

Séance n° 9:

Intervalle de confiance et tests d'hypothèses

exercice 9.1.

X = mesure effectuée = variable aléatoire

K = grandeur étudiée = valeur fixée / constante

Les mesures sont effectuées avec une erreur Y de loi $\mathcal{N}(0, \sigma)$
où σ est inconnue, c'est-à-dire que nous avons

$$X = K + Y \quad \text{et} \quad X \sim \mathcal{N}(K, \sigma)$$

Par conséquent, pour chaque mesure effectuée, X_i , $i=1, \dots, 10$,
nous déduisons que

$$X_i = K + Y_i \quad \text{iid } \mathcal{N}(K, \sigma)$$

Comme la moyenne K est connue et la variance σ^2 inconnue,
nous savons que

$$\frac{\bar{X} - K}{\frac{S}{\sqrt{n-1}}} \sim t_{n-1}$$

Donc,

$$P \left[-t_{n-1; \frac{\alpha}{2}} \leq \frac{\bar{X} - K}{\frac{S}{\sqrt{n-1}}} \leq t_{n-1; \frac{\alpha}{2}} \right] = 1 - \alpha.$$

$$\Leftrightarrow P \left[-(t_{n-1; \frac{\alpha}{2}}) \frac{S}{\sqrt{n-1}} \leq \bar{X} - K \leq (t_{n-1; \frac{\alpha}{2}}) \frac{S}{\sqrt{n-1}} \right] = 1 - \alpha$$

$$\Leftrightarrow P \left[(t_{n-1; \frac{\alpha}{2}}) \frac{S}{\sqrt{n-1}} \geq K - \bar{X} \geq -(t_{n-1; \frac{\alpha}{2}}) \frac{S}{\sqrt{n-1}} \right] = 1 - \alpha$$

$$\Leftrightarrow P \left[\bar{X} - \frac{S}{\sqrt{n-1}} t_{n-1; \frac{\alpha}{2}} \leq K \leq \bar{X} + \frac{S}{\sqrt{n-1}} t_{n-1; \frac{\alpha}{2}} \right] = 1 - \alpha$$

c'est-à-dire pour K au niveau de confiance $1 - \alpha$

est donc $\left[\bar{x} \pm \frac{s}{\sqrt{n-1}} t_{n-1; \alpha/2} \right]$.

Ici, $\bar{x} = 4,0759$, $s = 0,0039$, $n = 10$ et
 $\alpha = 1 - 0,95 = 0,05$. Par conséquent,

$$t_{n-1; \alpha/2} = t_{9; 0,025},$$

et comme $F_{t_9}(t_{9; 0,025}) = P[t_9 \leq t_{9; 0,025}] = 0,975$,

$$t_{9; 0,025} = 2,262.$$

L' I.C. pour μ au niveau de confiance 95% est

$$\begin{aligned} & \left[4,0759 - \frac{0,0039}{3} 2,262; 4,0759 + \frac{0,0039}{3} 2,262 \right] \\ & = [4,073; 4,0788] \end{aligned}$$

Exercice 3.2.

Soit X = temps de calcul pour résoudre le problème du plus court chemin dans un graphe à 50 sommets.
 $\sim \mathcal{N}(\mu, 100)$

Nous devons chercher l'effectif minimum pour que l'intervalle de confiance (à 95%) du temps de calcul moyen ait une longueur inférieure à 10 seconds.

comme la variance est connue et la moyenne inconnue, nous savons que

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Donc,

$$P\left[-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}\right] = 1 - \alpha$$

$$\Leftrightarrow P\left[-\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq \bar{X} - \mu \leq \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right] = 1 - \alpha$$

$$\Leftrightarrow P\left[\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right] = 1 - \alpha.$$

L'intervalle de confiance pour μ au niveau de confiance $1 - \alpha$ est $\left[\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right]$ et pour $\sigma = 100$ et $\alpha = 0,05$,

$$I.C. = \left[\bar{X} - \frac{100}{\sqrt{n}} z_{0,05/2} \leq \mu \leq \bar{X} + \frac{100}{\sqrt{n}} z_{0,05/2}\right].$$

Comme $F_{\mathcal{N}(0,1)}(z_{0,025}) = P[\mathcal{N}(0,1) \leq z_{0,025}] = 0,975$, nous deduisons que

$$z_{0,025} = 1,96.$$

Par conséquent, l'IC devient

$$IC = \left[\bar{x} - \frac{196}{\sqrt{n}} ; \bar{x} + \frac{196}{\sqrt{n}} \right]$$

Vous cherchez le naturel n le plus petit vérifiant

$$|IC| = \left(\bar{x} + \frac{196}{\sqrt{n}} \right) - \left(\bar{x} - \frac{196}{\sqrt{n}} \right) \leq 20$$

$$\Leftrightarrow 2 \cdot \frac{196}{\sqrt{n}} \leq 20$$

$$\Leftrightarrow n \geq \left(\frac{2 \cdot 196}{20} \right)^2 = 384,16$$



$$n \geq 385$$

Exercice 9.3.

Nous voulons construire dans le premier cas comme dans le second, un intervalle de sécurité à 95% pour la variance d'une population normale sachant l'effectif n et s^2 .

Nous savons que dans ce cas,

$$\frac{ns^2}{\sigma^2} \sim \chi^2_{n-1},$$

et donc,

$$P \left[\chi^2_{n-1; \alpha/2} \leq \frac{ns^2}{\sigma^2} \leq \chi^2_{n-1; 1-\alpha/2} \right] = 1-\alpha$$

$$\Leftrightarrow P \left[\frac{ns^2}{\chi^2_{n-1; 1-\alpha/2}} \leq \sigma^2 \leq \frac{ns^2}{\chi^2_{n-1; \alpha/2}} \right] = 1-\alpha$$

L'intervalle de confiance pour σ^2 au niveau de confiance $1-\alpha$ est donc

$$\left[\frac{ns^2}{\chi^2_{n-1; 1-\alpha/2}} ; \frac{ns^2}{\chi^2_{n-1; \alpha/2}} \right],$$

et comme $\alpha = 1 - 0,95 = 0,05$,

$$I.C. = \left[\frac{ns^2}{\chi^2_{n-1; 0,975}} ; \frac{ns^2}{\chi^2_{n-1; 0,025}} \right]$$

ici : $n = 14$, $\sigma^2 = 7,3$ et

$$\chi^2_{n-1; 1-\alpha/2} = \chi^2_{13; 0,975} = 24,74 \quad \chi^2_{n-1; \alpha/2} = \chi^2_{13; 0,025} = 5,01$$

L'I.C. devient finalement

$$\left[\frac{14 \cdot (7,3)}{24,74} ; \frac{14 \cdot (7,3)}{5,01} \right] = [4,131 ; 20,392]$$

2ème cas : $m = 80$ et $s^2 = 14$.

Comment évaluer

$$\chi^2_{m-1; 1-\alpha/2} = \chi^2_{79; 0,975} \quad \text{et} \quad \chi^2_{m-1; \alpha/2} = \chi^2_{79; 0,025} ?$$

Nous utilisons le résultat asymptotique suivant :

$$\sqrt{2 \chi^2_{m-1}} - \sqrt{2m-3} \approx \mathcal{N}(0,1),$$

qui est permis car $m=80 > 30$.

Donc,

$$\sqrt{\frac{2m s^2}{\sigma^2}} - \sqrt{2m-3} \sim \mathcal{N}(0,1),$$

et

$$P\left[-z_{\alpha/2} \leq \sqrt{\frac{2m s^2}{\sigma^2}} - \sqrt{2m-3} \leq z_{\alpha/2}\right] \approx 1-\alpha$$

$$\Rightarrow P\left[(\sqrt{2m-3} - z_{\alpha/2})^2 \leq \frac{2m s^2}{\sigma^2} \leq (\sqrt{2m-3} + z_{\alpha/2})^2\right] \approx 1-\alpha$$

$$\Rightarrow P\left[\frac{2m s^2}{(\sqrt{2m-3} + z_{\alpha/2})^2} \leq \sigma^2 \leq \frac{2m s^2}{(\sqrt{2m-3} - z_{\alpha/2})^2}\right] \approx 1-\alpha.$$

Ici, $m = 80$, $s^2 = 14$, ce qui implique que

$$\sqrt{2m-3} = \sqrt{157} = 12,53$$

$$2m s^2 = 160 \cdot 14 = 2240$$

$$z_{\alpha/2} = z_{0,025} = 1,96.$$

Une approximation de l' I.C. pour σ^2 au niveau de confiance 95% est

$$\left[\frac{2240}{(12,53 + 1,96)^2}, \frac{2240}{(12,53 - 1,96)^2} \right] = [10,6687; 20,0492]$$

Exercice 9.4.

L'expérience a pour objectif de déterminer la concentration moyenne μ d'un polluant. Pour $i = 1, \dots, 5$, nous avons

$X_i =$ mesures effectuées
 $\sim \mathcal{N}(\mu, \sigma)$ et indépendantes.

Nous avons obtenu les résultats suivants :

26,1 23,7 20,4 21,3 17,3

calcul de $\bar{x}$:
$$\bar{x} = \frac{1}{5} \sum_{i=1}^5 x_i = \frac{1}{5} (26,1 + 23,7 + 20,4 + 21,3 + 17,3)$$
$$= 21,76$$

calcul de s^2 :
$$s^2 = \frac{1}{5} \sum_{i=1}^5 (x_i - \bar{x})^2 = 8,9104$$

IC pour μ à 95%

Comme σ est inconnu, nous savons que :

$$\frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \sim t_{n-1}$$

Comme $\bar{x} = 21,76$, $s^2 = 8,9104$ et $n = 5$

$$\frac{21,76 - \mu}{1,4925} \sim t_{n-1} = t_4$$

Par conséquent,

$$P \left[-t_{4, \frac{0,05}{2}} \leq \frac{21,76 - \mu}{1,4925} \leq t_{4, \frac{0,05}{2}} \right] = 1 - 0,05 = 0,95$$

$$\Leftrightarrow P[(t_4; 0,025)(1,4925) \geq \mu - 21,76 \geq -(t_4; 0,025)(1,4925)] = 0,95$$

$$\Leftrightarrow P[21,76 + 1,4925 t_{4;0,025} \geq \mu \geq 21,76 - 1,4925 t_{4;0,025}] = 0,95$$

Donc, l'IC pour μ au niveau de confiance 0,95 est

$$[21,76 - 1,4925 t_{4;0,025}; 21,76 + 1,4925 t_{4;0,025}]$$

Comme $F_{t_4}(t_{4;0,025}) = P(t_4 \leq t_{4;0,025}) = 0,975$,

nous déduisons que $t_{4;0,025} = 2,776$.

Finalement,

$$IC = [17,61682, 25,90318]$$

IC pour σ à 95%

Comme μ est inconnu, nous savons que

$$\frac{ms^2}{\sigma^2} \sim \chi^2_{m-1}$$

Ici, $m=5$, $s^2 = 8,9104$, ce qui implique que

$$\frac{44,552}{\sigma^2} \sim \chi^2_4$$

$$P[\chi^2_{4;0,025} \leq \frac{44,552}{\sigma^2} \leq \chi^2_{4;0,975}] = 0,95.$$

$$\Leftrightarrow P[44,552 (\chi^2_{4;0,025})^{-1} \geq \sigma^2 \geq 44,552 (\chi^2_{4;0,975})^{-1}] = 0,95.$$

Donc, l'IC pour σ^2 au niveau de confiance 0,95 est

$$[44,552 (\chi^2_{4;0,975})^{-1}, 44,552 (\chi^2_{4;0,025})^{-1}]$$

Comme $\chi^2_{4;0,975} = 11,1$ et $\chi^2_{4;0,025} = 0,484$,

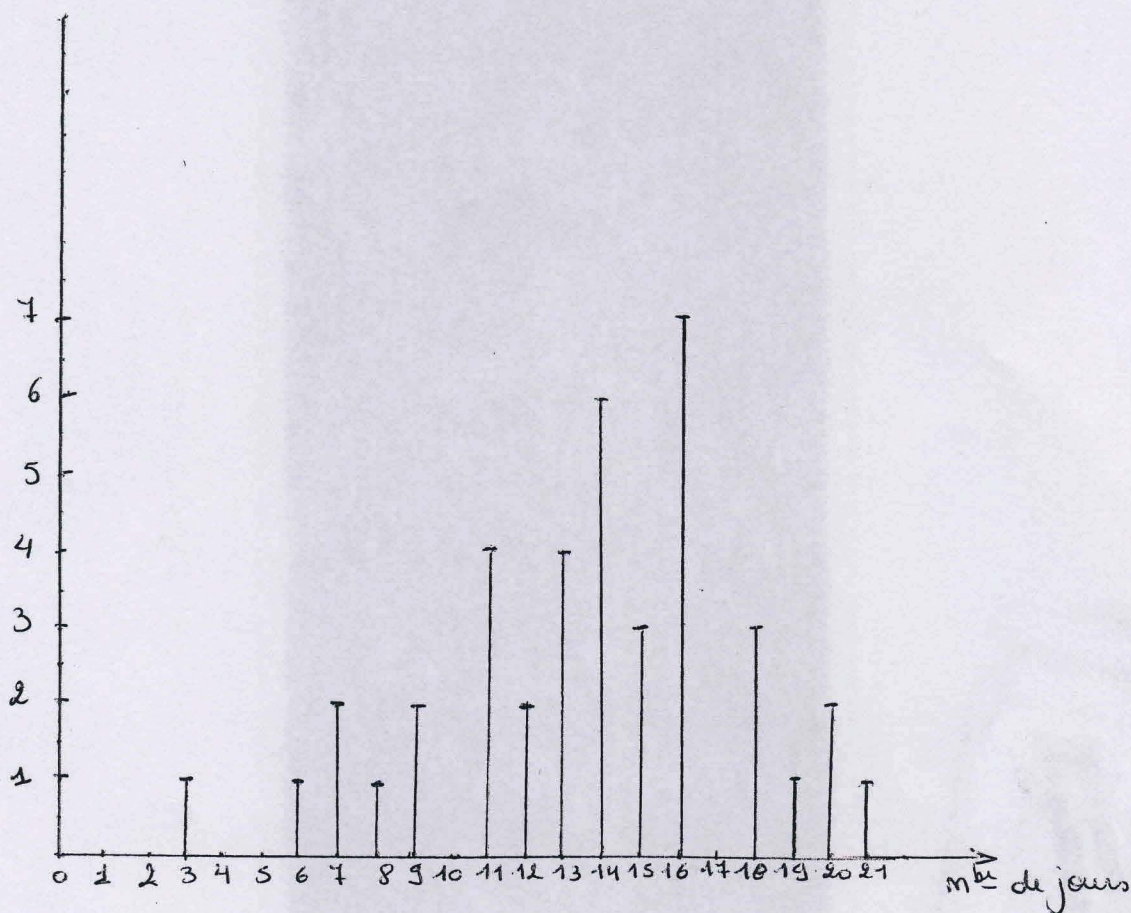
$$IC = [4,0137; 92,05]$$

Exercice 9.5.

X = temps (en jours) qui s'écoule entre l'exposition et l'apparition des rougeurs de la peau

$$\underline{x} = (6, 12, 14, 16, 13, 9, 15, 7, 20, \dots, 18, 16, 11, 13, 14, 16, 15, 15)$$

1.



Le graphique en bâtonnets permet de faire l'hypothèse que la variable X suit une loi normale.

2. Si $\sigma=4$, il faut construire un I.C. à 95% pour le temps moyen avant l'apparition des rougeurs.

Comme σ est connue, nous savons que

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

Etant donné que $n=40$, $\bar{x}=11,675$ et $\sigma=4$,

$$\frac{11,675 - \mu}{0,63245} \sim \mathcal{N}(0,1)$$

$$P \left[-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \right] = 1 - \alpha$$

$$\Rightarrow P \left[-z_{0,025} \leq \frac{11,675 - \mu}{0,63245} \leq z_{0,025} \right] = 0,95$$

$$\Rightarrow P \left[z_{0,025} \cdot 0,63245 \geq \mu - 11,675 \geq -z_{0,025} \cdot 0,63245 \right] = 0,95$$

$$\Rightarrow P \left[z_{0,025} \cdot 0,63245 + 11,675 \geq \mu \geq -z_{0,025} \cdot 0,63245 + 11,675 \right] = 0,95$$

Comme $z_{0,025}$ vérifie $P[\mathcal{N}(0,1) \leq z_{0,025}] = 0,975$,

$$z_{0,025} = 1,96,$$

et l'intervalle de confiance à 95% pour μ est donné par

$$\begin{aligned} \text{I.C.} &= \left[- (1,96)(0,63245) + 11,675; (1,96)(0,63245) + 11,675 \right] \\ &= [10,435; 12,91] \end{aligned}$$

3. Oui, nous serions surpris car au point 2. nous avons déterminé que la probabilité que la moyenne appartienne à l'intervalle $[10,435; 12,91]$ est de 0,95. Il y a donc une très faible probabilité que la moyenne soit en dehors de cet intervalle, et donc égale à 17.

Pour plus de rigueur, il aurait fallu tester l'hypothèse suivante:

$$\begin{cases} H_0: \mu = 17 \\ H_1: \mu \neq 17 \end{cases}$$

au niveau 95%. Nous aurions rejeté H_0 et donc obtenu la même réponse.