Correction des Travaux dirigés de Probabilités et Statistiques - Statistiques descriptives

Julian Tugaut*

*Si vous trouvez des erreurs de français ou de mathématiques ou bien si vous avez des questions et/ou des suggestions, envoyez-moi un mail à julian.tugaut@univ-st-etienne.fr

Table des matières

Page de garde **1**

Table des matières **3**

Exercice 84 **5**
Énoncé . . . 5
Correction . . . 5
Correction du **1)**, **2)**, **4)** et **5)** simultanément . . . 5
Correction du **3)** et **6)** simultanément . . . 6
Correction du **7)** . . . 7
Correction du **8)** . . . 7
Correction du **9)** . . . 7
Correction du **10)** . . . 7

Exercice 85 **9**
Énoncé . . . 9
Correction . . . 9
Correction du **1)**, **3)** et **4)** simultanément . . . 9
Correction du **2)** et **5)** simultanément . . . 9
Correction du **6)** . . . 11
Correction du **7)** . . . 11
Correction du **8)** . . . 11

Exercice 86 **13**
Énoncé . . . 13
Correction . . . 13
Correction du **1)** . . . 13
Correction du **2)** . . . 16
Correction du **3)** . . . 16
Correction du **4)** . . . 17
Correction du **5)** . . . 18

Exercice 87 **19**
Énoncé . . . 19
Correction . . . 19
Remarque . . . 21

Exercice 84

Énoncé

À la suite d'une interrogation portant sur 40 étudiants, on a relevé les notes suivantes :

$$14, 14, 6, 14, 11, 10, 15, 7, 14, 9, 9, 13, 12, 9, 10, 15,$$
$$8, 10, 16, 8, 12, 9, 7, 11, 10, 12, 11, 16, 11, 10, 10, 9,$$
$$7, 10, 13, 10, 9, 9, 16, 12\,.$$

1. Présenter cette série sous la forme d'une série à valeurs isolées.
2. Déterminer les fréquences
3. Construire le diagramme en bâtons des effectifs.
4. Déterminer les effectifs cumulés.
5. Déterminer les fréquences cumulées
6. Construire le diagramme en bâtons des effectifs cumulés.
7. Calculer la moyenne arithmétique.
8. Déterminer graphiquement la médiane, le premier quartile et le troisième quartile.
9. Déterminer le(s) mode(s).
10. Déterminer la variance et l'écart-type.

Correction

Correction du 1), 2), 4) et 5) simultanément

Voici directement le tableau :

Valeurs isolées	6	7	8	9	10	11	12	13	14	15	16
Effectifs	1	3	2	7	8	4	4	2	4	2	3
Fréquences	$\frac{1}{40}$	$\frac{3}{40}$	$\frac{2}{40}$	$\frac{7}{40}$	$\frac{8}{40}$	$\frac{4}{40}$	$\frac{4}{40}$	$\frac{2}{40}$	$\frac{4}{40}$	$\frac{2}{40}$	$\frac{3}{40}$
Effectifs cumulés	1	4	6	13	21	25	29	31	35	37	40
Fréquences cumulées	$\frac{1}{40}$	$\frac{4}{40}$	$\frac{6}{40}$	$\frac{13}{40}$	$\frac{21}{40}$	$\frac{25}{40}$	$\frac{29}{40}$	$\frac{31}{40}$	$\frac{35}{40}$	$\frac{37}{40}$	1

Correction du 3) et 6) simultanément

Figure 1 – Effectifs

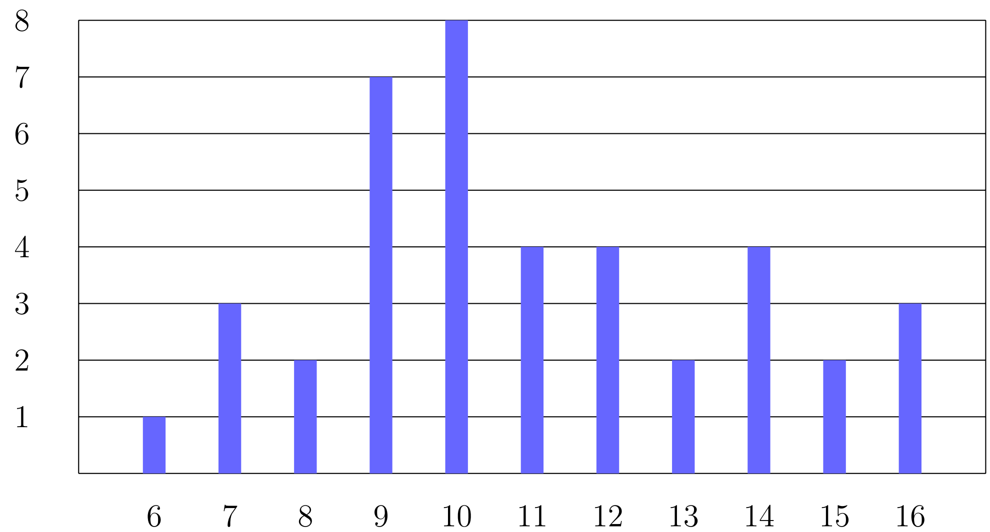
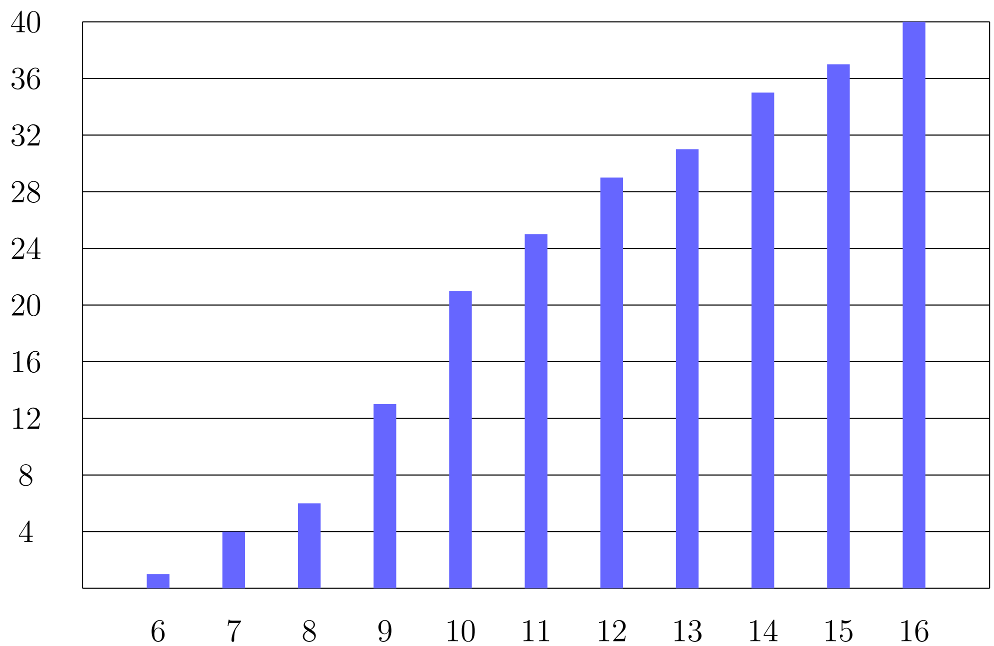


Figure 2 – Effectifs cumulés



Correction du 7)

La moyenne arithmétique vaut par définition :

$$\begin{aligned}\overline{x} &= \frac{1}{40}\times\Big\{6\times 1+7\times 3+8\times 2+9\times 7+10\times 8+11\times 4\\ &+ 12\times 4+13\times 2+14\times 4+15\times 2+16\times 3\Big\}\\ &= \frac{438}{40}=10.95\,.\end{aligned}$$

Correction du 8)

En traçant la droite d'équation $y=\frac{40}{2}=20$ sur le diagramme en bâton des effectifs cumulés, on voit que la médiane correspond à $q_{\frac{1}{2}}(x)=10$ puisque c'est la première note dont les effectifs cumulés sont en contact avec cette droite. De même, le premier quartile est $q_{\frac{1}{4}}(x)=9$ et le troisième quartile est $q_{\frac{3}{4}}(x)=13$.

Correction du 9)

Cette série statistique a trois modes : 7, 10 et 14 comme l'indique le diagramme en bâtons des effectifs.

Correction du 10)

Par définition, la variance de x est :

$$\begin{aligned}s_x^2 &= \overline{x^2}-\overline{x}^2\\ &= \frac{1}{40}\times\Big\{6^2\times 1+7^2\times 3+8^2\times 2+9^2\times 7+10^2\times 8+11^2\times 4\\ &+ 12^2\times 4+13^2\times 2+14^2\times 4+15^2\times 2+16^2\times 3\Big\}-\frac{438^2}{40^2}\\ &= \frac{5\,078\times 40-191\,844}{1600}=\frac{11\,276}{1600}\\ &\approx 7.05\,.\end{aligned}$$

Puis, $s_x\approx 2.65$.

Exercice 85

Énoncé

Dans un ensemble de 53 710 familles de huit enfants, on a observé le nombre de garçons par famille :

Nombre de garçons	Nombre de familles
0	215
1	1 485
2	5 331
3	10 649
4	14 959
5	11 959
6	6 678
7	2 092
8	342

1. Déterminer les fréquences.

2. Construire le diagramme en bâtons des effectifs.

3. Déterminer les effectifs cumulés.

4. Déterminer les fréquences cumulées.

5. Construire le diagramme en bâtons des effectifs cumulés.

6. Calculer la moyenne arithmétique.

7. Déterminer graphiquement la médiane, le premier quartile et le troisième quartile.

8. Déterminer la variance et l'écart-type.

Correction

Correction du 1), 3) et 4) simultanément

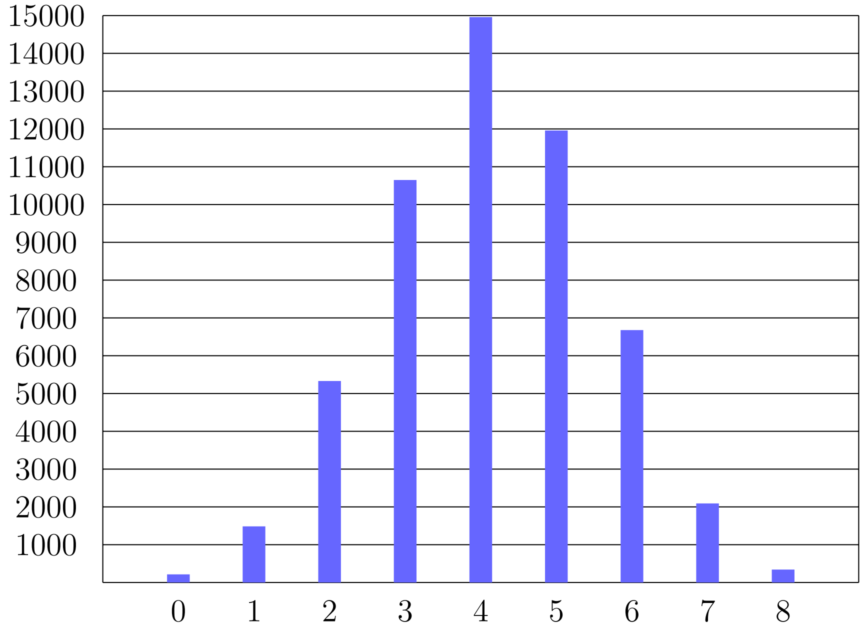
Voici directement le tableau :

i	0	1	2	3	4	5	6	7	8
n_i	215	1 485	5 331	10 649	14 959	11 959	6 678	2 092	342
f_i	$\frac{215}{53\,710}$	$\frac{1\,485}{53\,710}$	$\frac{5\,331}{53\,710}$	$\frac{10\,649}{53\,710}$	$\frac{14\,959}{53\,710}$	$\frac{11\,959}{53\,710}$	$\frac{6\,678}{53\,710}$	$\frac{2\,092}{53\,710}$	$\frac{342}{53\,710}$
N_i	215	1 700	7 031	17 680	32 639	44 598	51 276	53 368	53 710
F_i	$\frac{215}{53\,710}$	$\frac{1\,700}{53\,710}$	$\frac{7\,031}{53\,710}$	$\frac{17\,680}{53\,710}$	$\frac{32\,639}{53\,710}$	$\frac{44\,598}{53\,710}$	$\frac{51\,276}{53\,710}$	$\frac{53\,368}{53\,710}$	$\frac{53\,710}{53\,710}$

Correction du 2) et 5) simultanément

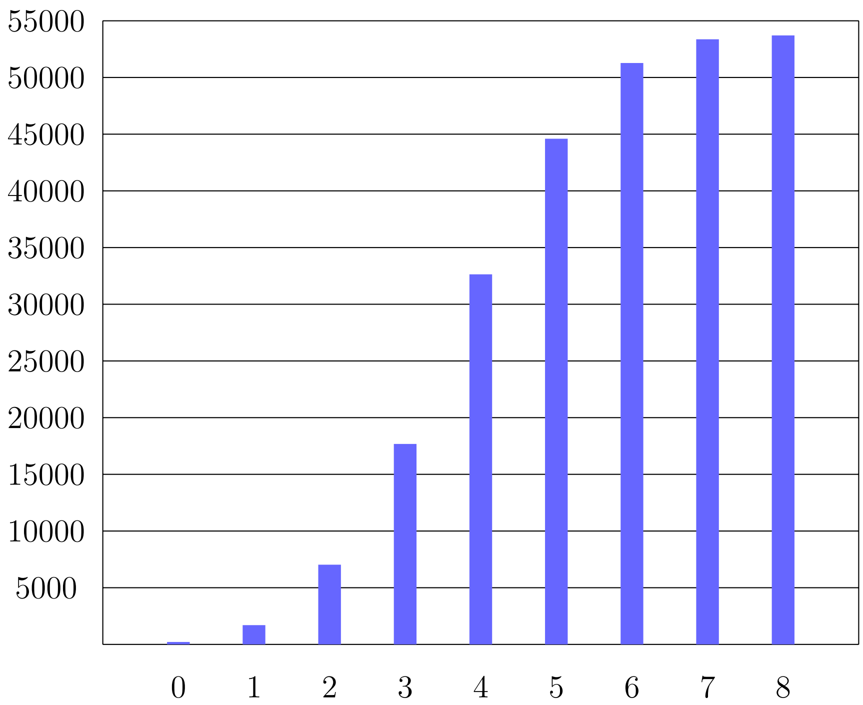
Voici les deux diagrammes :

Figure 3 – Effectifs



et

Figure 4 – Effectifs cumulés



Correction du 6)

La moyenne arithmétique vaut par définition :

$$\begin{aligned}\overline{x} &= \frac{1}{53\,710} \times \Big\{ 0 \times 215 + 1 \times 1\,485 + 2 \times 5\,331 + 3 \times 10\,649 + 4 \times 14\,959 \\ &+ 5 \times 11\,959 + 6 \times 6\,678 + 7 \times 2\,092 + 8 \times 342 \Big\} = \frac{221\,173}{53\,710} \approx 4.12\,.\end{aligned}$$

Correction du 7)

En traçant la droite d'équation $y = \frac{53\,710}{2} = 26855$ sur le diagramme en bâton des effectifs cumulés, on voit que la médiane correspond à $q_{\frac{1}{2}}(x) = 4$ puisque c'est la première note dont les effectifs cumulés sont en contact avec cette droite. De même, le premier quartile est $q_{\frac{1}{4}}(x) = 3$ et le troisième quartile est $q_{\frac{3}{4}}(x) = 5$.

Correction du 8)

Par définition, la variance de x est :

$$\begin{aligned}s_x^2 &= \overline{x^2} - \overline{x}^2 \\ &= \frac{1}{53\,710} \times \Big\{ 0^2 \times 215 + 1^2 \times 1\,485 + 2^2 \times 5\,331 + 3^2 \times 10\,649 \\ &+ 4^2 \times 14\,959 + 5^2 \times 11\,959 + 6^2 \times 6\,678 + 7^2 \times 2\,092 + 8^2 \times 342 \Big\} \\ &- \frac{221\,173^2}{53\,710^2} \\ &= \frac{1\,021\,773 \times 53\,710 - 48\,917\,495\,929}{2\,884\,764\,100} = \frac{5\,961\,931\,901}{2\,884\,764\,100} \\ &\approx 2.07\,.\end{aligned}$$

Puis, $s_x = \sqrt{s_x^2} \approx 1.44$.

Exercice 86

Énoncé

On a mesuré la taille (x) et le poids (y) de 300 adultes extraits par un tirage au hasard non exhaustif d'une population. Les données sont présentées dans le tableau de contingence suivant :

TABLE 1 – Tableau de contingence (poids/taille)

Taille (cm) / Poids (kg)	$[150; 160[$	$[160; 170[$	$[170; 180[$	$[180; 190[$	$[190; 200[$
$[45; 55[$	2	1	0	0	0
$[55; 65[$	7	8	4	2	0
$[65; 75[$	5	15	22	7	1
$[75; 85[$	2	12	63	19	5
$[85; 95[$	0	7	28	32	12
$[95; 105[$	0	2	10	20	7
$[105; 115[$	0	0	1	4	2

1) *Construire une représentation graphique des effectifs de chaque série marginale.*

2) *Calculer la moyenne et la variance de chaque série marginale.*

3) *Construire la courbe des moyennes conditionnelles du caractère y par rapport au caractère x dans cet échantillon.*

4) *Déterminer l'équation de la droite de régression au sens des moindres carrés du caractère y par rapport au caractère x dans cet échantillon.*

5) *Calculer le coefficient de corrélation linéaire des caractères x et y dans cet échantillon.*

Correction

Correction du 1)

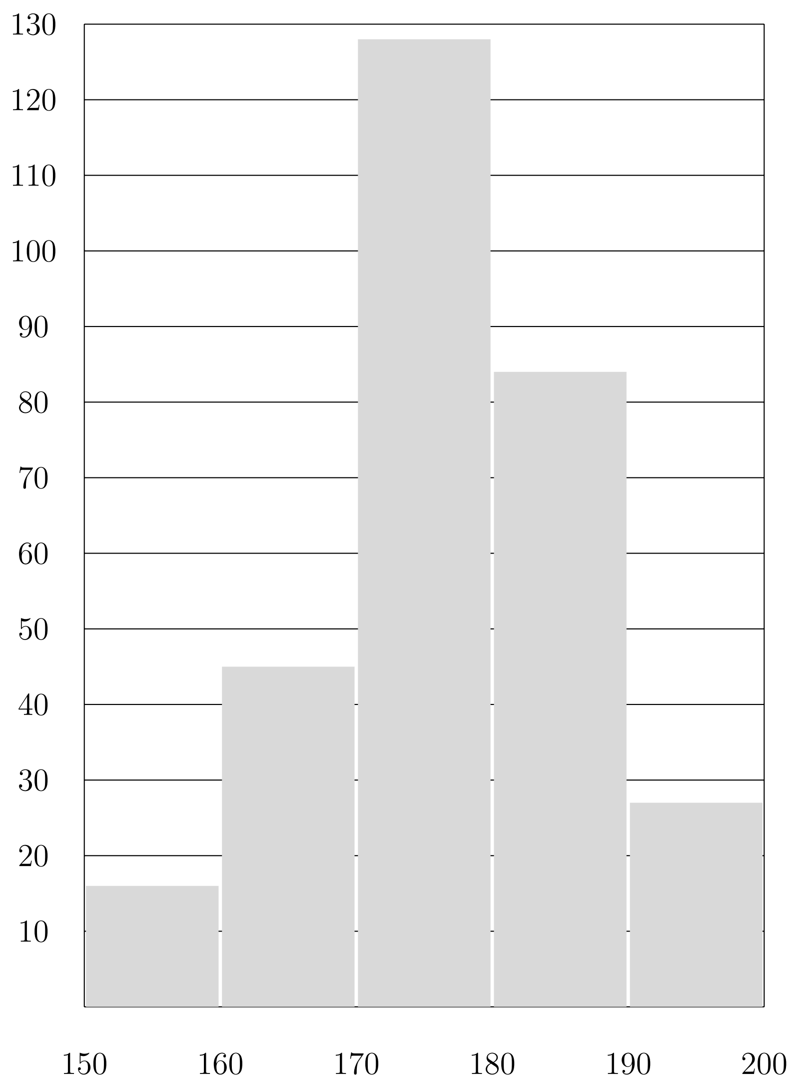
On note que les intervalles sont tous de même longueur. Il n'est donc a priori pas utile de faire les tableaux pour chaque série marginale. De même, les effectifs corrigés correspondent ici aux effectifs. Néanmoins, nous allons ici fournir les tableaux. D'abord, pour la série statistique x, on fait la somme sur chaque colonne. Puis, pour la série statistique y, on fait la somme sur chaque ligne.

Voici le tableau complet de la série statistique x (c'est-à-dire la taille) avec $u = 10$.

$[z_{i-1}; z_i[$	n_i	f_i	F_i	$k_i = \frac{z_i - z_{i-1}}{u}$	$\frac{n_i}{k_i}$
$[150; 160[$	16	$\frac{16}{300}$	$\frac{16}{300}$	1	16
$[160; 170[$	45	$\frac{45}{300}$	$\frac{61}{300}$	1	45
$[170; 180[$	128	$\frac{128}{300}$	$\frac{189}{300}$	1	128
$[180; 190[$	84	$\frac{84}{300}$	$\frac{273}{300}$	1	84
$[190; 200[$	27	$\frac{27}{300}$	1	1	27

On obtient ainsi l'histogramme des effectifs corrigés suivant :

FIGURE 5 – Histogramme de la taille

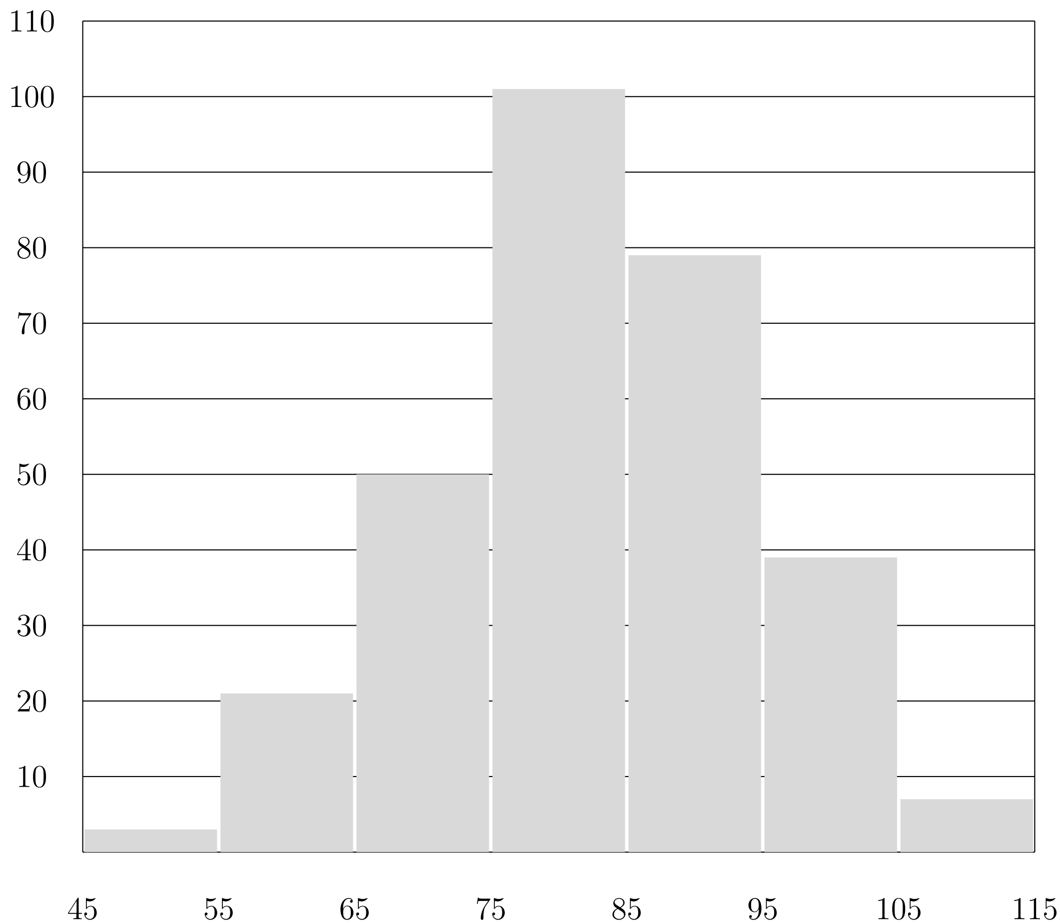


Voici le tableau complet de la série statistique y (c'est-à-dire le poids) avec $u = 10$.

$[z_{i-1}; z_i[$	n_i	f_i	F_i	$k_i = \dfrac{z_i - z_{i-1}}{u}$	$\dfrac{n_i}{k_i}$
$[45; 55[$	3	$\frac{3}{300}$	$\frac{3}{300}$	1	3
$[55; 65[$	21	$\frac{21}{300}$	$\frac{24}{300}$	1	21
$[65; 75[$	50	$\frac{50}{300}$	$\frac{74}{300}$	1	50
$[75; 85[$	101	$\frac{101}{300}$	$\frac{175}{300}$	1	101
$[85; 95[$	79	$\frac{79}{300}$	$\frac{254}{300}$	1	79
$[95; 105[$	39	$\frac{39}{300}$	$\frac{293}{300}$	1	39
$[105; 115[$	7	$\frac{7}{300}$	1	1	7

On obtient ainsi l'histogrammes des effectifs corrigés suivant :

Figure 6 – Histogramme du poids



Correction du 2)

On commence par calculer les moyennes en considérant les milieux de chaque classe :

$$\begin{aligned}\overline{x} =& \frac{155 \times 16 + 165 \times 45 + 175 \times 128 + 185 \times 84 + 195 \times 27}{300}\\ =& \frac{53\,110}{300} \approx 177.03\,,\end{aligned}$$

et

$$\begin{aligned}\overline{y} =& \frac{50 \times 3 + 60 \times 21 + 70 \times 50 + 80 \times 101 + 90 \times 79 + 100 \times 39 + 110 \times 7}{300}\\ =& \frac{24\,770}{300} \approx 82.57\,.\end{aligned}$$

On calcule maintenant la variance de la taille comme suit (en prenant les milieux de chaque classe et en ajoutant les variances intra-classes) :

$$\begin{aligned}s_x^2 =& \frac{1}{n}\sum_{i=1}^{r} n_i \left(\frac{z_{i-1}+z_i}{2}\right)^2 - \overline{x}^2 + \frac{1}{n}\sum_{i=1}^{r} n_i \frac{(z_i - z_{i-1})^2}{12}\\ =& \frac{155^2 \times 16 + \cdots + 195^2 \times 27}{300} - \overline{x}^2 + \underbrace{\frac{1}{300}\sum_{i=1}^{5} n_i \times \frac{(z_i - z_{i-1})^2}{12}}_{=\frac{100}{12}}\\ =& \frac{9\,431\,100}{300} - \frac{28\,206\,721}{900} + \frac{100}{12}\\ =& \frac{94\,079}{900} \approx 104.53\,.\end{aligned}$$

On peut maintenant calculer son écart-type (non demandé) : $s_x \approx 10.22$.

On fait de même avec la variance du poids :

$$\begin{aligned}s_y^2 =& \frac{50^2 \times 3 + \cdots + 110^2 \times 7}{300} - \overline{y}^2\\ &+ \frac{1}{12}\frac{1}{300}\left(3 \times (55-45)^2 + 21 \times (65-55)^2 + \cdots + 7 \times (115-105)^2\right)\\ =& \frac{2\,089\,100}{300} - \frac{6\,135\,529}{900} + \frac{100}{12}\\ =& \frac{139\,271}{900} \approx 154.75\,.\end{aligned}$$

On peut maintenant calculer son écart-type (non demandé) : $s_y \approx 12.44$.

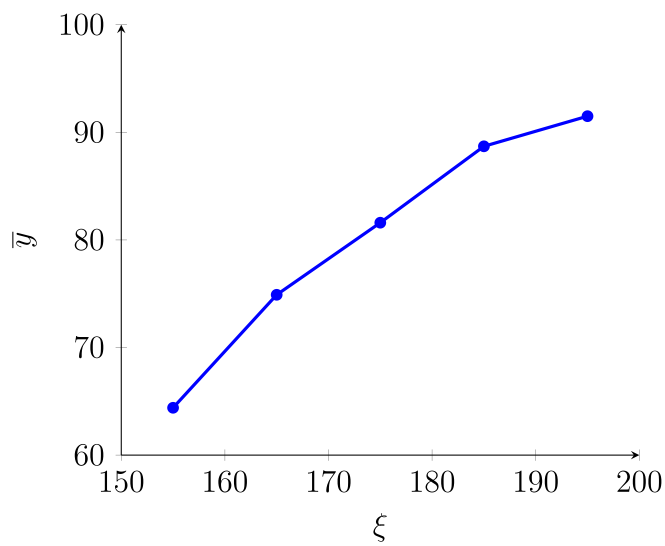
Correction du 3)

Pour construire cette courbe, on calcule d'abord les points de coordonnées $(\xi_i, \overline{y_i})$ où $\xi_i := \frac{z_{i-1}+z_i}{2}$ et $\overline{y_i}$ est la moyenne du poids pour les individus dans la classe $[z_{i-1}; z_i[$. Ici, i varie entre 1 et 5.

1. Le premier point a pour abscisse 155. On calcule alors le poids moyen pour les individus appartenant à cette classe : $\frac{50\times2+60\times7+70\times5+80\times2+90\times0+100\times0+110\times0}{2+7+5+2+0+0+0} = \frac{1\,030}{16} \approx 64.4$. Le premier point à placer est donc $(155; 64.4)$.
2. Le deuxième point a pour abscisse 165. On calcule alors le poids moyen pour les individus appartenant à cette classe : $\frac{50\times1+60\times8+70\times15+80\times12+90\times7+100\times2+110\times0}{1+8+15+12+7+2+0} = \frac{3\,370}{45} \approx 74.9$. Le deuxième point à placer est donc $(165; 74.9)$.
3. Le troisième point a pour abscisse 175. On calcule alors le poids moyen pour les individus appartenant à cette classe : $\frac{50\times0+60\times4+70\times22+80\times63+90\times28+100\times10+110\times1}{0+4+22+63+28+10+1} = \frac{10\,450}{128} \approx 81.6$. Le troisième point à placer est donc $(175; 81.6)$.
4. Le quatrième point a pour abscisse 185. On calcule alors le poids moyen pour les individus appartenant à cette classe : $\frac{50\times0+60\times2+70\times7+80\times19+90\times32+100\times20+110\times4}{0+2+7+19+32+20+4} = \frac{7\,450}{84} \approx 88.7$. Le quatrième point à placer est donc $(185; 88.7)$.
5. Le cinquième point a pour abscisse 195. On calcule alors le poids moyen pour les individus appartenant à cette classe : $\frac{50\times0+60\times0+70\times1+80\times5+90\times12+100\times7+110\times2}{0+0+1+5+12+7+2} = \frac{2\,470}{27} \approx 91.5$. Le cinquième point à placer est donc $(195; 91.5)$.

On obtient ainsi la courbe des moyennes conditionnelles :

FIGURE 7 – Courbe des moyennes conditionnelles



On remarque notamment que le poids n'évolue par linéairement par rapport à la taille.

Correction du 4)

On cherche à minimiser la fonction $(a, b) \mapsto \sum_{i=1}^{300} (y_i - ax_i - b)^2$. D'après le théorème correspondant, on connaît les valeurs de a_0 et b_0 telles que cette fonction y atteigne son minimum :

$$a_0 = \frac{s_{x,y}}{s_x^2} \quad \text{et} \quad b_0 = \overline{y} - a_0 \times \overline{x}\,.$$

Dans la question 2), on a calculé : $s_x^2 \approx 104.53$, $\overline{x} \approx 177.03$ et $\overline{y} \approx 82.57$. Il reste maintenant à calculer $s_{x,y}$. On pourrait utiliser le tableau de contingence comme suit :

$$\begin{aligned}
s_{x,y} =& \frac{1}{300}\Big((155 \times 50) \times 2 + (155 \times 60) \times 7 + \cdots + (155 \times 110) \times 0 \\
& \quad + (165 \times 50) \times 1 + (165 \times 60) \times 8 + \cdots + (165 \times 110) \times 0 \\
& \quad + \vdots \\
& \quad + (195 \times 50) \times 0 + (195 \times 60) \times 0 + \cdots + (195 \times 110) \times 2\Big) \\
& - \overline{x} \times \overline{y} \\
=& \frac{4\,404\,350}{300} - \overline{x} \times \overline{y} \approx 14\,681.17 - 177.03 \times 82.57 \approx 63.80\,.
\end{aligned}$$

On peut aussi utiliser la question 3) comme suit :

$$\begin{aligned}
s_{x,y} \approx& \frac{1}{300}\Big(16 \times 155 \times 64.4 + 45 \times 165 \times 74.9 + \cdots + 27 \times 195 \times 91.5\Big) - 177.03 \times 82.57 \\
\approx& 62.06\,.
\end{aligned}$$

On remarque ainsi que l'on commet plus d'erreurs mais le calcul est *beaucoup* plus rapide. Par conséquent, $a_0 \approx \frac{62.06}{104.53} \approx 0.59$. Puis, $b_0 \approx 82.57 - 0.59 \times 177.03 \approx -21.88$. L'équation de la droite de régression est donc $y = 0.59x - 21.88$.

Correction du 5)

Ce coefficient de corrélation linéaire entre les deux caractères est le quotient de $s_{x,y}$ par le produit des **écarts-types** de x et y :

$$r_{x,y} := \frac{s_{x,y}}{s_x s_y} \approx \frac{62.06}{10.22 \times 12.44} \approx 0.49\,.$$

Or, plus le coefficient est proche de 1 (ou de -1), plus il y a de corrélation linéaire. On en déduit donc que la taille et le poids sont faiblement corrélés dans l'échantillon.

Exercice 87

Énoncé

On a noté sur 6 800 individus les couleurs des cheveux et des yeux. On a construit le tableau de contingence :

Table 2 – Tableau de contingence

couleur des cheveux / couleur des yeux	blond	châtain	noir	roux
bleu	1 768	807	189	47
gris ou vert	946	1 387	746	53
marron	115	438	288	16

Calculer les fréquences conditionnelles des couleurs des yeux par rapport aux couleurs des cheveux dans cet échantillon. La couleur des yeux est-elle indépendante de celle des cheveux dans cet échantillon ?

Correction

On commence par poser les trois modalités pour les yeux :

— $a_1 := \{$La couleur des yeux de l'individu est : bleu$\}$.
— $a_2 := \{$La couleur des yeux de l'individu est : gris ou vert$\}$.
— $a_3 := \{$La couleur des yeux de l'individu est : marron$\}$.

On pose ensuite les quatre modalités pour les cheveux :

— $b_1 := \{$La couleur des cheveux est : blond$\}$.
— $b_2 := \{$La couleur des cheveux est : châtain$\}$.
— $b_3 := \{$La couleur des cheveux est : noir$\}$.
— $b_4 := \{$La couleur des cheveux est : roux$\}$.

On calcule ensuite le nombre d'individus pour chaque catégorie de cheveux.

— Il y a $q_1 := \sum_{i=1}^{3} n_{i,1} = 1\ 768 + 946 + 115 = 2\ 829$ individus qui ont la modalité b_1 pour les cheveux.
— Il y a $q_2 := \sum_{i=1}^{3} n_{i,2} = 807 + 1\ 387 + 438 = 2\ 632$ individus qui ont la modalité b_2 pour les cheveux.
— Il y a $q_3 := \sum_{i=1}^{3} n_{i,3} = 189 + 746 + 288 = 1\ 223$ individus qui ont la modalité b_3 pour les cheveux.
— Il y a $q_4 := \sum_{i=1}^{3} n_{i,4} = 47 + 53 + 16 = 116$ individus qui ont la modalité b_4 pour les cheveux.

Puis, l'on calcule les probabilités conditionnelles pour chaque couleur de cheveux comme suit :

b_1 $f(a_1|b_1) = \frac{n_{1,1}}{q_1} = \frac{1\ 768}{2\ 829} = 0.625$, $f(a_2|b_1) = \frac{n_{2,1}}{q_1} = \frac{946}{2\ 829} = 0.334$ et $f(a_3|b_1) = \frac{n_{3,1}}{q_1} = \frac{115}{2\ 829} = 0.041$. La distribution conditionnelle de la couleur des yeux par rapport aux

cheveux blonds est donc
$0.625\delta_{a_1} + 0.334\delta_{a_2} + 0.041\delta_{a_3}$.

b_2 $f(a_1|b_2) = \frac{n_{1,2}}{q_2} = \frac{807}{2\ 632} = 0.307$, $f(a_2|b_2) = \frac{n_{2,2}}{q_2} = \frac{1\ 387}{2\ 632} = 0.527$ et $f(a_3|b_2) = \frac{n_{3,2}}{q_2} = \frac{438}{2\ 632} = 0.166$. La distribution conditionnelle de la couleur des yeux par rapport aux cheveux châtains est donc
$0.307\delta_{a_1} + 0.527\delta_{a_2} + 0.166\delta_{a_3}$.

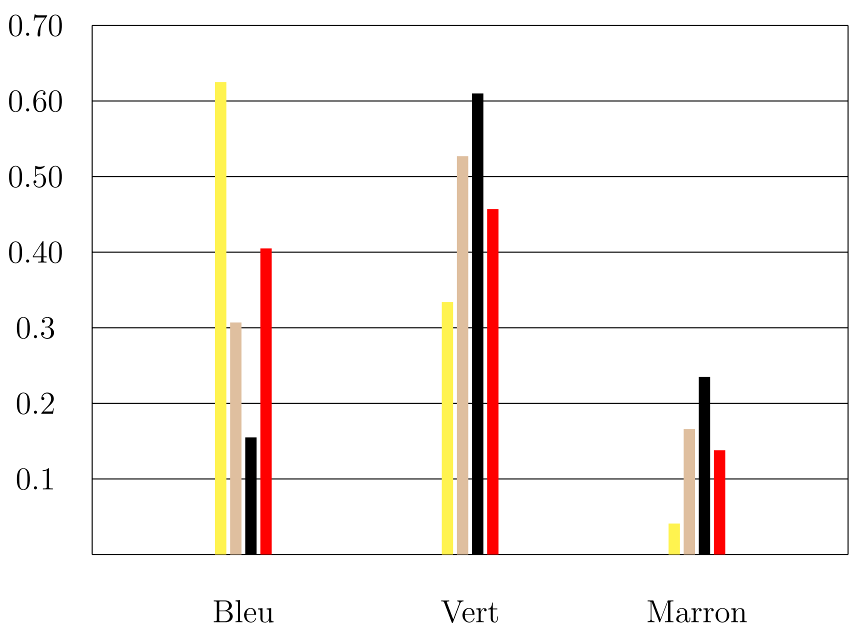
b_3 $f(a_1|b_3) = \frac{n_{1,3}}{q_3} = \frac{189}{1\ 223} = 0.155$, $f(a_2|b_3) = \frac{n_{2,3}}{q_3} = \frac{746}{1\ 223} = 0.610$ et $f(a_3|b_3) = \frac{n_{3,3}}{q_3} = \frac{288}{1\ 223} = 0.235$. La distribution conditionnelle de la couleur des yeux par rapport aux cheveux noirs est donc
$0.155\delta_{a_1} + 0.610\delta_{a_2} + 0.235\delta_{a_3}$.

b_4 $f(a_1|b_4) = \frac{n_{1,4}}{q_4} = \frac{47}{116} = 0.405$, $f(a_2|b_4) = \frac{n_{2,4}}{q_4} = \frac{53}{116} = 0.457$ et $f(a_3|b_4) = \frac{n_{3,4}}{q_4} = \frac{16}{116} = 0.138$. La distribution conditionnelle de la couleur des yeux par rapport aux cheveux roux est donc
$0.405\delta_{a_1} + 0.457\delta_{a_2} + 0.138\delta_{a_3}$.

Si la couleur des yeux était indépendante de la couleur des cheveux dans l'échantillon, les distributions conditionnelles seraient égales. Or, elles ne le sont pas. On en déduit que la couleur des yeux n'est pas indépendante de la couleur des cheveux dans l'échantillon considéré.

On peut visualiser cette non-indépendance par le diagramme en bâtons suivant où la distribution des cheveux blonds est en jaune, celle des cheveux châtains est en brun, celle des cheveux noirs est en noir et celle des cheveux roux est en rouge :

FIGURE 8 – Distribution de la couleur des yeux selon la couleur des cheveux



Remarque

Pour s'assurer que cette non indépendance n'est pas le fruit des fluctuations dans l'échantillon, on procède en général à un test statistique.