

DATE : ...12.../...1.../...2023.....à.....11h  
 EPREUVE: Statistiques et informatique appliquées aux Sciences sociales  
 ENSEIGNANT: ...Renaud Orain <[renaud.orain@univ-paris1.fr](mailto:renaud.orain@univ-paris1.fr)> .....  
 DIPLOME : Licence de Sciences sociales, 3<sup>e</sup> année  
 Étudiant-es de tous parcours, inscrit-es en contrôle continu...  
 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 sont regroupés en classe et la moyenne exacte de la classe ainsi que la somme de ces temps par classe pour permettre 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		10343	308437	17
Ensemble				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<sup>a</sup> et indiquez quelle est la classe modale. Évaluez<sup>a</sup> au moyen d'un diagramme cumulé croissant ou d'un calcul d'interpolation les quartiles dont la médiane. Enfin, dessinez la boîte à moustache de cette distribution<sup>a</sup> sachant que le maximum est de 41 mois.

<sup>a</sup> (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<sup>b</sup> 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<sup>c</sup> 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.

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

<sup>c</sup> 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	Somme 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 :** 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 serait inférieure à 0,0001. Vous pourrez également vous appuyer sur le tableau ci-dessous pour résumer les trajectoires d'emploi selon les niveaux d'études.

**Répartition des personnes ayant fini leurs études selon leur trajectoire d'emploi sur 4 ans résumée, détail des % en ligne selon le niveau d'étude atteint auparavant**

%ligne	→	EDI (rapide)	EDI (différé)	EDD	Inactivité	Chômage	Formation (s)	Ensemble
Niveau d'études	Effectif							
Collège ou -	282	7,1	29,4	7,5	23,1	25,2	7,8	100,0
CAP BEP	2341	11,3	45,1	9,1	13,3	16,2	5,0	100,0
Bac Pro	2944	12,0	47,2	9,2	14,1	12,8	4,7	100,0
Bac techno	332	12,7	33,1	12,7	16,6	16,0	9,0	100,0
Bac Gén	137	11,7	26,3	9,5	24,8	19,0	8,8	100,0
Bac+2	1625	13,7	53,7	9,4	11,5	6,6	5,2	100,0
Santé Social	2806	8,8	83,4	0,9	4,5	0,8	1,6	100,0
Lic Pro	989	20,0	58,8	6,0	9,4	2,6	3,2	100,0
Licence	1408	8,7	45,0	8,7	17,3	13,3	7,1	100,0
Aut Bac+3/4	136	12,5	62,5	11,0	7,4	5,9	0,7	100,0
Bac+5	1903	14,9	57,9	8,9	11,7	4,6	2,1	100,0
Ec Ing + Com	760	24,0	60,7	4,7	7,4	1,8	1,5	100,0
Doctorats	1990	13,6	63,1	8,6	9,0	3,8	1,9	100,0
Ensemble	17653	12,7	56,7	7,4	11,3	8,1	3,8	100,0

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 <sup>2</sup>	2296,067	
DL	60	
Prob>Khi <sup>2</sup>	<0,0001	
Significativité	5,00 %	1,00 %
Seuils	79,1	88,4
V de Cramer	0,1613	

## 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**  $Y = a \cdot X + b + U$ ;  $a = \frac{cov(X, Y)}{var(X)}$ ;  $b = \bar{Y} - a \cdot \bar{X}$   $R^2 = \frac{Var\,expl}{Var\,tot} = \frac{a^2 \cdot Var(X)}{Var(Y)} = correlat^2$



