LOIS D'ÉCHANTILLONNAGE

CARACTERE QUALITATIF

a) - Distribution de la fréquence

On considère une modalité d'un caractère qualitatif, de fréquence p dans une population $\mathcal{P}$. On constitue des échantillons de taille n avec remise.

F est la variable aléatoire qui, à chaque échantillon, associe la fréquence de la modalité dans l'échantillon.

- ❑ La loi de nF est la loi binomiale $\mathcal{B}(n\,;p)$.
- ❑ La loi de F est approchée par la loi normale $\mathcal{N}\left(p\,;\sqrt{\dfrac{p\,(1-p)}{n}}\right)$ pour $n \geq 30$, $np \geq 15$ et $np(1-p) > 5$.
- ❑ La loi de $\dfrac{F-p}{\sqrt{\dfrac{F\,(1-F)}{n}}}$ est approchée par la loi normale $\mathcal{N}(0\,;1)$ pour $n \geq 50$.
- ❑ La loi de $n\,F$ est approchée par la loi de Poisson P(np) pour $n \geq 30$, $p \leq 0{,}1$ et $n\,p < 15$.

b) - Distribution de la différence de deux fréquences

On considère une modalité du caractère ayant pour fréquences p_1 et p_2 respectivement dans deux populations $\mathcal{P}_1$ et $\mathcal{P}_2$.

Pour $i \in \{1\,;2\}$, F_i est la variable aléatoire qui, à chaque échantillons avec remise de taille n_i issu de $\mathcal{P}_i$, associe la fréquence de la modalité dans l'échantillon. F_1 et F_2 sont supposées indépendantes.

On pose $F = \dfrac{n_1\,F_1 + n_2\,F_2}{n_1 + n_2}$.

- ❑ **Si $p_1 = p_2$**, la loi de $\dfrac{F_1 - F_2}{\sqrt{F\,(1-F)\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$ est approchée par la loi normale $\mathcal{N}(0\,;1)$ si $\begin{cases} n_1 \text{ et } n_2 \text{ sont supérieurs ou égaux à } 30 \\ n_1 p_1 \text{ et } n_2 p_2 \text{ sont supérieurs ou égaux à } 15 \\ n_1 p_1 (1-p_1) \text{ et } n_2 p_2 (1-p_2) \text{ sont supérieurs à } 5 \end{cases}$

CARACTERE QUANTITATIF

a) - Distribution de la moyenne

On considère un caractère quantitatif de moyenne μ et d'écart type σ dans une population $\mathcal{P}$. On constitue des échantillons avec remise de taille n.

$\overline{X}$ et $\widehat{S}$ sont les variables aléatoires qui, à chaque échantillon, associent respectivement la moyenne et l'écart-type corrigé du caractère dans l'échantillon.

- ❑ La loi de $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$ est
  - la loi normale $\mathcal{N}(0\,,1)$ si le caractère est distribué normalement dans la population
  - approchée par la loi normale $\mathcal{N}(0\,;1)$ si $n \geq 30$
- ❑ La loi de $\dfrac{\overline{X}-\mu}{\dfrac{\widehat{S}}{\sqrt{n}}}$ est
  - la loi de Student à $n-1$ degrés de liberté si le caractère est distribué normalement dans la population
  - approchée par la loi normale $\mathcal{N}(0\,;1)$ si $n \geq 30$

b) - Autres distributions

On considère un caractère quantitatif ayant pour moyennes μ_1 et μ_2 et pour écarts-types σ_1 et σ_2 respectivement dans deux populations $\mathcal{P}_1$ et $\mathcal{P}_2$.

Pour $i \in \{1\,;2\}$, $\overline{X_i}$ (respectivement : $\widehat{S_i}$) est la variable aléatoire qui, à chaque échantillon avec remise de taille n_i issu de $\mathcal{P}_i$, associe la moyenne (respectivement : l'écart-type corrigé) du caractère dans l'échantillon.

X_1 et X_2 d'une part, $\widehat{S_1}$ et $\widehat{S_2}$ d'autre part, sont supposées indépendantes. On pose $\widehat{S} = \sqrt{\dfrac{(n_1-1)\,\widehat{S_1}^2 + (n_2-1)\,\widehat{S_2}^2}{n_1+n_2-2}}$.

Différence de deux moyennes

- ❑ **Si $\mu_1 = \mu_2$**, la loi de $\dfrac{\overline{X_1}-\overline{X_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$ est
  - la la loi normale $\mathcal{N}(0\,,1)$ si le caractère est distribué normalement dans les deux populations
  - approchée par la loi normale $\mathcal{N}(0\,;1)$ si n_1 et n_2 sont supérieurs ou égaux à 30

<u>**Si σ_1 et σ_2 sont égaux**</u> **:**

- ❑ **Si $\mu_1 = \mu_2$**, la loi de $\dfrac{\overline{X_1}-\overline{X_2}}{\widehat{S}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ est
  - la loi de Student à n_1+n_2-2 degrés de liberté si le caractère est distribué normalement dans les deux populations
  - approchée par la loi normale $\mathcal{N}(0\,;1)$ si le caractère est distribué normalement dans les deux populationset si $n_1+n_2-2 \geq 30$
  - approchée par la loi normale $\mathcal{N}(0\,;1)$ si n_1 et n_2 sont supérieurs ou égaux à 30

<u>**Si σ_1 et σ_2 sont différents**</u> **:**

- ❑ **Si $\mu_1 = \mu_2$**, la loi de $\dfrac{\overline{X_1}-\overline{X_2}}{\sqrt{\dfrac{\widehat{S_1}^2}{n_1}+\dfrac{\widehat{S_2}^2}{n_2}}}$ est approchée par la loi normale $\mathcal{N}(0\,;1)$ si n_1 et n_2 sont supérieurs ou égaux à 30.

Avec des variances

- ❑ **Si $\sigma_1 = \sigma_2$**, les lois de $\dfrac{\widehat{S_1}^2}{\widehat{S_2}^2}$ et de $\dfrac{\widehat{S_2}^2}{\widehat{S_1}^2}$ sont les lois F de Snédécor à respectivement $(n_1-1\,;n_2-1)$ et $(n_2-1\,;n_1-1)$ degrés de liberté si le caractère est distribué normalement dans les deux populations.
- ❑ **Si $\sigma_1 = \sigma_2$**, la loi de $\dfrac{\widehat{S_1}^2 - \widehat{S_2}^2}{\sqrt{2}\sqrt{\dfrac{\widehat{S_1}^4}{n_1}+\dfrac{\widehat{S_2}^4}{n_2}}}$ est approchée par la loi normale $\mathcal{N}(0\,;1)$ si n_1 et n_2 sont supérieurs ou égaux à 30.