

Feuille d'exercices n° 4

Séries statistiques doubles. Distributions marginales et conditionnelles.

Exercice 1 : Au cours d'une enquête portant sur des logements de construction récente dans les agglomérations de 20 000 habitants et plus, on a relevé, entre autres renseignements, le nombre de pièces principales (non compris cuisine, salle d'eau...) et la surface totale (y compris cuisine, ...) en m^2 . Voici ces renseignements en ce qui concerne 80 logements, numérotés de 1 à 80 (P : nombre de pièces principales ; S : surface en m^2).

N°	P	S	N°	P	S	N°	P	S	N°	P	S
1	5	73	21	3	78	41	5	72	61	4	47
2	4	75	22	4	69	42	3	68	62	3	63
3	3	66	23	4	73	43	5	89	63	5	84
4	4	51	24	6	94	44	5	71	64	3	34
5	3	58	25	10	163	45	5	78	65	4	68
6	4	62	26	3	58	46	6	117	66	4	71
7	3	93	27	4	62	47	5	62	67	3	59
8	5	110	28	5	121	48	1	24	68	6	90
9	4	84	29	4	79	49	4	76	69	4	65
10	4	64	30	4	98	50	4	60	70	4	67
11	4	61	31	2	46	51	4	66	71	4	71
12	4	68	32	3	70	52	3	59	72	4	66
13	5	66	33	3	51	53	4	76	73	4	72
14	5	92	34	6	78	54	5	95	74	5	94
15	5	91	35	4	66	55	3	77	75	6	104
16	5	80	36	3	47	56	4	138	76	5	88
17	5	84	37	4	63	57	4	67	77	9	133
18	4	91	38	5	161	58	2	55	78	7	92
19	3	53	39	3	62	59	5	79	79	2	32
20	4	87	40	6	80	60	4	82	80	7	112

- Télécharger dans le cours Arche le fichier Excel-feuille 4 contenant le tableau ci-dessus, et à l'aide d'un tableau croisé dynamique inséré sur une nouvelle feuille, afficher le nombre de logements par surface et par nombre de pièces (la surface S sera en champ de lignes et le nombre de pièces P en champ de colonnes). Copier ce tableau sur une feuille intitulée Ex1.1 (coller les valeurs).
- En modifiant les champs du tableau croisé dynamique dresser le tableau statistique de la distribution marginale des logements selon leur nombre de pièces : le champ P sera en champ de colonnes et les effectifs dans la 2ème ligne du tableau. Le copier dans une nouvelle feuille intitulée Ex1.2 (coller les valeurs).
Procéder de façon analogue pour dresser le tableau statistique de la distribution marginale des logements selon leur surface. Le champ S sera en champ de lignes. Le copier dans la feuille Ex1.2 également.
Déterminer la moyenne et la médiane de chacune de ces séries statistiques simples.
Donner une interprétation de la médiane dans une phrase citant les deux caractères en même temps.
- Copier la feuille contenant les données de l'exercice 1. La nommer Ex1.3. Insérer le nuage de points représentant la surface en fonction du nombre de pièces. Afficher la courbe de tendance linéaire, le coefficient de détermination ainsi que l'équation de cette droite.
Procéder de même avec le nuage de points représentant le nombre de pièces en fonction de la surface.
Peut-on dire que la corrélation linéaire est forte entre les caractères "S" et "P" ?
Le coefficient de détermination est le carré du coefficient de corrélation linéaire.

Exercice 2: Le tableau suivant donne le poids des feuilles (caractère X) et le poids des racines (caractère Y) de 1 000 *Cichorium antybus*, exprimés en grammes.

Poids des racines, Y	[40,120[[120,200[[200,280[[280,360[Total
Poids des feuilles, X					
[0,160[97	7	0	0	104
[160,320[403	171	7	0	581
[320,480[60	185	23	1	269
[480,640[1	18	19	1	39
[640,800[0	1	4	2	7
Total	561	382	53	4	1000

- Donner la série des poids des feuilles conditionnellement à la classe [40,120[du poids des racines : dans le tableau de l'énoncé, ne garder que la modalité [40,120[du caractère Y .
Faites de même avec chaque autre modalité du caractère Y .
- Calculer la variance et l'écart type de chacune de ces 4 séries. Conclusion ?