

Groupe 2

Exercice 1 (Un café pour commencer (*)). Une cafétéria souhaite remplacer son service par un ou plusieurs distributeurs automatiques de boisson. Afin de calibrer au plus juste le nombre de distributeurs nécessaires, elle mène donc une enquête pour connaître au mieux les besoins de ses utilisateurs.

1. Dans un premier temps, elle observe les temps *successifs* d'arrivée de clients dans son local, enregistrant ainsi un vecteur $T = (T_1, \dots, T_n)$. En considérant les intervalles de temps ainsi formés, et en admettant que les clients arrivent les uns après les autres selon un flux identique toute la journée, proposer un modèle statistique. On notera pour cela

$$\forall i \geq 1: \quad \Delta_i = T_i - T_{i-1},$$

avec la convention $T_0 = 0$. Discuter des éventuelles limitations de ce modèle.

$$0 = T_0 \leq T_1 \leq T_2 \leq T_3 \leq \dots$$

$$\Delta_i = T_i - T_{i-1}$$

$$\Delta_i \text{ est d.o.a. } \geq 0$$

Qu'est ce qu'un modèle stat ?

- espace mesurable (tribu etc...)

- famille de probas. $(P_\theta)_{\theta \in \Theta}$

$\hookrightarrow$ Ici, quel est l'espace mesurable, la tribu, la famille de probas ?

$\rightarrow \mathbb{R}$, voire $\mathbb{R}_+$ (les Δ_i sont ≥ 0)

$\rightarrow \mathcal{B}(\mathbb{R}_+)$

$\rightarrow \{E(\lambda), \lambda > 0\}$ famille de lois candidates

$\hookrightarrow$ Pour 1 obs.

$\rightarrow \{E(\lambda)^{\otimes n}, \lambda > 0\}$ n lois exp. iid.

Discuter des limitations : hyp : homogénéité du flux d'arrivée

$\hookrightarrow$ très discutable. Il faudrait $\lambda(t) \leadsto$ ça devient + compliqué.

2. Proposer un estimateur du temps moyen d'arrivée d'un nouveau client, noté $\hat{\Delta}_n$ ainsi qu'une approximation du nombre de clients arrivant sur un intervalle de temps donné de longueur $h > 0$ fixé.

Si on sait que $\mathbb{E}(A) = \frac{1}{\lambda}$, alors on va "estimer" cette espérance via la LGV :

$$\frac{\Delta_1 + \dots + \Delta_n}{n} := \hat{m}_n.$$

Nombre moyen de clients sur un temps de longueur h ?

$$\frac{h}{\hat{m}_n}$$

Objectif : Modèle translation/échelle. Cas Gaussien (questions 3/4) (**)

Exercice 3 (Modèle de translation et d'échelle). Soit g une densité par rapport à la mesure Leb. On considère le modèle statistique

$$(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \{p_{n,\theta} \cdot \text{Leb}^{\otimes n} : \theta \in \Theta := \mathbb{R} \times \mathbb{R}_*^+\})$$

où

$$p_{n,\theta}(x_1, \dots, x_n) := \sigma^{-n} \prod_{i=1}^n g\left(\frac{x_i - \mu}{\sigma}\right), \quad \theta := (\mu, \sigma).$$

On note $(X_1, \dots, X_n)$ les variables canoniques : pour tout $i \in \{1, \dots, n\}$ et $(x_1, \dots, x_n) \in \mathbb{R}^n$, on a $X_i(x_1, \dots, x_n) = x_i$.

1. Montrer que sous $p_{n,\theta} \cdot \text{Leb}^{\otimes n}$, les statistiques $(X_1, \dots, X_n)$ sont i.i.d. et identifier leur loi.

La structure du modèle statistique induit qu'on a n obs $(X_1, \dots, X_n)$ qui sont i.i.d : forme produit de la même densité.

Un modèle de translation / échelle est un modèle où on observe

$$Z = a + b \varepsilon, \quad \begin{array}{l} a: \text{paramètre de translation} \\ b: \text{_____ d'échelle.} \end{array}$$

ε étant la v.e. (en général centrée $\mathbb{E} \varepsilon = 0$)

$$\mathbb{E}[Z] = a \quad \text{_____ réduite Var } \varepsilon = 1$$

$$\text{Var}(Z) = b^2$$

Cas particulier std : $\varepsilon \sim \mathcal{D}^p(0, 1) \Leftrightarrow \mu + \sigma \varepsilon \sim \mathcal{D}^p(\mu, \sigma^2)$

Rem : Pour identifier des lois, il est vivement conseillé d'utiliser la méthode de la "fonction test", "nette". Soit φ continue bornée :

$$\mathbb{E}[\varphi(Z)] = \dots \int \varphi(y) \underbrace{p(y)}_{\text{densité de la v.a. } Z} dy.$$

Application pour nous : soit $Z = \mu + \sigma \varepsilon$, où $\varepsilon \sim g$.

Densité de Z ?

Soit φ continue bornée : je calcule $\mathbb{E}[\varphi(Z)]$:

$$\begin{aligned} \mathbb{E}[\varphi(Z)] &= \mathbb{E}[\varphi(\mu + \sigma \varepsilon)] \\ &= \int \varphi(\mu + \sigma t) g(t) dt && \text{On pose } s = \mu + \sigma t \\ &= \int \varphi(s) g\left(\frac{s - \mu}{\sigma}\right) \frac{ds}{\sigma} \end{aligned}$$

On obtient bien la forme de la densité pour $n = 1$.

2] C'est tautologique dès lors qu'on sait que $X = \mu + \sigma \varepsilon$ avec $\varepsilon \sim g$

3] On suppose g densité Gaussienne. $t \mapsto \frac{e^{-t^2/2}}{\sqrt{2\pi}}$

Estimateur de μ :

$$\hat{\mu}_n := \frac{S_n}{n}$$

$$\hat{\sigma}_n^2 := \frac{K_n}{n} \quad (\text{ou } \frac{K_n}{n-1})$$

Rem : ces objets $\hat{\mu}_n$ et $\hat{\sigma}_n^2$ peuvent être étudiés à "horizon fini" c-à-d avec n fini, ou bien asymptotiquement ($n \rightarrow +\infty$)

Pour n fini, les indicateurs de qualité d'estimation sont en général le biais et la variance.

Lorsque $n \rightarrow +\infty$, on s'intéresse aux "convergences".

- Conv. en proba : $Z_n \xrightarrow{\mathbb{P}} Z_\infty$
- Conv. en loi $Z_n \xrightarrow{\mathcal{L}} Z_\infty$
- ——— p.s. $Z_n \xrightarrow{\text{p.s.}} Z_\infty$

4. Déterminer le modèle statistique induit par les statistiques (S_n, K_n) . La démonstration du théorème de Gosset est donnée dans le polycopié de cours (Théorème IV-5-24).

$$S_n = \sum_{i=1}^n X_i \quad K_n = \sum_{i=1}^n \left(X_i - \frac{S_n}{n} \right)^2$$

S_n : somme de n lois gaussiennes $N(\mu, \sigma^2)$, donc

$$\mathbb{E} S_n = n\mu \quad \text{Var}(S_n) = n\sigma^2$$

De +, $S_n \sim N(n\mu, n\sigma^2)$

$$X_1 = \mu + \sigma \varepsilon_1 \quad X_2 = \mu + \sigma \varepsilon_2$$

$$X_1 + X_2 = 2\mu + \sigma (\varepsilon_1 + \varepsilon_2)$$

? (G)

Outil : fonction caractéristique

Transf de Laplace

$$\mathbb{E}[e^{-t\varepsilon}] = e^{-t^2/2}$$

Si $X \sim N(\mu, \sigma^2)$: $\mathbb{E}[e^{-tX}] = e^{-t\mu + \sigma^2 t^2/2}$ si $\varepsilon \sim N(0,1)$

$X_1 \sim N(\mu_1, \sigma_1^2) \quad X_2 \sim N(\mu_2, \sigma_2^2)$: ind:

$$\begin{aligned} \mathbb{E}[e^{-t(X_1+X_2)}] &= \mathbb{E}[e^{-tX_1}] \mathbb{E}[e^{-tX_2}] \\ &= e^{-t\mu_1 + \sigma_1^2 t^2/2} e^{-t\mu_2 + \sigma_2^2 t^2/2} \\ &= e^{-t(\mu_1+\mu_2)} e^{t^2(\sigma_1^2+\sigma_2^2)/2} \end{aligned}$$

$$K_m = \sum_{i=1}^m \left(X_i - \frac{S_m}{m} \right)^2$$

Ca sentirait un petit peu le χ^2
(n-1 ? DDL)

Def : Une v.a. $\chi^2(d)$ c'est la loi d'une somme de $\xi_1^2 + \dots + \xi_d^2$ où ξ_i sont i.i.d $N(0,1)$

Nous allons devoir utiliser les vecteurs Gaussiens.

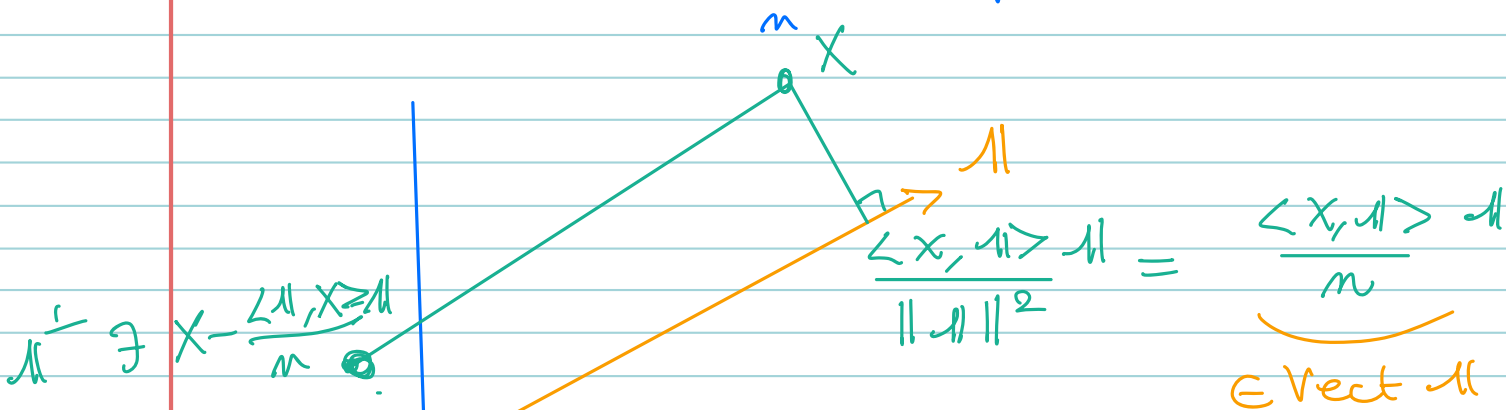
On dit que Z est $\perp \vec{G}$ si $\forall u : \langle u, Z \rangle$ est l.v.a.G.

Résultat : Si Z_1 et Z_2 sont $\perp \vec{G}$, $Z_1 \perp Z_2$ ou $\text{Cov}(Z_1, Z_2)$ est diagonale.

$$S_m = \sum_{i=1}^m X_i = \langle X, \mathbf{1} \rangle \text{ où } X = (X_1, \dots, X_m)$$

$$K_m = \left\| X - \frac{S_m}{m} \mathbf{1} \right\|^2$$

$$= \left\| X - \frac{\langle X, \mathbf{1} \rangle}{m} \mathbf{1} \right\|^2$$



Résultat  $\frac{S_m}{m} \mathbf{1}$ est le projeté orthogonal de X sur $\mathbf{1}$.

Résultat ! $X - \frac{\sum_{i=1}^n x_i}{n} \mathbf{1}$ est
le proj $\perp$ sur $\mathcal{H}^\perp$

Résultat ^{**} très ^{**} important sur les $\vec{N}G$.

Si X est $N(0, \sigma^2 \text{Id})$ et si

$$X = P(X) + [X - P(X)] \text{ où}$$

P est $\perp$ proj orthogonale sur $\perp$ ex E

Alors $P(X)$ et $X - P(X)$ sont indépendants

Démo : si $P(X)$ est le proj de X sur E , alors
 $\exists P$ tq $P(X) = P \cdot X$ avec $P^t P = P$
et $P^2 = P$.

$$X - P(X) = (\text{Id} - P) X$$

Je calcule le cor. de $P(X)$ avec $X - P(X)$:

$$\text{Cor}(P(X), X - P(X)) = \text{Cor}(P X, (\text{Id} - P) X)$$

$$\begin{aligned} \text{Cor}(A, B) &= E[A B^T] \\ &= P \underbrace{E[X X^T]}_{\sigma^2 I} (\text{Id} - P)^T \\ &= \sigma^2 P (I - P)^T \\ &= \sigma^2 P (I - P) \\ &= 0 \end{aligned}$$

Conclusion : Dans notre exercice, on a

$\frac{\langle X, 1 \rangle}{n}$ et $X - \frac{\langle X, 1 \rangle}{n} 1$ sont indépendants !

$$E = \text{Vect}(1) \quad \dim E = 1$$

$$\dim E^\perp = n - 1$$

$X - \frac{\langle X, 1 \rangle}{n} 1$ est la proj de X sur $E^\perp$,

$$\dim E^\perp = n - 1 \quad . \quad \text{Donc} \quad \left\| X - \frac{\langle X, 1 \rangle}{n} 1 \right\|^2 \text{ suit } \chi^2(n-1)$$

$$S_n \sim N(m\mu, m\sigma^2)$$

$$K_n \sim \chi^2(n-1) \times \sigma^2$$

$$S_n \perp K_n.$$

Exercice 4 : $Y_i = \beta (\langle X_i, \beta \rangle) + \sigma \varepsilon_i$

où ε_i sont iid de loi g .

Le modèle de régression linéaire est :

$$Y_i = \langle X_i, \beta \rangle + \sigma \varepsilon_i \quad \text{où } \varepsilon_i \sim N(0, 1)$$

iid

$$X_i = \begin{pmatrix} 1 \\ x_i \end{pmatrix}$$

$$Y_i = \beta_1 + \beta_2 x_i + \sigma \varepsilon_i$$