Estadística

Estudia las técnicas para recolectar, resumir y ordenar datos, y inferir resultados a partir de los datos recolectados.

Ejemplo: supongamos que interesa estudiar la altura de los estudiantes de la facultad. Se toman al azar n estudiantes y a cada uno se le mide la altura. Tenemos n datos tal que anotamos $x_1, x_2, \dots, x_n$.

Población → conjunto de todos los objetos de interés. Los estudiantes de la facultad

Muestra → una parte de la población. La muestra se toma al azar. (se eligen de la población n alumnos al azar como si se eligieran bolillas de una urna con N bolillas). Se hace un muestreo aleatorio simple.

La característica que importa estudiar es la altura de los alumnos. Se define X: "altura de un estudiante tomado al azar".

En estadística la distribución de X es desconocida. Los únicos datos disponibles son las mediciones $x_1, x_2, \dots, x_n$.

Estadística descriptiva: es la parte de estadística que estudia la muestra, esto es, estudia y ordena la muestra.

Estadística inferencial: estudia cómo inferir resultados a partir de la muestra sobre la población.

En el ejemplo nos puede interesar $\mu = E(X)$ que es desconocida. A partir de $x_1, x_2, \dots, x_n$ se trata de "estimar" un valor de μ. O sea de alguna forma obtener algún valor aproximado de μ

Entonces al tener una muestra, no obtendremos el verdadero valor de μ pues tenemos solamente n valores de la población.

Tenemos las mediciones $x_1, x_2, \dots x_n$. Podemos asumir que x_i es el resultado de medir la va X_i: "altura del alumno i" $i = 1, 2, \dots, n$

Si las mediciones se hacen bajo las mismas condiciones, podemos asumir que $X_1, X_2, \dots, X_n$ tienen la misma distribución y que es igual a la distribución de la va X original.

Además supondremos que $X_1, X_2, \dots, X_n$ son independientes.

Entonces se dice que $X_1, X_2, \dots, X_n$ es una muestra aleatoria de la va X.

Si interesa inferir algún valor para la $\mu = E(X)$ (desconocida), es razonable "estimar" μ con $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Se busca que la estimación de μ esté lo más cerca posible al verdadero valor de μ.

Si tomara otra muestra de n alumnos, y obtenemos otros valores $x_1', x_2', \dots, x_n'$ y entonces tendremos otra estimación $\bar{x}'$ para μ

Podemos considerar que $\bar{x}$ y $\bar{x}'$ son dos valores diferentes que tomó la va $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ donde ahora $X_1, X_2, \dots, X_n$ son las variables de una muestra aleatoria.

Decimos que $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ es un estimador puntual de μ y los valores $\bar{x}$ y $\bar{x}'$ son estimaciones de μ.

En general si tenemos una va que es una función de la muestra aleatoria, se dice que esa va es un estadístico.

$\bar{X}$ es un estadístico. Cuando se usa un estadístico para estimar un parámetro desconocido se dice que el estadístico es un estimador puntual.

En el ejemplo $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ es un estimador puntual de μ. Mientras que $\bar{x}$ y $\bar{x}'$ (valores que tomó $\bar{X}$) son estimaciones puntuales de μ.

Notación: se anota a cualquier parámetro desconocido con θ y con $\hat{\theta}$ al estimador puntual de θ. En el ejemplo: $\hat{\mu} = \bar{X}$

Podemos considerar cualquier estadístico como estimador puntual de un parámetro desconocido.
Entonces necesitamos criterios para considerar que estadístico es "mejor" estimador puntual.

Criterios para evaluar estimadores puntuales:

$\hat{\mu} = \bar{X} \to \bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n$

Estimador insesgado: Sea $\hat{\theta}$ un estimador puntual de θ, donde θ es un parámetro desconocido. Si $E(\hat{\theta}) = \theta$ se dice que $\hat{\theta}$ es un estimador insesgado de θ. (Cualquiera sea el valor de θ) $\to$ el valor verdadero

Ejemplo: sea $X_1, X_2, \dots, X_n$ muestra aleatoria de una va X donde $\mu = E(X)$ desconocida. Sea $\hat{\mu} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Veamos que $\hat{\mu} = \bar{X}$ es un estimador insesgado de μ, esto es: $E(\bar{X}) = \mu$ cualquiera sea el valor de μ.

$$E(\hat{\mu}) = E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \underbrace{E(X_i)}_{\mu} = \frac{1}{n}\sum_{i=1}^{n} \mu$$

$$\to \frac{1}{n} \cdot n\mu = \boxed{\mu = E(\bar{X})}$$ cualquiera sea el valor de μ.

Sea $X_1, X_2, \dots X_n$ muestra aleatoria de una va X con $\mu = E(X)$ y $\sigma^2 = V(X)$ ambos desconocidos. El estadístico que se usa regularmente para estimar σ^2 es:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2 \qquad \hat{\sigma}^2 = S^2 \text{ varianza muestral}$$

$\hat{\sigma}^2 = S^2$ es un estimador insesgado de σ^2 pues: Notar que

$$\boxed{\sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2} \to E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2\right)$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^{n} (X_i - \bar{X})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}\right) = \frac{1}{n-1}\left[E\left(\sum X_i^2\right) - nE(\bar{X}^2)\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)\right] \circledast$$

$E(\bar{X}^2) = V(\bar{X}) + (E(\bar{X}))^2 = \frac{\sigma^2}{n} + \mu^2$

$V(X) = E(X^2) - (E(X))^2$
$E(X^2) = V(X) + (E(X))^2 = \sigma^2 + \mu^2$

(*) $= \frac{1}{n-1}\left[\sum(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right] = \frac{1}{n-1}\left[n(\sigma^2+\mu^2) - \sigma^2 - n\mu^2\right]$

$= \frac{1}{n-1}\left[n\sigma^2 + \cancel{n\mu^2} - \sigma^2 - \cancel{n\mu^2}\right] = \frac{\sigma^2(n-1)}{n-1} = \sigma^2$

Observaciones

① si $X_1, X_2, X_3, \dots, X_n$ es una muestra aleatoria de una v.a. aleatoria X con $\mu = E(X)$ y $\sigma^2 = V(X)$ desconocida, entonces: $\hat{\mu}^2 = \bar{X}^2$ no es un estimador insesgado de μ^2 pues:

$$E(\hat{\mu}^2) = E(\bar{X}^2) = \underbrace{V(\bar{X})}_{\frac{\sigma^2}{n}} + (\underbrace{E(\bar{X})}_{\mu})^2 = \frac{\sigma^2}{n} + \mu^2 \neq \mu^2$$

Sesgo de un estimador puntual: $b(\hat{\theta}) = E(\hat{\theta}) - \theta$. Entonces en el ej anterior: $b(\hat{\mu}^2) = b(\bar{X}^2) = E(\bar{X}^2) - \mu^2$

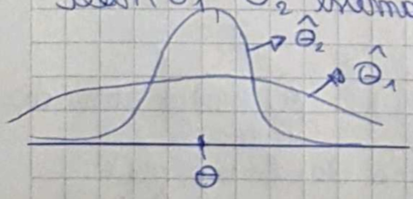
$$= \frac{\sigma^2}{n} + \mu^2 - \mu^2 = \boxed{\frac{\sigma^2}{n}}$$

Notar que el sesgo tiende a cero cuando el tamaño de la muestra tiende a infinito, entonces $\bar{X}^2$ es un estimador de μ^2 asintóticamente insesgado

En general: $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$ no es un estimador insesgado de σ.
Lo sería si viene de una $X \sim N(\mu, \sigma)$

→ si $\hat{\theta}$ es un estimador puntual de θ tal que $b(\hat{\theta}) = 0$ entonces $b(\hat{\theta}) = E(\hat{\theta}) - \theta = 0$, o sea $\hat{\theta}$ es un estimador insesgado de θ

Sean $\hat{\theta}_1$, $\hat{\theta}_2$ estimadores insesgados de θ. O sea $E(\hat{\theta}_1) = \theta$ y $E(\hat{\theta}_2) = \theta$



Consideramos mejor estimador puntual de θ a aquel que tenga menor varianza.
Es decir que si $V(\hat{\theta}_1) > V(\hat{\theta}_2)$, consideramos mejor estimador de θ a $\hat{\theta}_2$

Ej. $X_1, X_2, \dots, X_7$ muestra aleatoria de v.a. X con $\mu = E(X)$ y $\sigma^2 = V(X)$
Sean $\hat{\mu}_1 = \bar{X} = \frac{1}{7}\sum_{i=1}^{7} X_i$ y $\hat{\mu}_2 = \frac{X_1 + X_7}{2}$

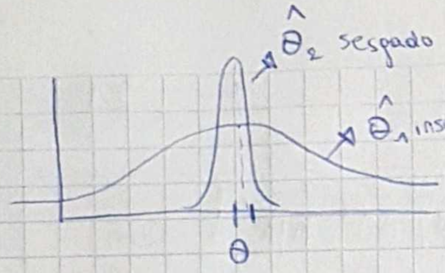
$E(\hat{\mu}_1) = E(\bar{X}) = \mu \qquad E(\hat{\mu}_2) = E\left(\frac{X_1+X_7}{2}\right) = \frac{1}{2}\left(E(X_1) + E(X_7)\right) = \frac{1}{2} 2\mu = \mu$

Ambos insesgados, se ven las varianzas.

$V(\hat{\mu}_1) = V(\bar{X}) = \boxed{\frac{\sigma^2}{7}} \qquad V(\hat{\mu}_2) = V\left(\frac{X_1+X_7}{2}\right) \underset{\text{indep.}}{=} \frac{1}{4}\left(V(X_1) + V(X_7)\right)$

$= \frac{1}{4}(\sigma^2 + \sigma^2) = \boxed{\frac{\sigma^2}{2}}$

Como $V(\hat{\mu}_1) < V(\hat{\mu}_2)$, se considera que $\hat{\mu}_1$ es mejor estimador puntual que $\hat{\mu}_2$



Error cuadrático medio: Sea $\hat{\theta}$ un estimador puntual de un parámetro θ. El ~~error~~

$$ECM(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$$

Observaciones:

① si $\hat{\theta}$ es un estimador insesgado de θ, entonces $E(\hat{\theta}) = \theta$, o sea
$ECM(\hat{\theta}) = E\left[(\hat{\theta} - E(\hat{\theta}))^2\right] = V(\hat{\theta})$ $\qquad V(X) = E\left[(X - E(X))^2\right] = E(X^2) - (E(X))^2$

② se puede probar que $ECM(\hat{\theta}) = V(\hat{\theta}) + (b(\hat{\theta}))^2$

En general si $\hat{\theta}_1$ y $\hat{\theta}_2$ son 2 estimadores puntuales de un parámetro θ, se considera mejor estimador de θ aquel que tenga menor ECM

Ej: Sea $X_1, X_2, \dots, X_7$ muestra aleatoria de una v.a. X con $E(X) = \mu$ y $\sigma^2 = V(X)$.
Consideramos $\hat{\mu}_1 = \frac{1}{7}\sum_{i=1}^{7} X_i = \bar{X}$ $\qquad \hat{\mu}_3 = \frac{2X_1 - X_7 + X_3}{3}$

$\rightarrow ECM(\hat{\theta}) = V(\hat{\theta}) + (b(\hat{\theta}))^2$

$E(\hat{\mu}_1) = E(\bar{X}) = \mu$ $\qquad \therefore b(\hat{\mu}_1) = 0$
$V(\hat{\mu}_1) = \frac{\sigma^2}{7}$ $\rightarrow ECM(\hat{\mu}_1) = V(\hat{\mu}_1)^2 + 0 = \frac{\sigma^2}{7}$

$$E(\hat{\mu}_3) = E\left(\frac{2X_1 - X_7 + X_3}{3}\right) = \frac{1}{3}\left(2E(X_1) - E(X_7) + E(X_3)\right)$$
$$= \frac{1}{3}(2\mu - \mu + \mu) = \frac{2}{3}\mu \rightarrow b(\hat{\mu}_3) = E(\hat{\mu}_3) - \mu = \frac{2}{3}\mu - \mu = \boxed{-\frac{\mu}{3}}$$

$$V(\hat{\mu}_3) = V\left(\frac{2X_1 - X_7 + X_3}{3}\right) = \left(\frac{1}{3}\right)^2\left(4V(X_1) + V(X_7) + V(X_3)\right) = \frac{1}{9}(4\sigma^2 + \sigma^2 + \sigma^2)$$
$$= \frac{2}{3}\sigma^2$$

$$\rightarrow ECM(\hat{\mu}_3) = \frac{2}{3}\sigma^2 + \left(-\frac{\mu}{3}\right)^2 = \frac{2}{3}\sigma^2 + \frac{\mu^2}{9}$$

Notar que $\frac{2}{3}\sigma^2 > \frac{1}{7}\sigma^2 \therefore \frac{2}{3}\sigma^2 + \frac{\mu^2}{9} > \frac{1}{7}\sigma^2$

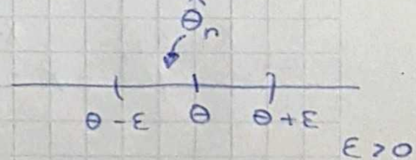
$\therefore ECM(\hat{\mu}_3) > ECM(\hat{\mu}_1) \therefore \hat{\mu}_1$ es mejor estimador de μ.

<u>Estimadores consistentes</u>: Sea $X_1, X_2, \dots X_n$ muestra aleatoria de X.
Sea $\hat{\theta}_n$ un estimador puntual de θ

Se dice que $\hat{\theta}_n$ es un estimador consistente de θ si

$$P\left(|\hat{\theta}_n - \theta| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0 \quad \forall \varepsilon > 0$$

$$\text{ó } P\left(|\hat{\theta}_n - \theta| \leq \varepsilon\right) \xrightarrow[n\to\infty]{} 1 \quad \forall \varepsilon > 0$$



Se puede probar el sig. teorema:

Teorema: Sea $\hat{\theta}_n$ un estimador puntual de θ. Si el sesgo de $b(\hat{\theta}_n) \xrightarrow[n\to\infty]{} 0$ (ó $E(\hat{\theta}_n) \xrightarrow[n\to\infty]{} \theta$) y $V(\hat{\theta}_n) \xrightarrow[n\to\infty]{} 0$ entonces $\hat{\theta}_n$ es un estimador consistente de θ

La recíproca <u>no</u> es válida

Se puede probar que S^2 es un estimador consistente de σ^2
$\bar{X}$ es un estimador consistente de μ.

Metodos para construir estimadores puntuales

Metodo de los momentos

Sea $X_1, \dots, X_n$ muestra aleatoria de una v.a X. Asumimos que se conoce la familia de distribuciones a la que pertenece X.

Sean $\theta_1 \dots \theta_r$ r parametros desconocidos de la distribución de X.

Definimos: $E(X^k)$ como el k-ésimo momento poblacional, donde $k=1,2,3,\dots$
si $k=1$, entonces tenemos $E(x)$

$\frac{1}{n}\sum_{i=1}^{n} X_i^k$ es el k-ésimo momento muestral con $k=1,2,3\dots$ ~~con~~, ~~tenemos~~

Si $k=1$, tenemos $\bar{X}$

Se puede probar que: $P\left(\left|\frac{1}{n}\sum_{i=1}^{k} X_i^k - E(X^k)\right| > \varepsilon\right) \xrightarrow[n\to\infty]{} 0$

Entonces el método de los momentos, para estimar $\theta_1, \theta_2, \dots, \theta_r$, consiste en plantear:

$$\begin{cases} \frac{1}{n}\sum_{i=1}^{n} X_i = E(X) \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 = E(X^2) \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} X_i^r = E(X^r) \end{cases}$$

queda un sist de r ecuaciones con r incógnitas

→ Se resuelve y encuentra estimadores para $\theta_1, \theta_2, \dots, \theta_r$ ($E(X), E(X^2), \dots, E(X^r)$ quedan expresados en función de $\theta_1, \theta_2, \dots, \theta_r$)

Ej. Sea $X_1, \dots X_n$ muestra aleatoria de una v.a X, donde $X \sim N(\mu, \sigma^2)$
Hallar estimadores puntuales para μ y σ^2 por el metodo de los momentos

$\theta_1 = \mu \quad \theta_2 = \sigma^2$

$$\begin{cases} \frac{1}{n}\sum_{i=1}^{n} X_i = E(X) \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 = E(X^2) \end{cases}$$

Sabemos que $E(X) = \mu$ y $E(X^2) = V(X) + (E(X))^2 = \sigma^2 + \mu^2$

$$\begin{cases} \frac{1}{n}\sum_{i=1}^{n} X_i = \mu & (1) \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \sigma^2 + \mu^2 & (2) \end{cases}$$

De (1) → $\boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}}$

De (2) → $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i - \bar{X}^2$

$$\boxed{\therefore \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2}$$

Observación:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = \frac{1}{n}\underbrace{\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)}_{\sum_{i=1}^{n}(X_i - \bar{X})^2} \quad \therefore \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Notar que

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \left(\frac{n-1}{n}\right)S^2$$

$$E(\hat{\sigma}^2) = E\left(\left(\frac{n-1}{n}\right)S^2\right) = \left(\frac{n-1}{n}\right)E(S^2) = \left(\frac{n-1}{n}\right)\sigma^2 \neq \sigma^2$$

Entonces el método de los momentos no produce, en general, estimadores insesgados.

Ej) Sea $X_1, \dots X_n$ muestra aleatoria de una va X, $X \sim P(\lambda)$. Hallar estimador puntual con el método de los momentos.

$\theta_1 = \lambda$

$$\frac{1}{n}\sum_{i=1}^{n} X_i = E(X) \rightarrow \frac{1}{n}\sum_{i=1}^{n} X_i = \lambda \quad \therefore \boxed{\hat{\lambda} = \bar{X}}$$

Ej): Sea $X_1, \dots X_n$ una muestra aleatoria de una va X, $X \sim N(0, \sigma^2)$. Hallar un estimador de σ^2 por el método de los momentos.

$$\frac{1}{n}\sum_{i=1}^{n} X_i = E(X) \rightarrow \frac{1}{n}\sum_{i=1}^{n} X_i = 0 \qquad X \text{ no sirve porque no aparece } \sigma^2$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 = E(X^2) = V(X) = \sigma^2 \quad \therefore \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2}$$

$V(X) = E(X^2) - (E(X))^2$
$V(X) = E(X^2) - 0^2$

Ej) Sea $X_1, \dots, X_n$ muestra aleatoria de una va X, donde la densidad de X está dada por:

$$f(x) = \begin{cases} \frac{\theta}{2^\theta} x^{\theta-1} & 0 < x < 2 \\ 0 & cc \end{cases}$$

Hallar estimador puntual de θ por el método de los momentos.

Observación: Supongamos que para hallar θ planteamos $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$1 = \int_0^2 \frac{\theta}{2^\theta} x^{\theta-1}\,dx = \frac{\theta}{2^\theta}\int_0^2 x^{\theta-1}\,dx = \frac{\theta}{2^\theta}\left(\frac{x^\theta}{\theta}\right)\Bigg|_0^2 = \frac{\theta}{2^\theta}\left(\frac{2^\theta}{\theta} - \frac{0^\theta}{\theta}\right) = \frac{\theta}{2^\theta}\cdot\frac{2^\theta}{\theta} = 1$$

Planteamos: $\frac{1}{n}\sum_{i=1}^{n} X_i = E(X)$ ①

Calculamos $E(X) = \int_0^2 x\left(\frac{2^\theta}{\theta} x^{\theta-1}\right)dx = \frac{\theta}{2^\theta}\int_0^2 x^\theta\,dx = \frac{\theta}{2^\theta}\left(\frac{x^{\theta+1}}{\theta+1}\right)\Bigg|_0^2$

$$= \frac{\theta}{2^\theta}\left(\frac{x^{2+1}}{2+1} - \frac{x^{\theta+1}}{\theta+1}\right) = \frac{\theta}{2^\theta}\left(\frac{2^{\theta+1}}{\theta+1} - 0\right) = \frac{2\theta}{\theta+1}$$

① $\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{2\theta}{\theta+1} \rightarrow (\theta+1)\bar{X} = 2\theta \rightarrow \theta\bar{X} + \bar{X} = 2\theta \rightarrow \bar{X} = 2\theta - \theta\bar{X}$

$$\rightarrow \boxed{\theta = \frac{\bar{X}}{2-\bar{X}}}$$

Método de máxima verosimilitud

Sea $X_1 \dots X_n$ muestra aleatoria de una va X. Asumimos conocida la familia de distribuciones a la que pertenece X.

Sea θ un parámetro desconocido de la dist de X.
Supongamos que X es una va discreta.
Supongamos que $x_1, x_2 \dots x_n$ es una observación de la muestra ($X_1, X_2 \dots X_n$ respectivamente).

Planteamos $P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \overset{\text{indep}}{=}$

$= P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \dots \cdot P(X_n = x_n) = L(x_1, x_2, \dots, x_n, \theta)$ función de verosimilitud

HOJA N° 23

FECHA

Notar que $P(X=x_i)$ queda escrita en función de θ
El método de máxima verosimilitud consiste en hallar aquel valor de θ que maximiza la función de verosimilitud $L(x_1, x_2, \dots, x_n, \theta)$ (la cual es una probabilidad)

Ej) Sea $X_1, X_2, \dots, X_n$ muestra de una va X, $X \sim P(\lambda)$. Hallar el EMV de λ.

Planteamos la función de verosimilitud $L(x_1, x_2, \dots, x_n, \lambda) = P(X_1 = x_1).P(X_2=x_2)\dots$
$P(X_n = x_n) = \frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda}\lambda^{x_2}}{x_2!} \cdot \dots \cdot \frac{e^{-\lambda}\lambda^{x_n}}{x_n!}$

$= \frac{(e^{-\lambda})^n \lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} = \frac{e^{-\lambda n}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$ Para facilitar se toma $\ln(L(x_1, x_2, \dots, x_n, \lambda))$

es estrictamente creciente

$\to \ln(L(x_1, x_2, \dots, x_n, \lambda)) = \ln\left(\frac{e^{-\lambda n}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}\right) = \ln(e^{-\lambda n}\lambda^{\sum_{i=1}^n x_i}) - \ln\left(\prod_{i=1}^n x_i!\right)$

$\to \ln(e^{-\lambda n}) + \ln(\lambda^{\sum_{i=1}^n x_i}) - \ln\left(\prod_{i=1}^n x_i!\right) = (-\lambda n).1 + \sum_{i=1}^n x_i \ln\lambda - \ln\left(\prod_{i=1}^n x_i!\right)$

$\frac{d}{d\lambda}\ln L(x_1, x_2, \dots, x_n, \lambda) \to -n + \left(\sum_{i=1}^n x_i\right)\frac{1}{\lambda} - 0 = 0$

$\therefore n = \left(\sum_{i=1}^n x_i\right)\frac{1}{\lambda} \to \lambda = \frac{1}{n}\sum_{i=1}^n x_i = \bar{x} \to$ se puede verificar que $L(x_1, x_2, \dots, x_n, \lambda)$ tiene en $\lambda = \bar{x}$ un máximo absoluto

$\therefore \boxed{\hat{\lambda} = \frac{1}{n}\sum_{i=1}^n x_i}$

Ej): Sea $X_1, X_2, \dots, X_n$ muestra aleatoria de una va X, $X \sim B(1, P)$

Supongamos que interesa estimar la proporción de artículos defectuosos fabricados por una máquina en 1 día.

Tomamos una muestra de n artículos de toda la producción. Para cada artículo se define

$X_i = \begin{cases} 1 & \text{si el artículo es defectuoso} \\ 0 & cc \end{cases}$ $i = 1, 2, \dots, n$

$\to X_i \sim B(1, P)$

Se quiere hallar EMV de P. $P(X=x_i) = \binom{1}{x_i} P^{x_i}(1-P)^{1-x_i} = P^{x_i}(1-P)^{1-x_i}$

$\to L(x_1, x_2, \dots, x_n, P) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$

$= P(X_1 = x_1).P(X_2 = x_2)\dots P(X_n = x_n)$

$= P^{x_1}(1-P)^{1-x_1} \cdot P^{x_2}(1-P)^{1-x_2} \cdot \dots \cdot P^{x_n}(1-P)^{1-x_n}$

$\to = P^{\sum_{i=1}^n x_i}(1-P)^{\sum_{i=1}^n (1-x_i)} = P^{\sum x_i}(1-P)^{n - \sum_{i=1}^n x_i}$

$\ln(L(x_1, x_2, \dots x_n, P)) = \ln(P^{\sum x_i}) + \ln((1-P)^{n-\sum x_i}) = \sum_{i=1}^n x_i . \ln(P) + (n - \sum x_i).\ln(1-P)$

regla de la cadena

$\frac{d}{dP}\ln(L(x_1, x_2, \dots, x_n, P)) = \left(\sum_{i=1}^n x_i\right)\frac{1}{P} + \left(n - \sum_{i=1}^n x_i\right)\frac{1}{1-P}(-1) = 0$

$\therefore \left(\sum_{i=1}^n x_i\right)\frac{1}{P} = \left(n - \sum_{i=1}^n x_i\right)\frac{1}{1-P} \to \frac{1-P}{P} = \frac{n - \sum x_i}{\sum x_i} \to \frac{1}{P} \cancel{-1} = \frac{n}{\sum x_i} \cancel{-1}$

$\rightarrow \frac{1}{p} = \frac{n}{\sum x_i} \rightarrow p = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x} \rightarrow \hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Ej) b) i) Se seleccionan $n=100$ chips fabricados por una compañía. 5 son defectuosos y $\sum_{i=1}^{100} X_i = 5$ ∴ la estimación de p es $\bar{x} = \frac{1}{100}\sum_{i=1}^{100} x_i = \frac{5}{100} = 0{,}05$

Caso continuo

Si $X_1, X_2, \dots, X_n$ es una muestra aleatoria de una va X continua con densidad $f(x)$ y si $x_1, x_2, \dots, x_n$ es una medición de la muestra aleatoria, entonces se toma:

$L(x_1, x_2, \dots, x_n, \theta) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n)$. Se toma como estimación de θ a aquel valor que maximiza $L(x_1, x_2, \dots, x_n, \theta)$

Ej) Sea $X_1, X_2, \dots, X_n$ muestra aleatoria de una va X, $X \sim Exp(\lambda)$

Hallar el EMV de λ

$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$ $\rightarrow L(x_1, x_2, \dots, x_n, \lambda) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n)$

$= \lambda e^{-\lambda x_1} \cdot \lambda e^{-\lambda x_2} \cdot \dots \cdot \lambda e^{-\lambda x_n}$

$= \lambda^n e^{-\sum_{i=1}^{n} -\lambda x_i} = \lambda^n e^{-\lambda \sum x_i}$

$\ln(L(x_1, x_2, \dots, x_n, \lambda)) = \ln(\lambda^n) + \ln(e^{-\lambda \sum x_i}) = n\ln(\lambda) + (-\lambda \sum_{i=1}^{n} x_i)\ln(e)$

$= n\ln(\lambda) - \lambda \sum_{i=1}^{n} x_i \rightarrow \frac{d}{d\lambda} \ln(L(x_1, x_2, \dots, x_n, \lambda)) = 0$

$\rightarrow \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \rightarrow \lambda = \frac{n}{\sum_{i=1}^{n} x_i} \rightarrow \frac{1}{\frac{\sum x_i}{n}} = \lambda = \frac{1}{\bar{x}}$

∴ $\hat{\lambda} = \frac{1}{\bar{X}}$

Se determina la duración en hs de cada una de 10 lámparas dando los siguientes datos:

$x_1 = 49{,}5$ $x_2 = 1015$ $x_3 = 1009{,}4$ $x_4 = 173{,}4$ $x_5 = 251{,}6$

$x_6 = 315$ $x_7 = 1301{,}4$ $x_8 = 732{,}8$ $x_9 = 660{,}6$ $x_{10} = 1456{,}5$

Se sospecha que la dur en hs tiene dist exponencial.

a) Hallar el EMV de λ.

b) Hallar la estimación de λ para la muestra dada.

X: "tiempo de vida en horas de una lámpara" $X \sim Exp(\lambda)$ λ desconocido

Se toma una muestra de 10 lámparas

X_i: "duración en horas de la lámpara i" $i = 1, 2, \dots, 10$

$X_1, X_2, \dots, X_{10}$ muestra aleatoria de la va X.

$x_1, x_2, \dots, x_{10}$ es una medición de la muestra aleatoria.

a) $\hat{\lambda} = \frac{1}{\bar{X}}$ b) $\lambda = \frac{1}{\bar{x}} = \frac{1}{696{,}52}$ es la estimación de λ para la muestra dada

FECHA

g) sea $X_1, X_2, \ldots, X_n$ muestra aleatoria de una v a X $X \sim N(\mu, \sigma^2)$

Hallar EMV para μ y σ^2 $f(x) = \frac{1}{\sqrt{\sigma^2 2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$

$L(x_1, x_2, \ldots, x_n, \mu, \sigma^2) = f(x_1) \cdot f(x_2) \cdot \ldots \cdot f(x_n)$

$= \frac{1}{\sqrt{\sigma^2 2\pi}} e^{-\frac{1}{2\sigma^2}(x_1-\mu)^2} \cdot \frac{1}{\sqrt{\sigma^2 2\pi}} e^{-\frac{1}{2\sigma^2}(x_2-\mu)^2} \cdot \ldots \cdot \frac{1}{\sqrt{\sigma^2 2\pi}} e^{-\frac{1}{2\sigma^2}(x_n-\mu)^2}$

$= \left(\frac{1}{\sqrt{\sigma^2 2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2} = \frac{1}{(\sigma^2 2\pi)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}(x_i-\mu)^2}$

$\ln(L(x_1, x_2, \ldots, x_n, \mu, \sigma^2)) = \ln\left(\frac{1}{(\sigma^2 2\pi)^{n/2}}\right) + \ln\left(e^{-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2}\right)$

$= 0 - \frac{n}{2}\ln(\sigma^2 2\pi) - \frac{1}{2\sigma^2}\sum(x_i-\mu)^2$

$\begin{cases} \frac{d}{d\mu}\ln L(x_1, x_2, \ldots, x_n, \mu, \sigma) & ① \\ \frac{d}{d\sigma}\ln L(x_1, x_2, \ldots, x_n, \mu, \sigma) & ② \end{cases}$ ① $-\frac{1}{2\sigma^2}\sum 2(x_i-\mu)(-1)$

② $\left(-\frac{n}{2}\right)\frac{1}{\sigma^2 2\pi}2\pi - \frac{1}{2}\sum(x_i-\mu)^2\left(-\frac{1}{\sigma^4}\right)$

① $-\frac{1}{\cancel{2}\sigma^2}\sum \cancel{2}(x_i-\mu)\cancel{(-1)} \rightarrow \frac{1}{\sigma^2}\sum(x_i-\mu) = 0$

$\rightarrow \sum(x_i-\mu) = 0 \rightarrow \sum x_i - \sum\mu = 0 = \sum x_i - n\mu = 0 \rightarrow \boxed{\hat{\mu} = \frac{\sum x_i}{n} = \bar{X}}$

② $\left(-\frac{n}{2}\right)\frac{1}{\sigma^2 2\pi}2\pi - \frac{1}{2}\sum(x_i-\mu)^2\left(-\frac{1}{\sigma^4}\right) \overset{0}{=} \Rightarrow -\frac{n}{\sigma^2} + \frac{1}{\sigma^4}\sum(x_i-\mu)^2 = 0$

$\rightarrow \frac{n}{\sigma^2} = \frac{1}{\sigma^4}\sum(x_i-\mu)^2 \rightarrow \frac{\sigma^4}{\sigma^2} = \frac{1}{n}\sum(x_i-\mu)^2 \Rightarrow \sigma^2 = \frac{1}{n}\sum(x_i-\bar{x})^2$

$\rightarrow \boxed{\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$ Importante acordarse los resultados

Propiedades de los EMV → son válidos si la muestra es grande

Bajo condiciones muy generales, se puede probar:

① Los EMV son asintóticamente insesgados

② Son consistentes los EMV

③ Para muestras muy grandes, son asintóticamente los de menor varianza

④ Propiedad de invarianza: sea $\hat{\theta}$ es el EMV de θ. Sea $g(x)$ una función inyectiva. Entonces el EMV de $g(\theta)$ es $g(\hat{\theta})$

En el ej anterior se encontró el EMV de σ^2. $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$.

Si queremos hallar el EMV de σ, usamos la propiedad de invarianza

$\sigma \overset{g(x)=\sqrt{x}}{=} \sqrt{\sigma^2} \rightarrow \therefore \hat{\sigma} = \sqrt{\hat{\sigma}^2} = \sqrt{\frac{1}{n}\sum(x_i-\bar{x})^2}$

NOTA

En el ej de la binomial: b) ii) Se pide el EMV de $(1-p)^6$ $\quad n=100$

$\sum_{i=1}^{100} x_i = 5$ $\quad \hat{p} = \bar{X}$. La estimación para p es $\hat{p} = \frac{5}{100} = 0,05$

$g(p) = (1-p)^6$ (o sea se toman 6 chips y ninguno es defectuoso)

el EMV de $(1-p)^6 = g(\hat{p}) = g(\bar{X}) = (1-\bar{X})^6$ y la estimación para $(1-p)^6$ es $(1-0,05)^6$

En el ej de las lámparas: X: duración en horas de una lámpara

$X_1, X_2, \dots, X_n$ m.a. de X. el EMV de $\lambda \rightarrow \hat{\lambda} = \frac{1}{\bar{X}}$

Hallar el EMV de $P(X < 1400) = 1 - e^{-\lambda 1400}$

$\therefore$ el EMV es $\widehat{1 - e^{-\lambda . 1400}} = 1 - e^{-\hat{\lambda} . 1400} = 1 - e^{-\frac{1}{\bar{X}} 1400}$

Intervalos de confianza

Sea $X_1, X_2, \dots, X_n$ una muestra aleatoria de una va X. Sea θ un parámetro desconocido de la distribución de X. Un intervalo de confianza para θ es un intervalo $(\hat{\theta}_1, \hat{\theta}_2)$ donde $\hat{\theta}_1$ y $\hat{\theta}_2$ son estadísticos, de manera tal que $P(\theta \in (\hat{\theta}_1, \hat{\theta}_2)) = 1 - \alpha$

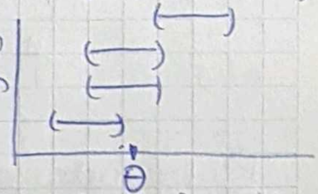
con $1-\alpha$ fijado de antemano. $(1-\alpha).100\%$ es el nivel de confianza del intervalo.

Si $x_1, x_2, \dots, x_n$ es una medición de la muestra aleatoria entonces $\hat{\theta}_1$ tomará el valor θ_1 y $\hat{\theta}_2$ tomará el valor θ_2

O sea para la medición $x_1, x_2, \dots, x_n$ tenemos un intervalo numérico $(\theta_1; \theta_2)$

Si $x_1', x_2', \dots, x_n'$ es otra medición de la muestra aleatoria entonces obtendremos otro intervalo numérico (θ_1', θ_2') $\quad \mu \in (-\infty, \infty)$

Si $1-\alpha = 0,95$, entonces si tenemos 100 mediciones de la m.a. vamos a tener 100 intervalos numéricos de los cuales aprox. 95 van a contener a θ



~~A mayor nivel de confianza, más precisión~~

La precisión de la estimación será mayor con un intervalo de longitud pequeña

→ Intervalo de confianza para la media de una normal con varianza conocida:

Sea $X_1, X_2, \dots, X_n$ m.a. de una v.a X donde $X \sim N(\mu, \sigma^2)$ con σ^2 conocido. Se quiere un intervalo de confianza para μ con nivel 95%

Utilizaremos el método del pivote

① Se toma un estimador puntual de μ $\quad \hat{\mu} = \bar{X}$

② Se construye a partir de $\hat{\mu} = \bar{X}$ un estadístico tal que

- debe contener a μ
- debe tener distribución conocida

NOTA

Consideramos el estadístico $Z = \dfrac{\bar{X} - \mu}{\sqrt{\frac{\sigma}{n}}}$ notar que $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ y entonces $Z \sim N(0,1)$

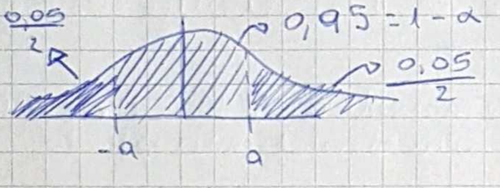
además Z contiene a μ. A Z se lo llama estadístico pivote.

Si conocemos la distribución de Z podemos plantear: hallar a tal que $0{,}95 = P(-a \leq Z \leq a)$. Si $Z \sim N(0,1)$:

$$P(-a \leq Z \leq a) = \Phi(a) - \Phi(-a) = 2\Phi(a) - 1$$

$$\rightarrow 0{,}95 = 2\Phi(a) - 1 \rightarrow \Phi(a) = \frac{1 - 0{,}95}{2} = 0{,}975 \therefore \boxed{a = 1{,}96} \text{ tabla}$$

$$a = z_{\frac{0{,}05}{2}} = 1{,}96$$



$$\rightarrow 0{,}95 = (-1{,}96 \leq Z \leq 1{,}96) = \left(-1{,}96 \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \leq 1{,}96\right)$$

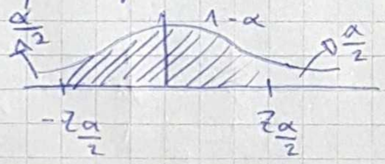
$$= P\left(-1{,}96\sqrt{\frac{\sigma^2}{n}} \leq \bar{X} - \mu \leq 1{,}96\sqrt{\frac{\sigma^2}{n}}\right) = P\left(-\bar{X} - 1{,}96\sqrt{\frac{\sigma^2}{n}} \leq -\mu \leq -\bar{X} + 1{,}96\sqrt{\frac{\sigma^2}{n}}\right)$$

$$= P\left(\bar{X} + 1{,}96\sqrt{\frac{\sigma^2}{n}} \geq \mu \geq -1{,}96\sqrt{\frac{\sigma^2}{n}} + \bar{X}\right) = P\left(\bar{X} - 1{,}96\sqrt{\frac{\sigma^2}{n}} \leq \mu \leq \bar{X} + 1{,}96\sqrt{\frac{\sigma^2}{n}}\right)$$

$$= P\left(\mu \in \left(\underbrace{\bar{X} - 1{,}96\sqrt{\frac{\sigma^2}{n}}}_{\hat{\theta}_1}, \underbrace{\bar{X} + 1{,}96\sqrt{\frac{\sigma^2}{n}}}_{\hat{\theta}_2}\right)\right) \therefore \left(\bar{X} - z_{\frac{0{,}05}{2}}\sqrt{\frac{\sigma^2}{n}}, \bar{X} + z_{\frac{0{,}05}{2}}\sqrt{\frac{\sigma^2}{n}}\right)$$

En general, un intervalo de confianza para μ de nivel $(1-\alpha)100\%$ será:

$$\boxed{\left(\bar{X} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}, \bar{X} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right)}$$



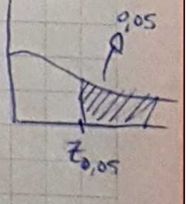
$z_{\frac{\alpha}{2}} \rightarrow$ valor crítico

Se usan tablas de valores críticos $\rightarrow$ del área a la derecha (el complemento de $\frac{0{,}05}{2}$) $P(X \geq x)$

Ejemplo. Se sabe que el límite de elasticidad de un tipo particular de barra de acero está normalmente distribuido con $\sigma = 100$. Se toman 25 barras al azar, a cada una se le mide el límite de elasticidad y resulta un $\bar{x} = 8439/6$. Calcule un intervalo de confianza de nivel 90% para la verdadera media de límite de elasticidad.

X: "límite de elasticidad de una barra de acero" $X \sim N(\mu, \sigma^2)$ $\sigma = 100$
X_i: "límite de elasticidad de barra de acero i" $i = 1, \ldots, 25$
$X_1, X_2, \ldots, X_{25}$ m.a. de la v.a. X. $\bar{x} = 8439/6$

$1 - \alpha = 0{,}9 \rightarrow \alpha = 0{,}1 \rightarrow \frac{\alpha}{2} = 0{,}05$ $\left(\bar{X} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}, \bar{X} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right)$
$z_{0{,}05} = 1{,}65$



$$\rightarrow \left(8439 - 1{,}65\sqrt{\frac{100^2}{25}}, 8439 + 1{,}65\sqrt{\frac{100^2}{n}}\right) = (8406, 8472)$$

$L = 8472 - 8406 = 66$

NOTA

Si tomamos $1-\alpha = 0,92$ $\quad \alpha = 0,08$ $\quad \frac{\alpha}{2} = 0,04$ $\quad z_{0,04} = 1,76$

$$\left(8439 - 1,76\sqrt{\frac{100^2}{n}} \,,\, 8439 + 1,76\sqrt{\frac{100^2}{n}}\right) = (8403,8 \,,\, 8474,2)$$

$L = 70,4 \longrightarrow$ aumentó la longitud

Observaciones: ① $\left(\bar{X} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}} \,,\, \bar{X} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right)$ está centrado en $\bar{X}$

② La longitud $L = \bar{X} + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}} - \left(\bar{X} - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}\right) = \boxed{2 z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}}}$

Notar que: si dejamos n fijo y aumenta $1-\alpha$ (disminuye α), entonces $z_{\frac{\alpha}{2}}$ aumenta $\therefore$ L aumenta

Si $1-\alpha$ se deja fijo y aumenta n entonces L disminuye (aumenta la precisión de la estimación)

Se puede plantear que tamaño n de muestra se necesita para que la longitud L sea menor a "l". ("l" se fija de antemano)

$$L = 2 z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}} \leq l \longrightarrow \left(\frac{2 z_{\frac{\alpha}{2}} \sigma}{l}\right)^2 \leq n$$

ej: que tamaño de muestra se necesita para que el intervalo de nivel 90% se reduzca a la mitad.

$$L = 2 \cdot 1,65\sqrt{\frac{100^2}{n}} \leq 33 \longrightarrow \underbrace{\left(\frac{2 \cdot 1,65 \cdot 100}{33}\right)^2}_{100} \leq n \quad \therefore \boxed{n \geq 100}$$

Si se estima puntualmente μ con $\bar{X}$, se comete un error E, y notar que ese $E \leq \frac{L}{2}$ con una confianza de $(1-\alpha)100\%$