Corrigé de l'examen de Mai 2006

Exercice 1.

1. Par substitution, on obtient

$$x = y - m - 2, \quad 2my = m^2 + m.$$

Ou bien $m = 0$: la deuxième équation est toujours réalisée et il reste donc comme ensemble de solutions, une droite affine

$$\mathcal{S} = \{(x, x+2) \mid \ x \text{ quelconque } \}.$$

Ou bien $m \neq 0$: la deuxième puis la première équation admette une unique solution en y puis en x. La solution est unique et est égale à

$$x = -\frac{m+3}{2}, \quad y = \frac{m+1}{2}.$$

2. On vérifie que

$$A^2 = \begin{bmatrix} m & m \\ m^2 - m & m^2 + 3m \end{bmatrix}$$

puis que $A^2 + mA - 2m\,\mathrm{Id} = 0$. On constate alors que $2m\,\mathrm{Id} = A(A + m\,\mathrm{Id})$. Si $m = 0$, on retrouve le fait que A ne peut pas être inversible. Si $m \neq 0$, on obtient

$$A^{-1} = \frac{1}{2m}(A + m\,\mathrm{Id}) = \begin{bmatrix} \frac{1+m}{2m} & -\frac{1}{2m} \\ \frac{1-m}{2m} & -\frac{1}{2m} \end{bmatrix}$$

Exercice 2.

1. On se reportera à la figure 1.
2. Idem.
3. La hauteur moyenne est donnée par la formule $h_{moyen} = \frac{1}{n}\sum_{i=1}^{r} n_i \xi_i$, où ξ_i est la milieu d'une classe et n_i est l'effectif de cette classe. On trouve

$$h_{moyen} = \frac{1}{50}(2 \times 150 + 3 \times 165 + \cdots) = 185.6 \text{ cm}.$$

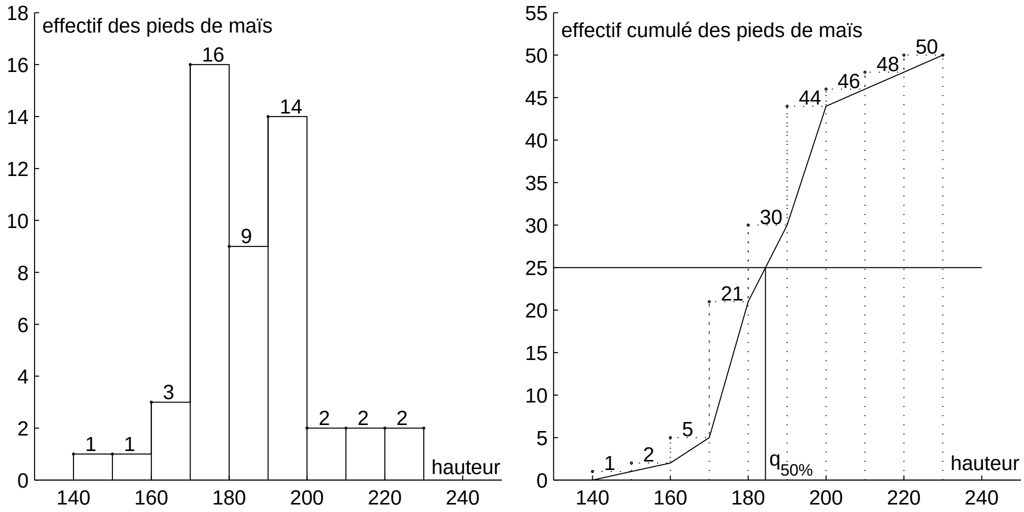

FIG. 1 – Histogramme et courbes des effectifs cumulés

4. La hauteur médiane est donnée par la formule

$$h_{median} = \xi_i + (\xi_{i+1} - \xi_i)\frac{N/2 - N_i}{N_{i+1} - N_i}$$

où $[\xi_i, \xi_{i+1}]$ représente la classe médiane et N_{i+1} l'effectif cumulé jusquà cette classe. On trouve

$$h_{median} = 180 + (190 - 180)\frac{25 - 21}{30 - 21} = 184.4.$$

Exercice 3. On note X_i la quantité du produit réglementé du ième paquet de lessive et n le nombre de paquets étudiés. On définit les estimateurs classiques de moyenne et de variance :

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2_{n-1} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

L'intervalle de confiance de μ est alors donné par $\mu = \bar{X} \pm t_\alpha \frac{S_{n-1}}{\sqrt{n}}$ où t_α est donné par la table de la loi de Student pour un degré de liberté $ddl = n - 1$. On trouve numériquement

$\mathbf{n} = 10$	$\bar{\mathbf{x}} = 54$	$\mathbf{ddl} = 9$
$\alpha = 0.05$	$\mathbf{s_{n-1}} = 4$	$\mathbf{t}_\alpha = 2.262$
$\mathbf{51 \leq \mu \leq 57}$		

Exercice 4.

1. Comme pour tout test d'hypothèse, la démarche générale doit être soigneusement décrite.
   (a) Il s'agit d'un test d'hypothèse de proportion.
   (b) L'hypothèse nulle est H_0 = « plus de 50% des étudiants sont contre la reprise ». On pourrait prendre un test bilatéral et les calculs suivants seraient différents mais il parait plus naturel de se demander si plus de 50% (ou moins de 50%) des étudiants sont contre la reprise.
   (c) On note $\hat{p}$, l'estimateur donnant la proportion d'étudiants contre la reprise. On appelle $p_0 = 50\%$, une proportion fixée à l'avance qu'on cherche à dépasser.
   (d) On choisit comme ensemble de rejet,

   $$R = \left\{ \sqrt{n}\frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)}} < q_\alpha \right\} = \left\{ \hat{p} < p_0 - q_{1-\alpha}\sqrt{\frac{p_0(1-p_0)}{n}} \right\},$$

   où $q_\alpha = -q_{1-\alpha}$ est le quantile d'ordre α de la loi normale centrée réduite. Les tables donnent plutôt les probabilités de dépassement en valeur absolue de la loi normale, soit $q_\alpha = -z_{2\alpha}$.
   (e) Calcul numérique

$n = 2507$	$p_0 = 0.5$	$\alpha = 0.05$
$q_{1-\alpha} = 1.645$	$p_{obs} = 0.46$	
$p_0 - q_{1-\alpha}\left(\frac{p_0(1-p_0)}{n}\right)^{1/2} = 0.48 > 0.46$		

   (f) Conclusion : On rejette H_0 ; moins de 50% des étudiants sont contre la reprise à 5% près d'erreur.
2. La p-valeur de ce test est le plus petit risque qu'on prend en rejetant H_0 à tort au vu des résultats. C'est donc la plus petite probabilité α telle que $q_\alpha > \sqrt{n}\frac{p_{obs}-p_0}{\sqrt{p_0(1-p_0)}}$, c'est-à-dire la probabilité

   $$\alpha = \mathbb{P}\left(N(0,1) < \sqrt{n}\frac{p_{obs} - p_0}{\sqrt{p_0(1-p_0)}} \right).$$

   On obtient numériquement $\alpha = 3\ 10^{-5}$.

Exercice 5.

1. A la première génération, le génotype de chaque tulipe est RJ. Le génotype de la tulipe obtenue par croisement est formé en prenant au hasard et sans distinction une allèle R ou J du génotype de chaque parent 1 ou 2 :

$1 \setminus 2$	**R**	**J**
R	RR	RJ
J	JR	JJ

Les génotypes de la tulipe « fille », RR, RJ, JR et JJ, arrivent avec probabilité $\frac{1}{4}$. Comme la tulipe ne se souvient pas de quel parent, chaque allèle peut provenir, JR et RJ représentent le même génotype. On obtient ainsi une distribution de génotypes : $p_{RR} = \frac{1}{4}$, $p_{RJ} = \frac{1}{2}$ et $p_{JJ} = \frac{1}{4}$.

2. La mise en œvre du test est la suivante :
   (a) Il s'agit d'un test d'hypothèse d'ajustement.
   (b) L'hypothèse nulle est H_0 = « la distribution des génotypes est conforme aux prévisions théoriques ».
   (c) La variable de décision est $D = \sum_{i=1}^{r}(n_i - np_i^0)^2/np_i^0$ où $r = 3$, n est la taille de l'effectif, i est un indice parcourant (RR,RJ,JJ), où $(p_1^0, p_2^0, p_3^0) = (p_{RR}, p_{RJ}, p_{JJ})$ et $(n_1, n_2, n_3) = (n_{RR}, n_{RJ}, n_{JJ})$ représentent les quantités de tulipes observées pour chaque génotype.
   (d) L'ensemble de rejet est $R = \{D > q_{1-\alpha}\}$, où q_β est le quantile d'ordre β de la distribution du chi-deux à $ddl = r - 1$ degrés de liberté.
   (e) Calcul numérique :

$n = 108$	$\alpha = 0.05$	$ddl = 2$	$q_{1-\alpha} = 5.992$
$D = 0.5 < 5.992$			

   (f) Conclusion : On accepte H_0 ; les résultats de l'agriculteur confirment le modèle théorique.