Université de Bourgogne 2023-2024 Statistique

durée du test: 30min. Correction.

Exercice : On cherche à déterminer si les enfants en bas âge ont une préférence pour une couleur. Pour cela, on fait une expérience avec un groupe d'enfants de 2 ans. Chaque enfant se voit proposer cinq seaux de couleurs différentes.
L'expérience donne le résultat suivant : sur les 22 enfants de l'échantillon, 7 ont choisi le seau rouge.

1. Peut-on conclure avec un risque inférieur à 5% que les enfants de bas âge préfèrent la couleur rouge à la couleur verte ?
2. Quelle est la p-value de ce test ?
3. Quelle est la puissance $\eta(0.4)$ du test ?

Correction:

1. Les hypothèses du test (2 points):
   (H_0) les enfants n'ont pas de préférence entre les couleurs. $p = 0.2$
   (H_1) les enfants de bas âge préfèrent le rouge $p > 0.2$.
   Ici p représente la probabilité pour un enfant de bas âge de choisir le seau rouge plutôt que les autres seaux.

2. le modèle statistique (1 point): il s'agit d'un petit échantillon. En notant X_{22} le nombre d'enfants qui choisissent le seau rouge dans un échantillon de taille 22. On sait que la loi utilisée est la loi binômiale:
   $$X_{22} \sim \mathcal{B}(22; 0.2)$$

3. Le niveau du test est choisi avec le risque de premier espèce $\alpha = 0.05$ (0.5 point)

4. La région de rejet (2 points):
   $$\mathbb{P}(X_{22} \geq 9) = 1 - \mathbb{P}(X_{22} \leq 8) = 1 - 0,9799 = 0,0201$$
   $$\mathbb{P}(X_{22} \geq 8) = 1 - \mathbb{P}(X_{22} \leq 7) = 1 - 0,9439 = 0,0561 > 0,05$$
   Il s'en suit donc que la région de rejet à 5% est donnée par
   $$K_{0.05} = \{9, 10, 11, \ldots, 21, 22\},$$
   et peut s'écrire aussi $K_{0.05} = \{X_2 \geq 9\}$.

5. La décision (1 point): la valeur de l'expérience est donnée par $X_{22}^{\text{exp}} = 7 \notin K_{0.05}$. La valeur n'est pas dans la région critique. On ne rejette donc pas l'hypothèse nulle: il n'y a donc pas de préférence pour la couleur rouge chez les enfants de bas âge.

6. la p-value (1,5 point): la p-value est la probabilité de la région de rejet qui admet à son bord la valeur de l'expérience c'est-à-dire 7.
   $$p - \text{value} = \mathbb{P}(X_{22} \geq X_{22}^{\text{exp}}) = \mathbb{P}(X_{22} \geq 7) = 1 - \mathbb{P}(X_{22} \leq 6) = 1 - 0,8670 = 0,1330$$
   La p-value est donc de $13,3\%$. Le test n'est pas vraiment significatif. Pour un risque de première espèce supérieur à $13,3\%$ on accepte (H_1) et pour un risque inférieur à $13,3\%$ on accepte (H_0).

7. La puissance du test (2 points). On veut calculer $\eta(0.4)$. On rappelle la définition de la puissance:

$$\begin{aligned}\eta(0.4) &= \mathbb{P}(\text{Accepter } (H_1) \mid p = 0.4) = \mathbb{P}(X_{22} \in K_{0.05} | \, p = 0.4) = \mathbb{P}(X_{22} \geq 9 \mid p = 0.4) \\ &= 1 - 0,4540 = 0,5460\end{aligned}$$

Le passage de $p = 0,2$ à $p = 0,4$ n'est remarqué que dans $54,6\%$ des cas: le test n'est pas suffisamment puissant.