

Bioestadística: Variables Aleatorias. Distribuciones de Probabilidad II

M. González

Departamento de Matemáticas. Universidad de Extremadura

Media Muestral y Varianza Muestral

3. El periodo de incubación de una determinada enfermedad se distribuye normalmente con un tiempo medio de 800 horas y una desviación típica de 60 horas. Calcula las siguientes probabilidades en una muestra de 16 pacientes contagiados:

(a) Que muestren una incubación media entre 790 y 810 horas.

$X = \text{tiempo de incubación} \sim N(800, 60^2)$

$n = 16$ $X_1, \dots, X_{16}$ v.a. i.i.d. X

$$\bar{X} = \frac{1}{16} \sum_{i=1}^{16} X_i \sim N(800, 15^2) \quad E[\bar{X}] = 800 \quad \text{Var}[\bar{X}] = \frac{60^2}{16}$$

$$P(790 < \bar{X} < 810) = P\left(\frac{790 - 800}{15} < \frac{\bar{X} - 800}{15} < \frac{810 - 800}{15}\right)$$

$$Z \sim N(0, 1) = P(-0.67 < Z < 0.67) = 0.4972$$

3. El periodo de incubación de una determinada enfermedad se distribuye normalmente con un tiempo medio de 800 horas y una desviación típica de 60 horas. Calcula las siguientes probabilidades en una muestra de 16 pacientes contagiados:

(d) Que ningún paciente muestre síntomas de la enfermedad antes de 830 horas.

$$X = \text{tiempo de incubación} \sim N(800, 60^2)$$

$$n = 16 \quad X_1, \dots, X_{16} \text{ v.a. i.i.d.} \quad X$$

$$P(X_1 > 830, \dots, X_{16} > 830) = P(X > 830)^{16} \simeq 0$$

Media Muestral y Varianza Muestral

3. El periodo de incubación de una determinada enfermedad se distribuye normalmente con un tiempo medio de 800 horas y una desviación típica de 60 horas. Calcula las siguientes probabilidades en una muestra de 16 pacientes contagiados:

(f) Que la desviación típica muestral en la duración de las incubaciones esté entre 50 y 65 horas.

Nota: $P(\chi_{15}^2 \leq 10.42) = 0.2075$, $P(\chi_{15}^2 \leq 15) = 0.5486$, $P(\chi_{15}^2 \leq 17.6) = 0.7157$.

$X = \text{tiempo de incubación} \sim N(800, 60^2)$

$$n = 16$$

$$X_1, \dots, X_{16} \text{ v.a. i.i.d.}$$

$$X$$

$$S^2 = \frac{1}{16-1} \sum_{i=1}^{16} (X_i - \bar{X})^2 \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \quad \frac{15 * S^2}{60^2} \sim \chi^2(15)$$

$$P(50 < S < 65) = P(50^2 < S^2 < 65^2) = P\left(\frac{15 * 50^2}{60^2} < \frac{15 * S^2}{60^2} < \frac{15 * 65^2}{60^2}\right)$$

$$\chi_{15}^2 = \frac{15 * S^2}{60^2} = P(10.42 < \chi_{15}^2 < 17.6) = 0.7157 - 0.2075 = 0.5082$$

Media Muestral y Varianza Muestral

7. Las edades en que se produce la muerte para una población concreta se distribuyen según una normal con media de 50 años y una varianza de 36 años².

(a) ¿Cuáles son las medias y las desviaciones típicas de las distribuciones muestrales de varianzas y medias con 19 individuos por muestra?

$$n \quad X_1, \dots, X_n \text{ v.a. i.i.d.} \quad X \sim N(\mu, \sigma^2)$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad E[\bar{X}] = \mu_{\bar{X}} = \mu \quad \text{Var}[\bar{X}] = \sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad E[S^2] = \mu_{S^2} = \sigma^2 \quad \text{Var}[S^2] = \sigma_{S^2}^2 = \frac{2\sigma^4}{n-1}$$

$$X = \text{edad de muerte} \sim N(50, 36) \quad n = 19$$

$$\mu_{\bar{X}} = 50 \quad \sigma_{\bar{X}} = \sqrt{\frac{36}{19}} = 1.38$$

$$\mu_{S^2} = 36 \quad \sigma_{S^2} = \sqrt{\frac{2 \cdot 36^2}{18}} = 12$$

Media Muestral y Varianza Muestral

7. Las edades en que se produce la muerte para una población concreta se distribuyen según una normal con media de 50 años y una varianza de 36 años².

(b) En una muestra elegida al azar, ¿qué valores de las medias y de las varianzas tienen una probabilidad del 1% de que los demás valores de la distribución sean mayores que ellos?

$$X = \text{edad de muerte} \sim N(50, 36) \quad n = 19 \quad \bar{X} \sim N(50, 36/19)$$

$$P(\bar{X} > a) = 0.01 \Leftrightarrow P\left(Z > \frac{a - 50}{\sqrt{\frac{36}{19}}}\right) = 0.01$$

$$\alpha/2 = 0.01 \Rightarrow \alpha = 0.02$$

$$\frac{a - 50}{\sqrt{\frac{36}{19}}} = z_{0.02} = 2.326 \Rightarrow a = 54.41$$

Media Muestral y Varianza Muestral

7. Las edades en que se produce la muerte para una población concreta se distribuyen según una normal con media de 50 años y una varianza de 36 años².

(b) En una muestra elegida al azar, ¿qué valores de las medias y de las varianzas tienen una probabilidad del 1% de que los demás valores de la distribución sean mayores que ellos?

$$X = \text{edad de muerte} \sim N(50, 36) \quad n = 19 \quad \frac{18 * S^2}{36} \sim \chi^2(18)$$

$$P(S^2 > b) = 0.01 \Leftrightarrow P\left(\frac{18 * S^2}{36} > \frac{18 * b}{36}\right) = 0.01$$

$$\text{TABLA V: } n, \alpha \rightarrow \chi^2_{\alpha}(n) \quad Y \sim \chi^2(n), \quad P(Y > \chi^2_{\alpha}(n)) = \alpha$$

$$\frac{18 * b}{36} = \chi^2_{0.01}(18) = 34.831 \Rightarrow b = 69.66$$

Aproximación de la Binomial por la Normal

$$X \sim B(n, p) \quad np(1-p) \geq 5 \quad X \approx Y \sim N(np, np(1-p))$$

$$P(X = k) = P(k - 0.5 \leq Y \leq k + 0.5), \quad k = 1, \dots, n - 1$$

$$P(X = 0) = P(Y \leq 0.5)$$

$$P(X = n) = P(Y \geq n - 0.5)$$

Aproximación de la Binomial por la Normal

12. Supongamos que en individuos con presión sanguínea alta, es igualmente probable que después de un cierto periodo de tiempo, la presión le haya bajado o no ligeramente. Por otro lado se ha comprobado que en individuos con presión sanguínea alta, que se encuentren bajo el efecto de un cierto medicamento H, la presión disminuye en el 80% de los casos.

Consideremos una muestra de 200 individuos con presión sanguínea alta:

(a) Si suponemos que no están afectados por ningún medicamento, calcula la probabilidad de que le baje la presión a más de 90 individuos.

X = número de individuos a los que les baja la presión entre 200 tomados al azar $\sim B(200, 0.5)$

$$np(1 - p) = 200 * 0.5 * 0.5 = 50 > 5 \quad X \approx Y \sim N(100, 50)$$

$$\begin{aligned} P(X > 90) &= P(X \geq 91) \simeq P(Y \geq 90.5) \\ &= P\left(Z \geq \frac{90.5 - 100}{\sqrt{50}}\right) = P(Z \geq -1.34) \\ &= P(Z \leq 1.34) = 0.9099 \end{aligned}$$

Aproximación de la Binomial por la Normal

12. Por otro lado se ha comprobado que en individuos con presión sanguínea alta, que se encuentren bajo el efecto de un cierto medicamento H, la presión disminuye en el 80% de los casos. Consideremos una muestra de 200 individuos con presión sanguínea alta:

(b) Si la muestra se encuentra bajo el efecto de H, calcula la probabilidad de que baje la presión en más de 172 casos ó en menos de 148.

X = número de individuos bajo los efectos de H a los que les baja la presión entre 200 tomados al azar $\sim B(200, 0.8)$

$$np(1 - p) = 200 * 0.8 * 0.2 = 32 > 5 \quad X \approx Y \sim N(160, 32)$$

$$\begin{aligned} P(\{X < 148\} \cup \{X > 172\}) &= P(X \leq 147) + P(X \geq 173) \\ &\simeq P(Y \leq 147.5) + P(Y \geq 172.5) \\ &= P(Z \leq \frac{147.5 - 160}{\sqrt{32}}) + P(Z \geq \frac{172.5 - 160}{\sqrt{32}}) \\ &= P(Z \leq -2.21) + P(Z \geq 2.21) \\ &= 2 * (1 - P(Z \leq 2.21)) = 2 * (1 - 0.9864) = 0.0272 \end{aligned}$$