ENSC 1A - Examen Stat. Descriptive - Janvier 2015

Exercice 1 :

① a) $\bar{x}_n = \frac{1}{n} \sum_i x_i = \frac{590}{60} \simeq 9,83$

b) $\sigma^2_{n,x} = \frac{1}{n} \sum_i x_i^2 - (\bar{x}_n)^2 \simeq \frac{6112}{60} - (9,83)^2 \simeq 5,24$

$\sigma_{n,x} = \sqrt{\sigma^2_{n,x}} \simeq \sqrt{5,24} \simeq 2,29$

② a) $\bar{y}_n = \frac{774,1}{60} \simeq 12,9$

b) $\sigma^2_{n,y} = \frac{10\,200}{60} - (12,9)^2 \simeq 3,59$ $\qquad \sigma_{n,y} = \sqrt{3,59} \simeq 1,89$

③ $c_n = \frac{1}{n} \sum_i x_i y_i - \bar{x}_n \bar{y}_n = \frac{7859}{60} - 9,83 \times 12,9 = 4,18$

④ $r_n = \frac{c_n}{\sigma_{n,x}\ \sigma_{n,y}} = \frac{4,18}{2,29 \times 1,89} \simeq 0,966$

Ce coef de corrélation linéaire est positif et très proche de 1, il existe donc une très bonne liaison linéaire croissante entre X et Y

⑤ a) $\beta_n = \frac{c_n}{\sigma^2_{n,x}} = \frac{4,18}{5,24} = 0,80$

$\alpha = \bar{y}_n - \beta_n \bar{x}_n = 12,9 - 0,80 \times 9,83 \simeq 5,04$

la droite de régression a pour équation : $Y = 5,04 + 0,80X$

b) Rendement attendu pour $X = 10$: $\hat{Y} = 5,04 + 0,80 \times 10 = 13,04$

Exercice 2 :

① $\bar{y}_n = \frac{n_I}{n} \bar{y}_I + \frac{n_N}{n} \bar{y}_N + \frac{n_R}{n} \bar{y}_R$ $\qquad$ où $n = n_I + n_N + n_R = 200$

$= \frac{30}{200} \times 75 + \frac{70}{200} \times 65 + \frac{100}{200} \times 60$

$= 64$

② $\sigma^2_{n,E} = \sum_k \frac{n_k}{n} (\bar{y}_k - \bar{y}_n)^2 = \frac{30}{200} (75-64)^2 + \frac{70}{200} (65-64)^2 + \frac{100}{200} (60-64)^2$

$= 26,5$

$$\sigma^2_{n,R} = \sum_k \frac{n_k}{n}\, \sigma^2_{n,k} = \frac{30}{200} \times 40 + \frac{70}{200} \times 30 + \frac{100}{200} \times 35$$

$$= 34$$

(3) $\sigma^2_{n,y} = \sigma^2_{n,E} + \sigma^2_{n,R} = 26,5 + 34 = 60,5$

(4) $s^*_{y|x} = \sqrt{\dfrac{\sigma^2_{n,E}}{\sigma^2_{n,y}}} = \sqrt{\dfrac{26,5}{60,5}} \simeq 0,66$

La valeur de ce coef n'est pas très élevée (proche de 0,7), ce qui indique qu'il n'y a pas un lien ~~très~~ fort entre le niveau de stress et le type de compétitions, même si une légère dépendance croissante est à noter.

Exercice 3

(1) (a) Profil ligne pour la région Nord

Produit	P_1	P_2	P_3
fréq (%)	$\frac{120}{410} \simeq 29,3$	$\frac{210}{410} \simeq 51,2$	$\frac{80}{410} \simeq 19,5\%$

(b) Profil-ligne pour la région Sud

Produit	P_1	P_2	P_3
fréq (%)	$\frac{70}{270} \simeq 25,9$	$\frac{140}{270} \simeq 51,9$	$\frac{60}{270} \simeq 22,2\%$

(c) Les 2 profils-lignes semblent très proches, ce qui laisse à penser que la région n'aurait pas d'impact sur les produits achetés.

(2) (a) Eff th $n^{th}_{ij} = \dfrac{n_{i\cdot}\, n_{\cdot j}}{n}$

R	P_1	P_2	P_3	$n_{i\cdot}$
Nord	$\frac{410 \times 190}{680} = 114,6$	$\frac{410 \times 350}{680} = 211$	84,4	410
Sud	75,4	139	55,6	270
$n_{\cdot j}$	190	350	140	$n = 680$

$$\chi^2_n = \sum_i \sum_j \frac{(n_{ij} - n_{ij}^{th})^2}{n_{ij}^{th}}$$

$$= \frac{(120-114,6)^2}{114,6} + \dots + \frac{(60-55,6)^2}{55,6} \simeq 1,23$$

(b) On a donc $\chi^2_n <$ seuil $= 5,99$.

Il semble donc y avoir indépendance entre la région et les produits achetés.

Exercice 4 :

$$r_n(Z,U) = \frac{C_n(Z,U)}{\sigma_{n,Z} \times \sigma_{n,U}}$$

or $\sigma_{n,Z} = \sigma_{n,aX+b} = |a|\, \sigma_{n,X}$

$\sigma_{n,U} = \sigma_{n,cY+d} = |c|\, \sigma_{n,Y}$

$C_n(aX+b, cY+d) = ac\, C_n(X,Y)$

d'où $$r_n(Z,U) = \frac{ac\, C_n(X,Y)}{|a|\,\sigma_{n,X}\ |b|\,\sigma_{n,Y}} = \frac{a}{|a|}\frac{c}{|c|}\, r_n(X,Y)$$

$$= \pm\, r_n(X,Y)$$

$\oplus$ si a et c sont de même signe

$\ominus$ si a et c sont de signes contraires.