

FEUILLE DE TRAVAUX DIRIGÉS N° 2

Enseignant-Formateur : H. El-Otmany

A.U. : 2019-2020

Exercice n°1 Ce document est issu des statistiques de l'INSEE et intitulé "La population française selon la nationalité et le lieu de naissance en 1999" (effectifs en milliers) :

Lieu de naissance \ Nationalité	France	Étranger
	France	Étranger
Français de naissance	51 340	1 560
Français par acquisition	800	1 560
Étrangers	510	2 750

1. Quelle est la population ? Sa taille ?
2. Quelles sont les variables étudiées ? Leurs types ?
3. Compléter la table de contingence précédente avec les distributions marginales.
4. Déterminer les distributions conditionnelles des deux variables. Laquelle nous permet de déterminer si le pourcentage de personnes nées en France est plus fort parmi les Français par acquisition ou parmi les étrangers ?
5. Représenter graphiquement la distribution conjointe des deux variables.
6. Représenter graphiquement les distributions conditionnelles aux modalités de chacune des deux variables.
7. Est-il plus fréquent d'acquérir la nationalité française lorsque l'on est né en France ou bien lorsque l'on est né à l'étranger ? (Quel graphique permet cette conclusion ?)
8. Déterminer l'indice du χ^2 et le coefficient C de Cramer. Commenter.

Exercice n°2 On s'intéresse aux tailles en mètre d'étudiants masculins de trois pays européens.

Pays	Tailles (x_i)	$\sum x_i$	$\sum x_i^2$
Allemagne	1.79 1.83 2.01 1.92 1.86 1.75 1.86 1.94 1.69 1.87 1.74 1.78 1.93 1.75 1.78 1.80 2.01 2.02 1.70 1.92	36.95	68.4665
France	1.76 1.84 1.93 1.72 1.70 1.91 1.82 1.75 1.67 1.72 1.86 1.72 1.96 1.69 1.90 1.71 1.87 1.69 1.81 1.83 1.68 1.97 1.88	41.39	74.6923
Italie	1.84 1.77 1.79 1.77 1.79 1.76 1.75 1.80 1.75 1.78 1.75 1.77 1.81 1.78 1.73 1.68 1.77 1.70	31.79	56.1687

1. Quels sont les individus étudiés ? les étudiants
2. Quelles sont les variables étudiées ? Préciser le type et leur domaine de définition ou modalités.
3. Compléter le tableau suivant relatif à la taille des étudiants :

	Allemagne	France	Italie	Total
Moyenne				
Variance				

4. Calculer la moyenne des variances conditionnelles Intra-classes.

5. Calculer la variance des moyennes conditionnelles Inter-classes.
6. Calculer le rapport de corrélation r_{XY} .
7. Interpréter le résultat.

Exercice n°3 Les données ci-dessous sont les offres de retransmission sportives par chaîne de télévision (TF1, A2, FR3, LA5) exprimées en millièmes pour différentes disciplines lors de l'année 1990 (Foot, Rugby, Tennis, Sport d'hiver, Cyclisme, Golf, Sport automobile, Athlétisme, Autres sports).

Sport	TF1	A2	FR3	LA5
Foot	595	139	104	58
Rugby	0	132	13	16
Tennis	0	294	547	547
Hiver	37	44	69	0
Vélo	0	171	29	153
Golf	0	14	35	0
Auto.	166	2	15	208
Athlé.	1	49	18	0
Autres	201	155	170	18

1. Le tableau ci-dessus est un "tableau de profils". Lequel ?
2. Quelles sont les variables étudiées ? Préciser le type et leur domaine de définition ou modalités.
3. Le tableau suivant est le tableau construit sous l'hypothèse d'indépendance. Compléter la première colonne.

Sport	TF1	A2	FR3	LA5
Foot				
Rugby		40.25		
Tennis			347.00	
Hiver				
Vélo				
Golf				12.25
Auto.				
Athlé.			17	
Autres				

4. Quelle particularité possède-t-il compte tenu de la question 1 ?
5. Le tableau suivant est le tableau des contributions de chaque croisement au χ^2 :

Sport	TF1	A2	FR3	LA5
Foot	?	32.2545	64.2857	123.0179
Rugby	40.2500	209.1444	18.4488	14.6102
Tennis	347.0000	8.0951	115.2738	115.2738
Hiver	0.0067	1.1267	26.4600	37.5000
Vélo	88.2500	77.5928	39.7797	47.5078
Golf	12.2500	0.2500	42.2500	12.25
Auto.	47.6528	93.7909	70.0518	124.3485
Athlé.	15.0588	60.2353	0.0588	17.0000
Autres	31.0662	2.6544	8.5000	102.3824

- Retrouver la valeur manquante ?
- Calculer χ^2 et l'indice de Cramer
- Peut-on considérer qu'il y a un lien les deux variables étudiées ? (justifier la réponse)

Exercice n°4 On donne pour les six premiers mois de l'année 1982 les nombres d'offres d'emploi (concernant les emplois durables et à plein temps) et le nombre des demandes d'emploi (déposées par des personnes sans emploi, immédiatement disponibles et à la recherche d'un emploi durable et à plein temps). Les données sont exprimées en milliers d'individus.

Offres (x)	61	66.7	75.8	78.6	82.8	87.2
Demandes (y)	2034	2003.8	1964.5	1928.2	1885.3	1867.1

- Représenter le nuage de points. Le nuage de points vous semble-t-il aligné le long d'une droite ?
- Trouver la droite de régression des demandes d'emploi en fonction des offres d'emploi et la tracer sur le graphique précédent.
- Calculer le coefficient de corrélation entre x et y . Commenter.

Exercice n°5 Une étude de marché a permis de relever le volume (y) des ventes (en milliers d'euros) d'un produit en fonction de son prix (x) de vente (en euros) :

x	95	130	146	210	250	280	330	350
y	104	58	37	22	12	12	9	7

- Donner le signe du coefficient de corrélation linéaire ρ_{xy} sans effectuer de calculs. En déduire le signe du coefficient de corrélation linéaire $\rho_{\ln(x)\ln(y)}$, toujours sans effectuer de calculs. Calculer le coefficient de corrélation de Spearman.
- On fait l'hypothèse d'un modèle de la forme $y = kx^{-m}$ ($k > 0$). Le coefficient m s'appelle, en économie, le coefficient d'élasticité. Estimer ce coefficient en se ramenant à un modèle linéaire. Prévoir le volume des ventes si le prix de vente est fixé à 360 euros.
- Quel aurait été l'effet sur l'estimation des coefficients k et m pour la même étude mais avec des prix exprimés en francs ?