Intervalos de confianza en poblaciones gausianas

N. Corral, C. Carleos, M.T. López

Departamento de Estadística
Universidad de Oviedo

14 de noviembre de 2020

Esperanza

Sea $\vec{x}$ un vector aleatorio de dimensión p.

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

Denotemos su esperanza así:

$$\vec{\mu} = E(\vec{x}) = \begin{pmatrix} E(x_1) \\ \vdots \\ E(x_p) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$$

$$\mu_j = E(x_j), \quad \forall\, j = 1, \ldots, p$$

Matriz de varianzas-covarianzas

$$\Sigma = Cov(\vec{x}) = E[(\vec{x}-\vec{\mu})(\vec{x}-\vec{\mu})']$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{pmatrix}$$

- Varianza de x_j.

$$\sigma_{jj} = E[(x_j-\mu_j)(x_j-\mu_j)] = \sigma_j^2, \quad j = 1, \dots, p$$

- Covarianza de x_j, x_k.

$$\sigma_{jk} = E[(x_j-\mu_j)(x_k-\mu_k)] \quad \forall j \neq k$$

$$\sigma_{jk} = 0 \Longleftrightarrow x_j \text{ y } x_k \text{ son linealmente independientes}$$

$$\text{independencia} \quad \substack{\Longrightarrow \\ \not\Longleftarrow} \quad \text{independencia lineal}$$

Transformaciones lineales

Sea A una matriz $q \times p$ de elementos fijos.
Se define una **trasformación lineal** de $\vec{x}$ como el vector aleatorio $\vec{y} = A\vec{x}$:

$$\vec{y} = A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1q} \\ a_{21} & a_{22} & \dots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qp} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} \vec{a}'_1 \\ \vec{a}'_2 \\ \vdots \\ \vec{a}'_q \end{pmatrix} \vec{x}$$

con $\vec{a}'_j = (a_{j1}, \dots, a_{jp})$ fila j de A.

$$y_j = \vec{a}'_j \vec{x} = \sum_{k=1}^{p} a_{jk} x_k, \quad j = 1, \dots, q$$

Esperanza de una transformación lineal

$$\vec{y} = A\vec{x} \iff y_j = \sum_{k=1}^{p} a_{jk} x_k, \quad j = 1, \ldots, q$$

$$E(y_j) = E\left(\sum_{k=1}^{p} a_{jk} x_k\right) = \sum_{k=1}^{p} a_{jk} E(x_k) = \sum_{k=1}^{p} a_{jk}\, \mu_k = \vec{a}_j'\, \vec{\mu}$$

Por tanto,

$$E(\vec{y}) = E(A\vec{x}) = AE(\vec{x}) = A\,\vec{\mu}$$

Covarianza de una transformación lineal

$$\vec{y} = A\vec{x} \iff y_j = \sum_{k=1}^{p} a_{jk} x_k$$

$$Cov(\vec{y}) = \begin{pmatrix} Cov(y_1, y_1) & Cov(y_1, y_2) & \dots & Cov(y_1, y_q) \\ Cov(y_2, y_1) & Cov(y_2, y_2) & \dots & Cov(y_2, y_q) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(y_q, y_1) & Cov(y_q, y_2) & \dots & Cov(y_q, y_q) \end{pmatrix}$$

$$\begin{aligned} Cov(y_i, y_j) &= E[(y_i - E(y_i))(y_j - E(y_j))] \\ &= E[\vec{a}_i'(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})'\vec{a}_j] \\ &= \vec{a}_i' E[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})']\vec{a}_j \\ &= \vec{a}_i' \Sigma \vec{a}_j \end{aligned}$$

$$\implies Cov(\vec{y}) = A \Sigma A'$$

Proyecciones ortogonales

- Sea Ω un subespacio vectorial de $\mathbb{R}^n$.
  Entonces $\forall \vec{y} \in \mathbb{R}^n \;\; \exists\, \vec{u} \in \Omega, \vec{v} \in \Omega^{\perp}$ únicos, $\;\vec{y} = \vec{u} + \vec{v}$.
- También $\exists$ una única matriz que verifica $P_{\Omega}\vec{y} = \vec{u}$.
  - $P_{\Omega} = X(X'X)^{-1}X'$, donde X es una base de Ω.
  - Es simétrica: $P'_{\Omega} = \left(X(X'X)^{-1}X'\right)' = X(X'X)^{-1}X' = P_{\Omega}$
  - Es idempotente:

$$P_{\Omega}P_{\Omega} = X\,(X'\,X)^{-1}X'\,X\;(X'\,X)^{-1}X' = X\,(X'\,X)^{-1}X' = P_{\Omega}$$

  - Sea $\vec{z} \in \Omega \iff \vec{z} = X\,\vec{a} \quad \exists\vec{a}$ por ser X base.

$$P_{\Omega}\vec{z} = X\,(X'\,X)^{-1}X'\,X\,\vec{a} = X\,\vec{a} = \vec{z}$$

  - Sea $\vec{w} \in \Omega^{\perp} \iff \vec{z}'\,\vec{w} = 0 \quad \forall\vec{z} \in \Omega \implies X'\vec{w} = 0$

$$P_{\Omega}\vec{w} = X\,(X'\,X)^{-1}X'\,\vec{w} = 0$$

  - P proyecta todo $\vec{y} \in \mathbb{R}^n$ ortogonalmente sobre Ω.

Matriz de proyección ortogonal

Sea A matriz simétrica e idempotente.
Sea $\vec{u}$ un vector propio de A.

- $A\vec{u} = \lambda\vec{u}$ por ser $\vec{u}$ propio.
- $A\vec{u} = AA\vec{u} = A\lambda\vec{u} = \lambda^2\vec{u}$ por ser A idempotente.
- $\lambda \geq 0$ por ser A simétrica.

En consecuencia, $\lambda \in \{0, 1\}$.

Sea A simétrica e idempotente de rango q. Entonces

$$A = U\Delta U' \qquad \text{con } \Delta \text{ diagonal y } U'U = I$$

$$traza(A) = traza(U\Delta U') = traza(\Delta U'U) = traza(\Delta) = q$$

ya que q es el número de valores propios no nulos de A.
Luego

$$traza(A) = rango(A)$$

Matriz de centrado H

- $H = I - \vec{1}(\vec{1}'\vec{1})^{-1}\vec{1}' = I - \frac{1}{n}\vec{1}\vec{1}'$

$$H\vec{x} = \left(I - \frac{1}{n}\vec{1}\vec{1}'\right)\vec{x} = \vec{x} - \frac{1}{n}\vec{1}\vec{1}'\vec{x} = \vec{x} - \bar{x}\,\vec{1} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}$$

- $HH = H$

$$\begin{aligned}(I - \tfrac{1}{n}\vec{1}\vec{1}')(I - \tfrac{1}{n}\vec{1}\vec{1}') &= I - \tfrac{1}{n}\vec{1}\vec{1}' - \tfrac{1}{n}\vec{1}\vec{1}' + \tfrac{1}{n^2}\vec{1}\vec{1}'\vec{1}\vec{1}' \\ &= I - \tfrac{1}{n}\vec{1}\vec{1}'\end{aligned}$$

- $rango(H) = traza(H) = n - 1$

Distribución gausiana p-dimensional

Si $\vec{x} = (x_1, \ldots, x_p)'$ sigue una $N_p(\vec{\mu}, \Sigma)$,
con Σ definida positiva, entonces

- La función de densidad es:

$$f(\vec{x}) = \frac{1}{\sqrt{2\pi|\Sigma|}} \exp\left(\frac{-1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})\right)$$

- $\vec{y} = A\,\vec{x}$, con $A_{q\times p}$, sigue una $N_q(A\,\vec{\mu}, A\Sigma A')$

$$E(\vec{y}) = A\,E(\vec{x}) = A\,\vec{\mu} \qquad Cov(\vec{y}) = A\,Cov(\vec{x})\,A' = A\Sigma A'$$

- Si Σ es diagonal ($\sigma_{jk} = 0$ para $j \neq k$),
  entonces todas las x_j son independientes entre sí.
- La distribución de x_j es $N(\mu_j, \sigma_{jj})$, para $j = 1, \ldots, p$.

Distribución de la varianza muestral

- Si $\vec{x} \equiv N_p(\vec{\mu}, \sigma^2 I)$ entonces

$$\frac{1}{\sigma^2}(\vec{x}-\vec{\mu})'(\vec{x}-\vec{\mu}) = \vec{z}'\vec{z} \equiv \chi^2_p$$

  con $\vec{z} \equiv N_p(\vec{0}, I)$.

- Si $\vec{x} \equiv N_n(\mu\vec{1}, I)$ y H es la matriz de centrado, entonces

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = \vec{x}'H\vec{x} \equiv \chi^2_{n-1}$$

Distribución de la varianza muestral

- Si $\vec{x} \equiv N_n(\mu\vec{1}, I)$ y H es la matriz de centrado, entonces

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \vec{x}' H \vec{x} \equiv \chi^2_{n-1}$$

pues

$$\begin{aligned}
\sum_{i=1}^{n}(x_i - \bar{x})^2 &= \sum_{i=1}^{n}[(x_i - \mu) - (\bar{x} - \mu)]^2 \\
&= \sum_{i=1}^{n}(y_i - \bar{y})^2 \\
&= \vec{y}' H \vec{y} = \vec{y}' U \Delta U' \vec{y} = \vec{z} \Delta \vec{z}
\end{aligned}$$

donde $\vec{z} = U'\vec{y} \equiv N(\vec{0}, I)$, ya que

$$Cov(\vec{z}) = U' Cov(\vec{y}) U = U'U = I$$

Distribución de la varianza muestral

Por ser H simétrica e idempotente de rango $n-1$,
Δ es una matriz diagonal con $n-1$ unos y un cero.

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

$$\vec{z}'\Delta\vec{z} = \sum_{i=1}^{n-1} z_i^2 \equiv \chi^2_{n-1}$$

Por tanto,

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 \equiv \chi^2_{n-1}$$

Distribución de la varianza muestral

- Si $\vec{x} \equiv N_n(\mu\vec{1}, I)$ y H es la matriz de centrado, entonces

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \vec{x}'H\vec{x} \equiv \chi^2_{n-1}$$

- Si $\vec{x} \equiv N_n(\mu\vec{1}, \sigma^2 I)$, p.ej. m.a.s. de $N(\mu, \sigma^2)$,

$$\frac{n\,S^2}{\sigma^2} = \frac{(n-1)\,\hat{S}^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2} =$$

$$\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sigma}\right)^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 \equiv \chi^2_{n-1}$$

con $\vec{y} = \frac{\vec{x}-\mu\vec{1}}{\sigma} \equiv N_n(\vec{0}, I)$

Intervalo de confianza para σ^2

- Sea $\vec{x}$ m.a.s. de $x \equiv N(\mu, \sigma^2)$, μ y σ desconocidas
- $n\,S^2/\sigma^2 \equiv \chi^2_{n-1}$ es una función pivote
- Sean a y b tales que

$$\begin{aligned} 1-\alpha &= P\left[a \leq \frac{n\,S^2}{\sigma^2} \leq b\right] = P\left[\frac{1}{a} \geq \frac{\sigma^2}{n\,S^2} \geq \frac{1}{b}\right] \\ &= P\left[\frac{1}{b} \leq \frac{\sigma^2}{n\,S^2} \leq \frac{1}{a}\right] = P\left[\frac{n\,S^2}{b} \leq \sigma^2 \leq \frac{n\,S^2}{a}\right] \end{aligned}$$

luego $\left[\frac{n\,S^2}{b}, \frac{n\,S^2}{a}\right]$ es un i.c. para σ^2 a nivel $1-\alpha$.

Independencia entre media y varianza

Sea $\vec{x} \equiv N(\vec{\mu}, I)$. Entonces $\bar{x}$ es independiente de $H\vec{x}$, pues

$$H\vec{x} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}$$

$$Cov(H\vec{x}, \bar{x}) = Cov\left(H\vec{x}, \frac{1}{n}\vec{x}'\vec{1}\right) = E\left(H(\vec{x} - \vec{\mu})\frac{1}{n}(\vec{x} - \vec{\mu})'\vec{1}\right)$$

$$= \frac{1}{n} H E[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})']\,\vec{1} = \frac{1}{n} H\,\vec{1} = \vec{0}$$

Consecuencia: $\bar{x}$ es independiente de toda $g(H\vec{x})$, en particular de

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \vec{x}'H'H\vec{x} = \vec{x}'HH\vec{x} = \vec{x}'H\vec{x}$$

Intervalo de confianza para μ

- Sea $\vec{x}$ m.a.s. de $x \equiv N(\mu, \sigma^2)$
- $\bar{x} \equiv N(\mu, \sigma^2/n)$
- $n\,S^2/\sigma^2 \equiv \chi^2_{n-1}$
- $\bar{x}$ y S^2 son independientes, luego

$$
\begin{aligned}
\frac{\bar{x}-\mu}{S/\sqrt{n-1}} &= \frac{\bar{x}-\mu}{\hat{S}/\sqrt{n}} \\
&= \frac{\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)\,\hat{S}^2/\sigma^2}{n-1}}} \equiv \frac{N(0,1)}{\sqrt{\frac{\chi^2_{n-1}}{n-1}}} = t_{n-1}
\end{aligned}
$$

Intervalo de confianza para μ

- Considérese el pivote

$$\frac{\bar{x}-\mu}{\hat{S}/\sqrt{n}} \equiv t_{n-1}$$

- Sean a y b tales que

$$\begin{aligned} 1-\alpha &= P\left[a \leq \frac{\bar{x}-\mu}{\hat{S}/\sqrt{n}} \leq b\right] = P\left[\frac{a\,\hat{S}}{\sqrt{n}} \leq \bar{x}-\mu \leq \frac{b\,\hat{S}}{\sqrt{n}}\right] \\ &= P\left[\bar{x}-\frac{b\,\hat{S}}{\sqrt{n}} \leq \mu \leq \bar{x}-\frac{a\,\hat{S}}{\sqrt{n}}\right] \end{aligned}$$

  luego $\left[\bar{x}-\frac{b\,\hat{S}}{\sqrt{n}}, \bar{x}-\frac{a\,\hat{S}}{\sqrt{n}}\right]$ es un i.c. para μ a nivel $1-\alpha$.
- Habitualmente $b=-a$ luego i.c. $= \bar{x} \pm \frac{b\,\hat{S}}{\sqrt{n}}$