Contrôle Terminal (2ème session) – 24/06/2022 – 2h00
Statistiques – L1 de psychologie

Réponses

Veuillez rendre ce sujet et votre copie.

Numéro d'anonymat :

Le formulaire et la calculatrice sont autorisés. Merci d'indiquer dans la case Numéro d'anonymat, *ci-dessus, un numéro que vous reporterez aussi sur votre copie. Vous rendrez l'énoncé et votre copie, et pouvez soit répondre sur l'énoncé, soit détailler certaines questions sur la copie si vous avez besoin de plus de place. Le soin de la rédaction entrera en compte dans la notation mais dans les questions où des détails ne sont pas explicitement demandés, un résultat correct, donné sans détails de calcul sera accepté.*

Exercice 1 : Genre et modes de vie

Une sociologue s'intéresse aux différences de modes de vie entre hommes et femmes, et décide de comparer "l'empreinte carbone" de ces modes de vie, c'est à dire le bilan des gaz à effet de serre que chaque personne émet ou qui ont été émis pour la production des produits que la personne achète (exprimés en Tonnes-équivalent-carbone par an).

Elle établit un bilan des émissions de 126 français·es au cours de l'année 2021. Elle obtient les données suivantes :

Empreinte carbone	[6 ; 7[	[7 ; 8[	[8 ; 9[	[9 ; 10[	[10 ; 11[	[11 ; 12[
Effectifs	3	23	48	33	15	4

1. Quel sont la moyenne et l'écart type des émissions de Gaz à effets de serre de ces 126 français·es ? *Dans cette question, on vous demande d'indiquer les calculs effectués.*

moyenne : $m(X) = \frac{\sum c_i\, n_i}{n} = \frac{6{,}5\times3+7{,}5\times23+8{,}5\times48+\cdots+11{,}5\times4}{126} = \frac{1\,117}{126} \simeq 8{,}87$ Tonnes-équivalent-carbone
$m(X^2) = \frac{\sum c_i^2\, n_i}{n} = \frac{6{,}5^2\times3+7{,}5^2\times23+8{,}5^2\times48+\cdots+11{,}5^2\times4}{126} = \frac{10\,049{,}5}{126}$
$Var(X) = m(X^2) - m(X)^2 = \frac{10\,049{,}5}{126} - \left(\frac{1\,117}{126}\right)^2 \simeq 1{,}17$
Écart-type : $s(X) = \sqrt{Var(X)} \simeq 1{,}08$ Tonnes-équivalent-carbone

2. En déduire une estimation des émissions moyennes des Français·es au cours de l'année 2021. *Établir un intervalle de confiance, pour la confiance 96%.*

Comme n= 126 >30, on cherche z_α tel que $F(z_\alpha) = \frac{0{,}96+1}{2} = 0{,}98$
On a $F(2{,}054) \simeq 0{,}98$ d'où $z_\alpha \simeq 2{,}054$ d'où $a_\alpha = z_\alpha \frac{s_e}{\sqrt{n-1}} \simeq 0{,}1984$
On estime donc que l'empreinte carbone moyenne est dans l'intervalle $[8{,}87 - 0{,}198\,4\,;\, 8{,}87 + 0{,}198\,4] \simeq [8{,}67\,;\, 9{,}07]$, avec la confiance $c = 0{,}96$.

3. À titre de comparaison, en 2020 (année où les transports et l'activité économique avaient été impactés plus lourdement par la crise sanitaire), l'empreinte carbone moyenne des Français·es (calculée selon la même méthodologie qu'en question précédente) était de 8,5 Tonnes-équivalent-carbone par personne. Pouvez vous affirmer, avec la confiance 96%, que les niveaux moyens d'émission des français·es étaient plus élevés en 2021 qu'en 2020 ?

Oui, on peut l'affirmer : toutes les valeurs de l'intervalle qu'on a obtenu sont au moins égales à 8,67 , donc plus grandes que l'empreinte carbone moyenne de 2020.

Pour étudier la différence d'émissions entre hommes et femmes, elle choisit ensuite de se restreindre aux personnes célibataires (afin de comptabiliser vraiment individuellement certaines émissions comme le chauffage). Elle constate que parmi les personnes dont elle a établi le bilan carbone, il y a 46 célibataires (dont les empreintes carbones sont les suivantes) :

empreinte	11,1	11	10,1	8,7	8,7	8,8	9,4	10,3	9,8	11,4	8,6	9,7
des hommes	9,4	9,2	8,2	9	10	10,2	8,7	7,6	8,2	9,3	10,1	

empreinte	9,7	9,5	8,8	9	9,7	7,3	8,1	9,1	8,1	7,8	11,4	7,5
des femmes	10,6	10,8	7,8	9,4	10,3	9,5	8,8	7,6	7,9	9,7	7,8	

1. Indiquer les moyenne et écart type des niveau d'émissions des 23 femmes célibataires de l'échantillon.

moyenne : $m(X) = \frac{\sum x_i}{n} = \frac{9{,}7+9{,}5+8{,}8+\cdots+7{,}8}{23} = \frac{206{,}2}{23} \simeq 8{,}97$ Tonnes-équivalent-carbone
$m(X^2) = \frac{\sum x_i^2}{n} = \frac{9{,}7^2+9{,}5^2+8{,}8^2+\cdots+7{,}8^2}{23} = \frac{1\,878{,}32}{23}$
$Var(X) = m(X^2) - m(X)^2 = \frac{1\,878{,}32}{23} - \left(\frac{206{,}2}{23}\right)^2 \simeq 1{,}29$
Écart-type : $s(X) = \sqrt{Var(X)} \simeq 1{,}14$ Tonnes-équivalent-carbone

2. Indiquer de même les moyenne et écart type des niveau d'émissions des hommes célibataires de l'échantillon.

moyenne : $m(X) = \frac{\sum x_i}{n} = \frac{11{,}1+11+10{,}1+\cdots+10{,}1}{23} = \frac{217{,}5}{23} \simeq 9{,}46$ Tonnes-équivalent-carbone
$m(X^2) = \frac{\sum x_i^2}{n} = \frac{11{,}1^2+11^2+10{,}1^2+\cdots+10{,}1^2}{23} = \frac{2\,078{,}01}{23}$
$Var(X) = m(X^2) - m(X)^2 = \frac{2\,078{,}01}{23} - \left(\frac{217{,}5}{23}\right)^2 \simeq 0{,}922$
Écart-type : $s(X) = \sqrt{Var(X)} \simeq 0{,}96$ Tonnes-équivalent-carbone

3. Estimer le niveau moyen d'émission des hommes célibataires. *Établir un intervalle de confiance, pour la confiance 0,96%.*

Comme $n = 23 \leqslant 30$, on cherche t_α à partir de la table inverse de Student avec $p = \frac{\alpha}{2} = 0{,}02$ et $n-1 = 22$ degrés de liberté (ddl)
On lit $t_\alpha \simeq 2{,}1829$ d'où $a_\alpha = t_\alpha \frac{s_e}{\sqrt{n-1}} \simeq 0{,}446\,78$.
On estime donc que l'empreinte carbone moyenne des hommes célibataires est dans l'intervalle $[9{,}46 - 0{,}446\,78\,;\, 9{,}46 + 0{,}446\,78] \simeq [9{,}01\,;\, 9{,}91]$ avec la confiance $c = 0{,}96$

4. Estimer le niveau moyen d'émission des femmes célibataires. *Établir un intervalle de confiance, pour la confiance 0,96%.*

Comme $n = 23 \leqslant 30$, on cherche t_α à partir de la table inverse de Student avec $p = \frac{\alpha}{2} = 0{,}02$ et $n-1 = 22$ degrés de liberté (ddl)
On lit $t_\alpha \simeq 2{,}1829$ d'où $a_\alpha = t_\alpha \frac{s_e}{\sqrt{n-1}} \simeq 0{,}530\,55$.
On estime donc que l'empreinte carbone moyenne des femmes célibataires est dans l'intervalle $[8{,}97 - 0{,}530\,55\,;\,8{,}97 + 0{,}530\,55] \simeq [8{,}44\,;\,9{,}5]$ avec la confiance $c = 0{,}96$

5. Conclure : peut on affirmer, avec la confiance 0,96% qu'en moyenne les hommes célibataires ont une empreinte carbone plus élevée que les femmes célibataires ?

Les intervalles que l'on a calculés se chevauchent, donc ces données ne permettent pas de conclure que les hommes aient une plus grande empreinte carbone que les femmes.
Remarque : Les données de cet énoncé sont fictives, mais des chercheurs ont montré avec des données réelles que les hommes ont en fait en moyenne une plus grande empreinte carbonne que les femmes, notamment à cause de leur plus grande consommation de viande (en moyenne).

Exercice 2 : Test cognitif

Un psychologue établit un test cognitif pour détecter les troubles du développement chez des enfants. Il souhaite que les scores obtenus au test suivent la loi normale $\mathcal{N}(100\ ;\ 15)$.

Après avoir établi et étalonné le test, il le fait passer à 75 enfants pour vérifier que les scores suivent bien la loi normale $\mathcal{N}(100\ ;\ 15)$.

Si les scores suivaient la loi normale $\mathcal{N}(100\ ;\ 15)$, quels seraient les effectifs théoriques des classes ci-dessous ?

Score	[50 ; 65[	[65 ; 75[	[75 ; 85[	[85 ; 95[	[95 ; 110[	[110 ; 120[	[120 ; 130[	[130 ; 140[
Effectif théorique	0,74	2,85	8,31	15,81	28,35	12,1	5,13	1,71

Il suffit de donner une réponse correcte pour avoir les points. Mais si vous souhaitez présenter des calculs intermédiaires, vous pouvez le faire ci-dessous.

Pour chaque intervalle, on calcule d'abord la probabilité de l'intervalle (en agrandissant juste le premier et le dernier intervalle). Puis on multiplie chaque probabilité par l'effectif total (75) pour obtenir l'effectif théorique :

- $\mathbb{P}[X \leqslant 65] \simeq 0{,}009\,815\,3$ d'où l'effectif théorique $75 \times 0{,}009\,815\,3 \simeq 0{,}736\,1$.
- $\mathbb{P}[65 \leqslant X < 75] \simeq 0{,}037\,975$ d'où l'effectif théorique $75 \times 0{,}037\,975 \simeq 2{,}848$.
- $\mathbb{P}[75 \leqslant X < 85] \simeq 0{,}110\,86$ d'où l'effectif théorique $75 \times 0{,}110\,86 \simeq 8{,}315$.
- $\mathbb{P}[85 \leqslant X < 95] \simeq 0{,}210\,79$ d'où l'effectif théorique $75 \times 0{,}210\,79 \simeq 15{,}809$.
- $\mathbb{P}[95 \leqslant X < 110] \simeq 0{,}378\,07$ d'où l'effectif théorique $75 \times 0{,}378\,07 \simeq 28{,}355$.
- $\mathbb{P}[110 \leqslant X < 120] \simeq 0{,}161\,28$ d'où l'effectif théorique $75 \times 0{,}161\,28 = 12{,}096$.
- $\mathbb{P}[120 \leqslant X < 130] \simeq 0{,}068\,461$ d'où l'effectif théorique $75 \times 0{,}068\,461 \simeq 5{,}135$.
- $\mathbb{P}[X \geqslant 130] \simeq 0{,}022\,75$ d'où l'effectif théorique $75 \times 0{,}022\,75 \simeq 1{,}706$.

Exercice 3 : Observance et efficacité des traitements

On s'intéresse à des patients atteints de troubles du comportement, qui se sont vu proposer un traitement médicamenteux. Pour 11 d'entre eux, des chercheurs ont mesuré l'intensité de leur troubles avant et après deux mois de traitement : ils notent Y la diminution d'intensité à l'issue du traitement.

De plus ils demandent à chacun s'il a respecté précisément la prescription. La réponse est résumée par un nombre X entre 0 et 3, selon la règle ci dessous :

"prescription suivie très précisément"	$\rightsquigarrow X = 3$
"prescription plutôt bien suivie"	$\rightsquigarrow X = 2$
"prescription peu suivie"	$\rightsquigarrow X = 1$
"aucune prise de médicament"	$\rightsquigarrow X = 0$

Ils obtiennent alors les données suivantes :

Individu	A	B	C	D	E	F	G	H	I	J	K
Observance (X)	2	3	1	3	2	0	2	0	3	0	2
Diminution (Y)	10	13	6	2	12	-10	7	-2	2	-2	3
x'_i	6,5	10	4	10	6,5	2	6,5	2	10	2	6,5
y'_i	9	11	7	4,5	10	1	8	2,5	4,5	2,5	6

On notera que comme Y est la diminution du score à l'issue du traitement, lorsqu'Y est négatif cela traduit une augmentation de l'intensité des troubles.

1. Calculer la moyenne et l'écart type de la variable Y.

moyenne : $m(Y) = \frac{\sum x_i}{n} = \frac{10+13+6+\cdots+3}{11} = \frac{41}{11} \simeq 3{,}73$
$m(Y^2) = \frac{\sum x_i^2}{n} = \frac{10^2+13^2+6^2+\cdots+3^2}{11} = \frac{623}{11}$
$Var(Y) = m(Y^2) - m(Y)^2 = \frac{623}{11} - \left(\frac{41}{11}\right)^2 \simeq 42{,}74$
Écart-type : $s(Y) = \sqrt{Var(Y)} \simeq 6{,}54$

Bien qu'on l'ait résumée par un nombre, la variable X est plutôt une variable qualitative ordinale, de sorte qu'il n'est pas très pertinent d'en calculer la moyenne et l'écart type.

2. Calculer la médiane de la variable X.

La médiane est la valeur numéro $\frac{11+1}{2} = 6$ (en ordonnant par ordre croissant). C'est donc 2.
En d'entre terme la médiane est « prescription plutôt bien suivie »

3. Afin de déterminer s'il y a un lien entre les variable X et Y, calculez le coefficient de corrélation des rangs de Spearman. Que peut on en conclure ?

On calcule tout d'abord les rangs x'_i et y'_i, entrés dans la table ci-dessus.
le coefficient de corrélation des rangs de Spearman est donc
$$1 - \left(6 \times \frac{6{,}25+1+9+30{,}25+12{,}25+\cdots+0{,}25}{11\left(11^2-1\right)}\right) \simeq 0{,}577\,3$$
Cela suggère un lien entre les deux variables, où Y augmente quand X augmente.

Exercice 4 : Technique d'incitation

On souhaite inciter de jeunes écoliers à bien jeter leur déchets dans la poubelle plutôt que de les laisser à leur place en quittant la pièce.

On a observé que si on demande aux enfants d'être propres et ordonnés, alors 55% d'entre eux jettent leurs déchets dans la poubelle (et ne les laissent pas à leur place).

1. On considère 33 écoliers auxquels on demande d'être propres et ordonnés. On leur distribue des papillotes pendant une journée de cours, et à la fin de la journée, on désigne par X le nombre d'écoliers qui ont laissé l'emballage des papillotes à leur place (par terre ou sur le bureau), c'est à dire le nombre d'écoliers qui n'ont pas été les jeter à la poubelle.
   (a) On considère pour simplifier que X suit une loi binomiale (c'est à dire qu'on fait comme si le tirage était avec remise, et comme si le comportement de chaque enfant n'était pas influencé par celui des autres enfants).
   Quels sont les paramètres de cette loi binomiale, et quels sont sa moyenne et son écart type ?

   $X \sim \mathcal{B}(33; 0{,}45)$.
   $m(X) = 33 \times 0{,}45 = 14{,}85$ $\quad s(X) = \sqrt{33 \times 0{,}45\ (1 - 0{,}45)} \simeq 2{,}858$

   (b) Si on veut calculer $\mathbb{P}[X = 20]$, quelle formule faut il utiliser ? Indiquer la formule, et la valeur qu'elle donne pour $\mathbb{P}[X = 20]$.

   $\mathbb{P}[X = 20] = \binom{33}{20} \times 0{,}45^{20} \times 0{,}55^{13} \simeq 0{,}028$

On considère désormais des enfants que l'on a félicités en disant qu'ils étaient propres et ordonnés (sans chercher à savoir s'ils l'étaient réellement). On constate que sur 75 écoliers auxquels on a distribué des papillotes après les avoir ainsi félicités, il y a en 57 qui ont laissé l'emballage des papillotes à leur place (tandis que les autres ont jeté l'emballage à la poubelle).

1. (a) En déduire une estimation de la proportion d'écoliers qui laissent leurs emballages sur place, si on les a préalablement félicités sur leur propreté. *Vous donnerez un intervalle de confiance, pour la confiance 95%.*

   On a $n\,p_e(1 - p_e) = 75 \times \frac{57}{75}\left(1 - \frac{57}{75}\right) = 13{,}68 > 10$, donc on peut utiliser la procédure du formulaire pour estimer cette proportion.
   On a $F(1{,}96) \simeq 0{,}975$ d'où $z_\alpha \simeq 1{,}96$
   d'où $a_\alpha = z_\alpha \sqrt{\frac{p_e(1-p_e)}{n}} \simeq 0{,}0967$ et $p_e - a_\alpha \simeq \frac{57}{75} - 0{,}0967 = 0{,}6633$ et $p_e + a_\alpha \simeq \frac{57}{75} + 0{,}0967 = 0{,}8567$.
   On estime donc que cette proportion est dans l'intervalle $[0{,}6633; 0{,}8567]$ avec la confiance $c = 0{,}95$.

   (b) Conclure : Peut on affirmer, avec la confiance 95%, que les écoliers laissent plus leurs emballages sur place si on leur demande d'être propres et ordonnés que si on les félicite en disant qu'ils sont propres et ordonnés ?

   Non : les valeurs de l'intervalle qu'on a calculé sont plus grandes que 45%, c'est à dire qu'il sont plus nombreux à laisser le papier à leur place si on les féléicte que si on leur demande d'être propres.