

Méthodes Statistiques

Corrigé de l'exercice 61

Un test de quotient intellectuel réalisé auprès de 200 enfants a donné la répartition suivante :

	[75, 85[[85, 95[[95, 105[[105, 115[[115, 125[
Effectifs observés	4	51	88	47	10

61-1) Calculer la moyenne et l'écart-type empiriques de l'échantillon.

Chaque intervalle d'amplitude 10 sera représenté par son milieu (par exemple, l'intervalle [75, 85[correspond à un *Q.I.* de 80). On obtient donc le tableau suivant :

	[75, 85[[85, 95[[95, 105[[105, 115[[115, 125[
Q.I.	80	90	100	110	120
Effectifs observés	4	51	88	47	10

On peut, avec ces valeurs, calculer la moyenne empirique :

$$\bar{X} = \frac{1}{200}(80 \times 4 + 90 \times 51 + 100 \times 88 + 110 \times 47 + 120 \times 10) = 100.4$$

On calcule ensuite la variance empirique modifiée s_n^2 :

$$s_n^2 = \frac{1}{199} \left((80-100.4)^2 \times 4 + (90-100.4)^2 \times 51 + (100-100.4)^2 \times 88 + (110-100.4)^2 \times 47 + (120-100.4)^2 \times 10 \right) = 76.84$$

D'où $s = 8.77$.

61-2) La distribution peut-elle être ajustée par une loi normale (au seuil $\alpha = 0,05$) ?

On formule l'hypothèse nulle :

H_0 : les résultats obtenus sont en adéquation avec une loi normale.

On utilise les valeurs estimées à la question précédente comme paramètres de la loi normale : ce sera la loi $\mathcal{N}(100.4; 8.77)$ de moyenne 100.4 et d'écart-type 8.77.

Calculons les probabilités théoriques des intervalles considérés. Par exemple, pour le premier intervalle, on cherche $P(75 \leq X \leq 85)$. On a :

$$\begin{aligned} P(75 \leq X \leq 85) &= P\left(\frac{75 - 100.4}{8.77} \leq T \leq \frac{85 - 100.4}{8.77}\right) \\ &= P(-2.8962 \leq T \leq -1.756) \\ &= F(-1.756) - F(-2.8962) \\ &= (1 - F(1.756)) - (1 - F(2.8962)) \\ &= (1 - 0.9605) - (1 - 0.9981) \\ &= 0.0377 \end{aligned}$$

En procédant de même avec les intervalles suivants, on trouve les probabilités théoriques et ensuite les effectifs théoriques en multipliant les probabilités par $N = 200$:

	[75, 85[[85, 95[[95, 105[[105, 115[[115, 125[
Probabilité	0.038	0.230	0.431	0.252	0.045
Effectifs théoriques	7.540	45.900	86.200	50.400	9.100

La statistique du test du χ^2 est

$$Y = \sum_{i=1}^n \frac{(O_i - C_i)^2}{C_i}.$$

où les O_i sont les valeurs observées et les C_i sont les valeurs calculées (ou valeurs théoriques).

On sait que, sous l'hypothèse H_0 , la variable aléatoire Y suit une loi du χ^2 à $\nu = n - 1 = 4$ degrés de liberté. On calcule la valeur de Y comme ceci :

O_i	4	51	88	47	10
C_i	7.540	45.900	86.200	50.400	9.100
$(O_i - C_i)^2$	12.53	26.01	3.24	11.56	0.81
$\frac{(O_i - C_i)^2}{C_i}$	1.66	0.57	0.04	0.23	0.09

En sommant la dernière ligne, on trouve $Y = 2.585$.

On cherche dans la table de la loi du χ^2 , le quantile u_c pour la probabilité $1 - \alpha = 95\%$ à 4 degrés de liberté. On trouve $u_c = 9.488$.

Comme $Y < u_c$, on ne peut pas rejeter, au seuil $\alpha = 5\%$, l'hypothèse H_0 .