

Fiche Analyse Statistique

cours 1

• Rappel:

$$\begin{array}{l|l} \rightarrow a^m \times a^n = a^{m+n} & \rightarrow \frac{a^m}{a^n} = a^{m-n} \\ \rightarrow (a^m)^n = a^{m \times n} & \rightarrow \left(\frac{a}{b}\right)^n = \frac{a^n}{b^n} \\ \rightarrow (ab)^n = a^n \times b^n & \end{array}$$

• Loi Gamma

• Soit X v.a. positive tq $X \sim \gamma_r$

• Densité:

$$f(x) = \frac{1}{\Gamma(r)} \exp(-x) x^{r-1}$$

avec, $\Gamma(x) = \underbrace{\int_0^{+\infty} \exp(-t) t^{x-1} dt}_{\text{fonction Gamma}} \quad \forall x > 0$

• Propriétés fonction Gamma:

$$\begin{array}{l|l} \textcircled{1} \Gamma(x+1) = x\Gamma(x) \quad (\text{IPP}) & \textcircled{4} \Gamma\left(k + \frac{1}{2}\right) = \frac{1 \times 3 \times 5 \times \dots \times (2k-1)}{2^k} \Gamma\left(\frac{1}{2}\right) \\ \textcircled{2} \Gamma(1) = 1 & \textcircled{5} \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \\ \textcircled{3} \Gamma(n+1) = n! & \end{array}$$

• $E(X) = V(X) = r$

• Loi Normale

• Soit X v.a. tq $X \sim \mathcal{N}(\mu, \sigma)$

• Densité:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2\right)$$

• $E(X) = \mu$; $V(X) = \sigma^2$

■ Loi Normale centrée réduite

• Soit $U = \frac{X - \mu}{\underbrace{\sigma}_{\text{var. définies page précédente}}}$ v.a. , $U \sim \mathcal{N}(0, 1)$

• Densité :

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right)$$

• $E(U) = 0$; $V(X) = 1$

• Moments de la loi normale centrée réduite :

$$\mu_k = E(U^k)$$

$\rightarrow$ si $k = 2p + 1 \Rightarrow \mu_{2p+1} = 0$ (car fonction impaire)

$\rightarrow$ si $k = 2p \Rightarrow \mu_{2p} = \frac{(2p)!}{2^p p!}$

■ Rappel :

$$\rightarrow e^{a+b} = e^a e^b$$

$$\rightarrow e^{-a} = \frac{1}{e^a}$$

$$\rightarrow e^{a-b} = \frac{e^a}{e^b}$$

$$\rightarrow (e^a)^m = e^{ma}$$

$$\rightarrow \ln(ab) = \ln(a) + \ln(b)$$

$$\rightarrow \ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b)$$

$$\rightarrow \ln\left(\frac{1}{b}\right) = -\ln(b)$$

$$\rightarrow \ln(a^m) = m \ln(a)$$

⚠ SAVOIR EFFECTUER CALCUL INTEGRALE AVEC CHANGEMENT DE VAR.
(cf. sur internet)

■ Fonctions caractéristiques

- Déf: Si X est une var. aléatoire, on appelle fonction caractéristique de X la fonction définie pour tout réel t par:

$$\varphi_X(t) = E(\exp(itX)) \quad (i \text{ complexe})$$

- A quoi ça sert ?

↳ déterminer si 2 v.a. X et Y sont ^{de} même loi $\left(\begin{array}{l} \text{si } \varphi_X(t) = \varphi_Y(t) \\ \Rightarrow X \text{ et } Y \text{ suivent la même loi} \end{array} \right)$

↳ déterminer la loi de la Σ de 2 v.a. indépendantes par exemple

↳ démontrer le théorème central limite par exemple

- Si X v.a. continue de densité f :

$$\varphi_X(t) = \int_{\mathbb{R}} \exp(itx) f(x) dx$$

Si X v.a. discrète:

$$\varphi_X(t) = \sum_k \exp(itk) P(X=k)$$

- Propriétés:

① $\varphi_{\lambda X}(t) = \varphi_X(\lambda t) \quad \forall \lambda, \text{ un scalaire}$

② $\varphi_{X+a}(t) = \exp(ita) \varphi_X(t) \quad \forall a, \text{ un scalaire}$

③ Si X v.a. tq $E(X) = \mu$ et $V(X) = \sigma^2$

et $U = \frac{X - \mu}{\sigma}$

alors $\varphi_{\frac{X-\mu}{\sigma}}(t) = \varphi_U(t) = \exp\left(-it \frac{\mu}{\sigma}\right) \varphi_X\left(\frac{t}{\sigma}\right)$

④ Si $X_1, \dots, X_n$ sont des v.a. indépendantes:

$$\varphi_{X_1 + \dots + X_n} = \varphi_{X_1} \times \dots \times \varphi_{X_n}$$

- Proposition: Soit X une v.a. de fonction caractéristique de répartition $\varphi_X(t)$.

$$\varphi_X(0) = 1 \quad \text{et} \quad \frac{d^k \varphi_X}{dt^k}(0) = \varphi_X^{(k)}(0) = t^k E(X^k)$$

- Si $\varphi_X(t)$ est indéfiniment dérivable $\rightarrow \varphi_X(t) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \varphi_X^{(k)}(0) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} i^k E(X^k)$ ③

FORMULE de TAYLOR - LAGRANGE

■ Convergences des suites de v.a.

simon LAGONE
(Concept important dans la théorie des probas
utilisée notamment en stats et ds l'étude de processus
stochastiques)

• Convergence en probabilité :

Si $\forall \varepsilon > 0, \eta > 0$ (arbitrairement petits), $\exists$ un entier m_0 tq

$$\forall n > m_0 \Rightarrow P(|X_n - X| \geq \varepsilon) < \eta$$

$$\text{càd } P(|X_n - X| \geq \varepsilon) \xrightarrow{n \rightarrow +\infty} 0$$

$$\text{On notera } (X_n) \xrightarrow{IP} X$$

• Inégalité de Bienaymé-Tchebychev (BT)

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2} \quad \forall \varepsilon > 0$$

Rq: Lorsque $E(X_n) \xrightarrow{n \rightarrow +\infty} a$, pour établir la convergence en proba de la suite (X_n) vers a , il suffit de montrer $V(X_n) \xrightarrow{n \rightarrow +\infty} 0$

$$\text{En effet, d'après BT: } P(|X_n - E(X_n)| > \varepsilon) < \frac{V(X_n)}{\varepsilon^2} \rightarrow 0$$

Donc en passant par la limite:

$$\lim_{n \rightarrow +\infty} P(|X_n - a| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0 \Rightarrow (X_n) \xrightarrow[n \rightarrow +\infty]{IP} a$$

• Convergence en moyenne quadratique

On suppose que $E(|X_n - X|^2)$ existe

Suite de v.a. (X_n) converge en moy. quadra. vers une var. X si

$$E(|X_n - X|^2) \xrightarrow{n \rightarrow +\infty} 0$$

$$\text{On notera } (X_n) \xrightarrow{m.q.} X$$

- Convergence en loi :

Suite (X_n) converge en loi vers la var. X de f.d.r. F si en tout point de continuité de F la suite (F_n) des f.d.r. des (X_n) converge vers F .

$$\text{Càd, } \lim_{n \rightarrow +\infty} F_n(x) = F(x) \quad \forall x \text{ point de continuité de } F$$

$$\text{On notera } (X_n) \xrightarrow{\mathcal{L}} X$$

Rq: Pour des var. discrètes la CV en loi est équivalente à

$$\lim_{n \rightarrow +\infty} P(X_n = k) = P(X = k)$$

Théorème: Si la suite des fonct caractéristiques $\varphi_{X_n}(t)$ CV vers $\varphi_X(t)$ alors $(X_n) \xrightarrow{\mathcal{L}} X$.

- Correction de continuité (cas lois discrètes)

(X_n) une suite de v.a. discrètes.

$$P(X = x) \approx P\left(x - \frac{1}{2} < X < x + \frac{1}{2}\right)$$

• Théorème de Moivre-Laplace

Soit (X_n) suite v.a. de loi $B(np, p)$,
 alors $\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$ } notation avec la fonction caractéristique de la loi $B(np, p)$

• Convergence en loi : Poisson vers loi Normale

Soit (X_λ) suite v.a. de loi de Poisson $\mathcal{P}(\lambda)$
 alors $\frac{X_\lambda - \lambda}{\sqrt{\lambda}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$

• Théorème Central Limite :

Soit (X_n) suite de v.a. iid d'espérance μ et d'écart-type σ
 alors $\frac{\left[\sum_{i=1}^n X_i \right] - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$

■ Estimateurs

Soit (X_n) un échantillon de X .

→ moyenne empirique : $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow[n \rightarrow +\infty]{IP} E(X^2)$$

$$E(\bar{X}_n) = \mu = E(X)$$

$$V(\bar{X}_n) = \frac{\sigma^2}{n} \xrightarrow[n \rightarrow +\infty]{} 0$$

Donc, d'après BT. $\bar{X} \xrightarrow[n \rightarrow +\infty]{IP} \underset{E(X)}{\mu}$ Loi Forte des Grands Nombres

→ variance empirique : $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2$ (d'après K-H)

$$S^2 \xrightarrow[n \rightarrow +\infty]{IP} V(X)$$

• Définition: estimateur

Population X distribuée suivant une loi de proba. qui dépend d'un param θ inconnu.

On prélève un échantillon (X_n) de X .

On appelle estimateur de θ , toute v.a. T_n fonction de l'échantillon:

$$T_n = f(X_1, \dots, X_n)$$

→ biais: $B(T_n) = E(T_n) - \theta$ | $B(T_n) = 0 \Leftrightarrow$ estimateur sans biais

• Définition:

Une suite (T_n) d'estimateurs de θ est asymptotiquement sans biais si:

$$\lim_{n \rightarrow +\infty} E(T_n) = \theta$$

→ risque quadratique: $R(T_n) = E((T_n - \theta)^2) = V(T_n) + (E(T_n) - \theta)^2$
si estimateur sans biais: $R(T_n) = V(T_n)$

• Définition: estimateur convergent

T_n est convergent s'il converge en proba. vers θ .

On écrira $T_n \xrightarrow[n \rightarrow +\infty]{IP} \theta$.

• Définition: vraisemblance de θ

C'est la densité de l'échantillon (X_n) :

$$- \mathcal{L}(x_1, \dots, x_n, \theta) = \prod_{i=1}^n P(X_i = x_i) \quad (\text{cas discret})$$

$$- \mathcal{L}(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i) \quad (\text{cas continu})$$

Rappel:

→ Théorème de transfert $E(g(X)) = \sum_k g(k) P(X=k)$

$$\rightarrow \sum_{k=0}^n \frac{x^k}{k!} = e^x$$

Fiche Analyse Statistique

COURS 4

- Définition: quantité d'informations de Fisher

La quantité positive apportée par un échantillon sur le paramètre θ notée:

$$I_n(\theta) = E\left(\left(\frac{\partial \log(\mathcal{L}(x_1, \dots, x_n, \theta))}{\partial \theta}\right)^2\right) = -E\left(\frac{\partial^2 \log(\mathcal{L}(x_1, \dots, x_n, \theta))}{\partial \theta^2}\right)$$

- Inégalité de FRECHET-DARTOIS-CRATER-RAO

Pour tout estimateur T sans biais de θ : $V(T) \geq \frac{1}{I_n(\theta)}$

→ estimateur T qualifié d'efficace si: $V(T) = \frac{1}{I_n(\theta)}$