

## Corrigé de l'examen de Mai 2005

**Exercice 1.** En réalisant successivement les opérations de ligne  $L_1 \rightarrow L_1$ ,  $L_2 \rightarrow L_2 - 2L_1$  et  $L_3 \rightarrow L_3 + L_1$ , on obtient

$$\begin{cases} x - 5y - z = -2 \\ \quad 9y + 4z = 5 \\ \quad -3y \quad = -3 \end{cases} \Rightarrow \begin{cases} x = 2 \\ y = 1 \\ z = -1 \end{cases}$$

**Exercice 2.**

1. L'histogramme des effectifs est donné par la figure 1

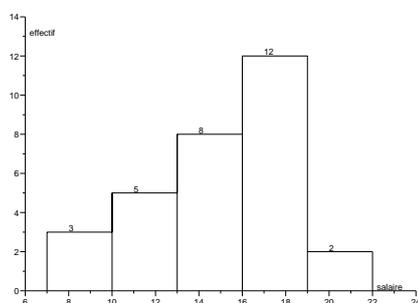


FIG. 1 – Histogramme des effectifs

2. La courbe des effectifs cumulés est donné par 2

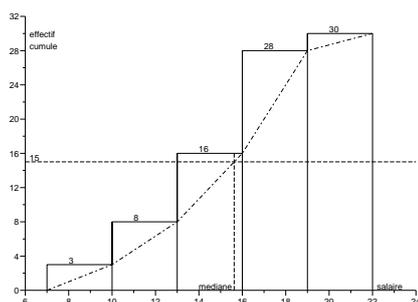


FIG. 2 – Courbe des effectifs cumulés

3. Le salaire médian est le plus grand salaire qu'un employé sur deux au moins reçoit. La classe médiane est la classe de salaires contenant le salaire médian ; elle est égale à  $[13, 16[$ . Le salaire médian vaut donc

$$\text{médiane} = 13 + (16 - 13) \frac{15 - 8}{16 - 8} = 15.6$$

### Exercice 3.

1. La figure 3 rassemble les 9 points  $(x_i, y_i)$  et la droite de régression

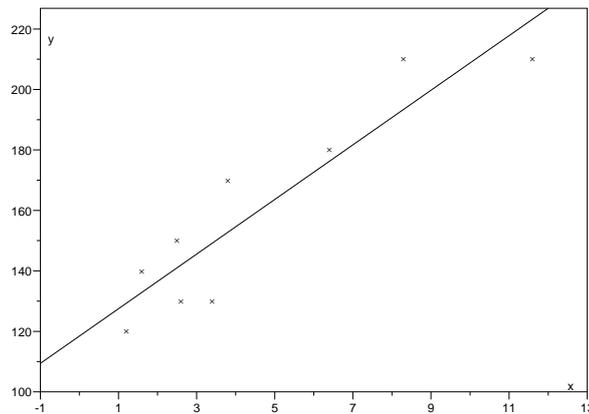


FIG. 3 – Droite de régression

2. La droite de régression a pour équation  $y = ax + b$  avec

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad b = \bar{y} - a\bar{x}$$

où

$$\begin{aligned} \bar{x} &= \frac{1}{9}(11.6 + 3.8 + \dots + 2.6 + 8.3) = 4.6 \\ \bar{y} &= \frac{1}{9}(210 + 170 + \dots + 130 + 210) = 160 \\ a &= 9.03, \quad b = 118.45 \end{aligned}$$

3. Pour  $x = 0$ , on trouve  $y = b = 118.45$ ; pour  $x = 5$ , on trouve  $y = 5a + b = 163.61$ .

### Exercice 4.

1. On réalise un test d'hypothèse sur la moyenne d'une série de dosages, connaissant a priori l'écart-type  $\sigma_0 = 0.05$  mg/litre. On note  $\mu_0 = 4.00$  mg/litre la concentration moyenne espérée. L'hypothèse nulle et la région critique de ce test sont données par

$$H_0 = \{\mu = \mu_0\}, \quad C_\alpha = \left\{ |\bar{x} - \mu_0| > z_\alpha \frac{\sigma_0}{\sqrt{n}} \right\}$$

où  $z_\alpha$  est la valeur telle que  $\mathbb{P}(|\mathcal{N}(0,1)| > z_\alpha) = \alpha$ . On propose un niveau d'erreur  $\alpha = 5\%$ ; on a donc  $z_\alpha = 1.96$ ,  $n = 5$ ,  $\bar{x} = 4.02$  et

$$\sqrt{n} \left| \frac{\bar{x} - \mu_0}{\sigma_0} \right| = 0.89 < 1.96.$$

On ne peut donc pas rejeter  $H_0$  : les résultats sont compatibles avec une moyenne égale à  $\mu_0$ .

2. L'écart-type n'est plus connu a priori ; on utilise alors la loi de Student pour déterminer un intervalle de confiance  $I_\alpha$  de  $\mu$  :

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

où  $s^2 = \frac{1}{n-1}((x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2)$  et  $t_\alpha$  est choisi tel que

$$\mathbb{P}(|\mathcal{T}(n-1)| > t_\alpha) = \alpha.$$

On choisit  $\alpha = 5\%$ ,  $n = 6$ , le degré de liberté est égal à  $n - 1 = 5$ , d'où  $t_\alpha = 2.571$ . De plus

$$\begin{aligned} \bar{x} &= \frac{1}{6}(2.97 + 3.01 + 2.98 + 2.94 + 3.03 + 2.95) = 2.98 \\ s^2 &= \frac{1}{5}((2.97 - 2.98)^2 + (3.01 - 2.98)^2 + \dots) = 0.0346 \end{aligned}$$

d'où l'intervalle de confiance  $\{2.94 < \mu < 3.02\}$ .

**Exercice 5.** On introduit deux v.a.  $X$  et  $Y$ , la première prenant deux valeurs (« atteint », « non atteint ») du cancer broncho-pulmonaire, la seconde prenant trois valeurs ( $A, B, C$ ). On cherche à savoir si c'est deux variables sont corrélées ou indépendantes. On réalise alors un test d'indépendance en prenant comme hypothèse nulle  $H_0 = \{X \text{ et } Y \text{ sont indépendantes}\}$ , c'est-à-dire  $H_0 = \{\text{l'usage du tabac ne favorise pas l'apparition du cancer}\}$ . La région critique est définie par

$$C_\alpha = \{D^2 > q_{1-\alpha}\}, \quad D^2 = \sum_i \sum_j \frac{(n_{ij} - n_{*j}n_{i*}/n)^2}{n_{*j}n_{i*}/n}.$$

$D^2$  suit une loi du chi-deux à  $ddl = (2 - 1)(3 - 1) = 2$  degrés de liberté. Pour  $\alpha = 5\%$ ,  $q_{1-\alpha} = 5.991$ . Le calcul du  $D^2$  se fait en remplissant le tableau suivant

<b>Personnes</b>	<b>A</b>	<b>B</b>	<b>C</b>	
<b>Atteintes</b>	25 77.5	243 272	334 252.5	602
<b>Non atteintes</b>	130 77.5	301 272	171 252.5	602
	155	544	505	1204

Par exemple  $77.5 = 155 * 602/1204$ . On trouve alors  $D^2 = 129.92$ . Comme  $129.92 \gg 5.991$ , on rejette  $H_0$  : l'usage du tabac favorise l'apparition du cancer.