

Méthodes Statistiques
Corrigé de l'exercice 44

Sur 4000 naissances dans une maternité de Chine, on a relevé 2075 garçons.

44-1) Cette observation est-elle conforme avec l'hypothèse que la probabilité théorique qu'il naisse un garçon est de 0.5? On prendra successivement $\alpha = 5\%$ et $\alpha = 1\%$. Commenter la différence de conclusion entre les deux cas.

On fait l'hypothèse H_0 suivante :

$$H_0 : p = p_0 = 0.5$$

La question posée nous conduit à faire un test bilatéral, autrement dit à considérer l'hypothèse H_1 suivante :

$$H_1 : p \neq 0.5$$

La statistique du test est :

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

où \hat{p} est la fréquence empirique.

Sous l'hypothèse H_0 , la variable Z suit asymptotiquement une loi normale $\mathcal{N}(0, 1)$. Comme on a un grand échantillon ($4000 > 30$), on peut utiliser ce résultat asymptotique.

On calcule

$$\hat{p} = \frac{2075}{4000} = 0.519$$

On obtient alors

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.51875 - 0.5}{\sqrt{\frac{0.5(1 - 0.5)}{4000}}} = 2.372$$

• **Calcul au seuil 5%**

Le quantile de la loi normale $\mathcal{N}(0, 1)$ au seuil 5% est 1.96.

Puisque $2.372 > 1.96$, on rejette l'hypothèse H_0 , au risque 5% de se tromper.

• **Calcul au seuil 1%**

Le quantile de la loi normale $\mathcal{N}(0, 1)$ au seuil 1% est 2.576.

Puisque $-2.576 < 2.372 < 2.576$, on accepte l'hypothèse H_0 .

Le fait de fixer le risque d'erreur à 1% plutôt que 5% conduit à prendre moins de risques de rejeter à tort : alors qu'à 5% on rejetait, à seulement 1%, on ne prend pas le risque de rejeter.

44-2) L'exercice 13 indiquait une proportion de 52% (et non pas 50%). Les données observées s'accordent-elles avec cette valeur au seuil 5% ?

L'hypothèse H_0 devient maintenant :

$$H_0 : p = p_0 = 0.52$$

On recalcule la statistique du test :

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.51875 - 0.52}{\sqrt{\frac{0.52(1-0.52)}{4000}}} = -0.158$$

Le quantile de la loi normale $\mathcal{N}(0, 1)$ au seuil 5% est 1.96.

Puisque $-1.96 < -0.158 < 1.96$, on accepte l'hypothèse H_0 . Les résultats observés sont plus conformes à un *sex-ratio* de 52% que de 50%.