Test d'hypothèse

Voir diapo de la semaine 11 pour une introduction.

Test d'hypothèse sur la valeur d'une moyenne

Soit μ la moyenne d'une population. On fait l'hypothèse nulle $H_0 : \mu = \mu_0$, où μ_0 est nombre fixé. On utilise la statistique de test $T_{stat} = \frac{X_1 + \cdots + X_n}{n}$, qui est l'estimateur de la moyenne.

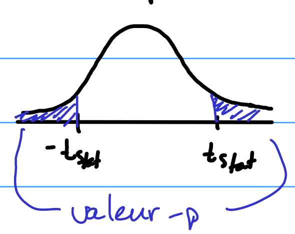
Soit σ^2 la variance de la population. Lorsque n est grand,

- si σ^2 est connue, on a $T_{stat} \approx N(\mu, \frac{\sigma^2}{n})$ et cette distribution est connue sous H_0.
- si σ^2 n'est pas connue, alors $\frac{T_{stat} - \mu}{\hat{\sigma}/\sqrt{n}} \approx Student(n-1)$ et cette distribution est connue sous H_0.

Dans ces cas, on calcule la valeur-p par

$$\text{valeur-p} = \mathbb{P}(|T_{stat}| \geq t_{stat} \mid H_0) = 2\,\mathbb{P}(T_{stat} \geq t_{stat} \mid H_0)$$

On met les valeurs absolues pour tenir compte des valeurs extrêmes $T_{stat} \geq t_{stat}$ et $T_{stat} \leq -t_{stat}$.



Ex. X_1, X_2, X_3 idd $\sim N(\mu, 16)$ et $H_0 : \mu = 5$, $H_1 : \mu \neq 5$.

$T_{stat} = \dfrac{\bar{X} - 5}{4} \sim N(0,1)$. On observe $x_1 = 6$, $x_2 = 4,5$, $x_3 = 6,5$.

$$\Rightarrow \bar{x} = \frac{6 + 4,5 + 6,5}{3} = 5,\bar{6}.$$

$$t_{stat} = \frac{\bar{x} - 5}{4} = 0,1\bar{6}.$$

$$p = \mathbb{P}(|T_{stat}| \geq t_{stat} \mid H_0)$$

$$= 2\mathbb{P}(T_{stat} \geq t_{stat} \mid H_0)$$

$$= 2\mathbb{P}(T_{stat} \geq 0,1\bar{6} \mid H_0) = 2(1 - \Phi(0,1\bar{6})) = 0,8676$$

Avec un seuil de $\alpha = 5\%$, on ne rejette pas H_0.

Ex. Reprenons le même exemple, mais supposons qu'on observe $x_1 = 7$, $x_2 = 8$, $x_3 = 8,5$.

$$\bar{x} = \frac{7 + 8 + 8,5}{3} = 7,8\bar{3}$$

$$t_{stat} = \frac{\bar{x} - 5}{4} = 0,708\bar{3}.$$

$$\text{valeur-}p = 2\mathbb{P}(T_{stat} \geq 0,708\bar{3} \mid H_0) = 0,4787$$

Pour un seuil $\alpha = 5\%$, on ne rejette pas H_0.

Ex La durée de vie d'un type d'ampoule suit une loi exponentielle de paramètre λ. On observe la durée de vie de 150 ampoules et on trouve que $\bar{x} = 2012$ (en heures) et $s = 50$ (en heures), où s^2 est la variance empirique non biaisée.
On voudrait prouver qu'en moyenne, ces ampoules durent plus de 2000 heures.

Sol. On fait l'hypothèse $H_0 : \mu = 2000$ et $H_1 : \mu \neq 2000$.

Soit $T_{stat} = \dfrac{\bar{X} - 2000}{\hat{\sigma}/\sqrt{n}}$, où $\hat{\sigma}^2$ est l'estimateur non biaisé de la variance et $n = 150$

Alors $T_{stat} \approx \text{Student}(149) \approx \mathcal{N}(0,1)$. (Lorsque le degré de liberté est grand, la de Student est à peu une loi normale)

$$t_{stat} = \frac{2012 - 2000}{50/\sqrt{150}} = 2{,}939$$

$$\text{valeur-p} = \mathbb{P}(|T_{stat}| \geq t_{stat} \mid H_0) = 2\mathbb{P}(T_{stat} \geq 2{,}939 \mid H_0)$$

$$\approx 2\left(1 - \Phi(2{,}939)\right) = 0{,}00329$$

Avec un seuil de 1%, on a bien valeur-p $\leq 1\%$, donc on rejette l'hypothèse nulle et on a prouvé $H_1 : \mu \neq 2000$. Comme $\bar{x} > 2000$, on conclut que $\mu > 2000$.