Esperanza

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

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

Esperanza.

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

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

Matriz de varianzas-covarianzas

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

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

donde

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

  Si $\sigma_{jk} = 0 \iff x_j$ y x_k son linealmente independientes
  La independencia implica la independencia lineal. Lo contrario no suele ser cierto

Transformaciones lineales

Sea A una matriz de elementos fijos con dimensiones (q,p) .
Se define una trasformación lineal de $\vec{x}$ como el vector aleatorio $\vec{y} = A\vec{x}$:

$$\vec{y} = A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1q} \\ a_{21} & a_{22} & ... & a_{2q} \\ ... & ... & ... & ... \\ a_{q1} & a_{q2} & ... & a_{qp} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_p \end{pmatrix} = \begin{pmatrix} \vec{a_1'} \\ \vec{a_2'} \\ ... \\ \vec{a_q'} \end{pmatrix} \vec{x}$$

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

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

Esperanza de una transformación lineal

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

$$E(y_j) = E(\sum_{k=1}^{p} a_{jk} x_k) = \sum_{k=1}^{p} a_{jk} E(x_k) = \sum_{k=1}^{p} a_{jk}\, \mu_j = \vec{a}_j^{\,\prime}\, \vec{\mu}$$

Por tanto

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

Covarianza de una transformación lineal

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

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

$$Cov(y_i, y_j) = E[(y_i - E(y_i))\,(y_j) - E(y_j))]$$

$$= E[\vec{a}_i'(\vec{x} - \vec{\mu})\,(\vec{x} - \vec{\mu})'\vec{a}_j] = \vec{a}_i' E[(\vec{x} - \vec{\mu})\,(\vec{x} - \vec{\mu})']\vec{a}_j = \vec{a}_i' \Sigma \vec{a}_j$$

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

Derivadas parciales

Sea $g(\vec{x}) = \vec{a'}\,\vec{x} = \sum_{k=1}^{p} a_k\, x_k$, con $\vec{a'} = (a_1, ..., a_p)$. Entonces

$$\frac{\partial g}{\partial \vec{x}} = \begin{pmatrix} \frac{\partial g}{\partial x_1} \\ \frac{\partial g}{\partial x_2} \\ \dots \\ \frac{\partial g}{\partial x_p} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_p \end{pmatrix} = \vec{a}$$

Derivadas parciales

Sea $g(\vec{x}) = \vec{x'} A \vec{x}$ una transformación cuadrática de X, con

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \dots & \dots & \dots & \dots \\ a_{p1} & a_{p2} & \dots & a_{pp} \end{pmatrix} = \begin{pmatrix} \vec{a'_1} \\ \vec{a'_2} \\ \dots \\ \vec{a'_p} \end{pmatrix} = \begin{pmatrix} a_{\vec{(1)}} \\ a_{\vec{(2)}} \\ \dots \\ a_{\vec{(p)}} \end{pmatrix}$$

donde:

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

$$a_{\vec{(k)}} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \dots \\ a_{pk} \end{pmatrix} \text{ columna k de } A$$

Derivadas parciales

$$g(\vec{x}) = \sum_{j=1}^{p} \sum_{k=1}^{p} a_{jk}\, x_j\, x_k = \sum_{j=1}^{p} a_{jj}\, x_j^2 + \sum_{j \neq k} a_{jk}\, x_j\, x_k$$

$$\frac{\partial g}{\partial x_1} = 2\, a_{11}\, x_1 + \sum_{j \neq 1} a_{j1}\, x_j + \sum_{k \neq 1} a_{k1}\, x_k$$

$$= \sum_{j=1}^{p} a_{j1}\, x_j + \sum_{k=1}^{p} a_{k1}\, x_k = \vec{a_1'}\, \vec{x} + \vec{a_{(1)}'}\, \vec{x} = (\vec{a_1'} + \vec{a_{(1)}'})\, \vec{x}$$

Por tanto,

$$\frac{\partial g}{\partial \vec{x}} = \begin{pmatrix} \vec{a_1'} + \vec{a_{(1)}'} \\ \vec{a_2'} + \vec{a_{(2)}'} \\ \cdots \\ \vec{a_p'} + \vec{a_{(p)}'} \end{pmatrix} \vec{x} = (A + A')\vec{x}$$

Proyecciones ortogonales

- Si Ω es un subespacio vectorial de $\mathbb{R}^n$ entonces todo vector $y \in \mathbb{R}^n$ se puede descomponer de forma única como $y = u + v,\ u \in \Omega, v \in \Omega^\perp$
- Si Ω es un subespacio vectorial de $\mathbb{R}^n$ existe una única matriz que verifique la condición: $P_\Omega y = u$.
  - $P_\Omega = X(X^tX)^{-1}X^t$, donde X es una base del subespacio.
  - Es simétrica e idempotente $P_\Omega P_\Omega = X\,(X'\,X)^{-1}X' \;\; X\,(X'\,X)^{-1}X' = X\,(X'\,X)^{-1}X'$
  - Sea $\vec{z} \in \Omega \Longleftrightarrow \vec{z} = X\,\vec{a}$

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

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

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

  - P proyecta cualquier vector de R^n ortogonalmente sobre Ω

Matriz de proyección ortogonal

- La matriz $Q_\Omega = I - P_\Omega$ también es simétrica e idempotente y proyecta ortogonalmente sobre $\Omega^\perp$. $rango(I - P) = n - p$
- La matriz $H = I - \vec{1}(\vec{1}\vec{1}')^{-1}\vec{1}' = I - \frac{1}{n}\vec{1}\vec{1}'$ se denomina matriz de centrado. $rango(H) = n - 1$

$$H\vec{x} = (I - \frac{1}{n}\vec{1}\vec{1}')\,\vec{x} = \vec{x} - \frac{1}{n}\vec{1}\vec{1}'\,\vec{x} = \vec{x} - \bar{x}\,\vec{1}$$

$$= \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \dots \\ x_n - \bar{x} \end{pmatrix}$$

Distribución normal p dimensional

Un vector aleatorio $\vec{x}$ con distribución Normal de media $\vec{\mu}$ y matriz de varianzas covarianzas Σ, $\Sigma > 0$, cumple:

- La función de densidad es:

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

- $\vec{y} = A\vec{x}$ sigue una distribución normal $N(A\vec{\mu}, A\Sigma A')$

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

- Si Σ es una matriz diagonal ($\sigma_{jk} = 0$ para $j \neq k$), entonces todos las componentes de $\vec{x}$ son independientes entre si.
- La distribución de x_j es $N(\mu_j, \sigma_{jj}), j = 1, ..., p$

Distribución normal p dimensional

- El cuadrado de una distribucion $N(0,1)$ sigue una distribución chi-cuadrado con un grado de libertad. $N(0,1)^2 \equiv \chi^2_1$
- Si $\vec{x} \equiv N_p(\vec{\mu}, \Sigma)$ entonces

$$(\vec{x} - \vec{\mu})' \, \Sigma^{-1} \, (\vec{x} - \vec{\mu}) \equiv \chi^2_p$$

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

$$(\vec{x} - \vec{\mu})' \, (\vec{x} - \vec{\mu}) \equiv \chi^2_p$$

- Si $\vec{x} \equiv N_p(\vec{\mu}, I)$ y H simétrica, idempotente con rango q, entonces

$$(\vec{x} - \vec{\mu})' \, H \, (\vec{x} - \vec{\mu}) \equiv \chi^2_q$$

t de Student y F de Fisher-Snedecor

- Dadas dos variables $X \equiv N(0,1)$ e $Y \equiv \chi^2_p$ independientes
$$t = \frac{X}{\sqrt{Y/p}}$$
sigue una distribución t de Student con p grados de libertad, t_p
- Dadas dos variables $X \equiv \chi^2_q$ e $Y \equiv \chi^2_p$ independientes
$$F = \frac{X/q}{Y/p} = \frac{p}{q}\frac{X}{Y}$$
sigue una distribución F con q, p grados de libertad, $F_{p,q}$

Estimación

Sea x una variable aleatoria con algún parámetro θ desconocido.

Objetivo: aproximar θ a partir de una muestra $(x_1, ..., x_n)$

Estimador: transforma la muestra en un valor del parámetro

$$T : (X_1, ..., X_n) \longrightarrow \Theta$$

Se desconoce $P(cara)$ en una moneda. Se lanza cien veces y se obtienen 52 caras.

$$\hat{P}(cara) = \frac{52}{100} = 0{,}52$$

Saber la estatura media de los habitantes de una ciudad. Se elige al azar una muestra de 1000 habitantes y se mide su estatura

$$\hat{\mu} = \frac{1}{1000} \sum_{i=1}^{1000} estatura_i = 1,78$$

Estimación

Métodos de estimación:

- Máxima verosimilitud: busca el valor del parámetro que hace máxima la verosimilitud de la muestra.

$$Max_{\theta \in \Theta} L(x_1, ..., x_n; \theta)$$

- Mínimos cuadrados: busca el valor de θ que minimiza el cuadrado de la distancia entre las muestras y valores teóricos

$$Min_{\theta \in \Theta} \sum_{i=1}^{n} (x_i - teor_i(\theta))^2$$

Propiedades deseables de un estimador:

- Ser centrado y tener varianza pequeña

$$E(T) = \theta; \quad Var(T) \simeq 0$$

- Tener un error cuadrático medio pequeño: $E(T - \theta)^2 \simeq 0$.