M1 Actuariat **2020-2021**

Econométrie
Contrôle Continu

Note : Pour tous les exercices, vous effectuerez les tests à un seuil de risque de 5%.

Exercice 1 :

Des études ont montré que l'absence d'un plancher dans les habitations (c'est-à-dire sol en terre battue) avait des effets négatifs sur la santé des enfants habitant ce logement (exemple : problèmes respiratoires, maladies de peau, etc.). Pour remédier à ce problème de santé publique, le Mexique a mis en place un programme gouvernemental destiné à remplacer le plancher en terre battue de milliers de logements précaires par un plancher en ciment. Des chercheurs ont tenté d'évaluer empiriquement l'efficacité de ce programme, en étudiant si dans les ménages qui avaient bénéficié du remplacement de leur plancher, la survenue de différentes maladies chez les enfants était plus faible que dans les ménages non bénéficiaires du programme. On s'intéresse ici aux résultats de l'étude concernant l'apparition de signes d'anémie chez les enfants, qui a été menée à partir d'une base de données contenant les variables suivantes :

— *anemie* : variable qui vaut 1 si l'enfant a souffert d'anémie au cours des 12 derniers mois et 0 sinon,
— *Programme* : variable qui vaut 1 si le ménage est bénéficiaire du programme et 0 sinon,
— *Age* : l'âge de l'enfant (en années),
— *Masculin* : si l'enfant est de sexe masculin et 0 sinon,
— *Revenu* : le niveau de revenu du ménage (en pesos),
— *NbPersonnes* : le nombre de personnes habitant le logement,
— *NbPieces* : le nombre de pièces du logement,
— *Electricite* : une variable qui vaut 1 si le logement est relié à l'électricité et 0 sinon,
— *Eaupotable* : une variable qui vaut 1 si le logement est raccordé à un système d'eau potable et 0 sinon.

En supposant que les termes d'erreurs suivent une loi normale, de moyenne nulle et de variance unitaire, les résultats par la technique d'estimation appropriée du modèle expliquant la variable *anemie* en fonction des autres variables de la base sont reportés ci-dessous :

Tableau 1

	Coefficient	p-value
Programme	-0,199	0,000
Age	-0,159	0,000
Masculin	0,095	0,024
Revenu	0,000004	0,563
Nb Personnes	0,012	0,245
Nb Pieces	-0,039	0,069
Electricité	0,274	0,195
Eau potable	0,025	0,560
Constante	-0,084	0,707
Nb Obs.	3756	

1. Calculer la probabilité p_i que l'enfant i ait souffert d'un problème d'anémie au cours des 12 derniers mois en fonction de ses caractéristiques (X_i) et la probabilité qu'il n'en ait pas souffert.
2. Ecrire la vraisemblance de ce modèle.
3. Interpréter l'ensemble des résultats de la régression du tableau 1.
4. Quelle est la probabilité $\hat{p}_i$ de souffrir d'un problème d'anémie pour une jeune fille de 15 ans, dont le ménage dispose d'un revenu de 565 pesos, qui est composé de 3 personnes, vivant dans un logement de 4 pièces, qui n'est ni relié à l'eau potable ni à l'électricité, et qui a bénéficié du programme gouvernemental.
5. Cette probabilité $\hat{p}_i$ est-elle fortement élastique au revenu du ménage ? Justifier votre réponse.
6. Le tableau ci-dessous confronte les prédictions du modèle aux observations réelles.
   (a) Expliquer comment a été construit ce tableau ?
   (b) Quelle(s) conclusion(s) pouvez-vous tirer de ce tableau ?

Tableau 2

```
           -------- True --------
Classified |         D            ~D  |      Total
-----------+--------------------------+-----------
     +     |       360           310  |        670
     -     |      1090          1996  |       3086
-----------+--------------------------+-----------
   Total   |      1450          2306  |       3756
```

Exercice 2 :

Dans la majorité des pays de l'OCDE, les parents qui sont en emploi peuvent, s'ils le souhaitent, prendre à la suite de la naissance de leur enfant, un congé parental afin de s'en occuper. Durant le congé, le parent ne perçoit plus son salaire mais il devient prestataire de l'assurance parentale et reçoit une allocation correspondant en général à un pourcentage de son salaire (50%, 70%, ... selon les pays). On souhaite analyser les déterminants de la durée des congés parentaux pris par les nouveaux parents. En particulier, on cherche à savoir si le salaire du parent prestataire avant le début du congé a tendance à diminuer la durée du congé prise par celui-ci. On suppose en effet que plus le salaire du parent est élevé, plus le coût d'opportunité du congé parental (perte de revenu) est élevé et donc plus le prestataire est incité à écourter la durée de son congé parental pour retrouver son niveau de revenu initial en retournant travailler. Pour ce faire, une régression linéaire multiple, à partir d'un échantillon de 35272 nouveaux parents, a été estimée par MCO. En particulier, on dispose des variables suivantes :

— *Duree* qui correspond à la durée du congé prise par le parent (exprimée en nombre de semaines) ;
— *revenu* qui correspond au revenu mensuel touché par le parent avant la naissance (en euros) ;
— *prestation* qui correspond au montant de l'allocation hebdomadaire perçue par le parent durant le congé pour remplacer son salaire (exprimée en euros) ;
— *sexe* qui vaut 1 si le prestataire est une femme, 0 sinon ;
— *age* qui vaut 1 si le prestataire a moins de 25 ans, 2 s'il a en 25 et 29 ans, 3 s'il a entre 30 et 34 et 4 s'il a 35 ou plus ;
— *region* qui vaut 1 si le prestataire vit dans la région Nord, 2 s'il vit dans la région Ouest, 3 s'il vit dans la région Est et 4 s'il vit dans la région Sud ;
— *statut_salarie* qui vaut 1 si le prestataire a le statut de salarié et 0 sinon ;
— *conge_seul* qui vaut 1 si le prestataire est le seul du couple à prendre un congé et 0 sinon.

Par ailleurs, des variables croisées ont également été construites. Ces variables correspondent aux termes d'interaction entre la variable de sexe et d'autres variables explicatives du modèle. Elles correspondent aux variables commençant par "*s_*" (exemple : $s\_prestation = sexe * prestation$).

Le tableau 1 reporte les résultats issus de l'estimation par le logiciel STATA :

Tableau 1

Source	SS	df	MS			
Model	79769.6129	16	4985.6008	Number of obs	=	35,272
Residual	499234.74	35,255	??????????	F(16, 35255)	=	??????
Total	579004.353	35,271	16.4158757	Prob > F	=	??????
				R-squared	=	0.1378
				Adj R-squared	=	0.1374
				Root MSE	=	3.7631

Duree	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
revenu	-.0386821	.0047738	-8.10	0.000	-.0480389	-.0293253
prestation	.0629598	.0084198	7.48	0.000	.0464568	.0794628
sexe	3.774912	.5207643	7.25	0.000	2.754198	4.795627
age_1	-3.25114	.1773869	-18.33	0.000	-3.598824	-2.903456
age_2	-1.951943	.161896	-12.06	0.000	-2.269264	-1.634622
age_3	-.4852684	.1440593	-3.37	0.001	-.7676291	-.2029076
region_1	-.2869399	.1537185	-1.87	0.062	-.588233	.0143532
region_2	-.0846219	.0944273	-0.90	0.370	-.2697025	.1004586
region_3	.22847	.0434962	5.25	0.000	.143216	.313724
statut_salarie	.0885853	.0914153	0.97	0.333	-.0905915	.2677621
conge_seul	1.640014	.0427179	38.39	0.000	1.556286	1.723743
s_revenu	.0281324	.0049829	5.65	0.000	.0183658	.0378991
s_prestation	-.0439611	.0087512	-5.02	0.000	-.0611137	-.0268084
s_age_1	3.044566	.18894	16.11	0.000	2.674238	3.414895
s_age_2	1.819358	.1766511	10.30	0.000	1.473116	2.1656
s_age_3	.4511563	.1622812	2.78	0.005	.1330801	.7692325
_cons	18.11694	.5005502	36.19	0.000	17.13585	19.09804

1. Tester si le modèle est globalement significatif.
2. Calculer la variance résiduelle.
3. Calculer l'intervalle de confiance associé à la variable *prestation* et interpréter.
4. Comme reporté dans le Tableau 1, on observe que les variables d'âge et de région décrites dans l'énoncé n'ont pas directement été introduites dans le modèle. Expliquer la démarche qui a conduit à introduire les variables *age*_1, *age*_2, *age*_3, *region*_1, *region*_2, et *region*_3. Quelles sont les conséquences pour l'interprétation des résultats ?
5. Que pouvez-vous dire de l'hypothèse faite sur l'effet du salaire à partir des résultats du modèle ?
6. Commenter littérairement l'ensemble des résultats du modèle.
7. Donner une estimation de la durée du congé parental prise par une mère âgée de 25 ans, qui habite la région Nord, qui a un statut non salarié, dont le conjoint a également pris un congé parental, dont le revenu mensuel est de 1550 euros et sa prestation hebdomadaire correspond à 50% de

son salaire hebdomadaire (soit 1/4 du salaire mensuel) plus une prime de 50 euros.

8. On cherche à tester $H_0 : \beta_{region\_1} = \beta_{region\_2}$. Grâce à la commande STATA appropriée, on obtient le résultat suivant :

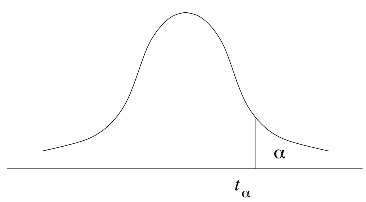
```
 ( 1)  region_1 - region_2 = 0

F(  1, 35259) =    0.34
Prob > F =    0.5609
```

Donner la formule du test qui a permis de calculer le F ? Préciser à quoi correspondent les éléments de la formule. Que pouvez-vous en conclure concernant l'hypothèse formulée ? Interpréter le résultat.

9. Expliquer en détails comment vous auriez mis en oeuvre le test de Chow dans le but de savoir si les déterminants de la durée du congé parental sont les mêmes pour les mères et les pères (régression(s) effectuée(s), variables explicative(s), statistique(s) de test, règle(s) de décision). A votre avis quel va être le résultat de ce test ? Justifier.

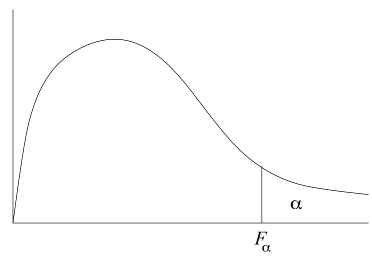
Tables statistiques

Loi de Student



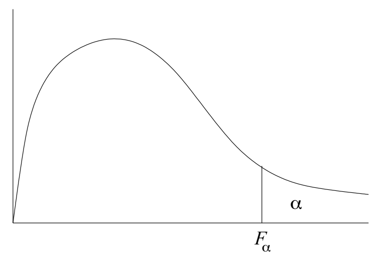
d.d.l. α	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
∞	1.282	1.645	1.960	2.326	2.576

Loi de Fisher : $\alpha = 0.05$



Degrés de liberté du dénominateur : ν_2	Degrés de liberté du numérateur : ν_1								
	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Loi de Fisher : $\alpha = 0.05$ (suite)



Degrés de liberté du dénominateur : ν_2	Degrés de liberté du numérateur : ν_1									
	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.30
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Exercice 1 :

1) On veut expliquer la va Y_i tq $Y_i = \begin{cases} 1 & \text{si l'enfant souffre d'anémie} \\ 0 & \text{sinon} \end{cases}$

Y_i est dichotomique

Donc on considère $Y_i = \begin{cases} 1 & \text{si } Y_i^* > 0 \\ 0 & \text{si } Y_i^* \leq 0 \end{cases}$

où $Y_i^* = X_i\beta + \varepsilon_i$ la va latente du modèle

$$\Rightarrow \mathbb{P}(Y_i = 1) = \mathbb{P}(Y_i^* > 0)$$
$$= \mathbb{P}(X_i\beta + \varepsilon_i > 0)$$
$$= \mathbb{P}(\varepsilon_i > -X_i\beta)$$
$$= 1 - \mathbb{P}(\varepsilon_i \leq -X_i\beta)$$
$$= F(X_i\beta)$$

où F est la fdr d'une $\mathcal{N}(0,1)$

2) $$\mathcal{L}(Y, X, \beta) = \prod_{i=1}^{n} F(X_i\beta)^{Y_i} (1 - F(X_i\beta))^{1-Y_i}$$

3) TCEEPA :

* Le fait de bénéficier du programme permet de diminuer la proba qu'un enfant souffre d'anémie
* L'âge diminue égalem^t la proba de souffrir d'anémie
* Être de sexe masculin augmente la proba
* Toutes les autres va n'ont pas d'impact car pvalue $> 5\%$

4) $X_i\beta = -0{,}199 \times 1 - 0{,}159 \times 15 + 0{,}035 \times 0 + 0{,}000004 \times 565 + 0{,}012 \times 3 - 0{,}039 \times 4$
$+ 0{,}274 \times 0 + 0{,}025 \times 0 - 0{,}084$

$= -2{,}7857$

$$\Rightarrow \mathbb{P}(Y_i = 1) = F(-2.7857) = 0{,}0026 = \int_{-\infty}^{-2.7857} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx$$

$\rightarrow$ Proba pour cette jeune fille de souffrir d'anémie est de 0.26%

5) Non car la VA revenu n'est pas significative.

6) a)

— A partir de l'estimation du modèle, on calcule pour chaque individu sa probabilité prédite de souffir d'anémie $P(\widehat{Y_i = 1})$. Si cette probabilité prédite est supérieure à 0.5, l'enfant est classé comme ayant souffert d'anémie durant les 12 derniers mois ($Y = 1$). Inversement, si la probabilité prédite est inférieure à 0.5, on classe l'enfant comme n'ayant pas souffert d'anémie ($Y = 0$). On obtient ainsi un nombre d'individus prédis en bonne et en mauvaise santé par le modèle.

— On confronte ensuite ce classement avec les vraies valeurs observées de la variable santé.

— On peut alors retrouver le nombre de personnes pour lesquelles le modèle prédit correctement/mal. Le taux de prédictions fausses correspond alors à : $\frac{\text{Nb de mal classés}}{\text{nb d'individus}}$. Si ce taux est supérieur à 50%, cela signifie que le modèle prédit encore plus mal que le hasard. Dans ce cas là le modèle ne pourra être utilisé.

b) Il y a 1090 + 310 mauvaises prédict° pour un total de 3756 obs

$\Rightarrow$ taux de prédict° fausses $= \frac{1090+310}{3756} = 0,373$

$\Rightarrow$ 37,3% < 50% $\Rightarrow$ on peut garder le modèle.

Exercice 2 :

1) Test de significativité du modèle :

$H_0 : \beta_1 = \ldots = \beta_k = 0$ contre $H_1 : \exists \beta_i \neq 0$

Stat de test : $F^* = \frac{R^2}{1-R^2} \frac{T-k}{k-1} \sim F(k-1, T-k)$

Règle de décis° : Rejet de H_0 si $F^* > F_\alpha(k-1, T-k)$

Ici $R^2 = 0.1378$
$T-k = 35255$
et $k-1 = 16$

$\Rightarrow F^* = 352,16$ et $F_{0.05}(16, 35255) = 1,67$

Donc $F^* > 1,67 \Rightarrow$ on rejette H_0

2) Variance résiduelle $= SCR/(n-k) = 499237.74/35255$
$= 14,16$

3) L'intervalle de conf associé à prestat° est

$[\hat{\beta}_{prestat^\circ} \pm t^{\alpha/2}_{(n-k)} \times \hat{\sigma}_{\hat{\beta}_{prestat^\circ}}]$

$\Rightarrow IC_{prestat^\circ} = [0.0629598 - 1,645 \times 0.0084198 ; 0.0629598 + 1,645 \times 0.0084198]$

$= [0.0491 ; 0.0768]$

Donc pour 1€ d'allocat° hebdomadaire supplémentaire la durée de congé parental augmente de 0.0491 semaines à 0.0768 semaines avec proba 0.95

4) Les va d'âges et région sont des va qualitatives à plusieurs modalités. Ce type de va ne peut être inclut dans un modèle économétrique directement

Ainsi, pour ces deux va, il a été nécessaire de créer autant de va dichotomiques que de modalités, pour les intégrer ensuite dans le modèle et en laissant une catégorie en référence

en référence. En particulier, pour la variable d'âge, on créé les 4 variables dichotomiques telle que : $age_1 = 1$ si $age = 1$, 0 sinon ; $age_2 = 1$ si $age = 2$, 0 sinon ; $age_3 = 1$ si $age = 3$, 0 sinon ; $age_4 = 1$ si $age = 4$, 0 sinon. On s'assure au préablable qu'il y a assez de monde dans chaque catégorie (au moins 5% de l'échantillon). On intègre ensuite toutes les variables dichotomique - 1 qui sera la référence. Les interprétations qui en découlent seront faites en fonction de la catégorie de référence. On fait la même chose pour la variable de région.

5) Elle est vraie car pour 1€ de revenu en + baisse de durée de 0,038 sem.

6) —

7) Durée $= -0{,}0386821 \times 1550 + 0.0629598 \times (\frac{1}{4} \times 1550) + 3,774912 \times 1$
$- 1,951943 \times 1 - 0.2869399 \times 1$

$+ 0.0281324 \times 1550 - 0.0439611 \times (\frac{1}{4} \times 1550) + 1.819358 \times 1$
$+ 18.11654$

$= 13,43$ semaines ~ 94 jours.

8) On peut écrire H_0 tq $R\beta = q$

avec $R = (0\ 0\ 0\ 0\ 0\ 0\ 1\ {-1}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$

$q = (0)$

$$\Rightarrow F^* = \frac{(R\hat{\beta} - q)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta} - q)/c}{S^2} \sim F(c, n-k)$$

$F^* = 0.34$ et p.value $= 0.5609 > 0.05$

$\Rightarrow$ on accepte H_0

$\Rightarrow$ Pas de $\neq$ Significative entre region 1 et region 2