

FICHE TD 2 - SUR LES LOIS CONDITIONNELLES

Exercice 1 (régression logistique). On considère une population constituée de deux groupes homogènes. On note $X = 0$ si l'individu appartient au premier groupe et $X = 1$ si l'individu est dans l'autre groupe. On s'intéresse à une grandeur numérique Y (par exemple la taille) et on modélise sa variabilité par les deux lois conditionnelles suivantes

$$\mu_{Y|X=0} \text{ loi normale } N(\mu_0, \sigma_0^2) \text{ et } \mu_{Y|X=1} \text{ loi normale } N(\mu_1, \sigma_1^2).$$

Notons p la proportion d'individus dans le groupe $X = 1$. Déterminer les lois conditionnelles de X sachant Y .

Exercice 2. (sur le conditionnement dans le cas d'un vecteur bi-gaussien)

On considère ici un vecteur $X = (X_1, X_2)^T$ gaussien bidimensionnel. On note $\mu = (\mu_1, \mu_2)^T$ sa moyenne et Γ sa matrice de covariance que l'on suppose inversible. On notera encore σ_1^2 et σ_2^2 les variances de X_1, X_2 et ρ le coefficient de corrélation linéaire.

1. Déterminer la projection orthogonale de X_2 sur l'espace vectoriel engendré par la constante 1 et X_1 . On la notera $E_L(X_2 | X_1)$.
2. En écrivant $X_2 = E_L(X_2 | X_1) + \varepsilon$, montrer que le résidu ε est indépendant de X_1 . En déduire que $E_L(X_2 | X_1) = E(X_2 | X_1)$ puis les lois conditionnelles de X_2 sachant $X_1 = x_1$. Commentaires.
3. Supposons $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2$ et $\rho = 1/2$. Tracer sur un même graphique les courbes d'isoprobabilité du vecteur (X_1, X_2) , la moyenne conditionnelle de X_2 sachant X_1 et les quantiles conditionnels à 2.5% et 97.5%. Que remarque-t-on? Expliquer.
4. Que mesure le module de ρ ? Quelle valeur de ρ^2 doit-on avoir pour réduire de 90% l'écart-type de X_2 à partir de la connaissance de X_1 . Commentaires.
5. Dans le cas général où (X_1, X_2) est un vecteur continu mais pas forcément gaussien, a-t-on que la variance conditionnelle de X_2 sachant $X_1 = x_1$ est toujours plus petite que la variance de X_2 ? Contre-exemple? Que peut-on dire néanmoins?

Exercice 3. (sur un problème de filtrage...)

Le problème du filtrage est de restaurer au mieux le signal émis X à partir du signal reçu $Y = X + \varepsilon$, ce dernier étant entaché d'un bruit de transmission.

1. Expliquer en quoi $E(X|Y)$ est optimal. A-t-on que $E(X|Y) = Y$ vu la relation $X = Y - \varepsilon$?
2. On considère la relation de décomposition orthogonale $X = \beta Y + Z$. Faire un dessin et déterminer la valeur de β pour que $Y \perp Z$.
3. Donner la relation de décomposition de la variance correspondante. Calculer encore le pourcentage de la variance expliquée (restaurée) en fonction du rapport $\frac{\sigma_X^2}{\sigma^2}$ (ou rapport signal/bruit).
4. Déterminer l'ensemble des lois conditionnelles de X sachant $Y=y$. Commentaires.