

EXAMEN DE ESTADÍSTICA (DESCRIPTIVA Y REGRESIÓN)

1º Farmacia

Modelo A

14 de octubre de 2019

Duración: 1 hora y 15 minutos.

- (4,5 pts.) 1. Se ha medido la tensión arterial sistólica (en mmHg) en dos grupos de 100 personas cada uno de dos poblaciones A y B , obteniendo los siguientes resultados:

Tensión sistólica	Num personas A	Num personas B
(80, 90]	4	6
(90, 100]	10	18
(100, 110]	28	30
(110, 120]	24	26
(120, 130]	16	10
(130, 140]	10	7
(140, 150]	6	2
(150, 160]	2	1

Se pide:

- ¿Cuál de las dos distribuciones de la tensión sistólica es menos asimétrica? ¿Cuál es más apuntada? ¿Pueden provenir estas muestras de poblaciones normales?
- ¿En cuál de los dos grupos es más representativa la media de la tensión sistólica?
- ¿Por encima de qué tensión sistólica se encuentra el 30 % de las personas del grupo de la población A ?
- ¿Qué tensión sistólica es relativamente más alta, 132 mmHg en el grupo de la población A , o 130 mmHg en el grupo de la población B ?
- Si a las 100 personas de la población A se les mide la tensión sistólica con otro tensiómetro, y la tensión obtenida (Y) está relacionada con la del primer tensiómetro (X) mediante la ecuación $y = 0,98x - 1,4$, ¿en cuál de las dos tensiones X o Y es más representativa la media? Justificar la respuesta.

Usar las siguientes sumas para los cálculos.

Grupo A : $\sum x_i n_i = 11520$ mmHg, $\sum x_i^2 n_i = 1351700$ mmHg², $\sum (x_i - \bar{x})^3 n_i = 155241,6$ mmHg³ y $\sum (x_i - \bar{x})^4 n_i = 16729903,52$ mmHg⁴.

Grupo B : $\sum x_i n_i = 11000$ mmHg, $\sum x_i^2 n_i = 1230300$ mmHg², $\sum (x_i - \bar{x})^3 n_i = 165000$ mmHg³ y $\sum (x_i - \bar{x})^4 n_i = 13632500$ mmHg⁴.

Solución

- Sean x e y las tensiones sistólicas de las poblaciones A y B respectivamente.
 $g_{1x} = 0,4024$ y $g_{1y} = 0,5705$, de manera que la distribución del grupo de la población A es más asimétrica ya que el coeficiente de g_1 está más lejos de 0.
- $g_{2x} = -0,2346$ y $g_{2y} = 0,3081$, de manera que la distribución del grupo de la población A es más apuntada que la de la población B ya que $g_{2x} > g_{2y}$.
- $\bar{x} = 115,2$ mmHg, $s_x^2 = 245,96$ mmHg², $s_x = 15,6831$ mmHg and $cv_x = 0,1361$.
 $\bar{y} = 110$ mmHg, $s_y^2 = 203$ mmHg², $s_y = 14,2478$ mmHg y $cv_y = 0,1295$.
 La media es más representativa en el grupo de la población A ya que el coeficiente de variación es menor.

- d) $g_{1x} = 0,4024$ y $g_{1y} = 0,5705$, de manera que la distribución de edades de los pacientes menores de 65 es menos simétrica.
- e) Las puntuaciones típicas son $z_x(132) = 1,0712$ y $z_y(130) = 1,4037$, de manera que 132 mmHg es relativamente menor en el grupo de la población A ya que su puntuación típica es menor.

-
- (1 pts.) 2. Se sabe que en una distribución de frecuencias simétrica la media vale 15, el primer cuartil vale 12 y el máximo valor es 25. Se pide:
- a) Dibujar el diagrama de caja y bigotes.
- b) ¿Podría considerarse como atípico en esta distribución un valor de 2?

Solución

- a) Diagrama de cajas
- b) Si porque la valla inferior es $v_1 = 3$.

-
- (4,5 pts.) 3. Un laboratorio está ensayando tres analgésicos diferentes y se quiere determinar si existe relación entre los tiempos que tardan en hacer efecto en un paciente. Para ello, realizan una prueba administrando los diferentes analgésicos a un grupo de 20 pacientes evaluando el tiempo que tarda en hacer efecto cada analgésico, denominándolos X , Y y Z a cada uno de los tiempos, obteniéndose los siguientes datos:

$$\begin{aligned}\sum x_i &= 668 \text{ min}, \sum y_i = 855 \text{ min}, \sum z_i = 1466 \text{ min}, \\ \sum x_i^2 &= 25056 \text{ min}^2, \sum y_i^2 = 42161 \text{ min}^2, \sum z_i^2 = 123904 \text{ min}^2, \\ \sum x_i y_i &= 31522 \text{ min}^2, \sum y_i z_i = 54895 \text{ min}^2.\end{aligned}$$

Se pide:

- a) ¿Existe relación lineal entre los tiempos X e Y ? ¿ Y entre los tiempos Y y Z ?
- b) Según un modelo lineal, ¿cuánto aumentaría el tiempo X por cada minuto que aumente el tiempo Y ?
- c) Si deseamos realizar una predicción del tiempo Y mediante un modelo lineal, ¿Cuál de los dos tiempos X o Z sería el más adecuado? ¿Por qué?
- d) Predecir según el mejor modelo lineal elegido en el apartado anterior el valor del tiempo Y para un valor del tiempo X o Z , según el modelo elegido, de 40 minutos.
- e) Si el coeficiente de correlación lineal entre los tiempos X y Z es $r = -0,69$, ¿Cuál es el modelo lineal que explica el tiempo X en función del tiempo Z ?

Solución
