Aprobado

Estadística - Contador Público Nacional

SEGUNDO EXAMEN PARCIAL

15 de octubre del 2015 - Tema 1

Ejercicio 1	Ejercicio 2	Ejercicio 3	Ejercicio 4	Ejercicio 5	Total
18	1	30	20	10	79

Ejercicio 1. (**20 puntos**) El consumo medio bimestral de energía eléctrica en una ciudad es de 59 Kwh, con una desviación típica de 6 Kwh. Se supone que se distribuye según una distribución normal. Al final del examen hay una tabla que le puede ayudar a realizar algunos cálculos.

a. ¿Cuántos Kwh tendría que consumir bimestralmente para pertenecer al 5% de la población que más consume?.Represente gráficamente en forma aproximada este resultado.

b. Si usted consume 47 Kwh ¿qué porcentaje de la población consume menos que usted?, represente gráficamente en forma aproximada este resultado.

c. Si se piensa que el consumo medio no es el que se indica anteriormente pero sí la desviación típica y la normalidad y se sabe que la proporción de consumos menores a 40Kwh es del 30%, calcule cuánto debe ser la media. 18

d. Si se sabe que la relación entre el consumo X e Y, el valor de la factura por el total del consumo, es de $\boxed{Y = 10X + 1}$, teniendo en cuenta el enunciado dado al comienzo, escriba la expresión que utilizaría para calcular con R la proporción de boletas de electricidad que pagan más 600 pesos. Justifique por qué la calcula de esta manera.

Ejercicio 2. (**10 puntos**) Pruebe que el coeficiente de correlación de las variables X, Y del Ejercicio (1) es igual a uno. Indique si ésto se debe a un resultado más general (en este caso indique cuál) o si es una casualidad.

1

Ejercicio 3. (**35 puntos**) En una caja hay 2 bolillas rojas, 3 blancas y 1 negra. Se sacan dos bolillas (sin reposición) y se observa su color. Considere las siguientes variables aleatorias (conjuntas) definidas en el espacio muestral correspondiente a este experimento aleatorio:

$$X \ : \ \text{"número de bolillas rojas seleccionadas"}$$
$$Y \ : \ \text{"número de bolillas negras seleccionadas"}$$

a. Escriba el espacio muestral ligado a este experimento aleatorio y el conjunto $(X, Y)(S)$. Compruebe, detallando el cálculo de uno cualquiera de sus valores, que la función densidad de probabilidad de este vector aleatorio está dada por,

$$f_{XY}(x, y) = \frac{2}{30} I_{\{(2,0)\}}(x, y) + \frac{4}{30} I_{\{(1,1)\}}(x, y) + \frac{6}{30} I_{\{(0,0),(0,1)\}}(x, y) + \frac{12}{30} I_{\{(1,0)\}}(x, y)$$

b. Calcule la probabilidad de que,
   i. la diferencia de ambas variables aleatorias sea igual a 0.
   ii. el número de bolillas rojas extraídas sea menor o igual que el de negras.
   iii. la suma de bolillas rojas y negras extraídas sea igual a dos.
   iv. no se extraiga ninguna bolilla negra

30

c. Son X e Y variables aleatorias independientes? Justifique la respuesta.

d. ¿Puede concluir a partir del resultado anterior si estas variables están correlacionadas? Justifique la respuesta.

Ejercicio 4. (**20 puntos**) Los miembros del departamento de imagen diagnóstica de un hospital de Buenos Aires dirigen un proyecto de mejoramiento para tratar de reducir el tiempo de procesamiento de las pruebas. El tiempo de procesamiento se define como el tiempo que pasa desde que se ordena la prueba hasta que el médico firma los resultados. Se sabe que el tiempo de procesamiento de un tipo de prueba es normal con una media de 68 hs y un desvío de 7 hs. Después de efectuar algunas mejoras se recolectó una muestra de 49 pruebas.

20

a. ¿Cuál es la probabilidad que la media muestral esté entre 66 y 69 horas?. Escriba además cómo calcularía con R esta probabilidad. Indique la variable aleatoria que considera, especifique su distribución y represente en un gráfico la probabilidad pedida (en forma aproximada).

b. Determine la probabilidad que el total de horas utilizadas sea menor a 3234 horas. Indique la variable aleatoria que considera y especifique su distribución.

c. ¿Cuál es la probabilidad que el desvío muestral sea mayor a 8 hs?. Justifique la distribución que usa. Indique el comando de R que le permitiría obtener el valor numérico.

d. ¿Cómo resolvería el inciso b si no puede suponer que la población es normal?. ¿Y el inciso c?.

Ejercicio 5. (**15 puntos**) Considere una muestra $(X_1, X_2, \ldots, X_{n_X})$ de la población representada por la variable aleatoria X y $(Y_1, Y_2, \ldots, Y_{n_Y})$ una muestra aleatoria (independiente de la anterior) de la población representada por Y. Pruebe que

a. $E(\bar{X} - \bar{Y}) = \mu_X - \mu_Y$

b. $var(\bar{X} - \bar{Y}) = \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}$ 10

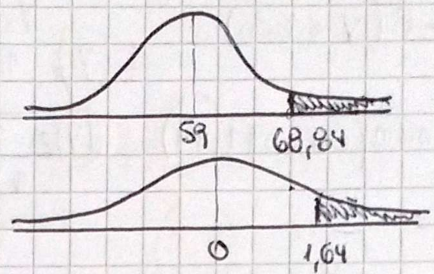
probabilidades	0.70	0.90	0.95	0.975
cuantiles z	0.52	1.28	1.64	1.96

ACE SAN RAFAEL

1) $M = 59$ $\sigma = 6$ x "Consumo medio Bimestral de energía eléctrica"

$X \sim N(59, 6^2)$

a) $X_{0,95} = \boxed{68,84}$



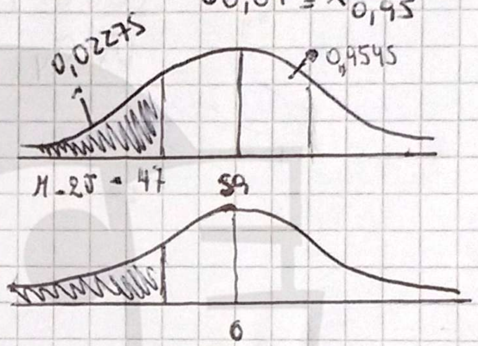
$1,64 = \frac{X_{0,95} - 59}{6}$

$1,64 \cdot 6 + 59 = X_{0,95}$

$68,84 = X_{0,95}$

5

b) $P(x < 47) = \boxed{0,02275}$



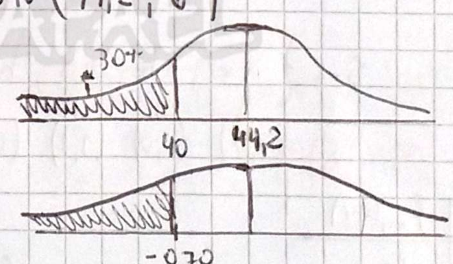
$0,9545 - 1 = 0,0455/2 = \boxed{0,02275}$

5

c) $P(M / P(x<40) = 30\%)$

$\boxed{M_x = 44,2}$ ✗

$X \sim N(44,2, 6^2)$



$X_{0,30}$

$-0,70 = \frac{x - M}{\sigma}$ ✗

$-0,70 = \frac{40 - M}{6}$

-0,52

$-0,70 \cdot 6 - 40 = -M$

$(-1) \cdot (-44,2) = M$

$\boxed{44,2 = M}$ ✗

4

(procedimiento bien)

d)
Valor Total = $\boxed{Y = 10X + 1}$ $\quad X \sim N(59, 6^2)$

X: "Consumo medio bimestral de energía"
Y: "Consumo total en Valores"

$P(Y > 600) = 1 - P(Y \leq 600)$ $\quad E(Y) = 10 . 59 + 1$ $\quad T \sim N(591, 19^2)$ X

$\downarrow$ Y

Programa R = $1 - pnorm(600, 591, 19)$ X $\quad \boxed{E(Y) = 591}$

Considero la fórmula $Y = 10X + 1$; como la formula de consumo total bimestral, y aplico sobre la Distribución Normal, con Parámetros $\mu = 59$ y $\sigma = 6$.

4

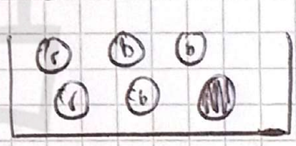
$\rho_{xy} = 1$ $\quad \rho_{xy} = \dfrac{COV(XY)}{\sigma_x . \sigma_y} = 1.114$ $\quad COV(XY) = E(XY) - E(X) . E(Y)$

$\rho_{xy} = \dfrac{10X+1}{59.10}$ $\quad E(XY) \sum \sum xy . f_{xy}$

? 1

Ejerc 3

X: "Nº Bolillas rojas seleccionadas"
Y: "Nº Bolillas Negras seleccionadas"



a) $S = \{rr, rb, rn, bb, br, bn, nr, nb\}$

$(X,Y)(S) = \{(2,0)\ (1,0)\ (1,1)\ (0,0)\ (1,0)\ (0,1)\ (1,1)\ (0,1)\}$

$(X\ Y)(S)$	$f_{xy}(xy)$
(0,0)	1/5
(0,1)	1/5
(1,0)	2/5
(1,1)	2/15
(2,0)	1/15
	$\boxed{1}$

$\boxed{\sum_{x=0}^{2} \sum_{y=0}^{1} f_{xy}(xy) = 1}$

$f_{xy}(0,0) = P\{b,b\} = P\{X=0 \wedge Y=0\}$

$f_{xy}(0,0) = \frac{3}{6} . \frac{2}{5} = 1/5$

5

$f_{xy}(0,1) = P\{bn \cup nb\} = P\{X=0 \wedge Y=1\}$

$f_{xy}(0,1)\ \frac{3}{6} . \frac{1}{5} + \frac{1}{6} . \frac{3}{5} = 1/10 + 1/10 = 1/5$

$f_{xy}(1,0) = P\{rb \cup br\} = P\{X=1 \wedge Y=0\}$

$f_{xy}(1,0) = 2/6 . 3/5 + 3/6 . 2/5 = 2/5$

$f_{xy}(2,0) = P\{rr\} = P\{X=2 \wedge Y=0\}$

$f_{xy}(2,0) = 2/6 . 1/5 = 1/15$

$f_{xy}(1,1) = P\{rn \cup nr\} = P\{X=1 \wedge Y=1\}$

$f_{xy}(1,1) = 2/6 . 1/5 + 1/6 . 2/5 = 2/15$

b)

i) $P(X-Y=0) = 1/3 \cong \boxed{0,33}$

Y \ X	0	1	2	f(y)
0	1/5	2/5	1/15	2/3
1	1/5	2/15	0	1/3
f(x)	2/5	8/15	1/15	1

ii) $P(X \leq Y) = 8/15 \cong \boxed{0,533}$

$f_{xy}(x,y)$
$(0,0) = 1/5$
$(1,1) = 2/15$
$(0,1) = 1/5$

$P(|X-Y|=0) = 1/3$

$(0,0) = 1/5$
$(1,1) = 2/15$
$1/3$

5

iii) $P(X+Y=2) = 1/5 = \boxed{0,20}$

(x,y)
$(1,1) = 2/15$
$(2,0) = 1/15$
$1/5$

5

iv) $P(Y=0) = 2/3 \cong \boxed{0,66}$

(x,y)
$(0,0) = 1/5$
$(0,1) = 2/5$
$(0,2) = 1/15$
$2/3$

5

c) $f_{xy}(x,y) \overset{?}{=} f(x) \cdot f(y)$

1

Por ej en $f_{xy}(0,0)$ $\quad f_{xy}(0,0) = f_x(0) \cdot f_y(0)$

$1/5 = 2/5 \cdot 2/3$

$1/5 \neq 4/15$

Por lo tanto las Variables X e Y NO son Independientes.

5

d) $\rho_{xy} = 0$, Al No Ser las V.A. Independientes su coeficiente de correlación es 0 (cero), y por lo tanto X e Y No estan Correlacionadas.

X

Ejer 4) X "Tiempo que pasa desde que se ordena la prueba hasta que el médico firma los rdos"

$X \sim N(68, 7^2)$ $\mu = 68hs$ $\sigma = 7.$ $n = 49$ pruebas

a) $P(66 < \bar{x} < 69)$

$\bar{x} \sim N\left(68, \frac{\sigma^2}{n}\right)$ $Var(\bar{x}) = \frac{Var(x)}{n} = \frac{49}{49} = \boxed{1}$

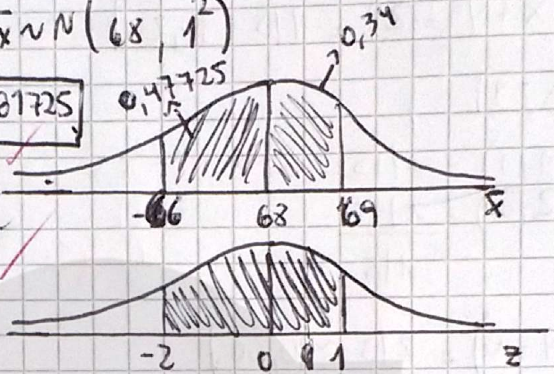
$\sigma_{\bar{x}} = 1$ $\sqrt{Var(x)} = 1 = \sigma_{\bar{x}}$

$P\left(\frac{66-68}{1} < \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} < \frac{69-68}{1}\right)$

$\bar{x} \sim N(68, 1^2)$

$P(-2 < Z < 1) = 0{,}47725 + 0{,}34 = \boxed{0{,}81725}$

pnorm(1,0,1) - pnorm(-2,0,1) = $\boxed{0{,}81725}$



$\bar{x} \sim N(68, 1^2)$

Porque al existir una muestra de tamaño mayor a 30, es decir $n > 30$, según el Teorema del Límite Central esta muestra también tiene distribución Normal

$z = \frac{\bar{x} - \mu}{\sigma}$

$1 . 1 + 68 = \bar{x}$

$69 = \bar{x}$

$-2 = \frac{\bar{x} - \mu}{\sigma}$

$-2 . 1 + 68 = \bar{x}$

$66 = \bar{x}$

b) $P(T < 3234)$

$T \sim N(n . \mu, n . \sigma^2)$ $Var(x) = 49 . 49 = 2401$

$T \sim N(3332, 49^2)$ $\sqrt{2401} = \boxed{49}$ σ

TCL

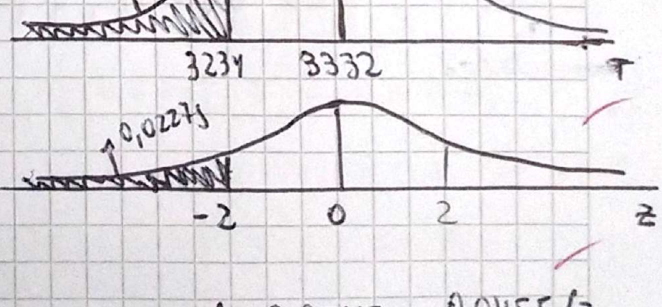
$n > 30$

$n \to \infty$

$P\left(\frac{T - \mu}{\sigma_x} < \frac{3234 - 3332}{49}\right)$

$P\left(Z < \frac{-98}{49}\right)$

$P(Z < -2) = \boxed{0{,}02275}$



• Considero la V. Aleatoria Total, y cumple con las condiciones del Teorema Central del Límite por lo tanto la va. va a tener distribución Normal.

$1 - 0{,}9545 = 0{,}0455/2$

$= \boxed{0{,}02275}$

Noguerol

PAGINA:
FECHA: / /

c) $P(S > 8)$

$$P\left(\frac{S^2(n-1)}{\sigma_x^2} > \frac{8^2 \cdot 48}{7^2}\right) = P\left(\chi^2_{48} > 62,69\right) = 1 - P\left(\chi^2_{48} < 62,69\right)$$

5

Programa R

$1 - pchisq(62,69 , 48)$

• Utilizo la Distribución Chi-cuadrado porque me permite encontrar los valores de la Varianza Muestral S^2 para una Variable Aleatoria.

d) $P(T < 3234)$ → En este caso del inciso b; si no supongo que viene de una población Normal; al tener un $n > 30$, utilizando el Teorema del Límite Central se aproximaría a la Dist. Normal.

→ El inciso c, no se podría resolver.

5

Ejerc. 5)

$$\left.\begin{array}{l}(x_1, x_2, \ldots, x_{nx}) \\ (y_1, y_2, \ldots, y_{ny})\end{array}\right\} \text{Existen Independencia entre } x \text{ e } y$$

b) $Var(\bar{x} - \bar{y}) = E(\bar{x} - \bar{y})^2 - (E(\bar{x} - \bar{y}))^2$

$= E(\bar{x}^2 - 2\bar{x}\bar{y} + \bar{y}^2) - [E(\bar{x}) - E(\bar{y})]^2$

$= E(\bar{x}^2 - 2\bar{x}\bar{y} + \bar{y}^2) - E(\bar{x})^2 - 2E(\bar{x})E(\bar{y}) + E(\bar{y})^2$

$\Rightarrow \underbrace{E(\bar{x}^2)} - 2 \cdot E(\bar{x}\bar{y}) + \overline{E(\bar{y}^2)} - \overline{E(\bar{x})^2} - 2 \cdot E(\bar{x})E(\bar{y}) + \underline{E(\bar{y})^2}$

$\Rightarrow Var(\bar{x}) + Var(\bar{y}) - \underbrace{2 \cdot cov(xy)}_{0}$ $\quad \exists$ Independencia $\Rightarrow cov(x,y) = 0$

lo que es lo mismo $\Rightarrow \boxed{\frac{\sigma_x^2}{nx} + \frac{\sigma_y^2}{ny}}$

$$E(\bar{x}-\bar{y}) = \sum_x \sum_y (\bar{x}-\bar{y}) \cdot f_{xy}(xy)$$

$f_{xy}(xy) = f(x) \cdot f(y)$ porque existe independencia

$$\Rightarrow \sum_x \sum_y (\bar{x}-\bar{y}) \cdot f_x(x) \cdot f_y(\bar{y})$$

$$\Rightarrow \sum_{\bar{x}} \sum_{\bar{y}=1}^{n} \bar{x} \; f_x(\bar{x}) - \sum_{\bar{y}} \sum_{\bar{x}=1}^{n} \bar{y}_i \cdot f_y(\bar{y})$$

lo vere al caso de v.a d

$$\Rightarrow E(\bar{x}) - E(\bar{y})$$

$$\Rightarrow \boxed{\mu_{\bar{x}} - \mu_{\bar{y}}}$$

o lo que sería lo mismo

? ?

$\mu_{\bar{x}} = \mu_x$ $\quad \mu_{\bar{y}} = \mu_y$

ACE SAN RAFAEL