

Exercice 1. On a relevé les températures corporelles d'un échantillon de 22 chats. Voici les résultats :

38	38,1	38,2	38,2	38,3	38,3	38,4	38,5	38,5	38,5	38,5
38,7	38,9	38,9	38,9	39	39,1	39,1	39,2	39,2	39,3	38,5

Calculer la moyenne, l'écart type et la médiane de cet échantillon.

Exercice 2. L'indice de masse corporelle (IMC) est défini par : $IMC = \frac{\text{masse en kg}}{(\text{taille en m})^2}$.

Il permet de mesurer la corpulence d'un adulte. L'Organisation Mondiale de la Santé (OMS) a défini les critères suivants :

- maigre (16,5 à 18,5)
- normal (de 18,5 à 25)
- risque de surpoids (de 25 à 30)
- obésité modérée (de 30 à 35)
- obésité sévère (35 à 40).

Remarque. Cet indice n'a qu'une valeur indicative : il ne prend pas en compte la proportion de masse musculaire, de masse osseuse, de masse grasse... en particulier les sportifs se retrouvent souvent en surpoids bien que leur forme physique est souvent meilleure que la moyenne des individus.

En deçà de 16,5 (dénutrition) et au-delà de 40 (obésité massive), les risques de mortalité sont élevés.

1. On donne ci-contre le poids et la taille d'un échantillon de 13 personnes.

$x =$ Poids en kg	70	65	95	58	42	75	45	89	77	83	62	48	59
$y =$ Taille en m	1,68	1,85	1,56	1,61	1,5	1,68	1,65	1,65	1,64	1,75	1,48	1,48	1,74

- a. Calculer leur IMC.
- b. Calculer la moyenne des tailles, des poids et des IMC, puis les écarts types de taille, de poids et d'IMC.
- c. Donner la taille, le poids et l'IMC médians.

2. Le tableau suivant donne la répartition (par classes) de l'IMC pour un échantillon de 215 individus :

- a. Calculer la moyenne et l'écart type de l'échantillon.
- b. Calculer les fréquences, les fréquences cumulées.
- c. Représenter graphiquement la courbe cumulative.
- d. Déterminer la médiane et les quartiles.
- e. Quelle proportion de l'échantillon a un IMC :
 - i. inférieur à 26 kg.m^{-2}
 - ii. supérieur à $33,5 \text{ kg.m}^{-2}$
 - iii. compris entre 26 kg.m^{-2} et $33,5 \text{ kg.m}^{-2}$.

	Classe	effectifs
Maigre	[16,5;18,5[	11
normal	[18,5;25[	114
risque de surpoids	[25;30[	65
obésité modérée	[30;35[	20
obésité sévère	[35;40[	5
	Total	215

Exercice 3. Les données suivantes précisent le taux d'hémoglobine dans le sang (par classes, en gl^{-1}) mesuré chez 60 hommes et 70 femmes présumés en bonne santé.

Classes	[105;115[	[115;125[	[125;135[	[135;145[	[145;155[	[155;165[	[165;175[	[175;185]
Effectifs hommes	0	0	3	4	18	19	12	4
Effectifs femmes	9	12	18	14	13	4	0	0

- Calculer les fréquences et les fréquences cumulées des hommes, des femmes et des deux réunis.
- Représenter sur même graphique les courbes cumulatives (ou courbes des fréquences cumulées) des hommes, des femmes et les deux réunis.
- Calculer approximativement les proportions :
 - a. d'hommes dont le taux d'hémoglobine est inférieur à 130 ;
 - b. d'hommes dont le taux d'hémoglobine est compris entre 130 et 152 ;
 - c. Reprendre les mêmes questions pour l'échantillon des femmes.
- Déterminer la médiane et les quartiles du taux d'hémoglobine des hommes, des femmes et des deux réunis.
- Calculer la moyenne, la variance et l'écart type du taux d'hémoglobine des hommes, des femmes et des deux réunis.

Exercice 2 (suite). On reprend les données de la question 1 de l'exercice 2.

Représenter graphiquement le nuage statistique et calculer la covariance $C(x, y)$ et le coefficient de corrélation $r(x, y)$. Déterminer la droite de régression de y en x et la représenter. Conclure.

Exercice 4. La série statistique suivante donne le nombre de stomates aérifères au mm^2 d'une feuille (variable y) en fonction du nombre de jours d'exposition au soleil (variable x).

Nombre de jours x	2	4	8	10	24	40	52
Nombre de stomates y	6	11	15	20	39	62	85

1. Représenter graphiquement ce nuage de points.
2. Calculer les moyennes $\bar{x}$ et $\bar{y}$.
3. Calculer les écarts types s_x et s_y la covariance $C(x,y)$ et le coefficient de corrélation $r(x,y)$.
4. Déterminer la droite de régression de y en x et la représenter. Conclure.
5. Combien de stomates prévoyez-vous à 30 jours ?

Exercice 5. Dans le cadre de travaux de recherche sur la durée de la saison de végétation en montagne, des stations météorologiques sont installées à différentes altitudes. La température moyenne (variable y en degrés Celsius) ainsi que l'altitude (variable x en mètres) de chaque station sont données dans le tableau ci-dessous :

Altitude x	1040	1230	1500	1600	1740	1950	2200	2530	2800	3100
Température y	7,4	6	4,5	3,8	2,9	1,9	1	-1,2	-1,5	-4,5

1. Mêmes questions 1 à 4 que dans l'exercice précédent.
2. Quelle température moyenne prévoyez-vous à 1300 m ? à 3000 m ?

Exercice 6. Une étude théorique de l'évolution d'une population en extinction conduit à penser que le nombre N d'individus varie avec le temps suivant une loi du type :

$$N(t) = ae^{-kt} \quad \text{c'est à dire} \quad \ln(N(t)) = -kt + \ln(a)$$

où k et a sont des constantes positives à déterminer expérimentalement. Voici les observations effectuées pendant 12 mois :

Mois : t	1	2	3	4	5	6	7	8	9	10	11	12
Survivants : N	180	152	145	125	109	95	82	76	58	51	35	31

1. Faire une représentation graphique du nuage de points.
2. Calculer les valeurs de $\ln(N)$ à la suite du tableau
3. Calculer les coefficients de la droite de régression de $\ln(N)$ par rapport à t . En déduire l'estimation exponentielle de N en fonction de t , et donner l'estimation de N lorsque $t = 18$ mois.

Exercice 7. On mesure le poids frais (x) et le poids sec (y) de 20 prélèvements de plancton. Les résultats sont les suivants (exprimés en g par $10m^3$ d'eau de mer) :

poids frais x	20	28	49	29	33	85	32	28	27	37	20,4	24	24	18	32	26	41	53	61	61
poids sec y	3,6	3,4	5,6	4,1	3,3	7,5	3,7	3,2	2,9	4,5	2,6	2,8	3,1	2,6	4,4	2,8	4,6	6	7,2	6,3

Mêmes questions 1 à 4 que dans l'exercice 4.

Exercice 8. Lors d'une étude sur la dynamique des populations de la tenthrede du pin (*Diprion fruterarum*), la capacité de reproduction (nombre y d'ovocytes par cocon) des individus de cet insecte a été étudiée en fonction de la longueur du cocon (x en mm).

1. Calculer les distributions marginales de x et y .
2. Calculer la moyenne et l'écart type de x et y .
3. Calculer la covariance et le coefficient de corrélation puis déterminer l'équation de la droite de régression de y en x .

$y \backslash x$	[20 ; 40[	[40 ; 60[	[60 ; 80[	[80 ; 100[
[7 ; 7,5[	8	2		
[7,5 ; 8[	11	15	1	1
[8 ; 8,5[		22	8	0
[8,5 ; 9[	1	9	10	0
[9 ; 9,5[		12	19	11
[9,5 ; 10[			10	20

Exercice 9. On a traité 80 parcelles de 0,4 hectares avec différentes quantités d'un engrais azoté pour la culture d'orge. Le tableau suivant donne la répartition de ces parcelles en fonction de la quantité d'engrais (x en kg) et du rendement obtenu (y en quintaux/hectare).

1. Calculer les distributions marginales de x et y .
2. Calculer la moyenne et l'écart type de x et y .
3. Calculer la covariance et le coefficient de corrélation puis déterminer l'équation de la droite de régression de y en x .

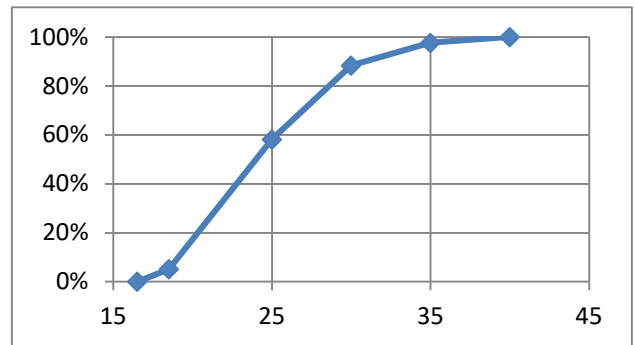
$y \backslash x$	50	60	70	80	90	100
[15;25[	1					
[25;35[	3	1				
[35;45[	1	3				
[45;55[		2	7		1	
[55;65[		4	8	8	6	
[65;75[			5	12	7	2
[75;85[			2	6	1	

Corrigé exercice 1. Moyenne 38,65 Écart type 0,387 Médiane 38,5

Corrigé exercice 2.

1.

	Moyenne	Écart type	Médiane
poids	66,76923	16,030	65
taille	1,636154	0,107	1,65
IMC	25,0079	6,089	24,80159



2.

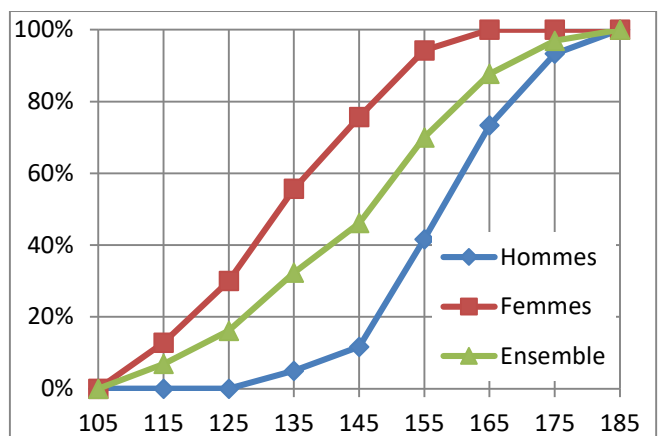
- a. $Q_1 = 20,9$ $Q_2 = 24,0$ $Q_3 = 27,8$
 b. $P(IMC \leq 26) = 64,19\%$
 c. $P(IMC \geq 33,5) = 5,12\%$
 d. $P(26 \leq IMC \leq 33,5) = 30,70\%$
 e. $m = 24,6$ $s = 4,4$.

	Classe	effectif	Fréq.	Fréq. cum.	Densité x100
Maigreur	[16,5;18,5[	11	5,12%	5,12%	1,5
normal	[18,5;25[	114	53,02%	58,14%	8,2
risque de surpoids	[25;30[	65	30,23%	88,37%	6,0
obésité modérée	[30;35[	20	9,30%	97,67%	1,9
obésité sévère	[35;40[	5	2,33%	100%	0,5
	Total	215	100%		

Corrigé exercice 3.

Classes	[105;115[	[115;125[	[125;135[	[135;145[	[145;155[	[155;165[	[165;175[	[175;185[
Eff. hommes n_i^H	0	0	3	4	18	19	12	4
Eff. femmes n_i^F	9	12	18	14	13	4	0	0
Ef. ensemble n_i^E	9	12	21	18	31	23	12	4
Fréq. hommes f_i^H	0%	0%	5,00%	6,67%	30,00%	31,67%	20,00%	6,67%
Fréq. femmes f_i^F	12,86%	17,14%	25,71%	20,00%	18,57%	5,71%	0,00%	0,00%
Fréq. ensemble f_i^E	6,92%	9,23%	16,15%	13,85%	23,85%	17,69%	9,23%	3,08%
Fréq. cum. hommes F_i^H	0%	0%	5,00%	11,67%	41,67%	73,33%	93,33%	100%
Fréq. cum. femmes F_i^F	12,86%	30,00%	55,71%	75,71%	94,29%	100%	100%	100%
Fréq. cum. ensemble F_i^E	6,92%	16,15%	32,31%	46,15%	70,00%	87,69%	96,92%	100%

Sexe	H	F	H	F	H	F
t	130	152				
$\mathbb{P}(X \leq t)$	2,5%	42,86%	32,7%	88,7%	30,2%	45,9%



	Effectif	Quartile 1	Quartile 2	Quartile 3	Moyenne	Variance	Écart type
Hommes	60	149,4	157,6	165,8	157,50	142,08	11,92
Femmes	70	122,1	132,8	144,6	133,14	204,41	14,30
Ensemble	130	130,5	146,6	157,8	144,38	323,08	17,97

Corrigé exercice 2 (suite).

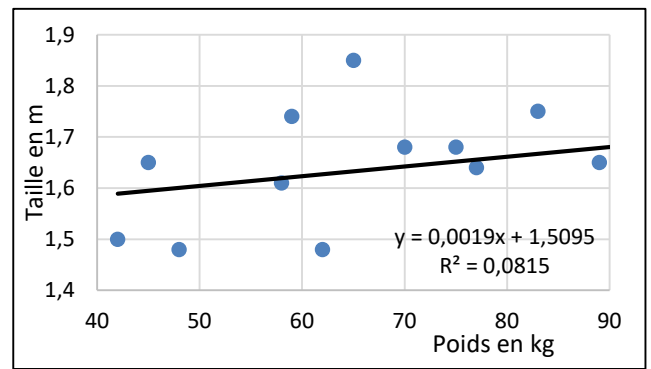
$$C(x, y) = 0,4875$$

$$r(x, y) = \frac{0,4875}{16,03 \times 0,1065} = 0,2856 \quad (r^2 = 0,0815).$$

penste $a = 1,898 \cdot 10^{-3}$ ordonnée origine $b = 1,509$.

$$y = 0,001898 \cdot x + 1,509.$$

La corrélation linéaire est faible, le nuage de point est très dispersé autour de la droite de régression. Il n'est pas raisonnable de faire des estimations de taille en fonction du poids en utilisant cette droite.



Corrigé exercice 4.

$$\bar{x} = 20 \quad \bar{y} = 34$$

$$v(x) = 323,43 \quad v(y) = 754,29$$

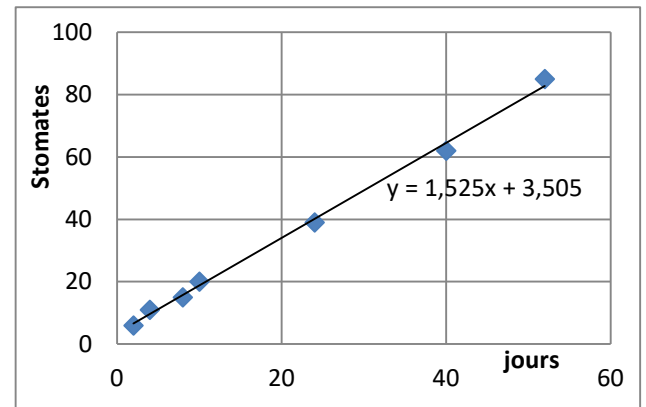
$$s(x) = 17,98 \quad s(y) = 27,46$$

$$c(x, y) = 493,14 \quad r(x, y) = 0,9984$$

penste $a = 1,525$ ordonnée origine $b = 3,505$.

$$y = 1,525x + 3,505.$$

La corrélation linéaire est forte : le nuage de point est concentré le long de la droite de régression. Il est raisonnable de faire des estimations de nombre de stomates en fonction du nombre de jours en utilisant cette droite. Pour 30 jours, on prévoit aux alentours de 49 stomates ($y = 1,525 \times 30 + 3,505 \approx 49,26$).



Corrigé exercice 5.

$$\bar{x} = 1969,0 \quad \bar{y} = 2,03$$

$$s(x) = 644,66 \quad s(y) = 3,48$$

$$c(x, y) = -2229,97 \quad r(x, y) = -0,9936$$

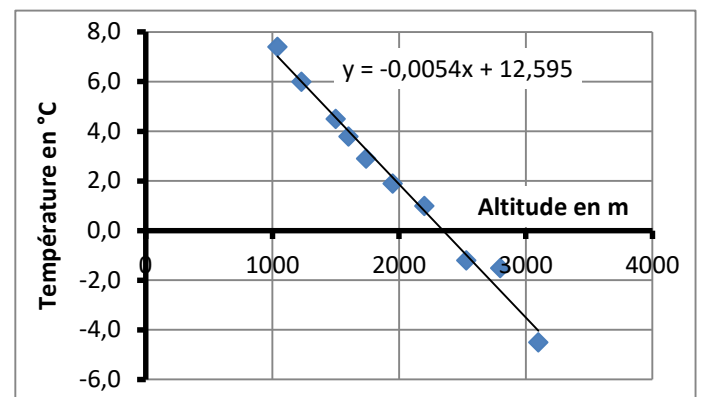
penste $a = -0,005365$

ordonnée origine $b = 12,595$.

$$y = -0,005365x + 12,595$$

Forte corrélation de penste négative.

Pour $x = 1300$, $y \approx 5,62$ et pour $x = 3000$, $y \approx -3,50$



Corrigé exercice 6.

Mois t	1	2	3	4	5	6	7	8	9	10	11	12
Survivants N	180	152	145	125	109	95	82	76	58	51	35	31
$y = \ln(N)$	5,19	5,02	4,98	4,83	4,69	4,55	4,41	4,33	4,06	3,93	3,56	3,43

$$\bar{t} = 6,5$$

$$\bar{y} = 4,416$$

$$s(t) = 3,452$$

$$s(y) = 0,5497$$

$$c(x, y) = -1,873$$

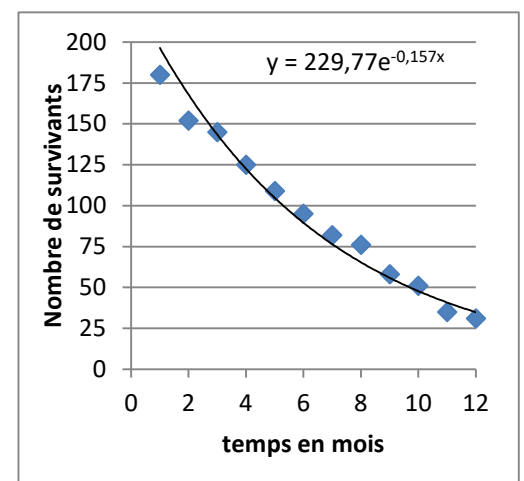
$$r(x, y) = -0,987$$

Penste : $-0,157$ ordonnée origine : $5,437$.

$$y = \ln(N) = -0,157t + 5,437$$

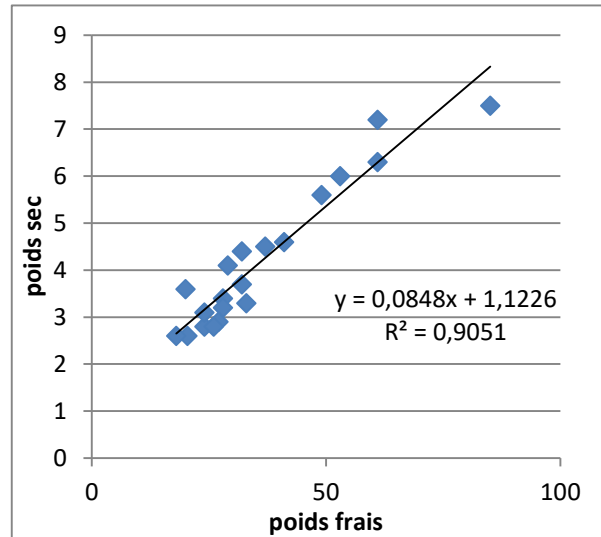
$$N = 229,77 e^{-0,157t}$$

Pour $t = 18$, $N = 13,6$



Corrigé exercice 7.

$\bar{x} = 36,42$ $\bar{y} = 4,21$
 $v(x) = 282,09$ $v(y) = 2,2399$
 $s(x) = 16,80$ $s(y) = 1,500$
 $c(x, y) = 23,91$
 $r(x, y) = 0,9513$
 pente $a = 0,08477$
 ordonnée origine $b = 1,123$.
 $y = 0,08477x + 1,123$.



Corrigé exercice 8.

$y \backslash x$	30	50	70	90	
7,25	8	2			10
7,75	11	15	1	1	28
8,25		22	8	0	30
8,75	1	9	10	0	20
9,25		12	19	11	42
9,75			10	20	30
	20	60	48	32	160

$m(x) = 8,71$ $m(y) = 61,5$
 $s(x) = 0,783$ $s(y) = 18,9$
 $c(x, y) = 11,32$ $r(x, y) = 0,764$
 pente $a = 18,438$ ordonnée origine $b = -99,02$
 $y = 18,438x - 99,02$

Corrigé exercice 9.

$m(x) = 75,25$ $m(y) = 61,25$
 $s(x) = 11,83$ $s(y) = 13,17$
 $c(x, y) = 102,19$ $r(x, y) = 0,6559$
 pente $a = 0,730$ ordonnée origine $b = 1,47$
 $y = 0,730x + 1,47$

$y \backslash x$	50	60	70	80	90	100	
[15;25[	1						1
[25;35[	3	1					4
[35;45[	1	3					4
[45;55[		2	7		1		10
[55;65[		4	8	8	6		26
[65;75[			5	12	7	2	26
[75;85[			2	6	1		9
	5	10	22	26	15	2	80