CY-Tech	Statistiques	Mathématiques
2020 - 2021	**Corrigé TD 2 (IDC)**	ING2-GI

Exercice 1

1) $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Les deux paramètres sont inconnus et on a un petit échantillon de taille $n = 9$.

IDC pour μ :

On cherche l'intervalle $[a, b]$ tel que : $P\ (a \leq \mu \leq b) = 1 - \alpha = 0.9$.

On utilise la moyenne $\overline{X}$ comme estimateur de μ. On sait que l'échantillon est gaussien avec σ^2 inconnu donc :

$W\ = \dfrac{\overline{X} - \mu}{S^*}\sqrt{n}\ \sim T_{n-1}$ loi de Student à $n-1$ degrés de liberté.

On cherche un intervalle symétrique centré en $\overline{X}$, d'où :

$$P\ (a \leq \mu \leq b) = P\ (\overline{X} - \ell \leq \mu \leq \overline{X} + \ell)\ =\ P(-\ell \leq \overline{X} - \mu \leq +\ell\ ) = 0.9$$

$$P\ \left(-\ell\frac{\sqrt{n}}{s^*} \leq \frac{\overline{X} - \mu}{S^*}\sqrt{n}\ \leq +\ell\frac{\sqrt{n}}{s^*}\ \right) = 0.9$$

La table de la loi de Student donne la valeur de t tel que $P(-t \leq T_8\ \leq t) = 0.9$

$t = 1.8595$, or $t = \ell\dfrac{\sqrt{n}}{s^*}$ donc $\ell = t\dfrac{s^*}{\sqrt{n}}$

L'IDC recherché est :

$$a = \overline{x} - \ell = 19.72 - 1.8595\frac{\sqrt{0.69}}{\sqrt{9}}\ = 19.21$$

$$b = \overline{x} + \ell = 19.72 + 1.8595\frac{\sqrt{0.69}}{\sqrt{9}}\ = 20.23$$

Donc la valeur moyenne μ de la pression sur toute la population des cylindres a 90% de chances de se trouver entre $19,21$ et $20,23$ kg/cm^2.

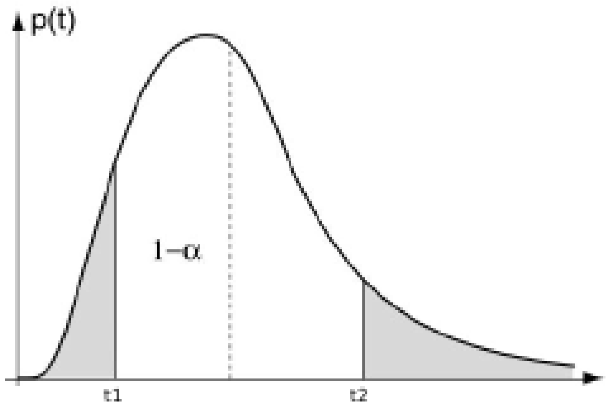
IDC pour σ^2 :

On cherche l'intervalle $[a, b]$ tel que : $P\ \left(a \leq \sigma^2 \leq b\right) = 1 - \alpha = 0.9$.

On utilise la variance empirique corrigée sans biais, S^{*2}, comme estimateur de σ^2.

L'échantillon est gaussien, et μ est inconnue donc :

$W = (n-1)\dfrac{S^{*2}}{\sigma^2}\ \sim \chi^2_{n-1}$ loi du khi-deux à $n-1$ degrés de liberté



On commence par chercher t_1 et t_2, tels que : $P(t_1 \leq W \leq t_2) = 0.9$
Par défaut, on suppose le risque symétrique, c à d : $\dfrac{\alpha}{2}$ en dessous de t_1 et $\dfrac{\alpha}{2}$ au dessus de t_2.
Mais la loi du chi-deux n'étant pas symétrique, on doit chercher les deux valeurs :

— $P(W > t_2) \ = \ 0.05 \iff P(W \leq t_2) \ = \ 0.95$
  La table de χ^2_8 donne $t_2 = 15.507$.

— $P(W \leq t_1) \ = \ 0.05$

La table de χ^2_8 donne $t_1 = 2.733$.
La réunion de ces deux informations donne :

$$P(t_1 \leq W \leq t_2) = 0.9 \iff P\left(t_1 \leq (n-1)\frac{s^{*2}}{\sigma^2} \leq t_2\right) = 0.9$$

En inversant, on obtient :

$$P\left(\frac{1}{t_2} \leq \frac{\sigma^2}{(n-1)s^{*2}} \leq \frac{1}{t_1}\right) = 0.9 \iff P\left(\frac{(n-1)s^{*2}}{t_2} \leq \sigma^2 \leq \frac{(n-1)s^{*2}}{t_1}\right) = 0.9$$

Les deux extrémités de l'IDC sont donc :

$$a = \frac{(n-1)s^{*2}}{t_2} = \frac{8 \times 0.69}{15.507} = 0.356$$

$$b = \frac{(n-1)s^{*2}}{t_1} = \frac{8 \times 0.69}{2.733} = 2.02$$

Donc σ^2 a 90% de chances de se trouver dans l'intervalle [0.356 ; 2.02] .
En prenant la racine carrée, on obtient : $\sigma \in [0.6 \ ; \ 1.42]$.

2) La variance de la population est connue, donc . On sait alors que

$$Z \ = \ \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \ \sim \mathcal{N}(0,1)$$

$$P\ (\overline{X} - \ell \leq \mu \leq \overline{X} + \ell) \ = \ P(-\ell \leq \overline{X} - \mu \leq +\ell\ ) = 0.9$$

$$P\ \left(-\ell\frac{\sqrt{n}}{\sigma} \leq \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \ \leq +\ell\frac{\sqrt{n}}{\sigma} \ \right) = 0.9$$

La table de la loi normale standard donne $t = 1.645$ comme valeur asurant :
$P(-t \leq Z \leq t) = 0.9$, on en déduit : $t = \ell\dfrac{\sqrt{n}}{\sigma}$ et donc $\ell = t\dfrac{\sigma}{\sqrt{n}} = 1.645\dfrac{\sqrt{0.69}}{\sqrt{9}} = 0.455$
L'intervalle de confiance obtenu a pour extrémités :

$$a = \overline{x} - \ell = 19.72 - 0.455 \ = 19.27$$
$$b = \overline{x} + \ell = 19.72 + 0.455 \ = 20.18$$

L'intervalle est plus resserré, donc plus précis car nous connaisons la valeur exacte d'un paramètre (σ) et ne nous contentons pas d'une estimation.

Exercice 2

1. Soit p la proportion de français favorables.
X est la variable aléatoire prenant la valeur 1 si la personne est favorable et 0 sinon. $X \sim \ \mathcal{B}(p)$, loi de Bernouilli. Donc $E(X) = p$ et $\sigma^2 = V(X) = p(1-p)$.
L'estimateur usuel de p est la fréquence empirique ,

$$F_n = \overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

La taille de l'échantillon est suffisamment grand ($n = 1502$) pour approcher la loi de l'estimateur par la loi normale

$$Z \ = \frac{F_n - p}{\sigma}\sqrt{n} \ \sim \mathcal{N}(0,1)$$

On cherche un intervalle centré en F_n, donc :

$$P\ (F_n - \ell \leq p \leq F_n + \ell) \ = \ P(-\ell \leq F_n - p \leq +\ell\ ) = 0.95$$

$$P\ \left(-\ell\frac{\sqrt{n}}{\sigma} \leq \frac{F_n - \mu}{\sigma}\sqrt{n}\ \leq +\ell\frac{\sqrt{n}}{\sigma}\ \right) = 0.95$$

La table de la loi normale standard donne $t = 1.96$ comme valeur asurant :
$P(-t \leq Z \leq t) = 0.95$, on en déduit : $t = \ell\frac{\sqrt{n}}{\sigma}$ et donc $\ell = t\frac{\sigma}{\sqrt{n}}$.
On remplace $\sigma = \sqrt{p(1-p)}$ par son estimation $\sqrt{f_n(1-f_n)}$
$\ell = t\frac{\sigma}{\sqrt{n}} = 1.96\frac{\sqrt{0.53(1-0.53)}}{\sqrt{1502}} = 0.025$
L'intervalle de confiance obtenu a pour extrémités :

$$a = f_n - \ell = 0.53 - 0.025 = 0.505$$
$$b = f_n + \ell = 0.53 + 0.025 = 0.555$$

Au final, il y a 95% de chance que la proportion de français favorables se situe entre $50,5\%$ et $55,5\%$
2. Toutes les valeurs étant supérieures à 50%, on peut dire que la majorité des français est favorable à cet accueil, avec un niveau de confiance de 95%.

Exercice 3

Le poids d'une pièce de 2 est représenté par la v.a. X.
On cherche un IDC à 99% pour $E(X) = \mu$ connaissant $Var(X) = \sigma^2 = 10^{-2}$.
1. La taille de l'échantillon est suffisamment grand ($n = 100$) pour approcher la loi de l'estimateur par la loi normale

$$Z \ = \frac{\overline{X_n} - \mu}{\sigma}\sqrt{n} \ \sim \mathcal{N}(0,1)$$

On cherche un intervalle centré en $\overline{X_n}$, donc :

$$P\ \left(\overline{X_n} - \ell \leq p \leq \overline{X_n} + \ell\right) \ = \ P\left(-\ell \leq \overline{X_n} - p \leq +\ell\ \right) = 0.99$$

$$P\ \left(-\ell\frac{\sqrt{n}}{\sigma} \leq \frac{\overline{X_n} - p}{\sigma}\sqrt{n}\ \leq +\ell\frac{\sqrt{n}}{\sigma}\ \right) = 0.99$$

La table de la loi normale standard donne $t = 2.576$ comme valeur asurant :
$P(-t \leq Z \leq t) = 0.99$, on en déduit : $t = \ell\frac{\sqrt{n}}{\sigma}$ et donc $\ell = t\frac{\sigma}{\sqrt{n}}$.
$\ell = t\frac{\sigma}{\sqrt{n}} = 2.576\frac{10^{-1}}{\sqrt{100}} = 0.026$
L'intervalle de confiance obtenu a pour extrémités :

$$a = \overline{x_n} - \ell = 8.55 - 0.026 = 8.52$$
$$b = \overline{x_n} + \ell = 8.55 + 0.026 = 8.58$$

Au final, il y a 95% de chance que $\mu \in [8.52\ ;\ 8.58]$.
2. La précision de l'intervalle précédent est de 0.03 (8.55 ± 0.03). Elle est donc donnée pat ℓ.
Pour avoir une précision de 10^{-2}, il faut avoir : $t\frac{\sigma}{\sqrt{n}} = 10^{-2} \iff \sqrt{n} = \frac{t\ \sigma}{10^{-2}}$
On trouve donc $n = \left(\frac{2.576\ 10^{-1}}{10^{-2}}\right)^2 = 663.6$
Il faut donc prendre un échantillon d'au moins 664 pièces.

Exercice 4

1. Les intervalles de confiance sont généralement centrés autour de l'estimateur usuel.
   Ici, ce sera autour de $\overline{X}$ dont la valeur dans cet échantillon est donc :
   $\overline{x} = \dfrac{11.13 + 12.33}{2} = 11.73$ qui est donc une estimation ponctuelle du nombre moyen de sièges vides par vol.
2. Lorsque l'écart-type est connu, donc la variance, l'IDC à 95% est donné par :
   $\overline{x} \pm t\dfrac{\sigma}{\sqrt{n}}$, avec $t = 1.96$.
   Par suite $1.96\dfrac{\sigma}{\sqrt{n}} = 11.73 - 11.13 = 0.6$ et $n = 225$.
   On en déduit : $\sigma = \dfrac{0.6\sqrt{225}}{1.96} = 4.59$

Exercice 5

$X \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Les deux paramètres sont inconnus et on a un petit échantillon de taille $n = 25$.

IDC pour μ :

On sait que l'échantillon est gaussien avec σ^2 inconnu donc :

$W = \dfrac{\overline{X} - \mu}{S^*}\sqrt{n} \sim T_{n-1}$ loi de Student à $n - 1$ degrés de liberté.

On cherche un intervalle symétrique centré en $\overline{X}$, d'où :

$$P\ (a \leq \mu \leq b) = P\ (\overline{X} - \ell \leq \mu \leq \overline{X} + \ell) = P(-\ell \leq \overline{X} - \mu \leq +\ell\ ) = 0.95$$

$$P\ \left(-\ell\frac{\sqrt{n}}{s^*} \leq \frac{\overline{X} - \mu}{S^*}\sqrt{n} \leq +\ell\frac{\sqrt{n}}{s^*}\right) = 0.95$$

La table de la loi de Student donne la valeur de t tel que $P(-t \leq T_{24} \leq t) = 0.95$

$t = 2.064$, or $t = \ell\dfrac{\sqrt{n}}{s^*}$ donc $\ell = t\dfrac{s^*}{\sqrt{n}}$

L'IDC recherché est :

$$a = \overline{x} - \ell = 6.7 - 2.064\frac{5.8}{\sqrt{25}} = 4.3$$

$$b = \overline{x} + \ell = 6.7 + 2.064\frac{5.8}{\sqrt{25}} = 9.09$$

Donc la durée moyenne μ de consultation d'internet chez les lecteurs a 90% de chances de se trouver entre 4.3 et 9.09 heures.

Exercice 4

Voir fichier **TD2-exo6.xls**