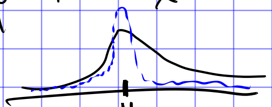
Inferencia sobre vector de medias

MLE: $\mu = \bar{X}$, $\Sigma = S_n$

Distr. muestral de $\bar{X}, S$

Si $X_1, \ldots, X_n$ es una muestra aleatoria de pob. normal con media μ y cov Σ. La distr. de

1) Si $p = 1$ $\quad \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$



; al crecer n la varianza decrece

en el caso general $\bar{X}$ tiene una distr. Normal multivariada

$$\bar{X} \sim N_p(\mu, \tfrac{1}{n}\Sigma)$$

Si $p = 1$ $\quad (n-1)S^2 = \sum_{j=1}^{n} (x_j - \bar{x})^2 \rightarrow$ distr. $\chi^2_{\underset{gl}{(n-1)}}$

$(n-1)S^2$ se distribuye como $\sigma^2 (\underbrace{Z_1}_{N(0,1)}{}^2 + \ldots + Z_{n-1}^2)$

en el caso general, la distr. de $(n-1)\overset{\rightarrow \text{matriz}}{S}$ es la distr. wishart con $(n-1)$ g.l.

$W_m(\cdot \mid \Sigma)$ es la distr. de $\sum_{i=1}^{m} \underbrace{Z_i}_{N(0,\Sigma)} Z_i'$

3) $\bar{X}$ y $\overset{\rightarrow \text{mat.}}{S}$ son indep.

es ser consistente

- $\bar{X}$ converge en probabilidad a μ
- S y S_n converge en prob. a Σ

Teorema del límite central → sean $X_1, \ldots, X_n$ observaciones indep. de una población con media μ y covarianza Σ. $\sqrt{n}(\bar{X} - \mu)$ es aprox. $N(0, \Sigma)$ si $n >> p$.

$$\bar{X} \approx N(\mu, \tfrac{1}{n}\Sigma)$$

No interesa la distr. de los X_i

obs: * usar S en lugar de Σ no afectará nuestros resultados seriamente.

* $n-p$ debe ser grande.

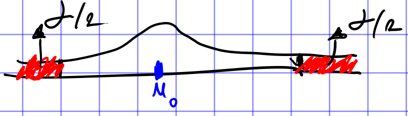
Inferencia sobre μ $\rightsquigarrow$ · pruebas de hipótesis
· Intervalos de confianza

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$$ $\rightsquigarrow$ se pueden hacer las pruebas de hipótesis por aparte pero es mejor hacerlo conjunto. porque es más potente.
↳ (Aumenta prob. de rechazar hipótesis nula).

Un recorderis...

P.H. respecto a una media poblacional μ_0

$$H_0: \mu = \mu_0 \qquad H_1: \mu \neq \mu_0$$
$$\leq$$



Z $\longrightarrow$ se usa cuando la muestra es grande y conozco σ^2.

T student $\rightarrow$ se usa cuando la muestra no es muy grande y cuando no conozco la var. poblacional σ^2.

$$t = \frac{(\bar{x} - \mu_0)}{\frac{s}{\sqrt{n}}}, \qquad s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2$$
$$\bar{x} = \frac{1}{n}\sum (x_i)$$

Si H_0 es cierta, t tiene una distr. t-student con $n-1$ g.l.

Rechazamos H_0 si $t \in RR$ i.e.
- $|t| > t_{\alpha/2}$ → 2 tail
- $t > t_\alpha$ → cola superior
- $t < -t_\alpha$ → cola inferior.

Rechazamos para $|t|$ grande $\equiv$ rechazar si t^2 es grande

$$t^2 = n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0)$$ → dist. estad. squared

Dada una muestra, se rechaza H_0 si:

$$n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0) > t^2_{n-1}(\alpha/2)$$

Si H_0 no se rechaza μ_0 es un valor posible para μ

<u>recuerde</u> recordemos que no rechazar H_0 a un nivel α es equiv. a que μ_0 esté en un I. Confi. con confianza $1-\alpha$.

- Valores de μ_0 en el I.C. son aquellos para los cuales no se rechaza $H_0: \mu = \mu_0$
- los limites de los intervalos de confianza son V.a's.
- probabilidad de que I.C. contenga a μ es $1-\alpha$ luego si tomo muchas muestras $(1-\alpha)$ (100)

Generalizando a p-variables

$$T^2 = (\bar{X} - \mu_0)' \left(\tfrac{1}{n} S\right)^{-1} (\bar{X} - \mu_0)$$
$$= n(\bar{X} - \mu_0)' (S^{-1})(\bar{X} - \mu_0)$$

donde $\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$, $S = \frac{1}{n-1}\sum (X_j - \bar{X})(X_j - \bar{X})'$

obs:) 1) T^2 se llama T^2 de Hotelling

2) $\frac{1}{n} S$ es la estimación de la covarianza de $\bar{X}$

3) si T^2 muy grande, $\bar{X}$ estará lejos de μ_0, luego se rechaza H_0.

4) en este caso, si H_0 es cierta

$$T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,n-p} \rightarrow \text{g.l.}$$

entonces

$$\alpha = P\left(T^2 > \overbrace{\frac{(n-1)}{n-p} \cdot F_{p,n-p}(\alpha)}^{\blacktriangle} \,\middle|\, H_0 \text{ cierta}\right)$$
$$= P\left(n(\bar{X} - \mu_0)'(S)^{-1}(\bar{X} - \mu_0) > \blacktriangle \,\middle|\, H_0 \text{ cierta}\right)$$

Si tenemos una P.H. de la forma

$$H_0: \mu = \mu_0 \quad , \quad H_1: \mu \neq \mu_0 \quad \text{nivel } \alpha.$$

rechazamos si

$$T^2 = n\underbrace{(\bar{X}-\mu_0)'}_{\substack{\text{Normal}\\ N_p(0,\Sigma)}}(S^{-1})(\bar{X}-\mu_0) > \frac{(n-1)}{n-p}\cdot F_{p,n-p}(\alpha)$$

ejemplo: → en el lb hay muchos ejemplos

Obs: T^2 es invariante bajo cambios de las unidades de las medidas de Y

$Y = C_{p\times p} X_{p\times 1} + d_{p\times 1} \rightarrow T^2$ es la misma para X e Y.

Tarea demostrar porqué

Ayuda: Dado que $\bar{Y} = C\bar{X}+d$, $S_Y = CS'C$

$$T^2 = n(\bar{Y}-\mu_{Y_0})(S_Y)^{-1}(\bar{Y}-\mu_{Y_0})$$

$$(\ldots)$$

$$= n(\bar{X}-\mu_0)(S^{-1})(\bar{X}-\mu_0)$$

Método de razón de verosimilitud (pruebas más potentes)

$$\max_{\mu,\Sigma} L(\mu,\Sigma) = \frac{1}{(2\pi)^{np/2}\cdot|\hat{\Sigma}|^{n/2}}\cdot e^{-np/2}$$

$$\hat{\Sigma} = \frac{1}{n}\sum(X_j-\bar{X})(X_j-\bar{X})'$$

$$\hat{\mu} = \bar{X}$$

Asumiendo $H_0: \mu = \mu_0$

$$L(\mu_0,\Sigma) = \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} e^{-\frac{1}{2}\sum_{j=1}^{n}(X_j-\mu_0)'\Sigma^{-1}(X_j-\mu_0)}$$

μ_0 fijo: Podemos variar Σ para encontrar valores más posibles con las obs se tiene

$$\max_{\Sigma} L(\mu_0,\Sigma) = \frac{1}{(2\pi)^{np/2}|\hat{\Sigma}_0|^{n/2}} e^{-np/2}$$

$$\hat{\Sigma}_0 = \frac{1}{n}\sum_j^{n}(X_j-\mu_0)(X_j-\mu_0)'$$

Para determinar si μ_0 es un valor posible de μ se comparan los máximos de $L(\mu_0, \Sigma)$ y $L(\mu, \Sigma)$

Nuevo estadístico: Λ (lambda de wilk)

$$\Lambda = \frac{\overbrace{\max_{\Sigma} L(\mu_0, \Sigma)}^{\text{max de todos los posibles } \Sigma}}{\underbrace{\max_{\mu, \Sigma} L(\mu, \Sigma)}_{\text{maximo de todos los posibles } \mu \text{ y } \Sigma}} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \right)^{n/2}$$

- Si Λ es pequeño $H_0: \mu = \mu_0$ es poco probable y H_0 se rechaza.

  (Si μ_0 está lejos de μ, $|\hat{\Sigma}_0|$ va a ser grande, más grande que $|\hat{\Sigma}|$)

- Más concretamente:

  H_0 se rechaza si $\Lambda < \underbrace{C_\alpha}_{\alpha \text{ percentil de la distribución de } \Lambda}$