MTH2302B - Probabilités et statistique

INTERVALLES DE CONFIANCE BILATÉRAUX USUELS (un paramètre)

Situation	Loi utilisée	Intervalle de confiance: niveau $1-\alpha$
Une moyenne μ		
σ^2 est connue et $X \sim N(\mu, \sigma^2)$, ou n grand	$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\overline{X} - z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}}$
σ^2 est inconnue et $X \sim N(\mu, \sigma^2)$	$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim T_{n-1}$	$\overline{X} - t_{\alpha/2;n-1} \cdot \dfrac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\alpha/2;n-1} \cdot \dfrac{S}{\sqrt{n}}$
σ^2 est inconnue et n est très grand	$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$	$\overline{X} - z_{\alpha/2} \cdot \dfrac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2} \cdot \dfrac{S}{\sqrt{n}}$
Une variance σ^2		
$X \sim N(\mu, \sigma^2)$ et μ et σ^2 sont inconnues	$(n-1)\dfrac{S^2}{\sigma^2} \sim \chi^2_{n-1}$	$\dfrac{(n-1)S^2}{\chi^2_{\alpha/2;n-1}} \leq \sigma^2 \leq \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2;n-1}}$
Un écart-type σ		
(cas approximatif) n est très grand ($n \geq 40$)	$\dfrac{S-\sigma}{\sigma/\sqrt{2n}} \sim N(0,1)$	$\dfrac{S}{1+\frac{z_{\alpha/2}}{\sqrt{2n}}} \leq \sigma \leq \dfrac{S}{1-\frac{z_{\alpha/2}}{\sqrt{2n}}}$
Une proportion p		
$X \sim$ Bernoulli de paramètre p et n est très grand	$\dfrac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$	$\hat{p} - z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

Les intervalles de confiance ci-dessus sont calculés en considérant un échantillon aléatoire $X_1, \ldots, X_n$ d'une variable X.
La moyenne et la variance de l'échantillon sont: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, et $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ respectivement.
Pour un niveau de confiance $1-\alpha$ donné, les nombres $z_{\alpha/2}$ $t_{\alpha/2;n-1}$ $\chi^2_{\alpha/2;n-1}$ et $\chi^2_{1-\alpha/2;n-1}$ sont tels que :

$$\begin{aligned} \Phi(z_{\alpha/2}) &= 1-\alpha/2 \\ P(T_{n-1} > t_{\alpha/2;n-1}) &= \alpha/2 \\ P(\chi^2_{n-1} > \chi^2_{\alpha/2;n-1}) &= \alpha/2 \\ P(\chi^2_{n-1} > \chi^2_{1-\alpha/2;n-1}) &= 1-\alpha/2. \end{aligned}$$

INTERVALLES DE CONFIANCE BILATÉRAUX USUELS (deux paramètres)

Situation	Loi utilisée	Intervalle de confiance: niveau $1-\alpha$
Différence $\mu_1 - \mu_2$		
σ_1^2 et σ_2^2 sont connues et $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$ ou alors n_1 et n_2 grands.	$\dfrac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$(\overline{X}_1 - \overline{X}_2) \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$
σ_1^2 et σ_2^2 sont inconnues avec $\sigma_1^2 = \sigma_2^2 = \sigma^2$ et $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2$	$\dfrac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim T_{n_1+n_2-2}$	$(\overline{X}_1 - \overline{X}_2) \pm t_{\alpha/2;n_1+n_2-2} \cdot S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ où $S_p = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}$
σ_1^2 et σ_2^2 sont inconnues avec $\sigma_1^2 \neq \sigma_2^2$ et $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$	$\dfrac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim T_\nu$	$(\overline{X}_1 - \overline{X}_2) \pm t_{\alpha/2;\nu} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ où $\nu = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$
σ_1^2 et σ_2^2 sont inconnues et n_1, n_2 sont grands $(n_1 \geq 30$, $n_2 \geq 30)$	$\dfrac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$	$(\overline{X}_1 - \overline{X}_2) \pm z_{\alpha/2} \cdot \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$
Les observations sont couplées $D = X_1 - X_2 \sim$ Normale $D_j = X_{1j} - X_{2j}$, $j = 1, \ldots, n$	$\dfrac{\overline{D} - (\mu_1 - \mu_2)}{S_D/\sqrt{n}} \sim T_{n-1}$	$\overline{D} \pm t_{\alpha/2;n-1} \cdot \dfrac{S_D}{\sqrt{n}}$ où $\overline{D} = \frac{1}{n}\sum_{j=1}^{n} D_j, S_D = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(D_j - \overline{D})^2}$
Rapport σ_1^2/σ_2^2		
$X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$	$\dfrac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2} \sim F_{n_2-1;n_1-1}$	$L \leq \dfrac{\sigma_1^2}{\sigma_2^2} \leq U$ $L = \dfrac{S_1^2}{S_2^2} \cdot F_{1-\alpha/2;n_2-1,n_1-1}$ $U = \dfrac{S_1^2}{S_2^2} \cdot F_{\alpha/2;n_2-1,n_1-1}$
Différence $p_1 - p_2$		
$X_1 \sim$ Bernoulli $X_2 \sim$ Bernoulli n_1 et n_2 sont grands	$\dfrac{\hat{p_1} - \hat{p_2} - (p_1 - p_2)}{\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}} \sim N(0,1)$	$(\hat{p_1} - \hat{p_2}) \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p_1}(1-\hat{p_1})}{n_1} + \dfrac{\hat{p_2}(1-\hat{p_2})}{n_2}}$

Les intervalles de confiance ci-dessus sont calculés en considérant deux échantillons aléatoires $X_{11}, \ldots, X_{1n_1}$ et $X_{21}, \ldots, X_{2n_2}$, indépendants (sauf dans le cas des observations couplées), provenant de deux variables (populations) X_1 et X_2.

On considère : $\overline{X}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i}$ $\quad \overline{X}_2 = \frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i}$ $\quad S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_{1i} - \overline{X}_1)^2$ $\quad S_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(X_{2i} - \overline{X}_2)^2$.

Pour un niveau de confiance $1-\alpha$ donné, les nombres $z_{\alpha/2}$ $\; t_{\alpha/2;d}$ $\; \chi^2_{\alpha/2;n-1}$ et $\chi^2_{1-\alpha/2;n-1}$ sont tels que :

$$\Phi(z_{\alpha/2}) = 1 - \alpha/2 \quad \text{et} \quad P(T_d > t_{\alpha/2;d}) = \alpha/2.$$

Pour un niveau de confiance $1-\alpha$ donné, les nombres $F_{\alpha/2;n_2-1,n_1-1}$ et $F_{1-\alpha/2;n_2-1,n_1-1}$ sont tels que:
$P(F_{n_2-1;n_1-1} > F_{\alpha/2;n_2-1,n_1-1} = \alpha/2$, et $P(F_{n_2-1;n_1-1} > F_{1-\alpha/2;n_2-1,n_1-1} = 1 - \alpha/2$.