STATISTIQUE DESCRIPTIVE

$\bar{y} = \frac{\sum_i y_i}{n}$; $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (y_i - \bar{y})^2 = \frac{\sum_{i=1}^n y_i^2}{n} - \bar{y}^2$; $s^2 = \frac{n}{n-1}\ \hat{\sigma}^2$; $r(x,y) = \frac{1}{n}\sum_{i=1}^n \frac{x_i - \bar{x}}{\hat{\sigma}_x}\ \frac{y_i - \bar{y}}{\hat{\sigma}_y}$

$\bar{y} = \frac{1}{n}\sum_{k=1}^K n_k y_k = \sum_{k=1}^K w_k y_k$; $\hat{\sigma}^2 = \sum_{k=1}^K w_k (y_k - \bar{y})^2 = \left(\sum_{k=1}^K w_k y_k^2\right) - \bar{y}^2$ avec $w_k = n_k/n$

QUELQUES NOTIONS DE PROBABILITE

$\mathrm{P}(AB|C) = \mathrm{P}(A|BC)\ \mathrm{P}(B|C) = \mathrm{P}(B|AC)\ \mathrm{P}(A|C)$; $\mathrm{P}(A+B|C) = \mathrm{P}(A|C) + \mathrm{P}(B|C) - \mathrm{P}(AB|C)$

Si ..., $\mathrm{P}(A) = \mathrm{P}(A|B_1)\mathrm{P}(B_1) + \ldots + \mathrm{P}(A|B_n)\mathrm{P}(B_n)$; $\mathrm{P}(A) = \mathrm{P}(A|B)\mathrm{P}(B) + \mathrm{P}(A|\bar{B})\mathrm{P}(\bar{B})$

Si $\mathrm{P}(A) > 0$, alors $\mathrm{P}(B|A) = \frac{\mathrm{P}(A|B)\mathrm{P}(B)}{\mathrm{P}(A)}$

Si $Y_1, \ldots, Y_n$ sont des v.a. indépendantes, alors $\mathrm{V}(Y_1 + \ldots + Y_n) = \mathrm{V}(Y_1) + \ldots + \mathrm{V}(Y_n)$

Si $Z = a + bY$: $\mathrm{E}(Z) = a + b\mu_Y$; $\sigma_Z^2 = b^2\sigma_Y^2$

<u>Variables aléatoires discrètes</u>:

$\mu_Y = \mathrm{E}(Y) = \sum_{y_i\ \in\ \mathcal{E}} p_i y_i$; $\sigma_Y^2 = \mathrm{V}(Y) = \mathrm{E}(Y-\mu)^2 = \sum_{y_i\ \in\ \mathcal{E}} p_i (y_i - \mu)^2 = \mathrm{E}(Y^2) - \mu^2$

Bernoulli: $\mathcal{E}_i = \{0,1\}$ et $p_0 = 1 - p$; $p_1 = p$; $\mathrm{E}(X) = p$; $\mathrm{V}(X) = p(1-p)$

<u>Variables aléatoires continues</u>:

$\mu_X = \mathrm{E}(X) = \int_{\mathcal{E}} x\ f(x)\ dx$; $\sigma_X^2 = \mathrm{V}(X) = \mathrm{E}(X-\mu)^2 = \int_{\mathcal{E}} (x - \mu_X)^2\ f(x)\ dx = \mathrm{E}(X^2) - \mu_X^2$

Distribution normale: $X \sim N(\mu, \sigma^2)$; $\mathrm{E}(X) = \mu$; $\mathrm{V}(X) = \sigma^2$; $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

SONDAGES

$K = \frac{N!}{n!\ (N-n)!}$; $\sum_{k=1}^K p(s_k) = 1$; $\mathrm{E}(\hat{\theta}) = \sum_{k=1}^K p(s_k)\ \hat{\theta}(s_k)$; Biais: $\mathrm{E}(\hat{\theta}) - \theta$

$V(\hat{\theta}) = \mathrm{E}(\hat{\theta} - \mathrm{E}(\hat{\theta}))^2 = \sum_{k=1}^K p(s_k)(\hat{\theta}(s_k) - \mathrm{E}(\hat{\theta}))^2$; EQM $= \mathrm{E}(\hat{\theta} - \theta)^2 = \mathrm{V}(\hat{\theta}) + (\mathrm{BIAIS})^2$

$\hat{\theta}(s) = \sum_{i\in s} W_i(s) Y_i$; $W_i(s)$: *poids de sondage* ; $P_i = \sum_{s:\ i\in s} p(s)$; $\sum_{i=1}^N P_i = n$

<u>Sondage aléatoire simple</u>:

$P_i = \frac{n}{N} = f =$ taux de sondage

Total: $T = \sum_{i=1}^N Y_i$; $\hat{T} = \sum_{i\in s} \frac{Y_i}{P_i} = \sum_{i\in s} \frac{N}{n}\ Y_i$; $\mathrm{V}(\hat{T}) \approx N^2(1-f)\ \frac{s^2}{n}$

Moyenne: $\bar{Y} = \frac{T}{N}$ et $\hat{\bar{Y}} = \bar{y} = \sum_{i\in s} \frac{Y_i}{n}$; $\mathrm{V}(\bar{y}) \approx (1-f)\ \frac{s^2}{n}$; $\mathrm{IC}(\bar{Y}) \approx \bar{y} \pm 2\ \sqrt{\mathrm{V}(\bar{y})}$

Proportion π: $\hat{\pi} = p = \frac{y}{n}$; $\mathrm{V}(p) \approx (1-f)\ \frac{p(1-p)}{n}$; $\mathrm{IC}(\pi) \approx p \pm 2\ \sqrt{\mathrm{V}(p)}$

<u>Sondage stratifié</u>:

Strate h: $\bar{Y}_h = \sum_{i\in G_h} \frac{Y_i}{N_h}$ estimée sans biais par $\bar{y}_h = \sum_{i\in s_h} \frac{Y_i}{n_h}$.

$\bar{Y} = \sum_{h=1}^H \frac{N_h}{N}\ \bar{Y}_h$ estimée sans biais par $\hat{\bar{Y}}_{\mathrm{st}} = \sum_{h=1}^H \frac{N_h}{N}\ \bar{y}_h$; $\hat{\mathrm{V}}(\hat{\bar{Y}}_{\mathrm{st}}) \approx \sum_{h=1}^H \left(\frac{N_h}{N}\right)^2\ (1-f_h)\frac{s_h^2}{n_h}$

Proportion estimée par $p_{\mathrm{st}} = \hat{\pi}_{\mathrm{st}} = \sum_{h=1}^H \frac{N_h}{N} p_h$; $\hat{\mathrm{V}}(p_{\mathrm{st}}) \approx \sum_{h=1}^H \left(\frac{N_h}{N}\right)^2 (1-f_h)\frac{p_h(1-p_h)}{n_h}$

Allocation proportionnelle: $\frac{n_h}{n} = \frac{N_h}{N} \Rightarrow P_i = \mathrm{Pr}(\text{Etre choisi} \mid \text{strate h}) = f_h = \frac{n_h}{N_h} = \frac{n}{N} = f$

Allocation proportionnelle avec budget C: $n = \frac{C}{\sum_{h=1}^H \frac{N_h}{N} c_h}$

Allocation optimale de Neyman: $n_h = \frac{N_h\ S_h}{\sqrt{c_h}}\ \frac{C}{\sum_{l=1}^H N_l S_l \sqrt{c_l}}$