Corrigé exercices statistiques

Exercice 1 :

La variable statistique "couleur de maisons d'un quartier" est-elle :

- [x] qualitative
- [ ] quantitative
- [ ] discrète
- [ ] continue

La variable statistique "revenu brut" est-elle :

- [ ] qualitative
- [x] quantitative
- [ ] discrète
- [x] continue

La variable statistique "nombre de maisons vendues par ville" est-elle :

- [ ] qualitative
- [x] quantitative
- [x] discrète
- [ ] continue

Exercice 2 : ① VRAI ② FAUX ③ FAUX ④ VRAI ⑤ VRAI
⑥ VRAI ⑦ VRAI ⑧ VRAI ⑨ VRAI

Exercice 3 : L'âge : bébé - enfant - adulte - senior.
Heure : matin - après-midi - soir

Exercice 4 : 1) Quantitative continue 2) Qualitative
3) Quantitative continue 4) Quantitative continue
5) Qualitative 6) Quantitative discret 7) Qualitative
8) Quantitative discret

Exercice 5 : 1) Une lampe, temps de validité, quantitative continue.
2) ouvrier, jours d'absences, quantitative discret
3) étudiant, mention au bac, qualitative
4) une intersection, nombre de collisions, quantitative discret.

Exercice 6 : 1) Variable statistique : groupe sanguin. Type : qualitatif.
2) 40 - 20 - 10 - 5 = 5.
3) Tuyaux d'orgue :

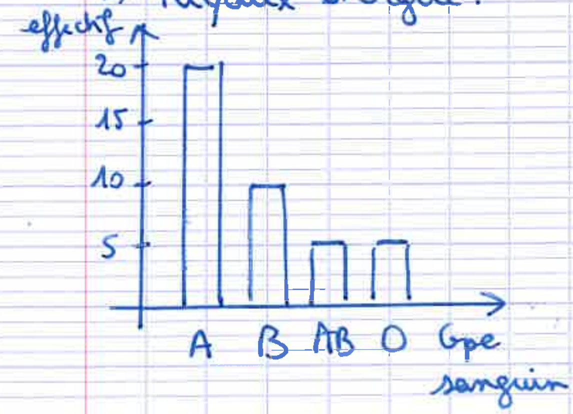
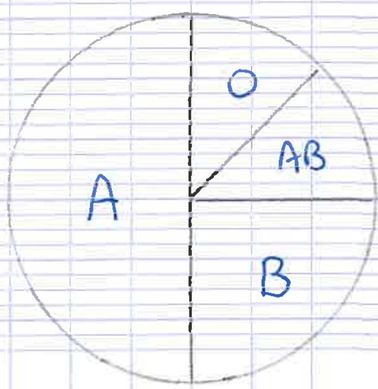


Diagramme circulaire :

Effectif	Angle
40	360°
20	180°
10	90°
5	45°

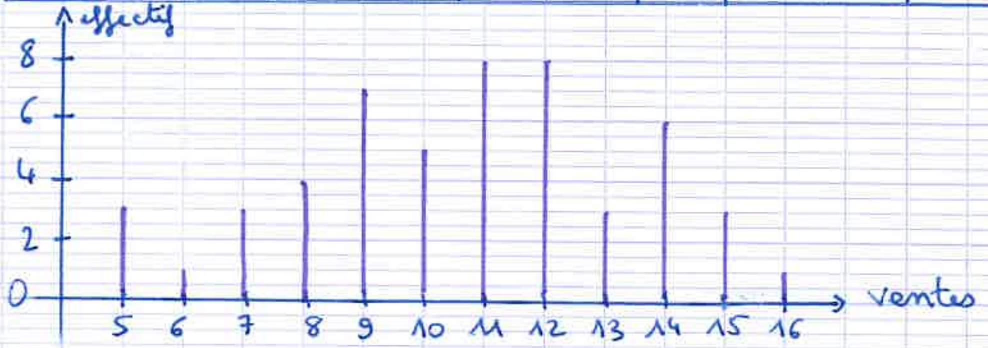


Exercice 7 : 1) Quantitatif discret.

2)

	5	6	7	8	9	10	11	12	13	14	15	16
Effectif	3	1	3	4	7	5	8	8	3	6	3	1
Fréquence	3/52	1/52	3/52	4/52	7/52	5/52	8/52	8/52	3/52	6/52	3/52	1/52
ECC	3	4	7	11	18	23	31	39	42	48	51	52
FCC	3/52	4/52	7/52	11/52	18/52	23/52	31/52	39/52	42/52	48/52	51/52	1

3)



4) $F(x) = 0$ si $x < 5$

$= \frac{3}{52}$ si $5 \leq x < 6$

$= \frac{4}{52}$ si $6 \leq x < 7$

$\vdots$

$= 1$ si $x \geq 16$

5) $M_o = 11$ et $M_o = 12$.

$$\bar{x} = \frac{5\times3 + 6\times1 + 7\times3 + \dots + 15\times3 + 16\times1}{52} = \frac{555}{52} \simeq 10{,}67.$$

6) $F(10) = \frac{23}{52} < 50\%$ et $F(11) = \frac{31}{52} > 50\%$ donc $M_e = 11$.

$F(8) = \frac{11}{52} < 25\%$ et $F(9) = \frac{18}{52} > 25\%$ donc $Q_1 = 9$.

$F(11) = \frac{31}{52} < 75\%$ et $F(12) = \frac{39}{52} = 75\%$ donc $Q_3 = 12$.

7) Etendue : $16 - 5 = 11$

Variance : $Var(X) = \frac{3\times5^2 + 1\times6^2 + 3\times7^2 + \dots + 3\times15^2 + 1\times16^2}{52} - \left(\frac{555}{52}\right)^2 \simeq 7{,}53$

Ecart-type : $\sigma = \sqrt{Var(X)} \simeq 2{,}74$.

Ecart interquartile : $Q_3 - Q_1 = 12 - 9 = 3$.

Exercice 8 :

		Groupe A	Groupe B
paramètres de position	Mode	8 et 9	6, 8 et 9
	Médiane	$F(8)<50\%$, $F(9)>50\% \Rightarrow M_e=9$	$F(6)<50\%$; $F(8)=50\% \Rightarrow M_e=8$
	Moyenne	$\bar{x}=\frac{8\times2+9\times2+10+11}{6}=\frac{55}{6}\approx 9{,}2$	$\bar{x}=\frac{6\times2+8\times2+9\times2+13+14}{8}=\frac{73}{8}=9{,}125$
	Q_1	$F(8)>25\% \Rightarrow Q_1=8$	$F(6)=25\% \Rightarrow Q_1=6$
	Q_3	$F(9)<75\%$, $F(10)>75\% \Rightarrow Q_3=10$	$F(9)=75\% \Rightarrow Q_3=9$
paramètres de dispersion	Etendue	$11-8=3$	$14-6=8$
	Variance	$\frac{8^2\times2+9^2\times2+10^2+11^2}{6}-\left(\frac{55}{6}\right)^2\approx 1{,}14$	$\frac{6^2\times2+8^2\times2+9^2\times2+13^2+14^2}{8}-\left(\frac{73}{8}\right)^2\approx 7{,}61$
	Ecart-type	$\sqrt{1{,}14}\approx 1{,}07$	$\sqrt{7{,}61}\approx 2{,}76$.
	Ecart interquartile	$10-8=2$	$9-6=3$

Le groupe A est légèrement meilleur que le groupe B (paramètres de position plus hauts) et surtout plus homogène (paramètres de dispersion plus faibles).

Exercice 9 : 1) Modalité 1 : $2339{,}50\times1{,}1=2\,573{,}45$ €.

Modalité 2 : $2\,339{,}50+200=2\,539{,}50$ €.

2) La modalité 1 car le salaire moyen est supérieur.

3) Paramètres de position : Mode $=1450$ et 1925.

Moyenne : $\bar{x}=\frac{1450\times15+1510\times10+1925\times15+5125\times10}{50}=2339{,}5$

Médiane : $F(1510)=50\% \Rightarrow M_e=1510$.

Quartiles : $F(1450)>25\% \Rightarrow Q_1=1450$; $\left.\begin{array}{l}F(1925)>75\%\\F(1510)<75\%\end{array}\right\}\Rightarrow Q_3=1925$

Paramètres de dispersion : Etendue : $5125-1450=3675$.

Variance : $\frac{1450^2\times15+\dots+5125^2\times10}{50}-2339{,}5^2\approx 1\,978\,322$

Ecart-type : $\sqrt{1\,978\,322}\approx 1\,406$

Ecart interquartile : $Q_3-Q_1=1925-1450=475$.

4) L'augmentation de 10% est intéressante pour les salaires > 2000€

Or $Q_3=1925$ € donc plus des $\frac{3}{4}$ des salariés ont intérêt à avoir une augmentation de 200 €. D'où le résultat du vote.

Exercice 10 :

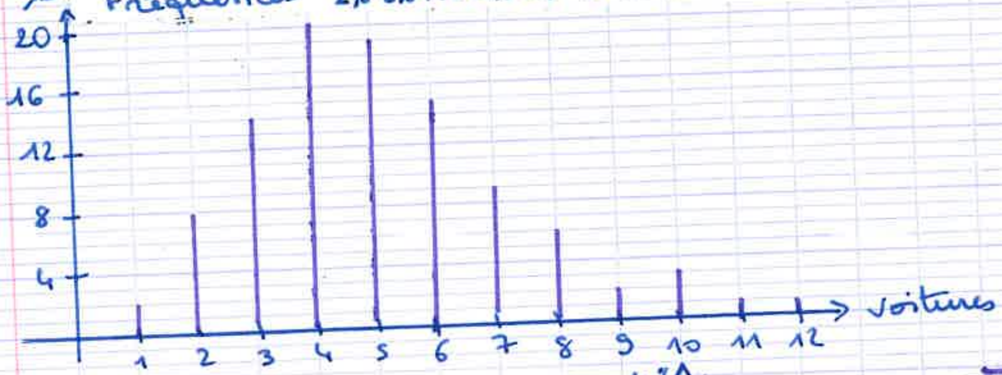
x_i	"CLEU"	"Studio"	"Résidence"	"Maison"	"Autre"	Total
%	4,8	16,5	38,6	28,6	11,6	100
eff	9	31	73	54	22	189

$\times\frac{189}{100}$

Exercice 11 : 1) Quantitatif discret.

2)

Nombre de voitures	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'observations	2	8	14	20	19	15	9	6	2	3	1	1
Fréquences	2%	8%	14%	20%	19%	15%	9%	6%	2%	3%	1%	1%



3) $F(x) = 0$ si $x < 1$

$= 0{,}02$ si $x \in [1;2[$

$= 0{,}10$ si $x \in [2;3[$

$= 0{,}24$ si $x \in [3;4[$

$= 0{,}44$ si $x \in [4;5[$

$= 0{,}63$ si $x \in [5;6[$

$= 0{,}78$ si $x \in [6;7[$

$= 0{,}87$ si $x \in [7;8[$

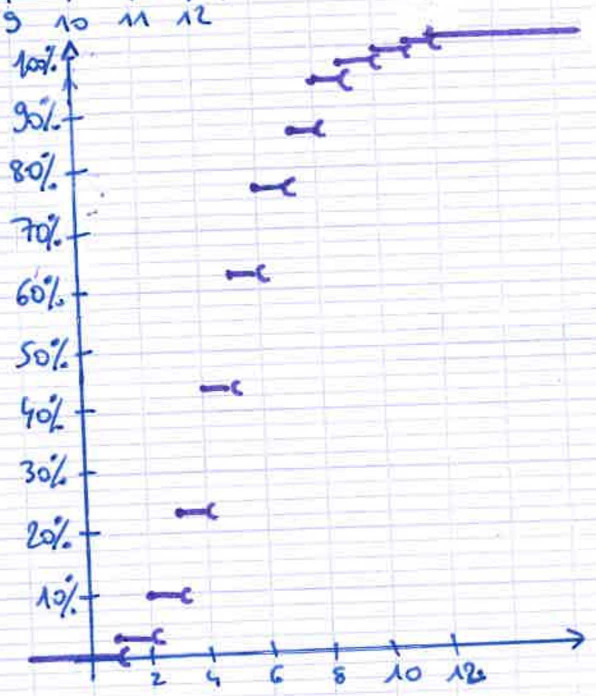
$= 0{,}93$ si $x \in [8;9[$

$= 0{,}95$ si $x \in [9;10[$

$= 0{,}98$ si $x \in [10;11[$

$= 0{,}99$ si $x \in [11;12[$

$= 1$ si $x \geqslant 12$



4) $M_o = 4$ et $\bar{x} = \dfrac{2\times1 + 8\times2 + 14\times3 + \dots + 11 + 12}{100} = \dfrac{507}{100} = 5{,}07$

5) $F(4) < 50\%$ et $F(5) > 50\% \Rightarrow M_e = 5$.

$F(3) < 25\%$ et $F(4) > 25\% \Rightarrow Q_1 = 4$ / $F(5) < 75\%$ et $F(6) > 75\% \Rightarrow Q_3 =$

6) Etendue : $12 - 1 = 11$

Variance : $\dfrac{2\times1^2 + 8\times2^2 + \dots + 11^2 + 12^2}{100} - 5{,}07^2 \simeq 4{,}77$.

Ecart-type : $\sqrt{4{,}77} \simeq 2{,}18$

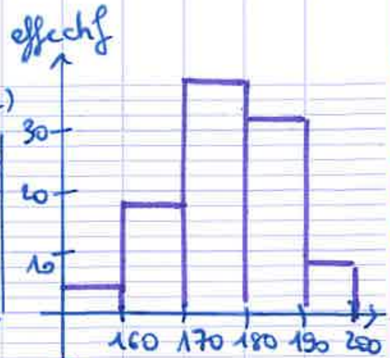
Ecart interquartile : $Q_3 - Q_1 = 6 - 4 = 2$.

Exercice 12 :

1)

X	[150;160[	[160;170[	[170;180[	[180;190[	[190;200]
Effectif	4	18	38	32	8
fréquence	0,04	0,18	0,38	0,32	0,08

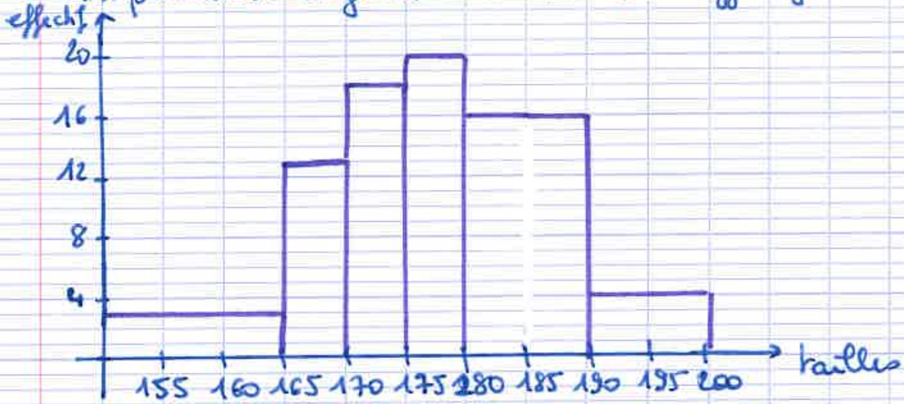
2)



3)

X	[150; 165[	[165; 170[	[170; 175[	[175; 180[	[180; 190[	[190; 200[
Effectif	9	13	18	20	32	8
Amplitude	15	5	5	5	10	10
Largeur rectangle	3	13	18	20	16	4

Amplitude × largeur = aire = 5 × effectif.



Exercice 13 :

1)

Salaires	[900;1200[	[1200;1400[	[1400;1600[	[1600;1800[	[1800;2000[	[2000;2400[
Effectifs	30	30	60	40	20	20
Fréquences	15%	15%	30%	20%	10%	10%

2) Mode : $M_o = [1400; 1600[$

Moyenne : $\bar{x} = \dfrac{1050\times30 + 1300\times30 + 1500\times60 + 1700\times40 + 1900\times20 + 2200\times20}{200}$

$\Rightarrow \bar{x} = 1552,5.$

Médiane : $M_e = [1400; 1600[$ / $Q_1 = [1200; 1400[$ et $Q_3 = [1600; 1800[$

Etendue : $2400 - 900 = 1500$

Variance : $\dfrac{30\times1050^2 + \dots + 20\times2200^2}{200} - 1552,5^2 = 106\,618,75$

Ecart-type : $\sqrt{106\,618,75} \approx 326,5$

Exercice 14 :

1) $\bar{x} = 10$, $Var(X) = 2,25 \Rightarrow \sigma = 1,5$, $M_e = 10$, $Q_1 = 8$, $Q_3 = 11$, $Q_3 - Q_1 = 3$

2) (a) $\bar{x} = 9,8$, $Var(X) = 6,96 \Rightarrow \sigma \approx 2,64$, $M_e = 10$, $Q_1 = 8$, $Q_3 = 12$, $Q_3 - Q_1 = 4$

(b)

Paramètres :	$\bar{x}$	σ	M_e	Q_1	Q_3
Ecart :	−2%	76%	0%	0%	9%

moins sensible

Exercice 16 :

1)

$X \setminus Y$	1	2	3	4	$n_{i\bullet}$
-1	0	1	2	2	5
0	1	1	0	1	3
1	0	1	1	0	2
$n_{\bullet j}$	1	3	3	3	10

2) $\bar{x} = \dfrac{-1\times5+0\times3+1\times2}{10} = -0,3$

$\bar{y} = \dfrac{1\times1+2\times3+3\times3+3\times4}{10} = 2,8$

$cov(X,Y) = \dfrac{1\times0\times1+1\times(-1)\times2+1\times0\times2+1\times1\times2+2\times(-1)\times3+1\times1\times3+2\times(-1)\times4+1\times0\times4}{10} - (-0,3)\times2,8$

$cov(X,Y) = -1,1-(-0,3)\times2,8 = -0,26.$

Exercice 17 : 1) $D(m/x)$: $m = ax+b$: $a = \dfrac{cov(x,m)}{Var(x)}$ et $b = \bar{m} - a\bar{x}$

$\bar{x} = 1$, $\bar{m} = 20$, $cov(x,m) = \dfrac{0\times0+0,5\times10+1,1\times20+1,5\times30+1,9\times40}{5} - 1\times20$

$\Rightarrow cov(x,m) = 9,6$

$Var(x) = \dfrac{0^2+0,5^2+1,1^2+1,5^2+1,9^2}{5} - 1^2 = 0,464$ $\Big\} \Rightarrow a \simeq 20,69$

$b = 20 - 20,69\times1 \approx -0,69.$

$\Rightarrow D(m/x)$: $m = 20,69\,x - 0,69$

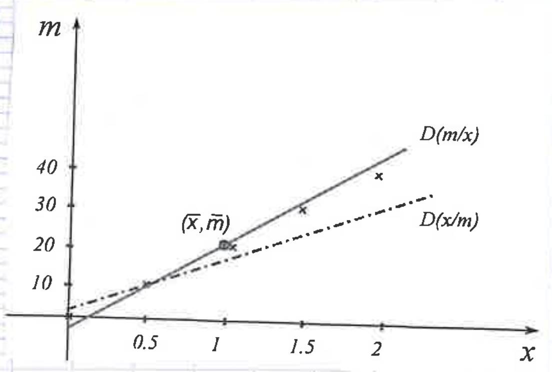
2) $D(x/m)$: $x = a'm + b'$: $a' = \dfrac{cov(m,x)}{Var(m)}$ et $b' = \bar{x} - a'\bar{m}$

$Var(m) = \dfrac{0^2+10^2+20^2+30^2+40^2}{50} - 20^2 = 200 \Rightarrow a' = 0,048$

$b' = 1 - 0,048\times20 = 0,04$

$\Rightarrow D(x/m)$: $x = 0,048m - 0,04$

3)



4) $\rho_{xm} = \dfrac{cov(x,m)}{\sigma_x\times\sigma_m} = \dfrac{9,6}{\sqrt{0,464}\times\sqrt{200}}$

$\approx 0,997$

Comme $\rho_{xm} \geq 0,7$,

l'ajustement linéaire est accepté donc on peut estimer x avec $0,048\times51,75 - 0,04 \simeq 2,44$

Exercice 18 : 1)

$X(\times100) \setminus Y$	[0, 8[	[8, 16[	[16, 24[	[24, 32[	Loi marginale
[20, 30[	5	6	1	0	12
[30, 40[	2	4	3	3	12
[40, 50[	0	2	4	10	16
Loi marginale	7	12	8	13	40

2) Comme

$40\times5 = 200$
et $12\times7 = 84$ $\Big] \neq$

$N\times n_{11} \neq n_{1\bullet}\times n_{\bullet1}$

$\Rightarrow$ X et Y ne sont pas indépendantes

3) X/y_3

X	[20;30[	[30;40[	[40;50[
$f_{\bullet/3}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{4}{8}$

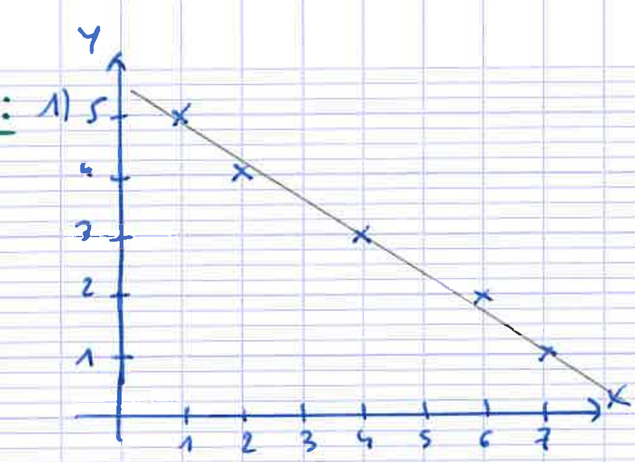
$\bar{x}_3 = 38,75$

Y/x_2

Y	[0;8[	[8;16[	[16;24[	[24;32[
$f_{\bullet/2}$	$\frac{2}{12}$	$\frac{4}{12}$	$\frac{3}{12}$	$\frac{3}{12}$

$\bar{y}_2 \approx 16,67.$

Exercice 19 : 1)



2) Comme les points semblent suivre une droite de pente négative, le coefficient de corrélation est négatif et proche de -1 (points presque alignés).

3) $\bar{x} = \frac{1+2+7+4+6}{5} = \frac{20}{5} = 4$ et $\bar{y} = \frac{5+4+1+3+2}{5} = \frac{15}{5} = 3$

$cov(X,Y) = \frac{1\times5+2\times4+7\times1+4\times3+6\times2}{5} - 4\times3 = 8,8 - 12 = -3,2$

$Var(X) = \frac{1^2+2^2+7^2+4^2+6^2}{5} - 4^2 = 5,2$ et $Var(Y) = \frac{5^2+4^2+1^2+3^2+2^2}{5} - 3^2 = 2$

$\Rightarrow \rho_{XY} = \frac{-3,2}{\sqrt{5,2\times2}} \approx -0,99 \geq 0,7 \Rightarrow$ l'ajustement linéaire sera accepté.

$D(Y/X)$: $a = \frac{cov(X,Y)}{Var(X)} = \frac{-3,2}{5,2} \approx -0,62$ et $b = \bar{y} - a\bar{x} \approx 5,46$.

$Y = -0,62X + 5,46$.

Exercice 20 : 1) $\bar{x} = \frac{2+3+52+2+4}{5} = 12,6$ 2) $\bar{y} = \frac{10+16+23+12+18}{5} = 15,8$

3) $Var(X) = \frac{2^2+3^2+52^2+2^2+4^2}{5} - 12,6^2 = 388,64 \Rightarrow \sigma_x = \sqrt{388,64} \approx 19,71$

4) $Var(Y) = \frac{10^2+16^2+23^2+12^2+18^2}{5} - 15,8^2 = 20,96 \Rightarrow \sigma_y = \sqrt{20,96} \approx 4,58$

5) $cov(X,Y) = \frac{2\times10+3\times16+52\times23+2\times12+4\times18}{5} - 12,6\times15,8 = 72,92$

6) $a = \frac{cov(X,Y)}{Var(X)} = \frac{72,92}{388,64} \approx 0,19$ et $b = \bar{y} - a\bar{x} = 15,8 - 0,19\times12,6 \approx 13,4$

$\Rightarrow Y = 0,19X + 13,4$.

7) $\rho_{XY} = \frac{cov(X,Y)}{\sigma_x\,\sigma_y} = \frac{72,92}{\sqrt{388,64}\times\sqrt{20,96}} \approx 0,81 \geq 0,7 \Rightarrow$ l'ajustement linéaire est accepté.

8) $0,19\times6 + 13,4 \approx 14,5 \Rightarrow$ le coût de production estimé est 14,50 €

Exercice 21 : 1) $\bar{x} = \frac{3+4+2+5+3}{5} = 3,4$ 2) $\bar{y} = \frac{12+14+8+19+11}{5} = 12,8$

3) $Var(X) = \frac{3^2+4^2+2^2+5^2+3^2}{5} - 3,4^2 = 1,04 \Rightarrow \sigma_x = \sqrt{1,04} \approx 1,02$

4) $Var(Y) = \frac{12^2+14^2+8^2+19^2+11^2}{5} - 12,8^2 = 13,36 \Rightarrow \sigma_y = \sqrt{13,36} \approx 3,66$

5) $cov(X,Y) = \frac{3\times12+4\times14+2\times8+5\times19+3\times11}{5} - 3,4\times12,8 = 3,68$

6) $a = \frac{cov(X,Y)}{Var(X)} = \frac{3,68}{1,04} \approx 3,54$ et $b = \bar{y} - a\bar{x} = 12,8 - 3,54\times3,4 \approx 0,77$

$\Rightarrow Y = 3,54X + 0,77$

7) $\rho_{XY} = \frac{cov(X,Y)}{\sigma_x\,\sigma_y} = \frac{3,68}{\sqrt{1,04}\times\sqrt{13,36}} \approx 0,99 \geq 0,7 \Rightarrow$ l'ajustement linéaire est accepté

Exercice 22 : 1) $10+8+5+8+9+4=44$.

2)

Marginale X	[0;6[	[6;12[	[12;18[	[18;24[	Total
Effectif	23	21	35	11	90

(somme des lignes)

Marginale Y	[20;25[	[25;30[	[30;35[	Total
Effectif	36	34	20	90

(somme des colonnes)

3)

X/Y=[25;30[	[0;6[	[6;12[	[12;18[	[18;24[
Fréquence	$\frac{8}{34}$	$\frac{9}{34}$	$\frac{11}{34}$	$\frac{6}{34}$

4) X et Y ne sont pas indépendants car $\frac{8}{90} \neq \frac{8}{34} \times \frac{8}{23}$ par ex.

5) $\bar{x} = \frac{23\times3+21\times9+35\times15+11\times21}{90} \approx 11,27$

$\bar{y} = \frac{36\times22,5+34\times27,5+20\times32,5}{90} \approx 26,61$

6) $Var(X) = \frac{23\times3^2+21\times9^2+35\times15^2+11\times21^2}{90} - 11,27^2 \approx 35,66$

$\Rightarrow \sigma_x \approx 5,97$

$Var(Y) = \frac{36\times22,5^2+34\times27,5^2+20\times32,5^2}{90} - 26,61^2 \approx 14,77$

$\Rightarrow \sigma_y \approx 3,84$

$cov(X,Y) = \frac{10\times3\times22,5+8\times9\times22,5+15\times15\times22,5+\dots+2\times21\times32,5}{90} - 11,27\times26,61$

$\Rightarrow cov(X,Y) \approx 1,01$.

$\rho_{XY} = \frac{cov(X,Y)}{\sigma_x \times \sigma_y} \approx \frac{1,01}{5,97\times3,84} \approx 0,044 \leq 0,7 \Rightarrow$ l'ajustement linéaire n'est pas accepté ici.

Donc la suite de l'exercice n'est pas pertinent.