STT 1000 - STATISTIQVES

ARTHVR CHARPENTIER

Loi Multinomiale

Loi binomiale, $X \sim \mathcal{B}(n, p)$,

$$\mathbb{P}[X = x] = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \ x \in \{0, 1, 2, \cdots, n-1, n\}.$$

Notons $X = (X_0, X_1)$ le nombre de 0 et de 1,

$$\mathbb{P}[X_0 = x_0, X_1 = x_1] = \frac{n!}{x_0! x_1!} p_1^{x_1} p_0^{x_0}, \text{ où } x_0 + x_1 = n,$$

où $p_0 = 1 - p_1$.

Loi Multinomiale

Loi multinomiale, à k catégories, de paramètre $\boldsymbol{p} = (p_1, p_2, \cdots, p_k)$, $p_1 + p_2 + \cdots + p_k = n$.
$X \in \{1, 2, \cdots, k\}$, $X = (X_1, X_2, \cdots, X_k)$ le nombre de 1, 2, ... et de k,

$$\mathbb{P}[X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k] = \frac{n!}{x_1!x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

où $x_1 + x_2 + \cdots + x_k = n$,

$$\mathbb{E}\big[X_j\big] = np_j,\ \text{Var}[X_j] = np_j(1 - p_j) \text{ et } \text{Cov}[X_i, X_j] = -np_i p_j.$$

Maximum de vraisemblance

Loi Multinomiale

D'après le théorème central limite,

$$\frac{x_j - np_j}{\sqrt{np_j(1-p_j)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \text{ lorsque } n \rightarrow \infty$$

$$\frac{(x_j - np_j)^2}{np_j(1-p_j)} \xrightarrow{\mathcal{L}} \chi^2(1), \text{ lorsque } n \rightarrow \infty$$

Si les variables $X_1, X_2, \cdots, X_m$ étaient dépendantes

$$\sum_{j=1}^{m} \frac{(x_j - np_j)^2}{np_j(1-p_j)} \xrightarrow{\mathcal{L}} \chi^2(m), \text{ lorsque } n \rightarrow \infty$$

... mais ce n'est pas vrai...

Si les x_j sont suffisamment grand,

$$\sum_{j=1}^{m} \frac{(x_j - np_j)^2}{np_j(1-p_j)} \xrightarrow{\mathcal{L}} \chi^2(m-1), \text{ lorsque } n \rightarrow \infty$$

Pourquoi passer de m à $m-1$?

Test sur une loi multinomiale

$$H_0 : p_j = p_j^0 \text{ pour tout } j = 1, \cdots, k$$

$$Q = \sum_{j=1}^{k} \frac{(X_j - np_j^0)^2}{np_j^0}$$

si H_0 est vraie, on doit espérer avoir $Q \sim \chi^2(k-1)$.

Test d'ajustement du χ^2

$$H_0 : p_j = p_j^0 \;\text{ pour tout } j = 1, \cdots, k$$

$$Q = \sum_{j=1}^{k} \frac{(X_j - np_j^0)^2}{np_j^0}$$

si H_0 est vraie, on doit espérer avoir $Q \sim \chi^2(k-1)$.
La règle de décision sera de rejeter H_0 si $Q_{obs} > q_{1-\alpha,k-1}$.

Test d'ajustement de loi

Application à un loi de Poisson $\mathcal{P}(\lambda)$.

Test d'indépendance

Considérons deux variables catégorielles.
On appelle tableau de contingence...