Statistiques Appliquées (L3)

Groupe n° 6 - Novembre 2006

Contrôle continu N°1 - Corrigé

Exercice 1 : Vrai ou Faux.

1. Faux, il peut y avoir plusieurs interprétations causales pour la même observation, "Corrélation n'est pas causalité".

2. Vrai cf table.

3. Faux, ceci n'est vrai que pour deux lois normales.

4. Faux cf exercice 3 TD1.

5. Vrai, l'inverse est fausse.

6. Faux, la convergence ne vaut que si $np(1-p) > 15$.

Exercice 2 : Sortie de route:

Démarche générale :

1) **Poser le problème**, deux routes, = deux durées D_1 et D_2, dans le premier cas, on cherche la probabilité que ces durées soient inférieures à 28.

$P(D_1 < 28)$ et $P(D_2 < 28)$

2) Quelle est la **loi de ces variables** et est ce que j'en connais les valeurs?

$D_1 \sim N(27, 2, 5^2)$, je centre et je réduis : $\frac{D_1-27}{2,5} \sim N(0,1)$ ainsi, je peux lire les valeurs dans la table

Je réécris les inégalités et les probabilités :

$P(D_1 < 28) = P(\frac{D_1-27}{2,5} < \frac{28-27}{2,5})$

3) **Lecture de la table** de la loi normale standard :

$P(T < 0,4) = 0,665$

Correction :

Soit D_i la variable aléatoire durée du trajet en minutes par la route n°i, $i = 1, 2$.

$\frac{D_1-27}{2,5} \sim N(0,1)$ et $\frac{D2-29}{1} \sim N(0,1)$

1.

$P(D_1 < 28) = P(\frac{D_1-27}{2,5} < \frac{28-27}{2,5}) = P(T < 0,4) = 0,665$

$P(D_2 < 28) = P(\frac{D_2-29}{1} < \frac{28-29}{1}) = P(T < -1) = 0,159$

Par la route n°1, il a 66% de chances d'arriver à l'heure et seulement 16% par la route n°2, il choisit donc la route n°1.

2.

$P(D_1 < 32) = P(\frac{D_1-27}{2,5} < \frac{32-27}{2,5}) = P(T < 2) = 0,9772$

$P(D_2 < 32) = P(\frac{D_2-29}{1} < \frac{32-29}{1}) = P(T < 3) = 0,9987$

Il choisit la route n°2.

Exercice 3 : Densités conjointes et marginales.

1. Par définition l'intégrale d'une densité sur toutes les valeurs possibles du domaine de définition, est égale à 1 :

$$\int_0^2 \int_0^1 kx^2ydxdy = 1$$

$$k\int_0^2 y\left(\int_0^1 x^2dx\right)dy = 1$$

$$k\int_0^2 y\left[\frac{x^3}{3}\right]_0^1 dy = 1$$

$$\frac{k}{3}\left[\frac{y^2}{2}\right]_0^2 = 1$$

$$k = \frac{3}{2}$$

2. Densités marginales :

$$f(x) = \int_0^2 kx^2ydy = \frac{3}{2}x^2 \times \frac{4}{2} = 3x^2$$

$$f(y) = \int_0^1 kx^2ydx = \frac{3}{2}y \times \frac{1}{3} = \frac{y}{2}$$

3. Indépendance (Rq : le produit et non la somme des densités!) :

$f(x) \times f(y) = 3x^2 \times \frac{y}{2} = \frac{3}{2}x^2y = f(x,y)$

Les deux variables sont donc indépendantes.

4.

$$E(X) = \int_0^1 xf(x)dx = \frac{3}{4}$$

$E(Y) = \int_0^2 yf(y)dx = \frac{4}{3}$

Les deux variables sont indépendantes la covariance est nulle.

Exercice 4 : Variable aléatoire et convergence (3 points)

Définition de la convergence en probabilité :

$\forall \varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0$

On pose : $X_n - X = Z_n$ alors :

$P(|X_n - X| > \varepsilon) = P(|Z_n| > \varepsilon) = P(|Z_n - E(Z_n)| > \varepsilon)$

Par l'inégalité de Bienaymé-Tchebychev (pas d'abréviation, si possible)

$P(|Z_n - E(Z_n)| > \varepsilon) < \frac{1}{(n^3-1)\varepsilon^2}$

On passe à la limite (en n'oubliant pas de dire qu'une probabilité est nécessairement **positive**!):

$0 \leq \lim_{n\to\infty} P(|Z_n - E(Z_n)| > \varepsilon) < \lim_{n\to\infty} \frac{1}{(n^3-1)\varepsilon^2}$

alors :

$\lim_{n\to\infty} P(|Z_n - E(Z_n)| > \varepsilon)$=0

Exercice 5 : Bayes : cf cours

La définition de la probabilité conditionnelle n'était pas suffisante.

Le cas considérant un ensemble complet d'évènement comprenant un évènement et son contraire était un peu particulier.

Exercice 6 :

1. Estimateur de m : $\overline{X} = \frac{\sum X_i}{n}$, sans biais, convergent, efficace.

$\overline{X} = 8773,3$ heures

2. On cherche un intervalle dans lequel "l'estimateur se situe à 95%"

a) On pose le problème :

$P(|\overline{X}| < u_0) = 0,95$

b) On ne peut connaître pas la valeur u_0 , $\overline{X}$ n'est pas une variable aléatoire tabulée, elle n'est pas centrée réduite. En effet,

on montre par le théorème central limite que :

$\overline{X} \sim N(m, \frac{\sigma^2}{n})$

c) on centre et réduit cette variable : $Y = \sqrt{n}\left(\frac{\overline{X}-m}{\sigma}\right) \sim N(0,1)$

On introduit Y dans la probabilité en retranchant m et en divisant par $\frac{\sigma}{\sqrt{n}}$ des deux côtés:
$P\left(\left|\sqrt{n}\left(\frac{\overline{X}-m}{\sigma}\right)\right| < \sqrt{n}\left(\frac{u_0-m}{\sigma}\right)\right) = 0,95$

d) en lisant la table de la loi normale standard on obtient que :
$\sqrt{n}\left(\frac{u_0-m}{\sigma}\right) = 1,96$
d'où :
$P\left(\left|\sqrt{n}\left(\frac{\overline{X}-m}{\sigma}\right)\right| < 1,96\right) = 0,95$

en remplaçant m, n et σ par leur valeurs on obtient que :
$P\left(8716 < \overline{X} < 8830\right) = 0,95$

Intervalle de confiance à 95% pour la moyenne de survie des ampoules : [8716 ; 8830]
3. Une durée de vie de 9 000 heures est incompatible avec l'intervalle de confiance.