Revista	1,62
Televisión	34,29
Trabajo	3,71
No respondió	3,78

Fuente: Encuesta realizada por el Centro de Asesorías y Proyectos Estadísticos (Ceape) de la Escuela de Estadística de la Universidad de Los Andes.

En relación al cuadro de al lado, responda lo siguiente:

a. ¿De qué tipo son los datos?
b. Grafique la información del cuadro. <u>Comente</u>.
c. Obtenga las medida(s) de posición que usted considere adecuadas para estos datos. <u>Interprétela(s)</u>.

(3 puntos)

3. 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 le suma la misma constante k la media de estos "nuevos datos" viene dada por $\bar{x}+k$. **(2 puntos)**

4. La Clínica Metropolitana de Mérida emplea 100 personas en su cuerpo de enfermeras. De ese personal, treinta son ayudantes de enfermera, veinte son enfermeras prácticas y cincuenta son enfermeras graduadas. Las primeras reciben un sueldo de Bs. 2000 por hora, las segundas ganan por hora Bs. 3500 y las últimas 5000 por hora
   a. ¿Cuál es el sueldo promedio por hora que paga la clínica a sus enfermeras?
   b. Se sabe que la desviación estándar de estos sueldos es Bs.1307,67. Si todas las enfermeras reciben un aumento del 15%:
      i. ¿Cuál es ahora el sueldo promedio que paga la clínica a sus enfermeras?
      ii. ¿Cuál es la varianza de los sueldos nuevos?
   c. ¿En donde existe mayor variabilidad, en los sueldos "viejos" o en los "nuevos" después del aumento? **(4 puntos)**

SOLUCIÓN EXAMEN I

FORMULARIO: $C_i = R/K$; $K = 1 + 3{,}3 * Log(n)$; $Q = (Q_3 - Q_1)/2$; $RQ = Q_3 - Q_1$; $\overline{AR}/\overline{AB} = \overline{RS}/\overline{BC}$;

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

$Md = LI_m + \left[\left(n/2 - Fa_m\right)/fm\right] * C_m$; $P_h = LI_p + \left\{\left[n * \left(\frac{h}{100}\right) - F_{ap}\right]/f_p\right\} * C_p$; $ASP = 3(\overline{x} - Md)/S_*$;

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

SOLUCIÓN EXAMEN I

1 a. Distribución de las estaturas de los estudiantes de Métodos Estadísticos I
Sección 02. Año 2012.

tallo	Hoja
1,5 I	2 2
1,5 S	5 6
1,6 I	0 0 2 4
1,6 S	5 5 5 5 5 5 7 8 8 8
1,7 I	0 0 0 2 3
1,7 S	5 5 6

Fuente: Encuesta realizada por la Cátedra de Métodos Estadísticos I
Mayo 2012.

b. Los datos (estaturas) presentan una asimetría negativa o por la izquierda

c.

i. En datos no agrupados

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{43,03}{26} \Rightarrow \boxed{\bar{x} = 1,66 \text{ mts}}$$

En datos agrupados:

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i \cdot f_i}{n} = \frac{43,08}{26} \Rightarrow \boxed{\bar{x} = 1,66} \text{ mts}$$

Interpretación:
La estatura promedio de los estudiantes de Métodos Estadísticos I Sección 02 es de 1,66 mts. Siendo este valor el punto de equilibrio o centro de gravedad de las 26 estaturas.

SOLUCIÓN EXAMEN I

ii. • Para datos no agrupados:

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = \frac{71,33}{26} - 1,66^2 \Rightarrow S^2 = 0,0043 \text{ mts}^2$$

Entonces, $S = \sqrt{S^2} = \sqrt{0,0043} \Rightarrow \boxed{S = 0,0654}$ metros

• Para datos agrupados:

$$S^2 = \frac{\sum_{i=1}^{k} m_i^2 \cdot f_i}{n} - \bar{x}^2 = \frac{71,50}{26} - (1,66)^2 = \boxed{0,0046 \text{ mts}^2 = S^2}$$

Entonces, $S = \sqrt{S^2} = \sqrt{0,0046 \text{mts}^2} \Rightarrow \boxed{S = 0,06782 \text{ metros}}$

Interpretación:
S = 0,06782 metros, es la dispersión de las estaturas con respecto a su media. Evidenciándose que la dispersión no es muy grande.

d. $P_{25} = Q_1 = LI_p + \left\{ \frac{\left(\frac{n \times h}{100}\right) - Fa_p}{f_p} \right\} \cdot C_p = 1,58 + \left(\frac{\frac{26 \times 25}{100} - 4}{2} \right) \cdot 0,04$

$\Rightarrow \boxed{P_{25} = 1,63 \text{ metros}}$

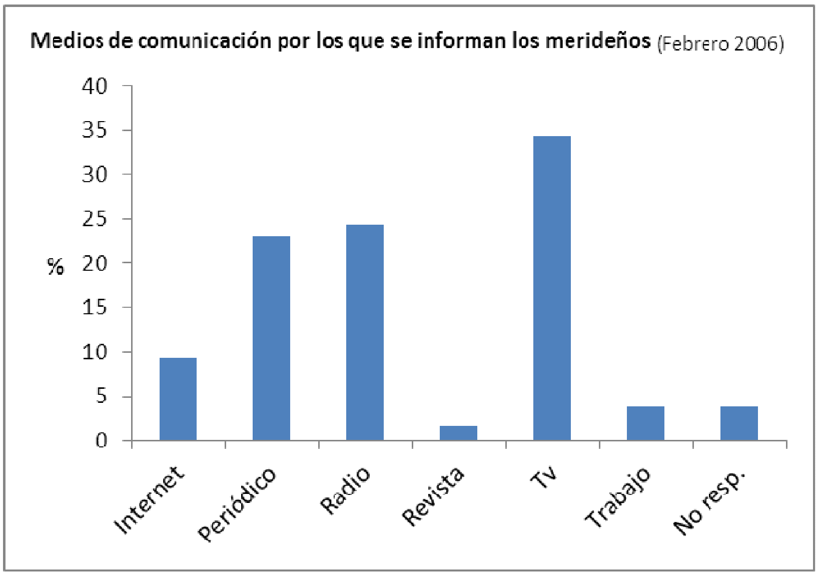
Por encima de 1,63 metros se encuentra el 75% de las estaturas de los estudiantes...

e. Dado que según se observa en el diagrama de tallo y hoja y en el valor de la desviación estándar, la dispersión de los datos no es muy grande. Y además la asimetría negativa que se observa también en el diagrama de tallo y hoja no es muy fuerte. Entonces se puede concluir que la media $\bar{X} = 1,66$ metros, es representativa de estos datos.

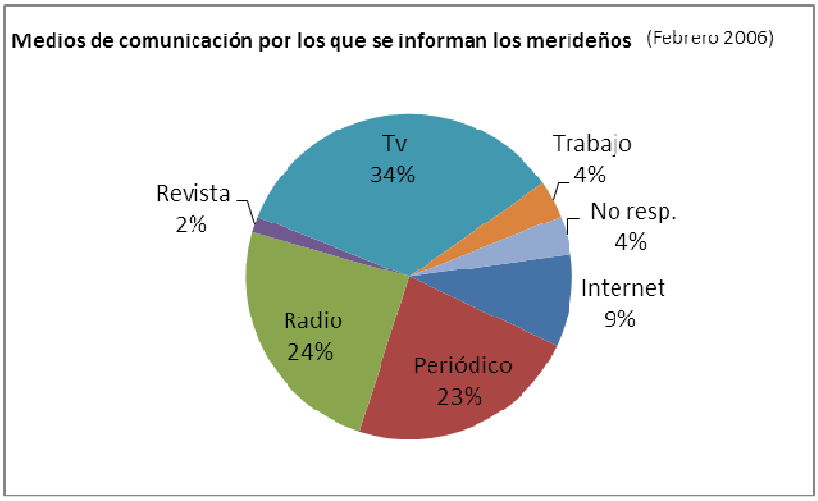
SOLUCIÓN EXAMEN I

2. a. Los datos son cualitativos nominales

b. Como los datos son cualitativos nominales se puede construir un grafico de barras o un grafico de Sectores!



Fuente: Encuesta realizada por el Centro de Asesorías y Proyectos Estadísticos (Ceape) de la Escuela de Estadística de la Universidad de Los Andes.



Fuente: Encuesta realizada por el Centro de Asesorías y Proyectos Estadísticos (Ceape) de la Escuela de Estadística de la Universidad de Los Andes.

SOLUCIÓN EXAMEN I

c. La única medida de posición que se puede calcular a datos cualitativos nominales es la Moda. Así,

$$Mo = \text{televisión}$$

Entonces se puede decir que el medio de comunicación más usado por los merideños para informarse, es la televisión

3. Se tiene $x_1, x_2, \ldots, x_n \rightarrow \bar{x}_{vieja}$

ahora cambia en que los "nuevos" datos se pueden escribir como: $x_1 + k, x_2 + k, \ldots, x_n + k$

La media "nueva" será:

$$\bar{X}_{nueva} = \frac{\sum_{i=1}^{n}(x_i + k)}{n} = \frac{\sum_{i=1}^{n} x_i}{n} + \frac{\cancel{n}k}{\cancel{n}} = \bar{x} + k.$$

$$\therefore \boxed{\bar{X}_{nueva} = \bar{X}_{vieja} + k.}$$

4. $n = 100$

$n_1 = 30$ (nro. de enfermeras ayudantes) | $\bar{x}_1 = 20$ Bs

$n_2 = 20$ (nro de " prácticas) | $\bar{x}_2 = 35$ Bs

$n_3 = 50$ (nro de " graduadas) | $\bar{x}_3 = 50$ Bs

a. $$\bar{\bar{X}} = \frac{\bar{x}_1 \cdot n_1 + \bar{x}_2 \cdot n_2 + \bar{x}_3 \cdot n_3}{n_1 + n_2 + n_3} = \frac{20 \times 30 + 35 \times 20 + 50 \times 50}{30 + 20 + 50}$$

$$\bar{\bar{X}} = \frac{38\cancel{00}}{1\cancel{00}} \Rightarrow \boxed{\bar{\bar{X}} = 38 \text{ Bs}}$$

SOLUCIÓN EXAMEN I

Así, el sueldo promedio por hora que paga la clínica a sus enfermeras es de Bs 38.

b. $S = 13{,}08$ Bs. aumento del 15 %.

i. $\bar{X}_{nueva} = 38 \times 1{,}15$

$$\boxed{\bar{X}_{nueva} = 43{,}7 \text{ Bs}}$$

ii. $S^2_{nueva} = (1{,}15)^2 \times S^2_{vieja}$

$= (1{,}15)^2 \times (13{,}08)^2$

$$\Rightarrow \boxed{S^2_{nueva} = 226{,}26 \text{ Bs}^2}$$

c. $S_{vieja} = 13{,}08$ Bs.

$\bar{X}_{vieja} = 38$ Bs

$CV_{viejo} = \dfrac{13{,}08}{38} \times 100\%$

$CV_{viejo} = 34{,}42\ \%$

$S_{nueva} = \sqrt{226{,}26 \text{ Bs}^2} = 15{,}04$ Bs

$\bar{X}_{nueva} = 43{,}7$ Bs.

$CV_{nuevo} = \dfrac{15{,}04}{43{,}7} \times 100\%$

$$\boxed{CV_{nuevo} = 34{,}42\ \%}$$

Entonces, existe la misma variabilidad en los sueldos "viejos" y los "nuevos".