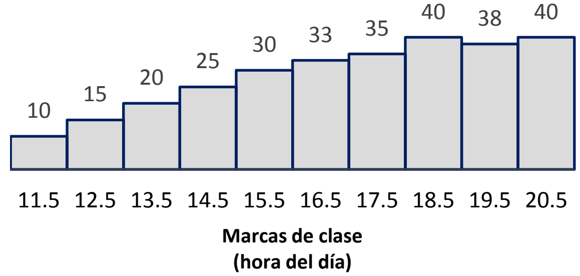
SOLUCIÓN PRIMER EXAMEN FINAL VESPERTINO "ESTADÍSTICA"

SEMESTRE 2018-1 — 6 de diciembre de 2017

1. En un boliche, se tomó una muestra aleatoria de los horarios de audiencia de los visitantes en día sábado y se elaboró el siguiente histograma de frecuencias:



Tomando en cuenta la información proporcionada por la gráfica:

a) Obtenga la hora promedio en que es visitado el boliche.
b) La hora a partir de la cual comienza a llegar el 50 % de la audiencia que visita el boliche en los horarios más tardíos.
c) Indique si los horarios de visita del boliche son insesgados y justifique su respuesta. **(25 ptos.)**

SOLUCIÓN:

La hora promedio en la que es visitado el boliche es $\bar{x} = \frac{1}{n}\sum_{i=1}^{m}(x_i)(f_i)$

n= número de datos

m= número de clases

x_i **= marca de clase**

f_i **= frecuencia de clase**

Por lo tanto :

$$\bar{x} = \frac{1}{286}[(11.5)(10) + (12.5)(15) + \cdots + (20.5)(40)] \approx 16.986$$

b) La hora a partir de la cual comienza a llegar el 50% de la audiencia que visita el boliche en los horarios más tardíos es el valor de la mediana $\tilde{x}$ de la muestra

El 50% de los datos es: $\frac{286}{2} = 143$

Dicho valor se encuentra ubicado entre la clase número 6 y la clase número 7.

La longitud de cada clase es de una hora, por lo que de acuerdo a la gráfica, a las 17 horas se tienen acumuladas 133 personas y a las 18 horas se tienen acumuladas 168 personas.

Por lo tanto el valor de $\tilde{x}$ es:

$$\tilde{x} = \frac{(18-17)}{(168-133)}(143 - 133) + 17 \approx 17.2857$$

c) En la gráfica se puede dejar ver que lo horarios de visita del boliche presentan sesgo hacia la izquierda, por lo que puede decirse que los horarios de visita del boliche son insesgados.

2. Se empacan 250 piezas de cierto componente electrónico en una caja. El peso de las piezas es una variable aleatoria con media de 125 g, y desviación estándar de 15 g. Se desea conocer la probabilidad de que el peso promedio de las piezas empacadas exceda 127 g de peso. **(15 ptos.)**

SOLUCIÓN:

Sea X el peso de las piezas, se sabe que: $\mu_x = 125 \ y \ \sigma_x = 15$

Sea $\bar{X}$ el peso promedio de las 250 piezas empacadas.

Como el número de piezas empacadas es mayor a 30, por el teorema central del límite, tenemos que: $\bar{X} \sim N\left(\mu_{\bar{x}} = 125, \ \sigma_{\bar{x}} = \frac{15}{\sqrt{250}}\right)$

La probabilidad de que el peso promedio de las piezas empacadas exceda 127 gr. de peso es:

$P(\bar{X} > 127) = P\left(Z > \frac{127-125}{\frac{15}{\sqrt{250}}}\right) = P(Z > 2.11) = 1 - F_Z(2.11) = 1 - 0.9826 =$ **0.0174**

3. Una variable aleatoria se distribuye de forma Normal. Se extraen muestras aleatorias simples de tamaño 4. Se tienen los siguientes estimadores del parámetro μ:

$$\hat{\mu}_1 = \frac{X_1 + 2X_2 + 3X_3}{6}, \qquad \hat{\mu}_2 = \frac{X_3 - 4X_2}{-3}$$

a) Comprobar si los estimadores son insesgados.
b) Indicar cuál de los dos estimadores es más eficiente. **(25 ptos.)**

SOLUCIÓN:

a) Los estimadores son insesgados si:

$E[\hat{\mu}_1] = \hat{\mu}_X$ **y** $E[\hat{\mu}_2] = \hat{\mu}_X$

Verificando:

$$E[\hat{\mu}_1] = E\left[\frac{X_1 + 2X_2 + 3X_3}{6}\right]$$

$$E[\hat{\mu}_1] = \frac{1}{6}[\mu_X + 2\mu_X + 3\mu_X]$$

$$E[\hat{\mu}_1] = \frac{1}{6}[6\mu_X]$$

$$\therefore E[\hat{\mu}_1] = \mu_X$$

$$E[\hat{\mu}_2] = E\left[\frac{X_3 - 4X_2}{-3}\right]$$

$$E[\hat{\mu}_2] = -\frac{1}{3}[\mu_X - 4\mu_X]$$

$$\therefore E[\hat{\mu}_2] = \mu_X$$

Por lo que: tanto $\hat{\mu}_1$, como $\hat{\mu}_2$ son estimadores insesgados.

b) El estimador más eficiente es aquel que tenga la menor variancia.
Se obtiene:

$$var[\hat{\mu}_1] = var\left[\frac{X_1+2X_2+3X_3}{6}\right]$$

$$var[\hat{\mu}_1] = \frac{1}{36}var[X_1 + 2X_2 + 3X_3]$$

$$var[\hat{\mu}_1] = \frac{1}{36}[\sigma_x^2 + 4\sigma_x^2 + 9\sigma_x^2]$$

$$var[\hat{\mu}_1] = \frac{14}{36}[\sigma_x^2] = \frac{7}{18}[\sigma_x^2] \approx 0.388\sigma_x^2$$

$$var[\hat{\mu}_2] = var\left[\frac{X_3-4X_2}{-3}\right]$$

$$var[\hat{\mu}_2] = \frac{1}{9}var[X_3 - 4X_2]$$

$$var[\hat{\mu}_2] = \frac{1}{9}[\sigma_x^2 + 16\sigma_x^2] = \frac{17}{9}[\sigma_x^2] \approx 1.88\sigma_x^2$$

$\therefore$ El estimador que es más eficiente es $\hat{\mu}_1$.

4. Una empresa fabrica resistencias de potencia cuya duración se distribuye de forma aproximadamente normal, con duración media igual a 800 horas y una desviación estándar de 40 horas. Si una muestra aleatoria de 40 resistencias tuvo una duración promedio de 788 horas, utilice un nivel de significancia del 0.04 y determine si se cumple que las resistencias no duran 800 horas en funcionamiento. **(20 ptos.)**

SOLUCIÓN:
Sea X la duración de las resistencias se sabe que:
X~ $N\,(\mu_x = 800,\ \sigma_x = 40)$
Se extrae una muestra de tamaño n=40, de la cual: $\bar{x} = 788$
Se busca probar la hipótesis H_0: μ_x =800
H_a: $\mu_x \neq 800$
Con un nivel de significancia de 0.04
Por provenir la muestra de una población normal y conocer el valor de σ_x el estadístico de prueba es:

$$Z_0 = \frac{\bar{x}-\mu_{\bar{x}}}{\frac{\sigma_x}{\sqrt{n}}} = \frac{788-800}{\frac{40}{\sqrt{40}}} \approx -1.897$$

El valor crítico $z_{\frac{\alpha}{2}}$ es: $Z_{0.02}$=2.06

Por lo que la región de NO rechazo de H_0 es $-2.06 \leq Z \leq 2.06$
Como el estadístico de prueba cae en la zona de NO rechazo; H_0, no se rechaza y se concluye que con un nivel de significancia del 4% es razonable pensar que la duración promedio de las resistencias es igual a 800 horas.

5. Los siguientes datos corresponden a las lecturas promedio de temperatura en grados centígrados y nivel de lluvia en milímetros, observadas en nuestro país, durante seis de los doce meses del año 2015.

Lluvia (mm)	37.5	27.2	69.6	27.4	53.7	119.2
Temperatura (°C)	16.2	17.8	19.4	23.3	23.9	25.7

Con base en los datos anteriores, realice una estimación de la cantidad de lluvia esperada para una temperatura de 20°. **(15 ptos.)**

SOLUCIÓN:
La estimación de la cantidad de lluvia esperada para una temperatura de 20° C, se obtiene mediante la obtención del modelo de regresión lineal construido a partir de los datos observados.

Donde:
X = temperatura
Y = cantidad de lluvia
El modelo que permite estimar la cantidad de lluvia a partir de la temperatura es: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

Empleando el método de mínimos cuadrados se tiene:

$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ y $\hat{\beta}_1 = \frac{SS\,xy}{SS\,xx}$

$$\boldsymbol{SSxy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n} = 7427.19 - \frac{(126.3)(334.6)}{6} = 383.86$$

$$\boldsymbol{SSxx} = \sum(x_i{}^2) - \frac{(\sum x_i)^2}{n} = 2730.23 - \frac{(126.3)^2}{6} = 71.615$$

$$\therefore \hat{\beta}_1 = \frac{383.86}{71.615} = 5.36$$

$$\hat{\beta}_0 = 55.7666 - (5.36)(21.05) = -57.06$$

Por lo que el modelo de regresión lineal obtenido es:

$$\hat{y} = 3.2491 + 2.4948x$$

Con dicho modelo se obtiene la estimación de lluvia para una temperatura de 20°C

Dicha estimación es:

$$\hat{y} = -57.06 + 5.36(20) = 50.14$$

Por lo tanto, la cantidad de lluvia esperada para una temperatura de 20°C es:
50.14 mm