Chp 2 : Les tests paramétriques

Test usuels.

I/ Comparaison ~~des~~ variances d'une V.a Gaussienne.

1/ ~~Tester H0 : σ = σ0~~

La comparaison des variances est un outil essentiel des statistiques, nous l'utiliserons intensivement en regression multiple et en analyse de la vari

1/ Comparaison d'une variance à une valeur déterministe

On veut comparer la variance obtenue à partir d'un échantillon que nous noterons $\hat{\sigma}^2$ à une valeur donnée (fixée) a priori noté σ_0^2

Test : $H_0 \ \sigma^2 = \sigma_0^2$ vs $H_1 : \sigma^2 \neq \sigma_0^2$

La regle de décision est : à un risque α fixé, on rejette H_0 si :

$$|\hat{\sigma}^2 - \sigma_0^2| > k_\alpha \quad \text{avec} \quad P_{H_0}(|\hat{\sigma}^2 - \sigma_0^2| > k_\alpha) = \alpha.$$

On considère l'estimateur sans biais de σ.

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2 \Rightarrow (n-1)\frac{\hat{\sigma}^2}{\sigma^2} \sim X^2_{n-1}$$

sous H_0 $(n-1)\frac{\hat{\sigma}^2}{\sigma_0^2} \sim X^2_{n-1}$

$$[|\hat{\sigma}^2 - \sigma_0^2| > k_\alpha] = \left[(n-1)\frac{\hat{\sigma}^2}{\sigma_0^2} > k'_{\alpha/2}\right]$$ avec k'_α donné tq

$$P(\chi^2_{n-1} > k_{\alpha/2}) = \alpha/2.$$

2/ Comparaison de deux variances de gaussiennes.

Soient $X_1^1, \ldots, X_{n_1}^1 \sim N(\mu_1, \sigma_1)$

$X_1^2, \ldots, X_{n_2}^2 \sim N(\mu_2, \sigma_2)$

Test : $H_0 : \sigma_1 = \sigma_2$ contre $H_1 : \sigma_1 \neq \sigma_2$

La regle de decision : à un risque α, on rejette H_0, si

$$|\hat{\sigma}_1 - \hat{\sigma}_2| > k_\alpha \quad \text{avec} \quad P(|\hat{\sigma}_1 - \hat{\sigma}_2| > k_\alpha) = \alpha$$

$$\hat{\sigma}_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_j^i - \bar{X}^i)^2$$

$$(n_1-1)\frac{\hat{\sigma}_1^2}{\sigma_1^2} \sim \chi^2_{n_1-1}$$

$$(n_2-1)\frac{\hat{\sigma}_2^2}{\sigma_2^2} \sim \chi^2_{n_2-1}$$

Sous H_0 $\sigma_1^2=\sigma_2^2 \Rightarrow \dfrac{(n_1-1)}{n_2-1}\dfrac{\hat{\sigma}_1^2/(n_1-1)}{\hat{\sigma}_2^2/(n_2-1)} = \dfrac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \sim F_{(n_1-1,n_2-1)}$

On rejette H_0 si $\left[\dfrac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} > k_{\alpha/2}\right]=W$

En effe $|\hat{\sigma}_1^2-\hat{\sigma}_2^2|>k_\alpha \Leftrightarrow \hat{\sigma}_1^2-\hat{\sigma}_2^2>k_\alpha$ ou $\hat{\sigma}_1^2-\hat{\sigma}_2^2<-k_\alpha$

$$\Rightarrow \frac{\hat{\sigma}_1^2}{\sigma_2^2}>k'_\alpha \quad \text{et} \quad \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}<-k'_\alpha$$

$$\alpha = P\left(F_{(n_1-1;n_2-1)}>k'_\alpha\right)+\underbrace{P\left(F_{(n_1-1,n_2-1)}<-k'_\alpha\right)}_{\alpha/2}$$

on trouve k'_α tq $P\left(F_{(n_1-1,n_2-1)}>k_\alpha\right)=\alpha/2$ $\quad k_\alpha=F^{-1}_{(n_1-1,n_2-1)}(\alpha/2)$

$$RC=W=\left[\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}>f^{\alpha/2}_{(n_1-1;n_2-1)}\right]$$

II | Comparaison portant sur les moyennes.

1 / Comparaison d'une moyenne à une valeur donnée

a/ On suppose que la variance de la population est connue.

$X_1,\dots,X_n$ un n-échantillon de $X \sim N(\mu,\sigma)$.

Test : $H_0: \mu=\mu_0$ vs $H_1: \mu\neq\mu_0$. test bilaterale

a/ On suppose que la variance de la population est σ_0^2 connue

La région ~~de decision~~ critique de test est ~~po~~ donnée par

$$W=\left\{x\in\mathbb{R}^n \,/\, \frac{|\bar{X}-\mu_0|}{\sqrt{\frac{\sigma_0^2}{n}}}>k_\alpha\right\} \text{ tq}$$

$$P_{H_0}(W)=1-F_{N(0,1)}(k_{\alpha/2})=\alpha.$$

Test unilateral : $H_0: \mu=\mu_0$; $H_1: \mu$

$RC=W=\left\{x\in\mathbb{R}^n \,/\, \dfrac{\bar{X}-\mu_0}{\sqrt{\frac{\sigma_0^2}{n}}}>k_\alpha\right.$

$P_{H_0}(N(0,1)>k_\alpha)=\alpha$

b/ On suppose la variance de la population est σ^2 inconnue.

Elle est donc estimée à partir de l'echantillon par $\hat{\sigma}^2$

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

Test bilateral : $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$.

La regle de decision : on rejette H_0 si $\underline{x} \in RC$.

$$RC = \left\{ \underline{x} \in \mathbb{R}^n \,/\, \frac{|\bar{X} - \mu_0|}{\sqrt{\frac{\hat{\sigma}_X^2}{n}}} \geq k_\alpha \right\}.$$

$$P_{H_0}(RC) = \alpha \iff P\left(\frac{\bar{X} - \mu_0}{\sqrt{\frac{\hat{\sigma}_X^2}{n}}} \geq k_\alpha \right) + P\left(\frac{\bar{X} - \mu_0}{\sqrt{\frac{\hat{\sigma}_X^2}{n}}} \leq -k_\alpha \right) = \alpha$$

T_{n-1}

$$P\left(\cancel{\mathcal{N}(0,1)} \geq k_\alpha \right) = \alpha/2 \qquad k_\alpha = F^{-1}_{\cancel{\mathcal{N}(0,1)}\, T_{n-1}}(1 - \alpha/2)$$

. Test unilateral $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$.

La regle de decision est de rejeter H_0 si $\underline{x} \in RC$

$$RC = W = \left\{ \underline{x} \in \mathbb{R}^n \,/\, \frac{\bar{X} - \mu_0}{\sqrt{\frac{\hat{\sigma}_0^2}{n}}} > k_\alpha \right\} \text{ tq}$$

$$P_{H_0}\left(\frac{\bar{X} - \mu_0}{\sqrt{\frac{\hat{\sigma}_0^2}{n}}} > k_\alpha \right) = P\left(T_{n-1} > k_\alpha \right) = \alpha.$$

$$k_\alpha = F^{-1}_{T_{n-1}}(1 - \alpha)$$

IV) Comparaison de deux moyennes.

Deux cas de figures se présente : soit les échantillons sont appariés, en d'autre terme les obs des deux échantillons sont realisés sur les mêmes inds, soit les deux échantillons sont indépendants

si les échantillons sont appariés, il faut calculer les moyennes de differences $Z_i = X_i - Y_i$. On se ramène au 1er cas.

Test : $H_0 : \mu_Z = 0$ vs $H_1 : \mu_Z \neq 0$ (ou $\mu_Z > 0$)

si les échantillons sont indépendants : Les variance σ_X et σ_Y sont egal (après test) ou $\sigma_X \neq \sigma_Y$.

a) $\sigma_X = \sigma_Y = \sigma$

Un est estimateur sans biais de σ^2 est $\hat{\sigma}^2 = \dfrac{(n-1)\hat{\sigma}_X^2 + (p-1)\hat{\sigma}_Y^2}{n+p-2}$

n taille de X et p taille de Y

le test de comparaison des moyennes est :

. test bilatéral : $H_0 : \mu_X - \mu_Y = \mu_0$ vs $H_1 : m_X - \mu_Y \neq \mu_0$.

la regle de decision, on rejette H_0 si

$$\left| \frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{1}{p}\right)}} \right| \geqslant t_{n+p-2}^{1-\alpha/2}$$

$$\bar{X} - \bar{Y} \sim N\left(\mu ; \sigma\left(\frac{1}{n} + \frac{1}{p}\right)\right) \quad U = \frac{\bar{X} - \bar{Y} - \mu}{\sigma\sqrt{\left(\frac{1}{n} + \frac{1}{p}\right)}}$$

$$(n-1)\hat{\sigma}_X^2 + (p-1)\hat{\sigma}_Y^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 + \sum_{i=1}^{p} (Y_i - \bar{Y})^2$$

$$V = \frac{(n-1)\hat{\sigma}_X^2 + (p-1)\hat{\sigma}_Y^2}{\sigma} \sim \chi^2_{n+p-2}$$

$$\frac{U}{\sqrt{V/_{n+p-2}}} = \frac{\bar{X} - \bar{Y} - \mu}{\frac{\cancel{(n+p-2)}\, \hat{\sigma} \sqrt{\frac{1}{n} + \frac{1}{p}}}{\cancel{n+p-2}}} = \frac{\bar{X} - \bar{Y} - \mu}{\hat{\sigma}^2 \sqrt{\frac{1}{n} + \frac{1}{p}}} \sim t_{n+p-2}$$

. Test unilateral : $H_0 : \mu_X = \mu_Y = \mu_0$ vs $H_1 : m_X - \mu_Y > \mu_0$

La regle de decision est de rejetter H_0 si

$$\frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{\hat{\sigma}^2 \left(\frac{1}{n} + \frac{1}{p}\right)}} \geqslant t_{n+p-2}^{1-\alpha}.$$

b – $\sigma_X \neq \sigma_Y$

si les variances des deux populations sont differents, on peut utiliser le test d'Aspin-Welch. Ce test est basé sur la statistique

$$S = \frac{\bar{X} - \bar{Y} - \mu_0}{\sqrt{\frac{\hat{\sigma}_X^2}{n} + \frac{\hat{\sigma}_Y^2}{p}}} \sim t_{d^\circ L}.$$

La regle de decision et la m^ que ds le cas d'egalités ds variance seulement le d°L et à changé !

$$dL = \frac{\left(\frac{\hat{\sigma}_X^2}{n} + \frac{\hat{\sigma}_Y^2}{p}\right)^2}{\left(\frac{\hat{\sigma}_X^2}{n}\right)^2 \Big/ (n-1) + \left(\frac{\hat{\sigma}_Y^2}{p}\right)^2 \Big/ p-1}$$

dL et tjs $\leq n+p-2$

diminue d°L, augment k_α