

DATE : ...12.../...1.../...2023.....à.....11h
 EPREUVE: Statistiques et informatique appliquées aux Sciences sociales
 ENSEIGNANT: ...Renaud Orain <renaud.orain@univ-paris1.fr>
 DIPLOME : Licence de Sciences sociales, 3^e année
 Étudiant-es de tous parcours, inscrit-es en contrôle terminal...
 DUREE : ...3. h
 DOCUMENT AUTORISE Calculatrice. Aucun document autorisé (cf. rappels en fin de sujet)

Les temps écoulés entre la fin des études et un premier emploi en mois se distribuent de la façon suivante pour les personnes ayant achevé leurs études en 2013, enquêtées en 2017. Les temps supérieurs à 12 mois sont regroupés en classe et la moyenne exacte de la classe ainsi que la somme de ces temps par classe pour le calcul de la moyenne.

Variable d'analyse : temps d'accès au premier emploi (TAPE)					
Nombre de mois (en classes)	Effectif	Moyenne exacte	Sommes des durées par classe	Sommes des carrés des durées par classe	Amplitudes des classes
0	9085	0	0	0	
1 à 3	3373	1,983	6687	15151	3
4 à 6	1571	4,887	7678	38614	3
7 à 12	1431	9,126	13059	123355	6
13 à 24	1026	17,211	17658	316184	12
Plus de 2 ans	352	29,384	10343	308437	17
Ensemble	16838		55425	801741	

Question 1 : Évaluez la moyenne et l'écart-type de cette distribution, en complétant si besoin le tableau. Tracez un histogramme de la distribution^a et indiquez quelle est la classe modale. Évaluez^a au moyen d'un diagramme cumulé croissant ou d'un calcul d'interpolation la médiane. On trouve de manière similaire les premier (1,72) et troisième (9,65) quartiles. Enfin, dessinez la boîte à moustache de cette distribution^a (en prenant 41 mois comme valeur maximale).

^a (toujours) en écartant les personnes déjà en emploi ou ayant mis moins d'un mois à en trouver un

Question 2 : On étudie la façon dont la durée des études en nombre d'années^b se répercute sur le temps d'accès au premier emploi. Les 815 personnes n'ayant pas (encore, au moment de l'enquête) accédé à un premier emploi ne sont pas incluses dans l'analyse.

Étudiez le tableau suivant puis évaluez la covariance et la corrélation entre le temps d'accès à l'emploi en mois et la durée des études.

Déduisez-en les coefficients^c de la régression linéaire de ce temps d'attente en fonction des notes au baccalauréat anticipé, ainsi que le coefficient de détermination de la régression.

^b Il s'agit d'une « proxy » de cette durée calculée en faisant la différence entre l'âge en 2013 et 6 ans

^c Vous pourrez en déduire la réduction moyenne du temps d'accès à l'emploi (en choisissant votre unité de temps) apportée par un an d'études en plus.

Variable	Sommes des produits		Statistiques simples					
	Dur études	Tps accès	N	Moyenne	Ecart-type	Somme	Min.	Max.
Durée études	801741	867954	16838	3,29166	6,06483	55425	0	41
Temps accès	867954	4928455	16838	16,75454	3,46184	282113	8	47

Champ : personnes ayant achevé leurs études en 2013 et ayant accédé au moins une fois à un emploi depuis ; Source : enquête du CEREQ « Génération » 2013.

Question 3 : On utilise à présent un résumé des trajectoires d'emploi après la fin des études et on souhaite étudier la façon dont le niveau d'étude se répercute sur cette trajectoire résumé.

Complétez les différents tableaux ci-dessous, calculez le PEM et le Khi2 ainsi que le V de Cramer du tableau initial et tracez le graphique des écarts à l'indépendance manquant.

Question 4 : Commentez l'ensemble de vos résultats. Pour la corrélation, vous pourrez utiliser le résultat du test de Student qui indique que la probabilité d'obtenir la corrélation observée si le lien était dû au seul hasard de l'échantillonnage statistique est inférieure à 0,0001.

Répartition des personnes ayant fini leurs études selon leur trajectoire d'emploi sur 4 ans résumée, ainsi que selon le niveau d'étude atteint auparavant

Effectifs	EDI (rapide)	EDI (différé)	EDD	Inactivité	Chômage	Formation(s)	Ensemble
Collège ou inférieur	20	83	21	65	71	22	282
CAP BEP	264	1056	214	312	379	116	2341
Bac Pro	352	1388	272	415	378	139	2944
Bac techno	42	110	42	55	53	30	332
Bac Gén	16	36	13	34	26	12	137
BTS DUT Bac+2	223	872	153	186	107	84	1625
Bac+2/3 Santé Social	248	2340	25	127	22	44	2806
Lic Pro	198	581	59	93	26	32	989
Licence	122	633	122	244	187	100	1408
Autres Bac+3/4	17	85	15	10	8	1	136
Bac+5	283	1101	169	223	88	39	1903
Ec Ing + Commerc	182	461	36	56	14	11	760
Doctorats	271	1256	171	179	76	37	1990
Total	2238	10002	1312	1999	1435	667	17653

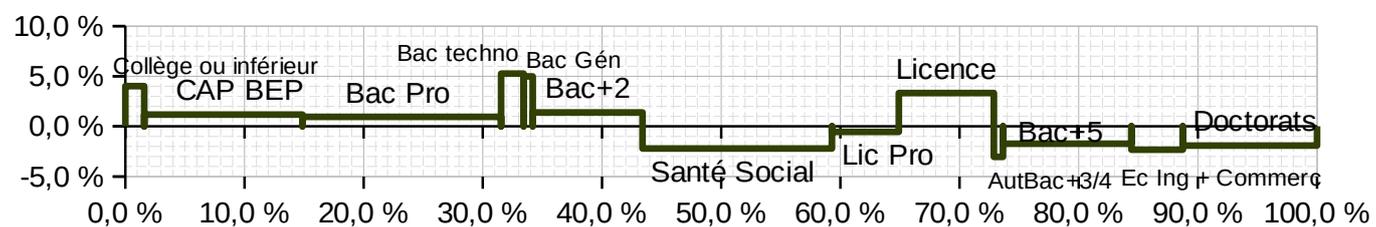
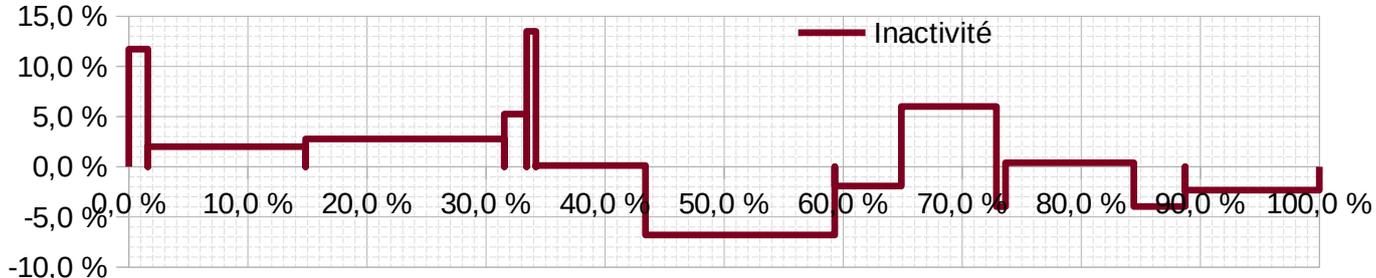
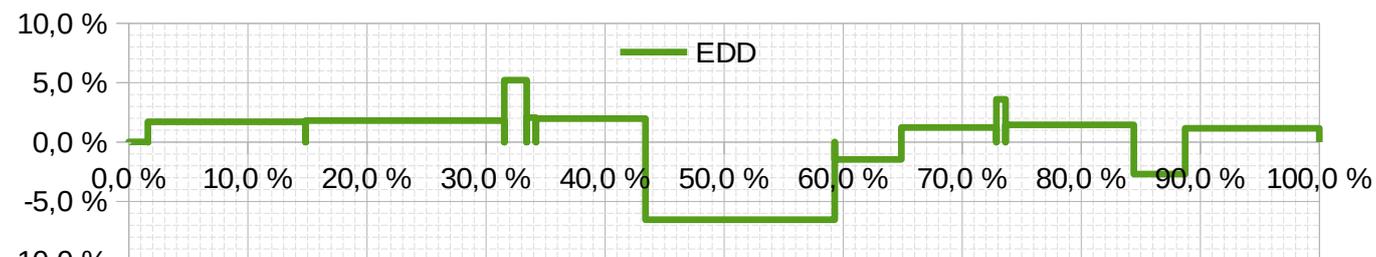
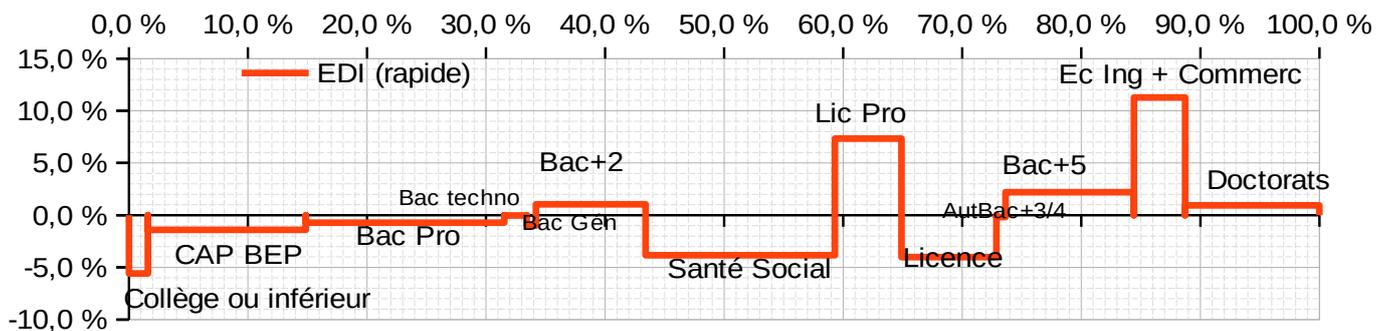
%ligne	EDI (rapide)	EDI (différé)	EDD	Inactivité	Chômage	Formation(s)	Ensemble
Collège ou -	7,1 %	29,4 %	7,4 %	23,0 %	25,2 %	7,8 %	100,0 %
CAP BEP	11,3 %	45,1 %	9,1 %	13,3 %	16,2 %	5,0 %	100,0 %
Bac Pro							100,0 %
Bac techno	12,7 %	33,1 %	12,7 %	16,6 %	16,0 %	9,0 %	100,0 %
Bac Gén	11,7 %	26,3 %	9,5 %	24,8 %	19,0 %	8,8 %	100,0 %
Bac+2	13,7 %	53,7 %	9,4 %	11,4 %	6,6 %	5,2 %	100,0 %
Santé Social	8,8 %	83,4 %	0,9 %	4,5 %	0,8 %	1,6 %	100,0 %
Lic Pro	20,0 %	58,7 %	6,0 %	9,4 %	2,6 %	3,2 %	100,0 %
Licence	8,7 %	45,0 %	8,7 %	17,3 %	13,3 %	7,1 %	100,0 %
Aut Bac+3/4	12,5 %	62,5 %	11,0 %	7,4 %	5,9 %	0,7 %	100,0 %
Bac+5	14,9 %	57,9 %	8,9 %	11,7 %	4,6 %	2,0 %	100,0 %
Ec Ing + Com	23,9 %	60,7 %	4,7 %	7,4 %	1,8 %	1,4 %	100,0 %
Doctorats	13,6 %	63,1 %	8,6 %	9,0 %	3,8 %	1,9 %	100,0 %
Ensemble	12,7 %	56,7 %	7,4 %	11,3 %	8,1 %	3,8 %	100,0 %

Effectifs théoriques	EDI (rapide)	EDI (différé)	EDD	Inactivité	Chômage	Formation(s)	Ensemble
Collège ou inférieur	35,8	159,8	21,0	31,9	22,9	10,7	282
CAP BEP	296,8	1 326,4	174,0	265,1	190,3	88,5	2341
Bac Pro							2944
Bac techno	42,1	188,1	24,7	37,6	27,0	12,5	332
Bac Gén	17,4	77,6	10,2	15,5	11,1	5,2	137
BTS DUT Bac+2	206,0	920,7	120,8	184,0	132,1	61,4	1625
Bac+2/3 Santé Social	355,7	1 589,8	208,5	317,7	228,1	106,0	2806
Lic Pro	125,4	560,4	73,5	112,0	80,4	37,4	989
Licence	178,5	797,8	104,6	159,4	114,5	53,2	1408
Autres Bac+3/4	17,2	77,1	10,1	15,4	11,1	5,1	136
Bac+5	241,3	1 078,2	141,4	215,5	154,7	71,9	1903
Ec Ing + Commerc	96,4	430,6	56,5	86,1	61,8	28,7	760
Doctorats	252,3	1 127,5	147,9	225,3	161,8	75,2	1990
Total	2238	10002	1312	1999	1435	667	17653

Écarts à l'indépendance	EDI (rapide)	EDI (différé)	EDD	Inactivité	Chômage	Formation(s)	Ensemble
Collège ou inférieur	-15,8	-76,8	0,0	33,1	48,1	11,3	0
CAP BEP	-32,8	-270,4	40,0	46,9	188,7	27,5	0
Bac Pro							0
Bac techno	-0,1	-78,1	17,3	17,4	26,0	17,5	0
Bac Gén	-1,4	-41,6	2,8	18,5	14,9	6,8	0
BTS DUT Bac+2	17,0	-48,7	32,2	2,0	-25,1	22,6	0
Bac+2/3 Santé Social	-107,7	750,2	-183,5	-190,7	-206,1	-62,0	0
Lic Pro	72,6	20,6	-14,5	-19,0	-54,4	-5,4	0
Licence	-56,5	-164,8	17,4	84,6	72,5	46,8	0
Autres Bac+3/4	-0,2	7,9	4,9	-5,4	-3,1	-4,1	0
Bac+5	41,7	22,8	27,6	7,5	-66,7	-32,9	0
Ec Ing + Commerc	85,6	30,4	-20,5	-30,1	-47,8	-17,7	0
Doctorats	18,7	128,5	23,1	-46,3	-85,8	-38,2	0
Total	0	0	0	0	0	0	0

Contributions au Khi ²	EDI (rapide)	EDI (différé)	EDD	Inactivité	Chômage	Formation(s)	Ensemble
Collège ou inférieur	6,94	36,89	0,00	34,24	100,83	12,08	190,98
CAP BEP	3,62	55,12	9,20	8,30	187,12	8,58	271,94
Bac Pro							
Bac techno	0,00	32,43	12,16	8,06	25,07	24,29	102,02
Bac Gén	0,11	22,32	0,78	22,03	19,84	8,99	74,07
BTS DUT Bac+2	1,40	2,58	8,60	0,02	4,77	8,32	25,69
Bac+2/3 Santé Social	32,63	353,95	161,54	114,51	186,22	36,28	885,13
Lic Pro	42,06	0,76	2,86	3,22	36,80	0,77	86,48
Licence	17,89	34,03	2,88	44,85	45,98	41,17	186,79
Autres Bac+3/4	0,00	0,82	2,37	1,89	0,84	3,33	9,26
Bac+5	7,22	0,48	5,37	0,26	28,75	15,06	57,15
Ec Ing + Commerc	76,14	2,15	7,43	10,50	36,95	10,93	144,09
Doctorats	1,39	14,64	3,61	9,53	45,47	19,40	94,04
Total	190,60	603,18	229,74	277,40	799,02	196,13	

Graphiques des écarts à l'indépendance



Trajectoires d'emploi (précision sur les modalités regroupées)
EDI (rapide) : accès rapide à un emploi à durée indéterminée
EDD : emploi à durée déterminée (accès rapide et récurrent)
Inactivité : accès tardifs à l'activité, sorties du Marché du travail, parcours durable hors de celui-ci ou inactivité durable
Chômage : sortie de l'emploi vers le chômage ou chômage persistant et récurrent
Formation(s) : différentes formes de retour en formation

Khi ²		
DL		
Prob>Khi ²	<0,0001	
Significativité	5,00 %	1,00 %
Seuils	79,1	88,4
V de Cramer		

Annexe : rappels de formules utiles pour les calculs demandés

Interpolation d'un quantile Données en classe, quantile (Qtl) compris dans l'intervalle [b.fin;b.sup], en ayant au préalable calculé les effectifs cumulés croissants pour chaque borne de chaque classe et pour le quantile considéré (médiane, quartile, etc.) :

$$Qtl = b.inf + \frac{n \uparrow Qtl - n \uparrow inf}{n \uparrow sup - n \uparrow inf} \cdot (b.sup - b.inf)$$

Coefficients d'une régression linéaire $Y = a \cdot X + b + U$; $a = \frac{cov(X, Y)}{var(X)}$; $b = \bar{Y} - a \cdot \bar{X}$ $R^2 = \frac{Var\,expl}{Var\,tot} = \frac{a^2 \cdot Var(X)}{Var(Y)} = correlat^2$

V de Cramer : $V = \sqrt{\frac{\chi^2}{n \cdot (k-1)}}$ k étant le minimum des nombres de lignes et de colonnes, n l'effectif du tableau, χ^2 le Khi² du tableau.

