

CC3 2024/2025 – Durée 1h30

Les documents et appareils électroniques (calculatrice, téléphone, ordinateur, ...) sont interdits. Il est sous-entendu que les estimateurs, intervalles de confiance et tests demandés doivent être non triviaux. Toutes les réponses doivent être justifiées.

Exercice 1 - Calibration de radar routier

Un radar routier est composé de deux éléments principaux : un appareil de mesure de la vitesse, et un flash qui se déclenche lorsque la vitesse mesurée dépasse un certain seuil. Pour un véhicule ayant une vitesse réelle μ , la vitesse mesurée est modélisée par une variable aléatoire de loi $\mathcal{N}(\mu, \sigma^2)$ (loi normale), où σ^2 représente la précision du radar.

Estimation de la précision du radar

Avant mise en fonction sur la voie publique, les experts de la DSR (délégation à la sécurité routière) veulent connaître la précision du radar fourni par le constructeur. Pour ce faire, ils observent la mesure donnée par le radar pour un véhicule dont ils connaissent la vitesse réelle ($\mu_0 = 100km/h$), et répètent l'expérience n fois de manière indépendante.

1. Donner le modèle de cette expérience.
2. Trouver un estimateur de σ^2 par la méthode des moments. On le notera $\hat{\sigma}_n^2$.
3. Montrer que le risque quadratique de $\hat{\sigma}_n^2$ vaut $2\sigma^4/n$. On pourra utiliser que si $N \sim \mathcal{N}(0, 1)$, $\mathbb{E}(N^4) = 3$.
4. En déduire un intervalle de niveau de confiance **non asymptotique** 95% sur σ^2 .
5. Montrer que $\hat{\sigma}_n^2$ est consistant.
6. Trouver le comportement asymptotique de $\hat{\sigma}_n^2$.
7. En déduire un intervalle de niveau de confiance **asymptotique** 95% sur σ^2 .
8. Comparer les deux intervalles de confiance obtenus sur la base de leurs longueurs (en supposant n assez grand). Lequel des deux vous semble le plus précis ?

Calibration du seuil de flash

Dans cette partie, on suppose la précision du radar connue (estimée précédemment), c'est à dire $\sigma_0^2 = 4$. Le radar mesure la vitesse d'un véhicule. Cette mesure est modélisée par une variable aléatoire de loi $\mathcal{N}(\mu, \sigma_0^2)$, où μ est la vitesse réelle du véhicule. Le radar veut détecter (flasher) si cette vitesse réelle dépasse les $50km/h$.

9. Donner le modèle de cette expérience.
10. Donner les hypothèses de test associées au déclenchement du flash du radar.
11. Construire un test de niveau 5% pour ces hypothèses.
12. **Application numérique** : À partir de quelle vitesse mesurée le flash se déclenche-t'il ?

Point de vue de l'association des motards de France : estimation du seuil du radar

L'association des motards de France ne connaît pas les détails techniques du radar, mais a accès aux vitesses de ses n adhérents s'étant fait flasher. On suppose que ces vitesses sont indépendantes, de même loi donnée par $\theta + \mathcal{E}(1/5)$ (loi d'une variable s'écrivant comme $\theta + E$, $E \sim \mathcal{E}(1/5)$), où θ est le seuil de déclenchement du radar (inconnu).

13. Donner un modèle pour cette expérience.
14. Trouver un estimateur de θ **par la méthode du maximum de vraisemblance**. On le notera $\hat{\theta}_n$.
15. Donner un intervalle de niveau de confiance 95% sur θ .
16. **Application numérique** : Sur 15 utilisateurs flashés, l'association observe une vitesse minimale de 54 kmh. On donne aussi la valeur $\log(20) = 3$. Jusqu'à quelle vitesse les motards sont-ils sûrs à 95% de pouvoir rouler sans se faire flasher ?

Solution 1 -

1. $(\mathbb{R}^n, (\mathcal{N}(\mu_0, \sigma^2)^{\otimes n})_{\sigma^2 > 0})$, avec $\mu_0 = 100 \text{ km/h}$.
2. Comme $\text{Var}_{\sigma^2}(X) = \sigma^2$ (qui est par ailleurs finie), un estimateur par la méthode des moments vérifie l'équation

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2.$$

3. On a

$$\begin{aligned} E_{\sigma^2}(\hat{\sigma}_n^2 - \sigma^2) &= \text{Var}_{\sigma^2} \left(\frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2 \right) \\ &= \frac{\text{Var}_{\sigma^2}(X_1 - \mu_0)^2}{n} \quad (\text{les } (X_i - \mu_0)^2 \text{ sont i.i.d.}) \\ &= \frac{\text{Var}(\sigma^2 N^2)}{n} = \frac{\sigma^4 \text{Var}(N^2)}{n}, \end{aligned}$$

pour $N \sim \mathcal{N}(0, 1)$. Or

$$\text{Var}(N^2) = \mathbb{E}(N^4) - (\mathbb{E}(N^2))^2 = 2.$$

On en déduit

$$E_{\sigma^2}((\hat{\sigma}_n^2 - \sigma^2)^2) = \frac{2\sigma^4}{n}.$$

4. Soit $t > 0$. L'inégalité de Bienaymé-Chebichev s'écrit

$$P_{\sigma^2}(|\hat{\sigma}_n^2 - \sigma^2| > t) \leq \frac{2\sigma^4}{nt^2}.$$

On en déduit alors que

$$P_{\sigma^2}\left(|\hat{\sigma}_n^2 - \sigma^2| > \sigma^2 \sqrt{\frac{40}{n}}\right) \leq 5\%.$$

Or,

$$\sigma^2(1 - \sqrt{40/n}) \leq \hat{\sigma}_n^2 \leq \sigma^2(1 + \sqrt{40/n}) \Leftrightarrow \frac{\hat{\sigma}_n^2}{1 + \sqrt{40/n}} \leq \sigma^2 \leq \frac{\hat{\sigma}_n^2}{(1 - \sqrt{40/n}) \vee 0},$$

en prenant pour convention $\hat{\sigma}_n^2/0 = +\infty$. On en déduit que $\left[\frac{\hat{\sigma}_n^2}{1 + \sqrt{40/n}}; \frac{\hat{\sigma}_n^2}{(1 - \sqrt{40/n}) \vee 0}\right]$ est un intervalle de niveau de confiance 95%.

5. D'après la question 3-, $E_{\sigma^2}((\hat{\sigma}_n^2 - \sigma^2)^2) \xrightarrow{n \rightarrow +\infty} 0$. On en déduit que $\hat{\sigma}_n^2 \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} \sigma^2$, donc que $\hat{\sigma}_n^2$ est consistant.

6. Comme $\text{Var}_{\sigma^2}((X_1 - \mu_0)^2) = 2\sigma^4 < +\infty$ (cf question 3-), le Théorème Central Limite donne

$$\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) \rightsquigarrow \mathcal{N}(0, 2\sigma^4).$$

7. On déduit de la question précédente que

$$\frac{\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2)}{\sqrt{2}\sigma^2} \rightsquigarrow \mathcal{N}(0, 1).$$

Soit q le quantile d'ordre 97,5% d'une loi normale standard. On a alors

$$P_{\sigma^2}\left(\left|\frac{\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2)}{\sqrt{2}\sigma^2}\right| \geq q\right) \xrightarrow{n \rightarrow +\infty} 5\%.$$

Or,

$$\left|\frac{\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2)}{\sqrt{2}\sigma^2}\right| \geq q \Leftrightarrow \frac{\hat{\sigma}_n^2}{1 + \sqrt{2q/n}} \leq \sigma^2 \leq \frac{\hat{\sigma}_n^2}{(1 - \sqrt{2q/n}) \vee 0},$$

avec la même convention qu'auparavant. On en déduit que $\left[\frac{\hat{\sigma}_n^2}{1 + \sqrt{2q/n}}; \frac{\hat{\sigma}_n^2}{(1 - \sqrt{2q/n}) \vee 0}\right]$ est un intervalle de niveau de confiance asymptotique 95%.

8. Comme $2q \leq 4 < 40$, la longueur du 2-ème intervalle est strictement inférieure à celle du premier. On préférera donc utiliser le deuxième (pour de grandes valeurs de n justifiant la validité de l'approche asymptotique).
9. $\mathbb{R}, (\mathcal{N}(\mu, \sigma_0^2))_{\mu \in \mathbb{R}}$.
10. Le radar veut prouver $\mu > 50$. Il choisit donc $H_1 : \mu > 50$, et par conséquent $H_0 : \mu = 50$.
11. En se basant sur la statistique de test X (mesure de la vitesse), on s'attend à de plus fortes valeurs de X sous H_1 que H_0 . La zone de rejet est donc du type $[t_\alpha; +\infty[$, et le test de type $\mathbb{1}_{X \geq t_\alpha}$ (le radar se déclenche si la vitesse mesurée dépasse un seuil).

Le seuil t_α doit vérifier

$$P_{50}(X \geq t_\alpha) \leq 5\% \quad \Leftrightarrow \quad \mathbb{P}(50 + 2\mathcal{N}(0, 1) \geq t_\alpha) \leq 5\%.$$

Si q désigne le quantile d'ordre 95% d'une loi normale standard, on en déduit $(t_\alpha - 50)/2 = q$, et donc que

$$T = \mathbb{1}_{X \geq 50 + 2q}$$

est un test de niveau 5% pour ces hypothèses.

12. Avec $q = 1.65$, le radar se déclenche pour des vitesses mesurées supérieures à 53.3 kmh.
13. $(]0; +\infty[^n, (\theta + \mathcal{E}(1/5))_{\theta > 0}^{\otimes n})$. La densité est donnée, pour $x_1, \dots, x_n > 0$, par

$$f_\theta(x_{1:n}) = \prod_{i=1}^n (1/5) e^{-(x_i - \theta)/5} \mathbb{1}_{x_i \geq \theta} = 5^{-n} e^{-n\bar{x}/5 + n\theta/5} \mathbb{1}_{m_n \geq \theta},$$

où $m_n = \min_{i=1, \dots, n} x_i$.

14. Soit $x_1, \dots, x_n > 0$. La vraisemblance s'écrit, pour $\theta > 0$,

$$V_{x_{1:n}}(\theta) = 5^{-n} e^{-n\bar{x}/5 + n\theta/5} \mathbb{1}_{m_n \geq \theta}.$$

$V_{x_{1:n}}$ est donc croissante sur $[0; m_n]$, à valeurs strictement positives, et nulle pour $\theta > m_n$. On en déduit que $V_{x_{1:n}}$ est maximale en m_n , et donc

$$\hat{\theta}_n = \min_{i=1, \dots, n} X_i.$$

15. Comme $\min_{i=1, \dots, n} (X_i - \theta) \sim \mathcal{E}(n/5)$, on en déduit que $\hat{\theta}_n \sim \theta + \mathcal{E}(n/5) \sim \theta + \frac{5}{n} \mathcal{E}(1)$.
Or

$$\mathbb{P}(\mathcal{E}(1) \geq -\log(5\%)) = 5\%.$$

On en déduit que

$$P_\theta \left(\theta \leq \hat{\theta}_n \leq \theta + \frac{5 \log(20)}{n} \right) = 95\%,$$

et donc que $\left[\hat{\theta}_n; \hat{\theta}_n + \frac{5 \log(20)}{n} \right]$ est un intervalle de niveau de confiance 95%.

16. L'intervalle de confiance devient $[53, 54]$. Les motards peuvent donc rouler jusqu'à 53 kmh.

TABLE DE QUANTILES : LOI NORMALE

t	2.33	1.96	1.65	1.28
$\mathbb{P}(\mathcal{N}(0, 1) \geq t)$	0.01	0.025	0.05	0.1