Université Claude Bernard Lyon 1
L2 MASS UE 42
Année 2011-2012

Proba-Stat : fiche n° 1

Exercice 1. On construit les intervalles de confiance (bilatères) de la moyenne d'un échantillon gaussien (relevant de la loi normale) de taille n fixée et de variance inconnue, pour un risque $\alpha_1 = 5\%$ puis pour un risque $\alpha_2 = 1\%$. Lequel des deux intervalles obtenus est-il le plus large ?

De façon générale, dans quel cas rejette-t-on plus facilement l'hypothèse H_0 (alors qu'elle est vérifiée par la loi dont relève l'échantillon) : pour un risque de 5% ou pour un risque de 1% ?

Exercice 2. On donne $\mathbf{P}(X \geq 1.96) = 0.025$ si X relève de la loi normale centrée réduite.

1. Expliciter l'intervalle I de prédiction (ou de confiance) de niveau 0.95 de la moyenne d'un échantillon de taille 100 relevant de la loi $\mathcal{N}(2, 4)$.
2. On simule 10000 échantillons de taille 100 relevant de la loi normale $\mathcal{N}(2, 4)$. Pour combien de ces échantillons environ a-t-on $\bar{x} \in I$?

Exercice 3. Soient X et Y des variables aléatoires de carré intégrable. On note, pour tout $\lambda \in \mathbf{R}$

$$R(\lambda) = \mathbf{E}((\lambda|X| + |Y|)^2).$$

1. Vérifier que, pour tout λ, $R(\lambda) \geq 0$.
2. En déduire l'inégalité de Cauchy-Schwarz

$$\mathbf{E}(|XY|) \leq \sqrt{\mathbf{E}(X^2)\mathbf{E}(Y^2)} \qquad (*)$$

3. Montrer également que $|\mathbf{E}(XY)| \leq \sqrt{\mathbf{E}(X^2)\mathbf{E}(Y^2)}$.
4. On suppose pour cette question que $Y = 1$. Comment s'écrit alors l'inégalité $(*)$? Que cela signifie-t-il ?
5. Montrer que $|\mathrm{cov}(X, Y)| \leq \sqrt{\mathrm{var}\,(X)\mathrm{var}\,(Y)}$.
6. Montrer que $\mathrm{cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}(X))Y] = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$.

Exercice 4. On considère $n \geq 2$ points (x_i, y_i) et on utilise les notations usuelles pour les moyennes et variances empiriques :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\bar{v}_x = \frac{1}{n}\sum_{i=1}^{n} x_i - \bar{x}^2 \qquad \bar{v}_y = \frac{1}{n}\sum_{i=1}^{n} y_i - \bar{y}^2$$

On définit également la covariance empirique du nuage de points par

$$\overline{c_{xy}} = \frac{1}{n}\sum_{i=1}^{n}(x_i y_i - \bar{x}\bar{y})$$

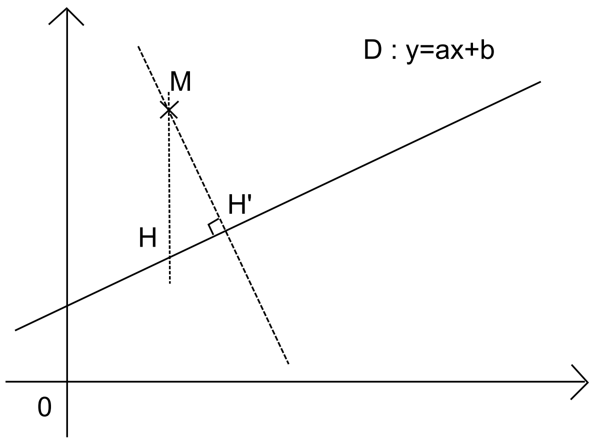
1. Montrer qu'en munissant l'ensemble $\{1, \ldots, n\}$ de la probabilité uniforme, $\overline{x}$ peut être vu comme l'espérance de la variable aléatoire $\tilde{x}$ définie pour tout $i \in \{1, \ldots, n\}$ par

$$\tilde{x}(i) = x_i.$$

2. Écrire de la même façon $\bar{y}$, $\bar{v}_x$, $\bar{v}_y$ et $\overline{c_{xy}}$ comme des espérances.
3. Réécrire dans ce cadre les égalités et inégalités obtenues à l'exercice précédent.

Remarque : $n\,\overline{c_{xy}}$ peut également s'écrire comme le produit scalaire sur $\mathbf{R}^n$ des vecteurs $u_x = (x_1 - \bar{x}, \ldots, x_n - \bar{x})$ et $u_y = (y_1 - \bar{y}, \ldots, y_n - \bar{y})$, et on a alors $n\bar{v}_x = \|u_x\|^2$ et $n\bar{v}_y = \|u_y\|^2$. L'inégalité de Cauchy-Schwarz se résume alors à $|\langle u_x, u_y\rangle| \leq \|u_x\|\|u_y\|$.

Exercice 5. On munit le plan $\mathbf{R}^2$ d'un repère orthonormé et on considère la droite D de ce plan, d'équation $y = ax + b$. On se donne un point M du plan, de coordonnées (x, y).



1. Graphiquement, comment détermine-t-on les paramètre a et b ?
2. Déterminer les coordonnées des points H et H' et les distances MH et MH'.