Fiche de Révision — Inférence Statistique by Yehor Korotenko

1. Modèle Statistique

Déf. (Modèle statistique)
$(\Omega, \mathcal{A}, \mathcal{P})$ avec $\mathcal{P} = \{P_\theta;\ \theta \in \Theta\}$.
- **Paramétrique**: $\exists p \in \mathbb{N}^*$, $\Theta \subset \mathbb{R}^p$
- **Non-paramétrique**: sinon

Déf. (Identifiabilité)
$\theta \mapsto P_\theta$ est **injective**.

Déf. (Estimateur)
$\hat{\theta} = h(X_1, \ldots, X_n)$ mesurable, indépendant de θ.

2. Qualité d'un Estimateur

Déf. (Biais)

$$B(\hat{\theta}, \theta) = \mathbb{E}[\hat{\theta}] - \theta$$

Sans biais $\Leftrightarrow B = 0$.

Déf. (Risque quadratique (MSE))

$$R(\hat{\theta}, \theta) = \mathbb{E}[(\hat{\theta} - \theta)^2]$$

Thm. (Décomposition Biais-Variance)

$$\boxed{R(\hat{\theta}, \theta) = B(\hat{\theta}, \theta)^2 + \mathrm{Var}(\hat{\theta})}$$

Preuve: développement de $(\hat{\theta} - \theta)^2$ autour de $\mathbb{E}[\hat{\theta}]$.

Déf. (Consistance)

$$\hat{\theta}_n \xrightarrow{\mathbb{P}} \theta \quad (n \to +\infty)$$

Forte: $\hat{\theta}_n \xrightarrow{\text{p.s.}} \theta$.
Outil: si $R(\hat{\theta}_n, \theta) \to 0 \Rightarrow$ consistant (Bienaymé-Tchebychev).

3. Méthode des Moments

Déf. (Moments)
- Théorique ord. k: $\mu_k = \mathbb{E}[X_i^k]$
- Empirique ord. k: $\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$

Par LGN: $\hat{\mu}_k \xrightarrow{\mathbb{P}} \mu_k$.

Prop. (Principe)
Si $\theta = \mathcal{L}(\mu_1, \ldots, \mu_k)$, alors

$$\hat{\theta}_{MM} = \mathcal{L}(\hat{\mu}_1, \ldots, \hat{\mu}_k)$$

est consistant (par LAC).

Exemples
- Bern(θ): $\hat{\theta} = \bar{X}$
- Exp(θ): $\hat{\theta} = 1/\bar{X}$
- $f_\theta(x) = \theta x^{\theta-1}\mathbf{1}_{[0,1]}$: $\hat{\theta} = \frac{\bar{X}}{1-\bar{X}}$
- **Variance empirique**: $\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$
  Biais: $-\frac{1}{n}\sigma^2$ (asymptotiquement sans biais)

4. Maximum de Vraisemblance

Déf. (Vraisemblance & log-vraisemblance)

$$L_n(\theta) = \prod_{i=1}^n f_\theta(X_i)$$

$$\ell_n(\theta) = \log L_n(\theta) = \sum_{i=1}^n \log f_\theta(X_i)$$

Déf. (EMV)

$$\hat{\theta}_{MV} = \arg\max_{\theta \in \Theta} \ell_n(\theta)$$

Équation de vraisemblance: $\ell_n'(\theta) = 0$

Prop. (Invariance de l'EMV)
Si $\hat{\theta}$ est EMV de θ, alors $g(\hat{\theta})$ est EMV de $g(\theta)$.

Exemples EMV
- Bern(θ): $\hat{\theta}^{MV} = \bar{X}$
- Exp(θ): $\hat{\theta}^{MV} = 1/\bar{X}$
- $f_\theta(x) = \frac{3}{\theta}x^2 e^{-x^3/\theta}$: $\hat{\theta} = \frac{1}{n}\sum X_i^3$

5. Modèle Régulier & Information de Fi

Déf. (Modèle régulier)
(P_θ) régulier si:
1. Θ ouvert, $\theta \mapsto f_\theta(x)$ est C^1
2. Supp(f_θ) indépendant de θ
3. Hypothèse (H): dérivation sous le signe $\int$

Déf. (Score)

$$S_n(\theta) = \frac{\partial}{\partial\theta}\ell_n(\theta) = \sum_{i=1}^n \frac{\partial}{\partial\theta}\log f_\theta(X_i)$$

Déf. (Information de Fisher)

$$I(\theta) = \mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f_\theta(X_1)\right)^2\right] = \int_S \frac{(\partial_\theta f_\theta)^2}{f_\theta}$$

$$I_n(\theta) = n\,I(\theta) = \mathrm{Var}_\theta(S_n(\theta))$$

Prop. (Score centré) Sous (H):

$$\mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(X_1)\right] = 0$$

Prop. (Dérivée seconde) Sous (H):

$$I_n(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\ell_n(\theta)\right]$$

Exemples $I(\theta)$
- Exp(θ): $I(\theta) = 1/\theta^2$
- Bern(θ): $I(\theta) = \frac{1}{\theta(1-\theta)}$
- $\mathcal{P}(\theta)$: $I(\theta) = 1/\theta$

6. Inégalité de Cramér–Rao

Thm. (Borne de Cramér–Rao)
Modèle régulier, $I(\theta) > 0$. Pour tout estimateur T sans biais de $g(\theta)$ avec $\mathbb{E}_\theta[T^2] < +\infty$:

$$\boxed{\mathrm{Var}_\theta(T) \geq \frac{[g'(\theta)]^2}{I_n(\theta)} = \frac{[g'(\theta)]^2}{n\,I(\theta)}}$$

Preuve: Cauchy-Schwarz sur $\langle T - g(\theta), \partial_\theta \log f_\theta \rangle_{P_\theta}$.

Déf. (Estimateur efficace)
T est **efficace** si $\mathrm{Var}_\theta(T) = \frac{[g'(\theta)]^2}{I_n(\theta)}$ (égalité dans CR).

7. Convergences

Déf. (Normalité asymptotique)
$\hat{\theta}_n$ est asymptotiquement normal si:

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(\theta))$$

Remarque: normalité asymptotique $\Rightarrow$ consistance.

Prop. (Lemme de Slutsky)
Si $X_n \xrightarrow{\mathcal{L}} X$ et $Y_n \xrightarrow{\mathcal{L}} c$ (constante), alors:
- $X_n + Y_n \xrightarrow{\mathcal{L}} X + c$
- $X_n Y_n \xrightarrow{\mathcal{L}} cX$
- $X_n / Y_n \xrightarrow{\mathcal{L}} X/c$ si $c \neq 0$

Thm. (δ-méthode)
Si $\sqrt{n}(Z_n - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2)$ et g dérivable en μ, $g'(\mu) \neq 0$:

$$\boxed{\sqrt{n}[g(Z_n) - g(\mu)] \xrightarrow{\mathcal{L}} \mathcal{N}(0, [g'(\mu)]^2\tau^2)}$$

Thm. (TLC (rappel))
(X_i) i.i.d., $\mathbb{E}[X_i] = \mu$, $\mathrm{Var}(X_i) = \sigma^2 < +\infty$:

$$\sqrt{n}(\bar{X} - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

Prop. (Convergence des couples)
$\begin{pmatrix} X_n \\ Y_n \end{pmatrix} \xrightarrow{\mathbb{P}} \begin{pmatrix} X \\ Y \end{pmatrix} \Leftrightarrow X_n \xrightarrow{\mathbb{P}} X$ et $Y_n \xrightarrow{\mathbb{P}} Y$
Attention: faux pour la convergence en loi!

Prop. (LAC (Lemme des app. continues))
Si $X_n \xrightarrow{\mathbb{P}} X$ et g continue, alors $g(X_n) \xrightarrow{\mathbb{P}} g(X)$.
Idem pour $\xrightarrow{\mathcal{L}}$ (et pour les couples).

-28. Fonction de Répartition Empirique

Déf. (F.R. empirique)

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{X_i \le x}$$

$n\,\hat{F}_n(x) \sim \mathrm{Bin}(n, F(x))$

Prop. (Propriétés de $\hat{F}_n$)

- $\hat{F}_n(x) \xrightarrow{\mathbb{P}} F(x)$ (LGN, consistance)
- $\sqrt{n}(\hat{F}_n(x) - F(x)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, F(x)(1 - F(x)))$ (TLC)
- **Glivenko-Cantelli**: $\sup_x |\hat{F}_n(x) - F(x)| \xrightarrow{\mathbb{P}} 0$

Déf. (Quantile empirique)
$\hat{q}_{n,\alpha} = \hat{F}_n^{-1}(\alpha) = X_{(\lceil n\alpha \rceil)}$
où $X_{(1)} \le \cdots \le X_{(n)}$ est l'échantillon ordonné.

-29. Intervalles de Confiance

Déf. (Intervalle de confiance)
$[B_{\mathrm{inf}, B_{\mathrm{sup}}]}$ (bornes aléatoires, ne dépendent pas de θ) tel que:

$$\mathbb{P}(\theta \in [B_{\mathrm{inf}, B_{\mathrm{sup}}]) \ge 1-\alpha}$$

- **Exact**: égalité $= 1 - \alpha$
- **Asymptotique**: $\mathbb{P}(\dots) \to 1 - \alpha$
- En pratique: $\alpha = 5\%$ ou 1%

Thm. (Méthode pivotale)
Si $\sqrt{n}(\hat{\theta}_n - \theta)/\hat{\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$, alors:

$$\boxed{IC_{1-\alpha} = \left[\hat{\theta}_n - \frac{q_{1-\alpha/2}\,\hat{\sigma}}{\sqrt{n}}\ ,\ \hat{\theta}_n - \frac{q_{\alpha/2}\,\hat{\sigma}}{\sqrt{n}}\right]}$$

où $q_\beta = \Phi^{-1}(\beta)$ quantile gaussien.
Symétrie: $q_{1-\alpha/2} = -q_{\alpha/2}$, donc largeur $= \frac{2q_{1-\alpha/2}\hat{\sigma}}{\sqrt{n}}$.

IC pour proportion $X_i \sim \mathrm{Bern}(\theta)$:
$\hat{\sigma}^2 = \hat{\theta}(1-\hat{\theta})$, pivot: $\frac{\sqrt{n}(\bar{X}-\theta)}{\sqrt{\hat{\theta}(1-\hat{\theta})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

IC pour $\mathrm{Exp}(\theta)$ ($\hat{\theta} = 1/\bar{X}$):
$\frac{\sqrt{n}(\hat{\theta}-\theta)}{\hat{\theta}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$ (Slutsky+δ-méthode)
$IC = \hat{\theta}\left[1 \pm \frac{q_{1-\alpha/2}}{\sqrt{n}}\right]$

Thm. (Pivot de Slutsky)
Si $\hat{\sigma}^2$ est consistant pour $\sigma^2(\theta)$:

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\hat{\sigma}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

Preuve: Slutsky + $\frac{\sigma(\theta)}{\hat{\sigma}} \xrightarrow{\mathbb{P}} 1$.

-210. Efficacité Asymptotique

Prop. (Asymptotiquement efficace)
$\hat{\theta}_n$ asymptotiquement efficace si:

$$n\,\mathrm{Var}(\hat{\theta}_n) \xrightarrow{n\to\infty} \frac{1}{I(\theta)}$$

L'EMV est généralement asymptotiquement efficace.

-2Aide-Mémoire Rapide

Objet	**Formule clé**
Biais	$\mathbb{E}[\hat{\theta}] - \theta$
MSE	$B^2 + \mathrm{Var}(\hat{\theta})$
I. Fisher	$\mathbb{E}[(\partial_\theta \log f)^2]$
I_n	$nI(\theta)$
BCR	$\mathrm{Var}(T) \ge [g'(\theta)]^2/I_n$
δ-méthode	$(g'(\mu))^2\tau^2$
IC asympt.	$\hat{\theta} \pm q_{1-\alpha/2}\hat{\sigma}/\sqrt{n}$

Quantiles $\mathcal{N}(0,1)$ usuels
$q_{0.975} = 1.96$, $q_{0.995} = 2.576$, $q_{0.95} = 1.645$