

Contrôle Terminal — **26 juin 2017** —
Statistiques – L1 de psychologie

Durée : 2 h
Réponses

Veuillez rendre ce sujet avec votre copie.

Numéro d'anonymat :

Le formulaire et la calculatrice sont autorisés. Vous êtes invités à répondre directement sur l'énoncé, mais si vous avez besoin de plus de place, vous pouvez détailler certaines questions sur une copie que vous rendrez avec l'énoncé. 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 des personnes souffrant de troubles alimentaires

Il y a en France environ 2 millions de personnes souffrant de troubles alimentaires. On cherche dans cet exercice à déterminer si ces troubles alimentaires touchent de manière égale les femmes et les hommes.

1. (a) Parmi un échantillon aléatoire de 70 français (hommes ou femmes) atteints de troubles alimentaires (choisis avec remise), on note X le nombre de femmes. Quelle serait la loi du nombre X s'il y avait autant d'hommes que de femmes parmi l'ensemble des français atteints de troubles alimentaires ?

$$X \sim \mathcal{B}(70; 0,5)$$

- (b) Justifier que l'on pourrait alors approcher X par la loi $\mathcal{N}(35; 4,18)$.

On a $n = 70 > 30$, et $np = 35 > 5$, et $n(1 - p) = 35 > 5$ donc on peut approximer X par $\mathcal{N}(35; 4,18)$ où $4,18 \simeq \sqrt{np(1-p)}$

- (c) Que vaudrait alors la probabilité $\mathbb{P}[X \geq 44]$. On demande d'effectuer une correction de continuité, et de détailler brièvement l'obtention du résultat

$$\begin{aligned}\mathbb{P}[X \geq 44] &= \mathbb{P}[X \geq 43,5] = \mathbb{P}\left[\frac{43,5-35}{4,18} < \frac{X-35}{4,18}\right] \\ &\simeq \mathbb{P}[2,03 < Z] \simeq 1 - F(2,03) \\ &\simeq 1 - 0,9788 \simeq 0,02\end{aligned}$$

- (d) Si l'on renouvelle les questions 1a, 1b et 1c en supposant désormais que l'échantillon est choisi sans remise, quels résultats obtiendrait-on alors ?

Une brève justification est demandée en plus des résultats

On obtiendrait les mêmes résultats car le coefficient d'exhaustivité vaut $\sqrt{\frac{2\,000\,000-70}{2\,000\,000-1}} \simeq 1$, de sorte que les résultats "avec" et "sans" remise sont extrêmement proches.

2. En interrogeant plusieurs psychologues sur leur patientelle, on parvient à savoir qu'ils ont 70 patients souffrants de troubles alimentaires, à savoir 52 femmes et 18 hommes.

- (a) Au vu des questions précédentes, vous semble-t-il vraisemblable qu'il y ait autant d'hommes que de femmes parmi l'ensemble des français atteints de troubles alimentaires ?

S'il y avait autant d'hommes que de femmes parmi l'ensemble des français atteints de troubles alimentaires, alors il y aurait peu de chances (environ 2% d'après la question 1c) d'avoir au moins 44 femmes dans un échantillon de 70 personnes atteintes de troubles alimentaires. Or dans cette patientelle il y a au moins 44 femmes, donc il est peu vraisemblable qu'il y ait autant d'hommes que de femmes parmi l'ensemble des français atteints de troubles alimentaires.

- (b) Estimer, pour la confiance 95%, la proportion de femmes, parmi l'ensemble des français atteints de troubles alimentaires.

On déterminera, par un court calcul, un intervalle de confiance correspondant à la confiance 95%.

On a $n = 70 > 30$, et $p_e = \frac{52}{70} \simeq 0,743$, d'où $np_e = 52 > 5$, et $n(1 - p_e) = 18 > 5$ donc on peut utiliser la procédure du formulaire pour estimer la proportion p .

On a $F(1,96) \simeq 0,975$ d'où $z_\alpha \simeq 1,96$

d'où $a_\alpha = z_\alpha \sqrt{\frac{p_e(1-p_e)}{n}} \simeq 0,1024$.

On estime donc que p est dans l'intervalle $[0,6405; 0,8453]$ avec la confiance $c = 0,95$.

Exercice 2 : Anxiété chez le dentiste

Des chercheurs cherchent à déterminer le niveau d'anxiété des patients dans la salle d'attente du dentiste. Ils utilisent une échelle de mesure de l'anxiété, pour laquelle un échantillon de 58 patients fournit les données ci-dessous.

Niveau d'anxiété (X)	[78 ; 90[	[90 ; 96[	[96 ; 102[	[102 ; 108[	[108 ; 114[	[114 ; 120[	[120 ; 126[	[126 ; 138[
Effectif	8	7	11	9	8	7	6	2
Fréquences f_i	0,138	0,121	0,19	0,155	0,138	0,121	0,103	0,034
Fréq, cum, F_i	0,138	0,259	0,449	0,604	0,742	0,863	0,966	1

1. Au sein de cet échantillon combien de patients ont un niveau de stress inférieur à 96 ?

Dans cet échantillon, il y a $8 + 7 = 15$ habitants dont le niveau de stress est inférieur à 96.
(on ajoute les effectifs des classes [78 ; 90 [et [90 ; 96 [)

2. Calculer les fréquences et les fréquences cumulées des différentes classes.

Vous les remplirez dans des lignes vierges du tableau ci-dessus.

3. Déterminer la moyenne et l'écart type de l'anxiété au sein de cet échantillon.

$$\text{moyenne : } m(X) = \frac{\sum_i c_i n_i}{n} = \frac{84 \times 8 + 93 \times 7 + 99 \times 11 + 105 \times 9 + 111 \times 8 + 117 \times 7 + 123 \times 6 + 132 \times 2}{58} = \frac{6066}{58} \simeq 104,59$$

$$m(X^2) = \frac{\sum_i c_i^2 n_i}{n} = \frac{84^2 \times 8 + 93^2 \times 7 + 99^2 \times 11 + 105^2 \times 9 + 111^2 \times 8 + 117^2 \times 7 + 123^2 \times 6 + 132^2 \times 2}{58} = \frac{644040}{58}$$

$$Var(X) = m(X^2) - m(X)^2 = \frac{644040}{58} - \left(\frac{6066}{58}\right)^2 \simeq 165,86$$

$$\text{Écart-type : } s(X) = \sqrt{Var(X)} \simeq 12,88$$

4. Déterminer de même la médiane et le troisième quartile.

(a) Médiane :

On demande de justifier le résultat par un cours calcul

Classe de la médiane : $[102; 108[$

$$\text{Méd} \simeq a_i + \frac{a_{i+1} - a_i}{F_X(a_{i+1}) - F_X(a_i)}(0,5 - F_X(a_i)) \simeq 102 + \frac{108 - 102}{0,604 - 0,449}(0,5 - 0,449) \simeq 103,97$$

(b) Troisième quartile :

Classe du troisième quartile : $[114; 120[$

$$Q_3 \simeq a_i + \frac{a_{i+1} - a_i}{F_X(a_{i+1}) - F_X(a_i)}(0,75 - F_X(a_i)) \simeq 114 + \frac{120 - 114}{0,863 - 0,742}(0,75 - 0,742) \simeq 114,4$$

5. Combien y a-t-il environ de patients, au sein de l'échantillon, dont le niveau d'anxiété dépasse 115 ?

On demande de justifier le résultat par un cours calcul

- On commence par calculer la proportion $\mathbb{P}_r[X \leq 115]$, c'est à dire $F_X(115)$ dans les notations du formulaire :

$115 \in [114; 120[$ donc on pose $a_i = 114$ et $a_{i+1} = 120$:

$$F_X(115) = F_X(a_i) + \frac{F_X(a_{i+1}) - F_X(a_i)}{a_{i+1} - a_i}(115 - a_i) = 0,742 + \frac{0,863 - 0,742}{120 - 114}(115 - 114) \simeq 0,762$$

- Donc $\mathbb{P}_r[X > 115] = 1 - 0,762 = 0,238$. Pour en déduire le nombre de patients, on multiplie cette proportion par l'effectif total : on obtient $58 \times 0,238 \simeq 14$ patients.

6. En utilisant cet échantillon, estimer, avec la confiance 98%, le niveau moyen d'anxiété de l'ensemble des patients en salle d'attente des dentistes.

On déterminera, par un court calcul, un intervalle de confiance correspondant à la confiance 98%.

Comme $n = 58 > 30$, on cherche z_α tel que $F(z_\alpha) = \frac{c+1}{2}$

On a $F(2,326) \simeq 0,99$ d'où $z_\alpha \simeq 2,326$ d'où $a_\alpha = z_\alpha \frac{s_e}{\sqrt{n-1}} \simeq 3,968$

On estime donc que μ est dans l'intervalle $[104,59 - 3,968; 104,59 + 3,968] \simeq [100,62; 108,56]$, avec la confiance $c = 0,98$.

7. Si on supposait que le niveau d'anxiété des patients en salle d'attente suive la loi $\mathcal{N}(105; 10)$, quel serait l'effectif théorique de la classe $[102; 108[$?

- proportion théorique :

$$\mathbb{P}[102 \leq X \leq 108] = \mathbb{P}\left[\frac{102-105}{10} < \frac{X-105}{10} < \frac{108-105}{10}\right] = \mathbb{P}[-0,3 < Z < 0,3]$$

$$= F(0,3) - F(-0,3) \simeq 0,6179 - (1 - 0,6179) \simeq 0,6179 - 0,3821 \simeq 0,236$$
- effectif théorique : $0,236 \times 58 \simeq 13,7$

Exercice 3 : Insonorisation et anxiété

On cherche désormais à savoir si l'insonorisation de cabinets dentaires a un impact sur le niveau d'anxiété des patients en salle d'attente. On considère un nouvel échantillon de patients, dont on mesure que les niveaux d'anxiété, dans une salle d'attente insonorisée, sont les suivants : 103, 95, 112, 89, 81, 84, 105, 102, 99, 104, 97, 77, 81, 77 et 111.

1. Quelle est la taille de cet échantillon ?

L'échantillon comporte 15 individus.

2. Quels sont la moyenne et l'écart type du niveau d'anxiété au sein de l'échantillon ?

$$\begin{aligned} \text{moyenne : } m(X) &= \frac{\sum x_i}{n} = \frac{103+95+\dots+111}{15} = \frac{1417}{15} \simeq 94,47 \\ m(X^2) &= \frac{\sum x_i^2}{n} = \frac{103^2+95^2+\dots+111^2}{15} = \frac{135911}{15} \\ \text{Var}(X) &= m(X^2) - m(X)^2 = \frac{135911}{15} - \left(\frac{1417}{15}\right)^2 \simeq 136,78 \\ \text{Écart-type : } s(X) &= \sqrt{\text{Var}(X)} \simeq 11,7 \end{aligned}$$

3. Estimer, à partir de cet échantillon, l'anxiété moyenne de l'ensemble des patients attendant dans une salle d'attente insonorisée.

On déterminera, par un court calcul, un intervalle de confiance correspondant à la confiance 98%.

Comme $n = 15 \leq 30$, on cherche t_α à partir de la table inverse de Student avec $p = \frac{\alpha}{2} = 0,01$ et $n - 1 = 14$ degrés de liberté (ddl)
On lit $t_\alpha \simeq 2,6245$ d'où $a_\alpha = t_\alpha \frac{s_e}{\sqrt{n-1}} \simeq 8,2067$.
On estime donc que μ est dans l'intervalle $[86,26; 102,7]$ avec la confiance $c = 0,98$

4. En comparant à la question 6 de l'exercice 2, peut-on affirmer (avec la confiance 98%) que le niveau moyen d'anxiété est plus faible parmi les patients qui attendent dans une salle d'attente insonorisée ?

Les intervalles calculés se chevauchent, on ne peut donc pas conclure dans laquelle des deux situation le stress moyen est le plus élevé