Estadística

Curso 2016-2017/1

Fecha: 9 de Enero de 2017

Examen Parcial 2

1. `[2.5 puntos]` Un proceso industrial está formado por n pasos sucesivos independientes. La duración de cada paso X_i, con $i = 1, 2, \ldots, n$, varía de manera aleatoria según una distribución de media 10 minutos y desviación típica de 1 minuto. Sea $X = X_1 + X_2 + \cdots + X_n$ la duración total (en minutos) del proceso.

(a) `[1.25 puntos]` Si el número de pasos es $n = 100$, calcula la probabilidad que el proceso total tenga una duración superior a 1025 minutos.

(b) `[1.25 puntos]` Si sabemos que la duración total del proceso cumple $P(X \leq 655.68) = 0.975$, ¿cuál es, en este caso, el número de pasos del proceso?

Solución

La durada total del procés és $X = \sum_{i=0}^{n} X_i$, on X_i=[durada del pas i-èsim], amb $\mu_i = 10$ i $\sigma_i = 1, i = 1, \cdots, n$. Si $n > 30$, segons el Teorema Central del Límit, la durada total X té una distribució aproximadament normal amb $\mu_x = 10n$ i $\sigma_x = \sqrt{n}$.

(a) Si $n = 100$, la durada total X té distribució aproximadament normal amb paràmetres $\mu_x = 1000$ i $\sigma_x = 10$.

$$P(X > 1025) = P\left(\frac{X - \mu_x}{\sigma_x} > \frac{1025 - 1000}{10}\right) = P(Z > 2.5) = 1 - P(Z \leq 2.5)$$
$$= 1 - \Phi(2.5) = 1 - 0.99379 = 0.00621$$

(b) Suposant que $n > 30$, es compleix

$$P(X \leq 655.68) = 0.975 \quad \Rightarrow \quad P\left(\frac{X - \mu_x}{\sigma_x} \leq \frac{655.68 - 10n}{\sqrt{n}}\right) = 0.975$$

obtenim

$$P(Z \leq z_0) = 0.975, \text{ amb } z_0 = \frac{655.68 - 10n}{\sqrt{n}} \quad \Rightarrow \quad \frac{655.68 - 10n}{\sqrt{n}} = 1.96$$

resolent $10n + 1.96\sqrt{n} - 655.68 = 0$, obtenim $\sqrt{n} = 8$ i, per tant, $n = 64$.

2. `[2.5 puntos]` El tiempo de vida útil (T) de un dispositivo es una variable aleatoria continua con la siguiente función de densidad:

$$f(t) = \begin{cases} \lambda\, e^{-\lambda(t-\tau)}, & t \geq \tau \\ 0, & \text{otros} \end{cases}$$

donde $\lambda > 0$ y τ son parámetros. Sabemos que su valor esperado es

$$E(T) = \tau + \frac{1}{\lambda}$$

Sea $T_1, T_2, \ldots, T_n$ una muestra aleatoria simple de T:

(a) `[1.5 puntos]` Supón τ conocida y determina el estimador ($\hat{\lambda}_V$) de máxima verosimilitud del parámetro λ.

(b) `[0.5 puntos]` Supón τ conocida, determina el estimador ($\hat{\lambda}_M$) de λ por el método de los momentos.

(c) `[0.5 puntos]` Supón λ conocida, determina el estimador ($\hat{\tau}_M$) de τ por el método de los momentos.

Solución

(a) Conegut τ, considerem la funció de versemblança

$$L(t_1, \cdots, t_n, \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda(t_i-\tau)}$$

El seu logaritme és

$$\ln(L) = \sum_{i=1}^{n} \ln(\lambda) + \sum_{i=1}^{n} \ln(e^{-\lambda(t_i-\tau)}) = n\ln(\lambda) - \lambda \sum_{i=1}^{n} t_i + n\,\lambda\,\tau$$

Igualant la derivada a zero obtindrem el valor del paràmetre λ que maximitza la funció de versemblança.

$$\frac{d}{d\lambda}(\ln(L)) = \frac{n}{\lambda} + n\,\tau - \sum_{i=1}^{n} t_i = 0 \quad \Rightarrow \lambda = \frac{n}{\sum_{i=1}^{n} t_i - n\,\tau} = \frac{1}{\overline{T} - \tau}$$

on $\overline{T}$ és la mitja mostral. Així doncs, l'estimador de màxima versemblança és

$$\hat{\lambda}_V = \frac{1}{\overline{T} - \tau}$$

(b) Per estimar un paràmetre pel mètode dels moments, s'ha d'igualar el valor esperat de la variable a la mitja mostral:

$$E(T) = \overline{T} \quad \Rightarrow \quad \tau + \frac{1}{\lambda} = \overline{T}$$

Suposant τ conegut, obtenim l'estimador de λ:

$$\hat{\lambda}_M = \frac{1}{\overline{T} - \tau}$$

Observem que coincideix amb el de màxima versemblança.

(c) Suposant λ conegut, procedint com a l'apartat anterior, obtenim l'estimador de τ: $\hat{\tau}_M = \overline{T} - \frac{1}{\lambda}$

3. `[2.5 puntos]` Para calibrar una balanza de precisión, se pesa repetidas veces un *peso patrón* que tiene un valor real conocido de 10 gramos. Sabemos que los valores obtenidos en la medición siguen una distribución normal con media μ desconocida. Por estudios anteriores, también sabemos que la desviación típica σ se mantiene constante y es igual a 0.0002 gramos.

(a) `[1 punto]` Se pesa el patrón cinco veces, y la media aritmética de las cinco pesadas es igual a 10.0023 gramos. Determina el intervalo que permite estimar μ con un nivel de confianza del 98%.

(b) `[1 punto]` ¿Cuantas pesadas tendríamos que realizar para conseguir un intervalo de confianza que nos proporcionase una estimación de μ con un margen de error ± 0.0001 y un nivel de confianza del 98%?

(c) `[0.5 puntos]` ¿Cuáles serían los valores de μ y σ si el instrumento de medida fuera perfecto?

Solución

(a) L'interval de confiança al 98% per a la mitjana poblacional amb variància coneguda $\sigma^2 = 0.0002^2$ correspon a:

$$I_{0.98}(\mu) = (\overline{x} - \frac{\sigma}{\sqrt{n}} z_{0.99}, \overline{x} + \frac{\sigma}{\sqrt{n}} z_{0.99}) = (10.0023 - \frac{0.0002}{\sqrt{5}} 2.325, 10.0023 + \frac{0.0002}{\sqrt{5}} 2.325) =$$

$$= (10.00209205, 10.00250795).$$

(b) El marge d'error al considerar la mitjana mostral com aproximació de la mitjana poblacional és $\pm \frac{\sigma}{\sqrt{n}} z_{0.99}$, per tant hem de resoldre $\frac{\sigma}{\sqrt{n}} z_{0.99} \leq 0.0001$, d'on s'obté:

$$n \geq \left(\frac{\sigma \cdot z_{0.99}}{0.0001}\right)^2 = \left(\frac{0.0002 \cdot 2.325}{0.0001}\right)^2 = 21.6225.$$

Hem de realitzar, llavors, un mínim de 22 pesades.

(c) Si l'instrument de mesura fos perfecte, a cada pesada del patró retornaria el seu valor real, 10 grams i, per tant, aquest seria el valor de la mitjana poblacional, $\mu = 10$. I la seva variància seria nula $\sigma^2 = 0$.

4. `[2.5 puntos]` En un centro médico, se realizan tratamientos para superar la adicción al tabaco. Según los datos científicos disponibles, con el tipo de terapia aplicado, la tasa de recaídas debiera ser inferior al 35%. El director del centro médico sospecha que la tasa de recaídas real es superior a la indicada. Para verificar su sospecha, selecciona una muestra aleatoria de 150 individuos y observa que se han producido un total de 61 casos de recaída. Realiza un contraste de hipótesis, con nivel de significación $\alpha = 0.05$, para decidir si los datos muestrales obtenidos confirman la sospecha del director del centro médico. Estructura la respuesta en los dos apartados siguientes:

(a) `[1.5 puntos]` Hipótesis del test, estadístico de contraste, criterio de decisión.

(b) `[1 punto]` Resultado del test y conclusiones.

Solución

(a). Test de proporciones. Muestra grande $N = 150$, con $N\hat{p} = 61 > 5$ y $N(1 - \hat{p}) = 89 > 5$. Podemos aplicar una aproximación normal

Test unilateral superior. Hipótesis del test

$$\begin{cases} H_0 : \pi = 0.35 \\ H_1 : \pi > 0.35 \end{cases}$$

Estadístico de contraste

$$Z = \frac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{N}}} = \frac{\hat{p} - 0.35}{0.038944}$$

Si la hipótesis nula es cierta, el estadístico Z tiene una distribución aproximada $N(0, 1)$

Criterio de decisión (basado en z_{crit})

Valor crítico del estadístico $Pr(Z > z_{\text{crit}}) = 0.05 \Longrightarrow z_{\text{crit}} = F_z^{-1}(0.95) \approx 1.645$

$$\begin{cases} \text{si } z_{\text{obs}} \leq 1.645, \text{ mantenemos } H_0 : \pi = 0.35 \\ \text{si } z_{\text{obs}} > 1.645, \text{ aceptamos } H_1 : \pi > 0.35 \end{cases}$$

$$z_{\text{obs}} = \frac{\hat{p}_{\text{obs}} - 0.35}{0.038944}$$

Criterio de decisión (basado en el p-value)

$p\text{-value} = Pr(Z > z_{\text{obs}})$

$$\begin{cases} \text{si } p\text{-value} \geq 0.05, \text{ mantenemos } H_0 : \pi = 0.35 \\ \text{si } p\text{-value} < 0.05, \text{ aceptamos } H_1 : \pi < 0.35 \end{cases}$$

(b) Resultado del test. Tenemos

$\hat{p}_{\text{obs}} = 61/150 = 0.40667 \Longrightarrow z_{\text{obs}} = (0.40667 - 0.35)/0.038944 = 1.45507$

$p\text{-value} = Pr(Z > 1.45507) > F_z(1.46) = 1 - F_z(1.46) = 1 - 0.92785 = 0.072150$

Mantenemos la hipótesis nula. Los datos obtenidos no proporcionan una evidencia suficientemente sólida para afirmar que la tasa de recaídas es superior al 35%.