

5 Statistiques

Exercice 5.1 (Séries normalement distribuées) Une entreprise analyse l'ancienneté de son personnel. Le résultat est représenté par le tableau suivant, où n_i représente le nombre d'individus ayant entre x_i et x_{i+1} années d'ancienneté.

$]x_i, x_{i+1}]$	$]0,5]$	$]5,15]$	$]15,25]$	$]25,35]$	$]35,40]$
n_i	5	20	25	20	5

- (i) Tracer l'histogramme de la série X , ainsi que le graphe de sa fonction de répartition.
- (ii) Calculer la moyenne, la médiane et le(s) mode(s).
- (iii) Calculer l'écart-type de X et donner une estimation du pourcentage de la population qui ont plus de $\bar{x} - 2\sigma_X$ année d'ancienneté.
- (iv) Comparer la politique de recrutement de cette entreprise avec une autre entreprise dont l'ancienneté du personnel est de 24.5 ans et d'écart-type de 5.13 ans.

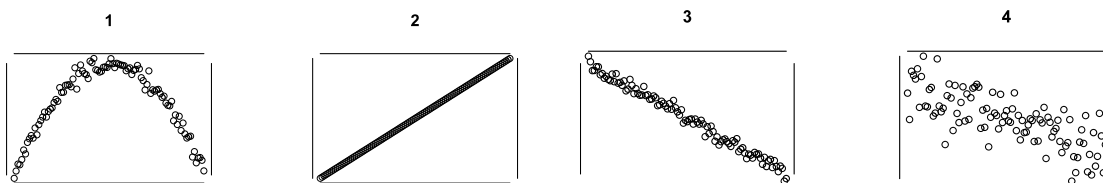
Indication: on supposera que la série Y du nombre d'années d'ancienneté du personnel de la seconde entreprise est normalement distribuée; ce qui permettra d'estimer à 97.72% le pourcentage de la population ayant plus de $\bar{y} - 2\sigma_Y$ années d'ancienneté.

Exercice 5.2 On souhaite comparer le taux de criminalité entre les états des États-Unis ayant ou non aboli la peine de mort. On donne les séries statistiques en annexe.

- (i) Quelles colonnes sont intéressantes ?
- (ii) Donner le calcul de la quatrième colonne.
- (iii) En moyenne, quel groupe a le taux de criminalité le plus élevé ?
- (iv) Tracer côte à côte les boîtes de distributions de chaque groupe.
- (v) Commenter. Peut-on établir un lien de cause à effet entre l'abolition de la peine de mort et la criminalité ?
- (vi) Si le District of Columbia abolissait la peine de mort, les commentaires précédents changeraient-ils fondamentalement ?

Exercice 5.3 Attribuer à chaque nuage de points de la figure ci dessous son coefficient de corrélation en justifiant votre choix.

$$r_1 = 0.034 \quad r_2 = -0.99 \quad r_3 = -0.741 \quad r_4 = 1.$$



Exercice 5.4 Dans une usine de peinture, des machines produisent des pots de manière continue à la sortie d'une cuve alimentée en composants nécessaires.

Une peinture est principalement composée de quatre éléments : le liant, le solvant, les pigments, la charge. Ce dernier élément donne l'opacité et donc le pouvoir couvrant de la peinture. On utilise ici de la craie mise en poudre conditionnée en sac de 50 kg.

Suite à un défaut de pureté, des agrégats calcaires se sont retrouvés dans un sac de craie. Ce sac a alimenté une machine.

Sur cette machine, on relève la quantité d'agrégats calcaires moyenne par cm^3 au cours du temps en minutes.

temps	t	0	1	2	3	4	6	8	15	20	25
agrégats	y	8	6,25	4,77	3,65	2,81	1,7	1,06	0,32	0,23	0,21
changement de variable	$z = \ln(y - 0,2)$	2.05	1.8	1.52	1.24	0.96	0.41	-0.15	-2.12	-3.51	-4.61

On souhaite modéliser l'évolution de la quantité d'agrégats par cm^3 dans les pots de peinture en fonction du temps exprimé en minutes.

- (i) Tracer rapidement le nuage de points (t, y) . On renonce à une régression linéaire, expliquer pourquoi. On effectue le changement de variable $z = \ln(y - 0,2)$.
- (ii) À l'aide de la fonction PENTE d'Excel, on obtient le coefficient de la régression linéaire de z sur t par la méthode des moindres carrés. Il vaut $-0,27$. Trouver l'équation de la droite de régression.
- (iii) À l'aide de la fonction COEFFICIENT.CORRELATION d'Excel, on obtient la valeur du coefficient de corrélation linéaire qui est $-0,996$. Qu'en déduire sur la légitimité du changement de variable ?
- (iv) Donner l'expression de y en fonction de t .

State	Population (en milliers)	Total murders (2009)	Nombre de meurtres pour 10000 habitants	Peine de Mort abolie en 2009
Vermont	622	7	0,113	oui
Iowa	3 008	34	0,113	oui
Minnesota	5 266	72	0,137	oui
North Dakota	647	9	0,139	oui
Hawaii	1 295	21	0,162	oui
Maine	1 318	26	0,197	oui
Wisconsin	5 655	144	0,255	oui
Rhode Island	1 053	31	0,294	oui
Alaska	698	22	0,315	oui
New Jersey	8 708	319	0,366	oui
West Virginia	1 820	76	0,418	oui
Michigan	9 970	625	0,627	oui
New Mexico	2 010	144	0,717	oui
Florida	18 538			
New Hampshire	1 325	10	0,075	
Utah	2 785	37	0,133	
South Dakota	812	11	0,135	
Idaho	1 546	22	0,142	
Wyoming	544	11	0,202	
Oregon	3 826	83	0,217	
Nebraska	1 797	40	0,223	
Washington	6 664	169	0,254	
Massachusetts	6 594	169	0,256	
Montana	975	28	0,287	
Connecticut	3 518	107	0,304	
Colorado	5 025	167	0,332	
Illinois	12 910	479	0,371	
Kentucky	4 314	170	0,394	
New York	19 541	779	0,399	
Kansas	2 819	118	0,419	
Ohio	11 543	502	0,435	
Virginia	7 883	347	0,440	
Indiana	6 423	293	0,456	
Delaware	885	41	0,463	
Arizona	6 596	328	0,497	
Mississippi	2 952	151	0,512	
North Carolina	9 381	480	0,512	
Pennsylvania	12 605	658	0,522	
California	36 962	1 972	0,534	
Texas	24 782	1 325	0,535	
Georgia	9 829	543	0,552	
Nevada	2 643	156	0,590	
Arkansas	2 889	171	0,592	
Oklahoma	3 687	225	0,610	
South Carolina	4 561	286	0,627	
Missouri	5 988	381	0,636	
Alabama	4 709	318	0,675	
Tennessee	6 296	461	0,732	
Maryland	5 699	438	0,768	
Louisiana	4 492	486	1,082	
District of Columbia	600	144	2,401	

Les états sont triés par abolition puis par nombre moyen de meurtres