# Correction des Travaux dirigés de Probabilités et Statistiques - Tests statistiques

Julian Tugaut[1]

1. Si vous trouvez des erreurs de Français ou de mathématiques ou bien si vous avez des questions et/ou des suggestions, envoyez-moi un mail à julian.tugaut@univ-st-etienne.fr

# Table des matières

# Exercice 88

## Énoncé

On observe le nombre de filles dans 320 fratries de cinq enfants :

| Nombre de filles | Nombre de fratries |
|---|---|
| 0 | 18 |
| 1 | 56 |
| 2 | 110 |
| 3 | 88 |
| 4 | 40 |
| 5 | 8 |

**1)** *En supposant que le genre d'un enfant suit une loi de Bernoulli de paramètre* $\frac{1}{2}$, *quelle est la loi du nombre de filles dans une fratrie de cinq enfants ?*
**2)** *La distribution du nombre de filles suit-elle une loi binomiale de paramètres* $m := 5$ *et* $p := 0.5$ *au seuil de risque de* 5% *? Conclure.*

## Correction

### Correction du 1)

On sait que la somme de cinq variables aléatoires indépendantes qui suivent la loi de Bernoulli de paramètre 0.5 suit une loi binomiale de paramètres $m := 5$ et $p := 0.5$.

### Correction du 2)

Comme les effectifs théoriques sont tous supérieurs à 5 (et même supérieurs à 10), le test du khi-deux va pouvoir être utilisé.
On procède au test du $\chi^2$ comme suit :

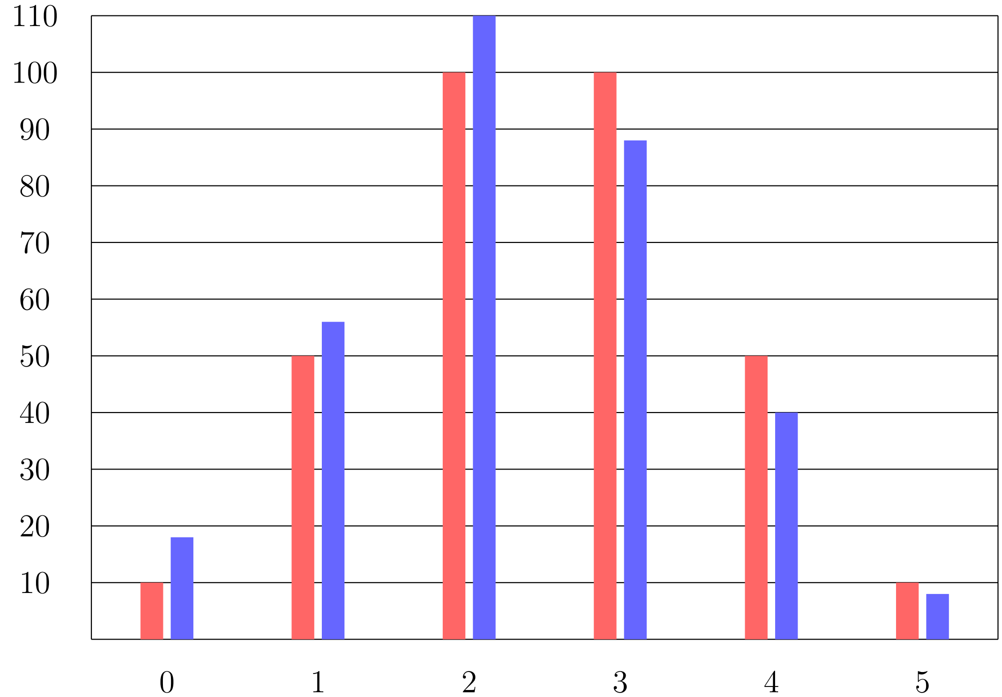
TABLE 1 – Test du $\chi^2$

| Nombre de filles | 0 | 1 | 2 | 3 | 4 | 5 | Somme |
|---|---|---|---|---|---|---|---|
| $n_i$ | 18 | 56 | 110 | 88 | 40 | 8 | 320 |
| $np_i$ | 10 | 50 | 100 | 100 | 50 | 10 | 320 |
| $\frac{(n_i-np_i)^2}{np_i}$ | 6.4 | 0.72 | 1 | 1.44 | 2 | 0.4 | $\chi^2_{obs} = 11.96$ |

Le nombre de degrés de liberté est $6 - 1 = 5$. Au seuil de risque de 5%, d'après la table, $\chi^2_c \approx 11.1$, donc on refuse l'hypothèse selon laquelle la distribution suit la loi binomiale $\mathcal{B}(5; 0.5)$. Conséquemment, l'hypothèse suivant laquelle le genre est équidistribué est rejetée.

Comparons graphiquement les effectifs théoriques et les effectifs réels :

Figure 1 – Comparaison des effectifs réels (en bleu) et théoriques (en rouge)



# Exercice 89

## Énoncé

Au départ d'une course de chevaux, il y a habituellement huit positions de départ et la position numéro un est la plus proche de la palissade. On soupçonne qu'un cheval a plus de chances de gagner quand il porte un numéro faible, c'est-à-dire quand il est plus proche de la palissade intérieure. Voici les données de 144 courses :

| Numéro de départ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nombre de victoires d'un cheval ayant ce numéro | 31 | 17 | 16 | 27 | 15 | 12 | 13 | 13 |

L'hypothèse nulle $(H_0)$ est l'hypothèse d'équiprobabilité des numéros de départ dans les places de premier à l'arrivée.

Nous prendrons comme hypothèse alternative $(H_1)$ la non équiprobabilité des numéros de départ.

Nous supposons que l'échantillon de 144 courses est un échantillon aléatoire.

Tester l'hypothèse $(H_0)$ contre l'hypothèse $(H_1)$ au risque d'erreur 5%.

## Correction

Dans l'hypothèse d'équiprobabilité, chaque numéro de départ a la même probabilité que les autres de conduire à une place de premier, soit $\frac{1}{8}$.

Dans cette hypothèse, les nombres de victoires d'un cheval devraient être :

Table 2 – Effectifs théoriques

| Numéro de départ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Nombre de victoires d'un cheval ayant ce numéro | 18 | 18 | 18 | 18 | 18 | 18 | 18 | 18 |

Comme les effectifs théoriques sont tous supérieurs à 5 (et même supérieurs à 10), le test du khi-deux va pouvoir être utilisé.

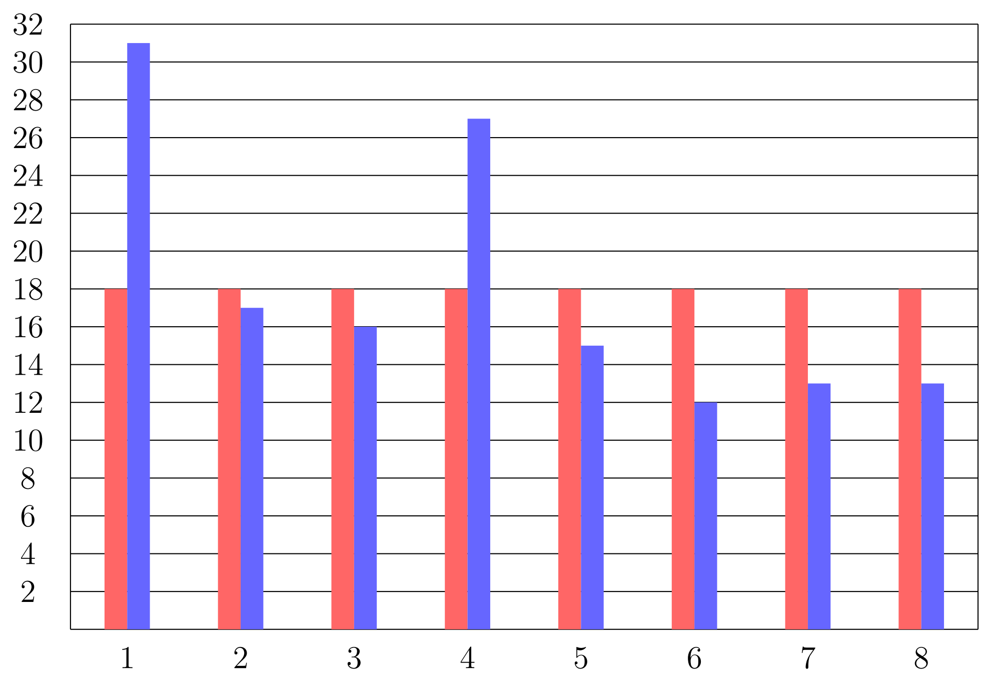
Le khi deux observé est donc

$$\begin{aligned}
\chi^2 &= \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i} \\
&= \frac{(31-18)^2}{18} + \frac{(17-18)^2}{18} + \cdots + \frac{(13-18)^2}{18} \\
&= \frac{1}{18}\left(13^2 + 1^2 + 2^2 + 9^2 + 3^2 + 6^2 + 5^2 + 5^2\right) \\
&= \frac{169 + 1 + 4 + 81 + 9 + 36 + 25 + 25}{18} \\
&= \frac{350}{18} \\
&\approx 19.44\,.
\end{aligned}$$

Le nombre de degrés de liberté est $8 - 1 = 7$. La table du khi-deux pour $n = 7$ et pour $1 - 5\% = 1 - 0.05 = 0.95$ nous donne une valeur égale à 14.1. Comme $19.44 > 14.1$, la différence observée permet de rejeter l'hypothèse d'équiprobabilité au risque d'erreur 5%.

Comparons graphiquement les effectifs théoriques et les effectifs réels :

Figure 2 – Comparaison des effectifs réels (en bleu) et théoriques (en rouge)



# Exercice 90

## Énoncé

On a noté sur 6 800 individus les couleurs des cheveux et des yeux. On a construit le tableau de contingence :

TABLE 3 – Tableau de contingence

| couleur des cheveux / couleur des yeux | blond | châtain | noir | roux |
|---|---|---|---|---|
| bleu | 1 768 | 807 | 189 | 47 |
| gris ou vert | 946 | 1 387 | 746 | 53 |
| marron | 115 | 438 | 288 | 16 |

Effectuer un test du khi-deux d'indépendance pour vérifier que la couleur des yeux n'est pas indépendante de la couleur des cheveux.

## Correction

Yeux bleus : $p_1 = 2\ 811$ ; Yeux verts : $p_2 = 3\ 132$ ; Yeux marrons : $p_3 = 857$.

Cheveux blonds : $q_1 = 2\ 829$ ; Cheveux châtains : $q_2 = 2\ 632$ ; Cheveux noirs : $q_3 = 1\ 223$ ; Cheveux roux : $q_4 = 116$.

On calcule donc la quantité suivante :

$$\begin{aligned}
\chi^2 =& \frac{1}{6\ 800} \times \sum_{i=1}^{3}\sum_{j=1}^{4} \frac{(6\ 800 \times n_{i,j} - p_i q_j)^2}{p_i q_j} \\
&= \frac{1}{6\ 800}\left(\frac{(6\ 800 \times 1\ 768 - 2\ 811 \times 2\ 829)^2}{2\ 811 \times 2\ 829} + \frac{(6\ 800 \times 807 - 2\ 811 \times 2\ 632)^2}{2\ 811 \times 2\ 632}\right. \\
&+ \frac{(6\ 800 \times 189 - 2\ 811 \times 1\ 223)^2}{2\ 811 \times 1\ 223} + \frac{(6\ 800 \times 47 - 2\ 811 \times 116)^2}{2\ 811 \times 116} \\
&+ \frac{(6\ 800 \times 946 - 3\ 132 \times 2\ 829)^2}{3\ 132 \times 2\ 829} + \frac{(6\ 800 \times 1\ 387 - 3\ 132 \times 2\ 632)^2}{3\ 132 \times 2632} \\
&+ \frac{(6\ 800 \times 746 - 3\ 132 \times 1\ 223)^2}{3\ 132 \times 1\ 223} + \frac{(6\ 800 \times 53 - 3\ 132 \times 116)^2}{3\ 132 \times 116} \\
&+ \frac{(6\ 800 \times 115 - 857 \times 2\ 829)^2}{857 \times 2\ 829} + \frac{(6\ 800 \times 438 - 857 \times 2\ 632)^2}{857 \times 2\ 632} \\
&+ \left.\frac{(6\ 800 \times 288 - 857 \times 1\ 223)^2}{857 \times 1\ 223} + \frac{(6\ 800 \times 16 - 857 \times 116)^2}{857 \times 116}\right) \\
&\approx 1\ 073.5
\end{aligned}$$

On dispose de $(3-1)(4-1) = 6$ degrés de liberté. Alors, vu le nombre obtenu, on rejette l'hypothèse d'indépendance de la couleur des yeux et des cheveux, même avec un seuil de risque de 0.5%.