Corrigé de l'examen de Septembre 2005

Exercice 1.

1. Le calcul donne

$$B = A^2 - 6A + 10\,\mathrm{Id} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \qquad AB = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} = 4\,\mathrm{Id}\,.$$

D'où la matrice inverse $A^{-1} = \frac{1}{4}B$:

$$A^{-1} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Exercice 2.

1. On commence par établir le tableau des effectifs par classe :

$[2200, 2600[$	$[2600, 3000[$	$[3000, 3400[$	$[3400, 3800[$	$[3800, 4200[$
$n_1 = 3$	$n_2 = 5$	$n_3 = 6$	$n_4 = 8$	$n_5 = 2$

Le total des effectifs est égal à $n = 24$. On prend comme poids moyen de chaque classe, le milieu de celles-ci, c'est-à-dire $\xi_1 = 2400$, $\xi_2 = 2800$, $\xi_3 = 3200$, $\xi_4 = 3600$ et $\xi_5 = 4000$. Le poids moyen m d'un bébé est donc égale à

$$m = \frac{n_1}{n}\xi_1 + \cdots + \frac{n_5}{n}\xi_5 = \frac{3}{24}2400 + \frac{5}{24}2800 + \cdots = 3217 \text{ gr}$$

2. La courbe des poids cumulés est donnée par la figure 1.
3. La classe médiane des poids est obtenu en tracant la droite horizontale à l'ordonnée égale au milieu des effectifs, soit 12. La classe médiane est donc $[3000, 3400[$. Le poids médian $q_{50\%}$ se calcule par la formule

$$q_{50\%} = 3000 + (3400 - 3000)\frac{12 - 8}{14 - 8} = 3267 \text{ gr.}$$

Exercice 3.

1. L'estimateur de la proportion théorique est donné par $\hat{p} = \frac{k}{n}$, où k désigne le nombre de conducteurs ayant subi des blessures graves. L'intervalle de confiance est donné par

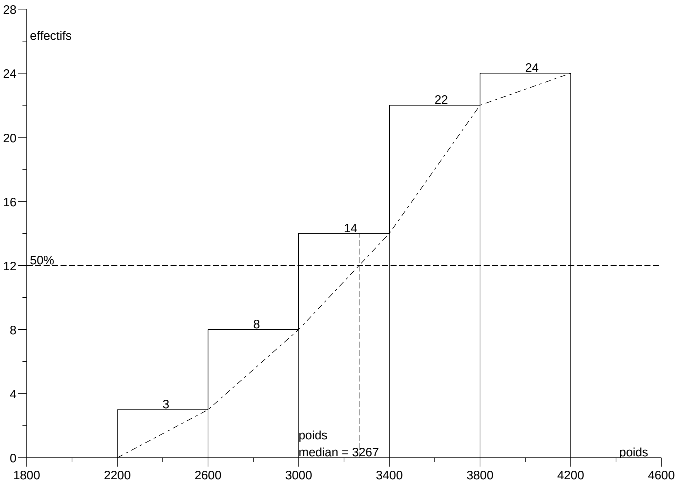

FIG. 1 – Courbe des poids cumulés

$$I = \left]\hat{p} - z_\alpha\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_\alpha\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right[$$

z_α est calculé par $\mathbb{P}(|\mathcal{N}(0,1)| > z_\alpha) = 5\% = \alpha$. La table donne $z_\alpha = 1.96$; et on obtient $\hat{p} = 5\ 10^{-3}$ et l'intervalle

$$I = \left]0.6\ 10^{-3},\ 9.4\ 10^{-3}\right[$$

2. On note p_0, resp. p, la proportion théorique des conducteurs gravement blessés et portant une ceinture, resp. ne portant pas de ceniture. On note $\hat{p}$ la proportion estimée dans le cas où ils ne portent pas de ceinture. On réalise un test de proportion. L'hypothèse $H_0 = \{p = p_0\}$ désigne le cas où le port de ceinture ne modifie pas sensiblement la proportion de blessés graves. La question précédente suggère de prendre $p_0 = 0.005$. On note $H_1 = \{p \neq p_0\}$ l'hypothèse complémentaire démontrant l'efficacité du port de la ceinture. La région critique de ce test est donnée par

$$C_\alpha = \left\{|\hat{p} - p_0| > z_\alpha\sqrt{\frac{p_0(1-p_0)}{n}},\right\}$$

où z_α est calculé par $\mathbb{P}(|\mathcal{N}(0,1)| > z_\alpha) = 0.001 = \alpha$. On trouve dans la table $z_\alpha = 3.291$. La proportion estimée est égale à $\hat{p} = 0.013$. On trouve

$$\left|\frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}\right| = 3.586.$$

On rejette donc l'hypothèse H_0 : le port de la ceinture est nécessaire pour diminuer le nombre des blessés graves.

Exercice 4.

1. Chaque câble d'acier est une v.a. X_i de même loi mais inconnue. On s'intéresse ici à estimer sa moyenne μ en supposant connu son écart-type $\sigma_0 = 25$ kg. On mesure $\bar{x} = \frac{1}{n}(X_1 + \cdots + X_{100})$ sur un échantillon de taille $n = 100$. Asymptotiquement $\sqrt{n}(\bar{x}-\mu)/\sigma_0$ est une variable normale $\mathcal{N}(0,1)$. On obtient donc l'intervalle de confiance

$$I_\alpha = \left]\bar{x} - z_\alpha \frac{\sigma_0}{\sqrt{n}},\ \bar{x} + z_\alpha \frac{\sigma_0}{\sqrt{n}}\right[$$

   où z_α est calculé à partir de la loi normale. Numériquement, $\alpha = 5\%$, $z_\alpha = 1.96$, d'où un intervalle $I_\alpha =]2625 \text{ kg}, 2635 \text{ kg}[$.

2. Dans le cas où l'écart-type σ est aussi une inconnue, $\sqrt{n}(\bar{x}-\mu)/s_{n-1}$ est asymptotiquement une variable de Student. On obtient maintenant comme intervalle de confiance

$$I_\alpha = \left\{\bar{x} - t_\alpha \frac{s_{n-1}}{\sqrt{n}},\ \bar{x} + t_\alpha \frac{s_{n-1}}{\sqrt{n}}\right\}$$

   où t_α est calculé à partir de la loi de Student.

   (a) Numériquement, $\alpha = 5\%$, $\bar{x} = 2660$ kg, $s_{n-1} = 20$ kg et $n = 30$. La table de Student nous donne (en prenant 30 au lieu de $n-1 = 29$) $t_\alpha = 2.042$. On obtient alors l'intervalle $I_\alpha =]2652 \text{ kg}, 2668 \text{ kg}[$.

   (b) Pour un même écart-type observé s_{n-1}, la résistance moyenne sera connue à une erreur près $\pm t_\alpha s_{n-1}/\sqrt{n}$. Numériquement, pour $\alpha = 1\%$, on calcule successivement l'erreur pour différentes valeurs de n dans le tableau suivant

$\boldsymbol{n}$	30	35	40	45	50
$\boldsymbol{t_\alpha}$	2.750	2.724	2.704	2.690	2.678
erreur	10.04	9.21	8.55	8.02	7.57

   On obtient alors une taille minimale environ égale à $n = 45$.

Exerice 5. On cherche à réaliser un test d'indépendance entre deux caractères : la persistance des unions et le lieu de résidence. On introduit donc l'hypothèse nulle

$$H_0 = \{ \text{ La persistance est indépendante du lieu } \}$$

et on rejetera ce test si le critère suivant sera réalisé

$$D^2 = \sum_{i,j} \frac{(n_{i,j} - n_{i,*}n_{*,j}/n)^2}{n_{i,*}n_{*,j}/n} > q_{1-\alpha}$$

où n est le nombre total d'unions et $q_{1-\alpha}$ est le quantile de la loi du chi-deux à $r = (2-1)(3-1) = 2$ degrés de liberté. Le calcule de D^2 se fait en remplissant le tableau suivant

Union	**rurale**	**moyenne**	**grande**	
mariés	430 413	1686 1644	3121 3180	$n_{1,*} = 5237$
séparés	27 44	133 175	397 338	$n_{2,*} = 557$
	$n_{*,1} = 457$	$n_{*,2} = 1819$	$n_{3,*} = 3518$	$n = 5794$

On trouve alors

$$D^2 = \frac{(430-413)^2}{413} + \frac{(1686-1644)^2}{1644} + \cdots = 29.62 \simeq 30.$$

Par ailleurs, pour $\alpha = 5\%$, $q_{1-\alpha} = 5.991$. le fait que $D^2 >> q_{1-\alpha}$ montre que les deux caractères sont fortement corrélés, qu'il faut rejeter H_0. On constate donc que le lieu géographique d'un ménage peu influencer le taux de divorce après 5 ans de mariage.