Université de Picardie Jules Verne

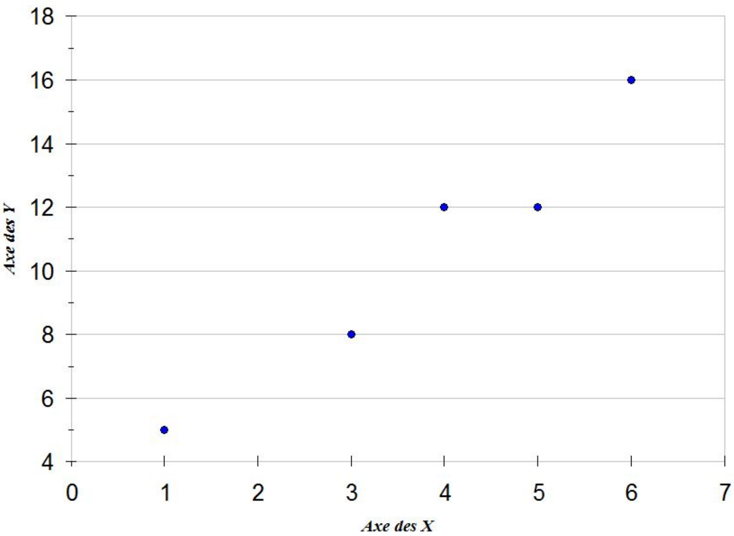
UFR d'économie et de gestion

Statistiques

Licence 1 - Semestre 1

Exercices d'entrainement **Feuille 4** **Corrigés**

Correction 1

1. 

2. $\operatorname{cov}(X,Y) = E(XY) - E(X)E(Y)$ où $E(Z)$ désigne la moyenne arithmétique de Z.

   On complète le tableau suivant :

	X_i	Y_i	X_iY_i	X_i^2	Y_i^2
	1	5	5	1	25
	3	8	24	9	64
	4	12	48	16	144
	5	12	60	25	144
	6	16	96	36	256
Somme	19	53	233	87	633
Moyenne	3.8	10.6	46.6	17.4	126.6

   $\operatorname{cov}(X,Y) \approx 46,6 - 3,8 \times 10,6 \approx 6,32$

3. $r_p = \dfrac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \dfrac{\operatorname{cov}(X,Y)}{\sqrt{V(X)V(Y)}}$.

   Plus le coefficient est proche de -1 ou 1, plus la corrélation est forte.

   $V(X) = E(X^2) - E(X)^2 \approx 17,4 - (3,8)^2 \approx 2,96$

   $V(Y) = E(Y^2) - E(Y)^2 \approx 126,6 - (10,6)^2 \approx 14,24$

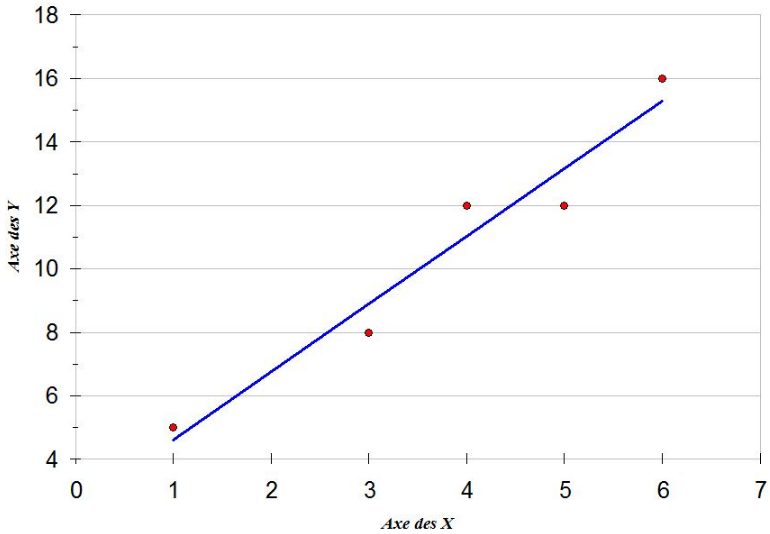
   $r_p \approx \dfrac{6,32}{\sqrt{2,96 \times 14,24}} \approx 0,97$

4. La droite de régression de Y en X est la droite qui passe par le point moyen et de coefficient directeur $\dfrac{\text{cov}(X,Y)}{V(X)}$.

   Son équation est $y = \dfrac{\text{cov}(X,Y)}{V(X)}(x - E(X)) + E(Y)$ ou encore $y = ax + b$ avec :

   $a = \dfrac{\text{cov}(X,Y)}{V(X)} \approx \dfrac{6,32}{2,96} \approx 2,14$ et $b = E(Y) - a \times E(X) \approx 10,6 - 2,14 \times 3,8 \approx 2,47$

   Le point moyen est $G(3,8\,;\,10,6)$. On a juste besoin d'un autre point pour tracer la droite. En prenant $x = 1$, par exemple, on obtient $y = 1 \times 2,14 + 2,47 = 4,61$.

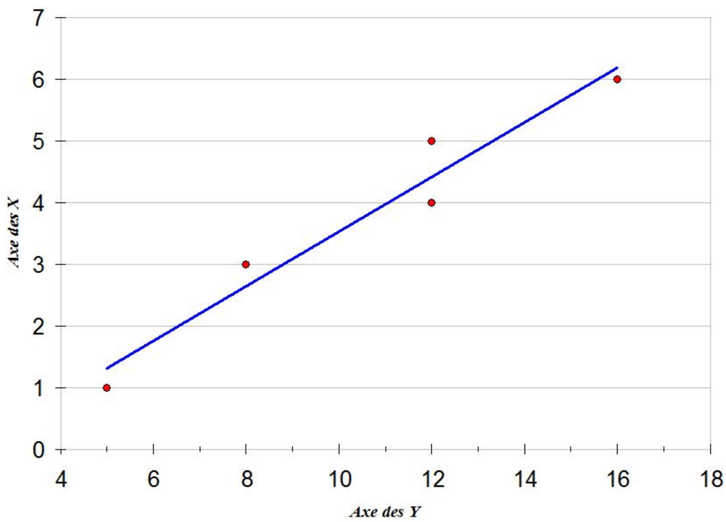


   La droite de régression de X en Y est la droite qui passe par le point moyen et de coefficient directeur $\dfrac{\text{cov}(X,Y)}{V(Y)}$.

   Son équation est : $y = \dfrac{\text{cov}(X,Y)}{V(Y)}(x - E(Y)) + E(X)$ ou encore $y = ax + b$ avec :

   $a = \dfrac{\text{cov}(X,Y)}{V(Y)} \approx \dfrac{6,32}{14,24} \approx 0,44$ et $b = E(X) - a \times E(Y) \approx 3,8 - 0,44 \times 10,6 \approx -0,86$

   Le point moyen est $G(3,8\,;\,10,6)$. On a juste besoin d'un autre point pour tracer la droite. En prenant $x = 5$, par exemple, on obtient $y = 5 \times 0,44 - 0,86 = 1,34$.

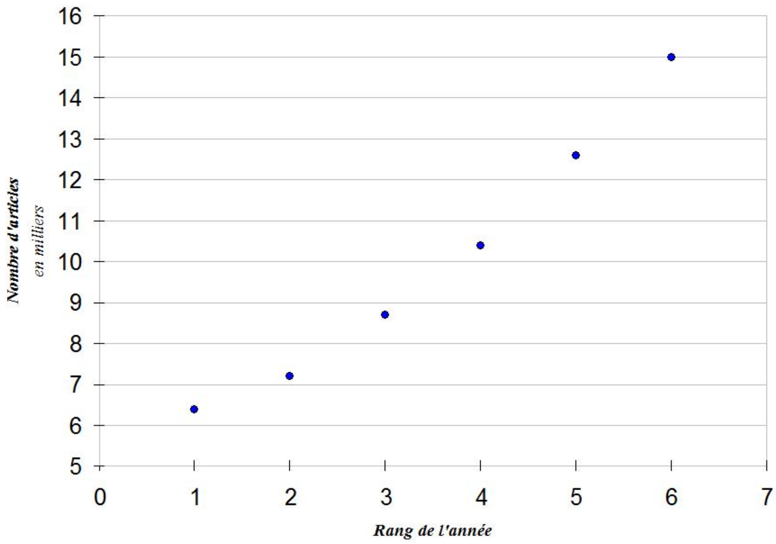


5. On sépare la série en 2 demi-séries d'effectifs égaux (ou à peu près égaux dans le cas impair). La droite de Mayer de Y en X est la droite qui passe par les points moyens de la première demi-série et de la deuxième demi-série.

	Demi-série 1		Demi-série 2	
	X	Y	X	Y
	1	5	5	12
	3	8	6	16
	4	12		
Moyenne	$8/3$ $\approx 2,66$	$25/3$ $\approx 8,33$	5,5	14

Les points sont donc $G_1(2,66; 8,33)$ et $G_2(5,5; 14)$.
Le coefficient directeur est $a = \dfrac{14 - 25/3}{5,5 - 8/3} = \dfrac{17}{8,5} = 2$
Et l'ordonnée à l'origine est $b = 25/3 - 2 \times 8/3 = 3$.

Correction 2

1. 

2. Rappel : $r_p = \dfrac{\operatorname{cov}(t,Q)}{\sigma_t \sigma_Q} = \dfrac{\operatorname{cov}(t,Q)}{\sqrt{V(t)V(Q)}}$

	t_i	Q_i	t_iQ_i	t_i^2	Q_i^2
	1	6 400	6 400	1	40 960 000
	2	7 200	14 400	4	51 840 000
	3	8 700	26 100	9	75 690 000
	4	10 400	41 600	16	108 160 000
	5	12 600	63 000	25	158 760 000
	6	15 000	90 000	36	225 000 000
Somme	21	60 300	241500	91	660 410 000
Moyenne	3,5	10 050	40 250	15,17	110 068 333,33

$\operatorname{cov}(t,Q) = E(tQ) - E(t)E(Q) \approx 40250 - 3,5 \times 10\,050 \approx 5\,075$

$V(t) = E(t^2) - E(t)^2 \approx 15,17 - (3,5)^2 \approx 2,92$

$V(Q) = E(Q^2) - E(Q)^2 \approx 110\,068\,333,33 - (10\,050)^2 \approx 9\,065\,833,33$

$r_p \approx \dfrac{5\,075}{\sqrt{2,92 \times 9\,065\,833}} \approx 0,98$

3. Rappel : La notation log correspond au logarithme décimal.

   On a $\log x = \log_{10}(x) = \dfrac{\ln x}{\ln 10} = \dfrac{\mathrm{Log}\, x}{\ln 10}$.

	t_i	$\log(Q_i)$	$t_i \log(Q_i)$	t_i^2	$(\log(Q_i))^2$
	1	3.81	3.81	1	14.49
	2	3.86	7.71	4	14.88
	3	3.94	11.82	9	15.52
	4	4.02	16.07	16	16.14
	5	4.1	20.5	25	16.81
	6	4.18	25.06	36	17.44
Somme	21	23.9	84.97	91	95.28
Moyenne	3,5	3,9828	14,161	15,17	15,8792

$r_p = \dfrac{\mathrm{cov}\,(t, \log Q)}{\sigma_t \sigma_{\log Q}} = \dfrac{\mathrm{cov}\,(t, \log Q)}{\sqrt{V(t)V(\log Q)}}$.

$\mathrm{cov}\,(t, Q) = E(t\log(Q)) - E(t)E(\log(Q)) \approx 14,161 - 3,5 \times 3,9828 \approx 0,2212$

$V(t) = E(t^2) - E(t)^2 \approx 15,1667 - (3,5)^2 \approx 2,9167$.

$V(\log(Q)) = E(\log^2(Q)) - E(\log(Q))^2 \approx 15,8792 - (3,9828)^2 \approx 0,0169$

$r_p \approx \dfrac{0,2212}{\sqrt{2,9167 \times 0,0169}} \approx 0,9963$

4. L'équation de la droite de régression de $\log(Q)$ en t est $y = ax + b$ où

   $a = \dfrac{\mathrm{cov}\,(t, \log Q)}{V(t)} \approx \dfrac{0,2212}{2,9167} \approx 0,076$

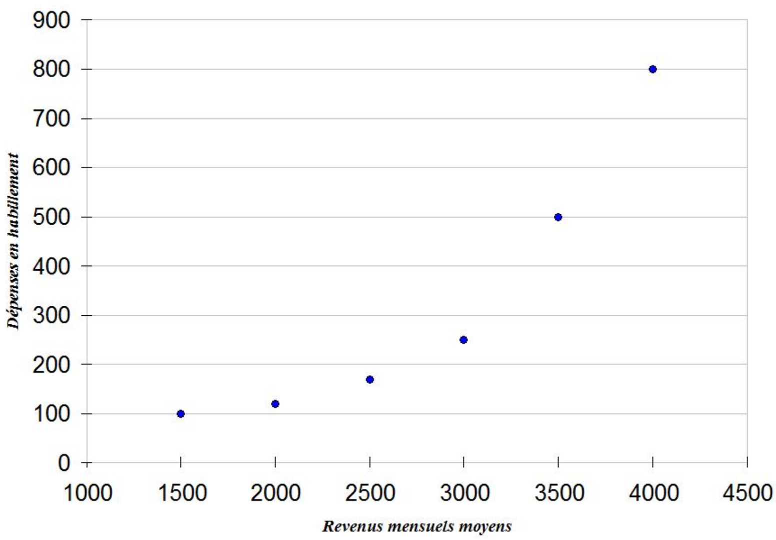
   $b = E(\log(Q)) - a \times E(t) \approx 3,9828 - 0,076 \times 3,5 \approx 3,7174$

   On a $\log(Q) = at + b$ c'est-à-dire $Q = 10^{at+b}$. L'année 2016 correspond à $t = 7$.

   Pour $t = 7$, on obtient $\log(Q) \approx 4,25$ et donc $Q \approx 17782$.

Correction 3

1.



2.

	X_i ($\times 10^3$)	Y_i	X_iY_i ($\times 10^3$)	X_i^2 ($\times 10^6$)	Y_i^2
	1,5	100	150	2,25	10
	2	120	240	4	14,4
	2,5	170	425	6,25	28,9
	3	250	750	9	62,5
	3,5	500	1 750	12,25	250
	4	800	3 200	16	640
Somme	165	1 940	6,515	49 750	1 005,8
Moyenne	2,750	323,33	1 085,83	8,2917	167,63

$\mathrm{cov}\,(X,Y) = E(XY) - E(X)E(Y) \approx 1\,085\,833 - 2\,750 \times 323,33 \approx 196\,666,67$

$V(X) = E(X^2) - E(X)^2 \approx 8\,291\,666,67 - (2\,750)^2 \approx 729\,166,67$

$V(Y) = E(Y^2) - E(Y)^2 \approx 167\,633,33 - (323,33)^2 \approx 63\,088,89$

$r_p \approx \dfrac{196\,666,67}{\sqrt{729\,166,67 \times 63\,088,89}} \approx 0,917$

3. Rappel : $\log x = \log_{10}(x) = \dfrac{\ln x}{\ln 10} = \dfrac{\mathrm{Log}\, x}{\ln 10}$.

	$\log(X_i)$	$\log(Y_i)$	$\log(X_i)\log(Y_i)$	$(\log(X_i))^2$	$(\log(Y_i))^2$
	3,1761	2,0000	6,3522	10,0876	4,0000
	3,3010	2,0792	6,8634	10,8968	4,3230
	3,3979	2,2304	7,5789	11,5460	4,9749
	3,4771	2,3979	8,3379	12,0904	5,7501
	3,5441	2,6990	9,5653	12,5604	7,2844
	3,6021	2,9031	10,4571	12,9748	8,4279
Somme	20,4983	14,3096	49,1549	70,1560	34,7604
Moyenne	3,4164	2,3849	8,1925	11,6927	5,7934

$r_p = \dfrac{\mathrm{cov}\,(\log X, \log Y)}{\sigma_X \sigma_{\log Y}} = \dfrac{\mathrm{cov}\,(\log X, \log Y)}{\sqrt{V(\log X)V(\log Y)}}$.

$\mathrm{cov}\,(\log X, \log Y) = E(\log X \log Y) - E(\log X)E(\log Y) \approx 0,0446$

$V(\log X) = E((\log X)^2) - E(\log X)^2 \approx 11,6927 - (3,4164)^2 \approx 0,0210.$

$V(\log Y) = E(\log Y)^2) - E(\log(Y))^2 \approx 5,7934 - (2,3849)^2 \approx 0,1055$

$r_p \approx 0,9486$

4. L'équation de la droite de régression de $\log(Y)$ en X est $y = ax + b$ où

$a = \dfrac{\mathrm{cov}\,(X, \log Y)}{V(X)} \approx 2,1271$

$b = E(\log(Y)) - a \times E(X) \approx 2,3849 - 2,1271 \times 3,4164 \approx -4,8822$

On a $\log(Y) = a\log(X) + b$ c'est-à-dire $Y = 10^{a\log(X)+b}$.

Pour $X = 5\,000$, on obtient $\log(Y) \approx 2,98$ et donc $Y \approx 968$.