Exercice 1

jeudi 27 avril 2017 11:09

On note Y_{ij} le sexe du j-ème enfant de la i-ème famille

$$X_i = \sum_{i=1}^{5} Y_{ij}\,,\ (X_i)_{i=1,\dots,n} \text{ indépendants}$$

Si les $\left(Y_{ij}\right)_{i,j}$ sont indépendants alors :

$$\forall i, X_i \sim B\left(5, \frac{1}{2}\right) : p(X_i = j) = \binom{5}{j}\left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^{5-j} = \frac{5!}{j!\,(5-j)!}\frac{1}{2^5} = p_j$$

Test de H_0 : "$X_i \sim \mathcal{B}\left(5, \frac{1}{2}\right)$", H_1 : "X_i ne suit pas $\mathcal{B}\left(5, \frac{1}{2}\right)$"

$$W = \left\{\sum_{l=0}^{5} \frac{(N_l - np_l)^2}{np_l} > F^{-1}_{\chi^2_5}(1-\alpha)\right\}$$

$$\sum_{l=0}^{5} \frac{(N_l - np_l)^2}{np_l} \approx F^{-1}\left(1 - \underbrace{0{,}025}_{p\text{-valeur}}\right)$$

Aux seuils 5%, 2,5%, ... on décide H_1.

i	0	1	2	3	4	5
$\frac{(N_i - np_i)^2}{np_i}$	6.4	0.72	1	1.44	0.18	2.5

$N_i = 18 > np_i = 10$

$N_i = 15 > np_i = 10$

L'inadéquation à la loi $\mathcal{B}\left(5, \frac{1}{2}\right)$ provient d'une sur-représentation des familles à 0 ou 5 filles.

b) On a 813 filles, 787 garçons. Mais les variables $\left(Y_{ij}\right)_{\substack{i=1,\dots,320\\ j=1,\dots,5}}$ ne sont pas indépendantes. On ne peut pas utiliser le test H_0 : "$p = \frac{1}{2}$", H_1 : "$p \neq \frac{1}{2}$"

Exercice 2

jeudi 4 mai 2017 10:13

On veut tracer une droite des moindres carrés passant par l'origine : $Y_i = ax_i + \varepsilon_i$

Outils :

$$s_x^2 = \operatorname{Var} X = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}_n^2$$

$$\operatorname{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

$$\Rightarrow C_{xY} = \frac{1}{n}\sum_{i=1}^{n} x_i Y_i - \bar{x}_n \bar{Y}_n$$

On a :

$$\delta^2(\beta) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta x_i)^2$$

$$\Rightarrow \frac{d\delta^2}{d\beta}(\beta) = -\frac{1}{n}\sum_{i=1}^{n} 2x_i(y_i - \beta x_i) = \frac{-2}{n}\left(\sum_{i=1}^{n} x_i y_i - \beta \sum_{i=1}^{n} x_i^2\right)$$

$$\frac{d\delta^2}{d\beta}\left(\hat{\beta}_n\right) = 0 \Rightarrow \widehat{\beta_n} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^2} = \frac{C_{xy} + \bar{x}_n \bar{y}_n}{s_x^2 + (\bar{x}_n)^2}$$

$$E\left[\hat{\beta}_n\right] = \frac{1}{\sum_{j=1}^{n} {x_j}^2} E\left[\sum_{i=1}^{n} x_i Y_i\right] = \frac{1}{\sum_{j=1}^{n} x_k^2}\sum_{i=1}^{n} x_i \underbrace{E[Y_i]}_{\beta x_i}$$

$$= \frac{\beta \sum_{j=1}^{n} x_j^2}{\sum_{j=1}^{n} x_j^2} = \beta \Rightarrow \hat{\beta}_n \text{ sans biais}$$

$$\operatorname{Var}\hat{\beta}_n = \sum_{i=1}^{n} \frac{x_i^2}{\left(\sum_{j=1}^{n} x_j^2\right)^2} \underbrace{\operatorname{Var} Y_j}_{\sigma^2} = \frac{\sigma^2}{\frac{n}{n}\sum_{j=1}^{n} x_j^2} = \frac{\sigma^2}{n\overline{(x^2)_n}}$$

Les Y_j sont indépendantes, gaussiennes, donc :

$$\hat{\beta}_n \sim \mathcal{N}\left(\beta, \frac{\sigma^2}{n\overline{(x^2)_n}}\right)$$