MPSI - Fermat
2017-2018

Le 19.06.18

CONCOURS BLANC

Durée de l'épreuve : 4 heures
La calculatrice est interdite

Le devoir est composé d'un exercice et d'un problème de probabilité.
Lorsqu'une question est jugée, a priori, plus difficile, elle est précédée du symbole $(*)$ voire $(**)$.
La notation tiendra particulièrement compte de **la qualité de la rédaction, la précision des raisonnements et l'énoncé des formules utilisées**.

BON COURAGE

Problème

D. Optimisation pour des inconnues aléatoires

On considère deux variables aléatoires X et Y discrètes, définies sur un espace $(\Omega, \mathcal{A}, \mathbf{P})$ admettant chacun un moment d'ordre 2.
On se place sur l'espace préhilbertien réel $L^2(\Omega)$, muni du produit scalaire : $\langle Z_1, Z_2\rangle = \mathbf{E}(Z_1Z_2)$.
On cherche à mesurer l'éventuelle influence d'une variable aléatoire X sur une variable aléatoire Y.

1. On suppose dans cette première question que l'on peut décrire Y sous forme d'une fonction affine de X, notée $Z = aX + b$.
   On cherche à minimiser $f(a,b) = \|Y - Z\|^2 = \mathbf{E}((Y-Z)^2)$ avec $a, b \in \mathbb{R}$.
   On note $F = \text{vect}(1, X)$ sous-espace vectoriel de $L^2(\Omega)$.

   (a) Notons, pour commencer que $\dim F = 2$, car X et non constante donc $(1, X)$ libre et forme une base de F.
   Puis : $X^* = X - \mathbf{E}(X) = X - \mathbf{E}(X)1 \in F$.
   Ensuite, comme X^* est centré :

$$\langle X^*, 1\rangle = \mathbf{E}(X^*) = 0$$

   Donc $(1, X^*)$ est une famille orthogonale donc libre.
   Ainsi $(1, X^*)$ est une base (orthogonale) de F Enfin,

$$\|1\|^2 = \mathbf{E}(1) = 1 \qquad \|X^*\|^2 = \mathbf{E}((X^*)^2) = \mathbf{V}(X^*) + \mathbf{E}(X)^2 = 1 + 0 = 1$$

   Donc

$$\boxed{(1, X^*) \text{ est une base orthonormée de } F.}$$

   (b) D'après le cours, avec une base orthonormée de F :

$$\boxed{p_F^\perp : Y \mapsto \langle Y, 1\rangle 1 + \langle Y, X^*\rangle X^* = \mathbf{E}(Y) + \mathbf{E}(YX^*)X^*}$$

   Puis en remplaçant X^* par sa valeur : $\dfrac{X - \mathbf{E}(X)}{\sigma(X)}$;

$$\begin{aligned}
p_F^\perp(Y) &= \mathbf{E}(Y) + \frac{1}{\mathbf{V}(X)}\mathbf{E}(Y(X - \mathbf{E}(X))(X - \mathbf{E}(X)) \\
&= \mathbf{E}(Y) + \frac{1}{\mathbf{V}(X)}[\mathbf{E}(YX) - \mathbf{E}(X)\mathbf{E}(Y)](X - \mathbf{E}(X)) \\
&= \mathbf{E}(Y) + \frac{1}{\mathbf{V}(X)}\text{cov}(X,Y)(X - \mathbf{E}(X)) = \frac{\text{cov}(X,Y)}{\mathbf{V}(X)}X + \frac{\mathbf{E}(Y)\mathbf{V}(X) - \text{cov}(X,Y)\mathbf{E}(X)}{\mathbf{V}(X)} \\
&= \frac{\text{cov}(X,Y)}{\mathbf{V}(X)}X + \frac{\mathbf{E}(Y)(\mathbf{E}(X^2) - \mathbf{E}(X)^2) - (\mathbf{E}(XY)\mathbf{E}(X) - \mathbf{E}(X)^2\mathbf{E}(Y))}{\mathbf{V}(X)}
\end{aligned}$$

$$\boxed{p_F^\perp(Y) = \frac{\text{cov}(X,Y)}{\mathbf{V}(X)}X + \frac{\mathbf{E}(Y)\mathbf{E}(X^2) - \mathbf{E}(XY)\mathbf{E}(X)}{\mathbf{V}(X)}}$$

(c) On va tout simplement faire le calcul, mais à partir de la première formule.
Par linéarité de $\mathbf{E}$, comme $\mathbf{E}(X^*)=0$:

$$\mathbf{E}(p_F^\perp(Y)) = \mathbf{E}\left(\mathbf{E}(Y)+\mathbf{E}(YX^*)X^*\right) = \mathbf{E}(Y)+\mathbf{E}(YX^*)\mathbf{E}(X^*) = \mathbf{E}(Y)$$

Puis, par linéarité et orthogonalité de $Y-p_F^\perp(Y)$ avec $p_F^\perp(Y)$

$$\begin{aligned}\langle Y-p_F^\perp(Y), p_F^\perp(Y)-\mathbf{E}(Y)\rangle &= \langle Y-p_F^\perp(Y), p_F^\perp(Y)\rangle - \langle Y-p_F^\perp(Y), \mathbf{E}(Y)\rangle \\ &= 0+\mathbf{E}\big((Y-p_F^\perp(Y))\mathbf{E}(Y)\big) = \mathbf{E}(Y)[\mathbf{E}(Y)-\mathbf{E}(p_F^\perp(Y))] = 0\end{aligned}$$

car $\mathbf{E}(p_F^\perp(Y))=\mathbf{E}(Y)$ On a donc montrer :

$$\boxed{\mathbf{E}(p_F^\perp(Y))=\mathbf{E}(Y) \text{ et } \langle Y-p_F^\perp(Y), p_F^\perp(Y)-\mathbf{E}(Y)\rangle = 0}$$

(d) En fait, il s'agit d'un problème de minimisation sur d'une distance à un sous-espace vectoriel.
Cela se résout avec une projection orthogonale sur F. Comme $f(a,b)=\|Y-(aX+b)\|^2$,

$$\begin{aligned}\inf_{a,b\in\mathbb{R}} f(a,b) &= \inf_{Z\in F}\|Y-Z\|^2 = \|Y-p_F^\perp(Y)\|^2 = \langle Y-p_F^\perp(Y), Y-p_F^\perp(Y)\rangle \\ &= \langle Y-p_F^\perp(Y), Y-\mathbf{E}(Y)\rangle - \langle Y-p_F^\perp(Y), p_F^\perp(Y)-\mathbf{E}(Y)\rangle \\ &= \langle Y, Y-\mathbf{E}(Y)\rangle - \langle p_F^\perp(Y), Y-\mathbf{E}(Y)\rangle - 0 \\ &= \langle Y, Y-\mathbf{E}(Y)\rangle - \langle p_F^\perp(Y), Y-\mathbf{E}(Y)\rangle - \langle p_F^\perp(Y), Y-p_F^\perp(Y)\rangle \\ &= \mathbf{E}(Y^2)-\mathbf{E}(Y)^2 - \langle p_F^\perp(Y), Y\rangle + \langle p_F^\perp(Y), \mathbf{E}(p_F^\perp(Y)\rangle\end{aligned}$$

par multiple linéarité.

$$= \|Y\|^2 - \|p_F^\perp(Y)\|^2$$

La dernière égalité est la conséquence du théorème de Pythagore.

d'après le En déduire que $\min_{a,b\in\mathbb{R}} f(a,b) = \mathbf{V}(Y)-\mathbf{V}(p_F^\perp(Y))$.
Le coefficient de corrélation de (X,Y) est donné par $\rho(X,Y)=\dfrac{\mathrm{cov}(X,Y)}{\sigma(X)\sigma(Y)}$,

$$\mathbf{V}(p_F^\perp(Y)) = \mathbf{E}\big((p_F^\perp(Y)-\mathbf{E}(p_F^\perp(Y)))^2\big) = \mathbf{E}\big((p_F^\perp(Y)-\mathbf{E}(Y))^2\big)$$

Or on a vu que $p_F^\perp(Y)=\mathbf{E}(Y)+\mathbf{E}(YX^*)X^*$, donc

$$\begin{aligned}\mathbf{V}(p_F^\perp(Y)) &= \mathbf{E}\big((\mathbf{E}(YX^*)X^*)^2\big) = [\mathbf{E}(YX^*)]^2-\mathbf{E}(X^{*2}) \\ &= \frac{1}{\sigma(X)^2}\big([\mathbf{E}(Y(X-\mathbf{E}(X)))]^2-\mathbf{E}((X-\mathbf{E}(X))^2)\big)\end{aligned}$$

Conclure que $\min_{a,b\in\mathbb{R}} f(a,b) = (1-\rho^2(X,Y))\mathbf{V}(Y)$,
où $\rho(X,Y)$ est le coefficient de corrélation de X et Y.

2. On suppose toujours $X,Y\in L^2(\Omega)$.
On suppose que X est une variable aléatoire discrète connue.
En revanche, Y nous est connue que par la valeur de son espérance sous condition $X=x$, pour tout $x\in X(\Omega)$:

$$\forall\, x\in X(\Omega), \qquad \mathbf{E}(Y|X=x) := \sum_{y\in Y(\Omega)} y\mathbf{P}_{X=x}(Y=y)$$

On cherche donc Z, une variable aléatoire définie à partir de X, la plus proche possible de Y.
On suppose que $Z\in\mathrm{vect}\big((\mathbf{1}_{X=x})_{x\in X(\Omega)}\big)$ et $\|Y-Z\|$ minimal.

(a) Montrer que la famille $\big((\frac{\mathbf{1}_{X=x}}{\sqrt{\mathbf{P}(X=x)}})_{x\in X(\Omega)}\big)$ est une famille orthonormale.

(b) Calculer $\langle Y, \frac{\mathbf{1}_{X=x}}{\sqrt{\mathbf{P}(X=x)}}\rangle$.
On suppose donc que $Z=\displaystyle\sum_{x\in X(\Omega)}\mathbf{E}(Y|X=x)\mathbf{1}_{X=x}$.

(c) Montrer que $\mathbf{E}(Z)=\mathbf{E}(Y)$.
On admettra qu'on peut appliquer le lemme de Fubini ici : $\displaystyle\sum_{x\in X(\Omega)}\left(\sum_{y\in Y(\Omega)} a_{x,y}\right) = \sum_{y\in Y(\Omega)}\left(\sum_{x\in X(\Omega)} a_{x,y}\right)$

(d) On reprend les notations de la partie B.
On considère $X=S_1(=X_1)$ et $Y=X_2$.
   i. Montrer que $\forall\, k\leqslant n$, $\mathbf{E}(Y|X=k)=\mathbf{E}(X_2|S_1=k)=(n-k)p$
   ii. Exprimer alors la variable aléatoire Z et calculer $\mathbf{E}(Z)$.
   iii. Retrouver alors $\mathbf{E}(X_2)$ de deux façons (avec Z et S_2-S_1)

La première méthode présentée ici correspond, en statistiques, à la recherche par la méthode des moindres carrées.
La seconde méthode correspond à une définition naturelle de l'espérance conditionnelle : meilleure approximation de Y sur l'espace des combinaisons linéaires des indicatrices constituant X.