Estimation par Maximum de Vraisemblance

Exemples

Plan

Plan

Loi normale – Présentation

- La loi normale est omniprésente en statistique :
  - Erreurs de mesure en physique, métrologie
  - Taille, poids et caractéristiques biologiques
  - Scores de tests psychométriques
  - Rendements financiers (en première approximation)
  - Distribution limite par le théorème central limite
- C'est la distribution de référence pour la théorie de l'estimation : c'est souvent le cas où les calculs sont les plus explicites.

Loi normale – Modèle

- $X_i \sim \mathcal{N}(\mu, \sigma^2)$ avec $\mu \in \mathbb{R}$ et $\sigma^2 > 0$:

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

- Moments :

$$\mathbb{E}[X] = \mu, \qquad \mathbb{V}[X] = \sigma^2$$

- On peut paramétrer par (μ, σ^2) ou par (μ, σ). Nous utilisons (μ, σ^2).

Loi normale – Log-vraisemblance

- Vraisemblance :

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

- Log-vraisemblance :

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

- On peut réécrire le dernier terme :

$$\sum_{i=1}^{n}(x_i-\mu)^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2$$

Loi normale – Estimateur de μ

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} x_i = n\mu$$

- Solution :

$$\boxed{\hat{\mu}_{MV} = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}}$$

- C'est la moyenne empirique.

Loi normale – Estimateur de σ^2

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0$$

- En remplaçant μ par $\hat{\mu} = \bar{x}$:

$$\frac{n}{2\sigma^2} = \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

- Solution :

$$\boxed{\hat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

- C'est la variance empirique (avec diviseur n, pas $n-1$).

Loi normale – Distribution exacte de $\hat{\mu}$

- Puisque $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$ est une combinaison linéaire de variables normales indépendantes :

$$\hat{\mu} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

- Propriétés :
  - $\mathbb{E}[\hat{\mu}] = \mu$: estimateur sans biais
  - $\mathbb{V}[\hat{\mu}] = \dfrac{\sigma^2}{n}$: variance exacte
  - Distribution exactement normale pour tout n

Loi normale – Distribution exacte de $\hat{\sigma}^2$

- On montre que :
$$\frac{n\hat{\sigma}^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$
- Puisque $\mathbb{E}[\chi^2_{n-1}] = n - 1$:
$$\mathbb{E}\left[\frac{n\hat{\sigma}^2}{\sigma^2}\right] = n - 1 \quad \Rightarrow \quad \mathbb{E}[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2$$
- L'estimateur MV est biaisé : il sous-estime σ^2.
- Biais : Biais$(\hat{\sigma}^2) = -\dfrac{\sigma^2}{n}$

Loi normale – Estimateur sans biais de σ^2

- L'estimateur sans biais de σ^2 est :

$$\boxed{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n}{n-1}\hat{\sigma}^2_{MV}}$$

- On a $\mathbb{E}[S^2] = \sigma^2$.

- Le diviseur $n-1$ (et non n) correspond aux degrés de liberté : on perd un degré de liberté en estimant μ par $\bar{X}$.

- C'est l'estimateur utilisé par défaut dans la plupart des logiciels statistiques.

Loi normale – Variance de $\hat{\sigma}^2$

- Puisque $\mathbb{V}[\chi^2_k] = 2k$, on a $\mathbb{V}[\chi^2_{n-1}] = 2(n-1)$.
- Variance de l'estimateur MV :

$$\mathbb{V}\left[\frac{n\hat{\sigma}^2}{\sigma^2}\right] = 2(n-1) \quad \Rightarrow \quad \frac{n^2}{\sigma^4}\mathbb{V}[\hat{\sigma}^2] = 2(n-1)$$

$$\boxed{\mathbb{V}[\hat{\sigma}^2] = \frac{2(n-1)\sigma^4}{n^2}}$$

- Pour n grand : $\mathbb{V}[\hat{\sigma}^2] \approx \dfrac{2\sigma^4}{n}$

Loi normale – Variance de S^2

- Puisque $S^2 = \frac{n}{n-1}\hat{\sigma}^2$:

$$\mathbb{V}[S^2] = \frac{n^2}{(n-1)^2}\mathbb{V}[\hat{\sigma}^2] = \frac{n^2}{(n-1)^2} \cdot \frac{2(n-1)\sigma^4}{n^2}$$

$$\boxed{\mathbb{V}[S^2] = \frac{2\sigma^4}{n-1}}$$

- Pour n grand, $\mathbb{V}[\hat{\sigma}^2_{MV}] < \mathbb{V}[S^2]$.
- L'estimateur MV (biaisé) a une variance légèrement plus faible que l'estimateur sans biais.

Loi normale – Indépendance de $\hat{\mu}$ et $\hat{\sigma}^2$

Théorème de Cochran

Pour un échantillon i.i.d. de loi normale, $\bar{X}$ et $\sum_{i=1}^{n}(X_i - \bar{X})^2$ sont indépendants.

- Conséquence : $\hat{\mu} = \bar{X}$ et $\hat{\sigma}^2$ sont indépendants.
- Cette indépendance est cruciale pour construire la statistique de Student :

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

où $S = \sqrt{S^2}$ est l'écart-type empirique corrigé.

Loi normale – Information de Fisher

- Dérivées secondes :

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{n}{\sigma^2}$$

$$\frac{\partial^2 \ell}{\partial (\sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\frac{\partial^2 \ell}{\partial \mu \partial \sigma^2} = -\frac{1}{\sigma^4} \sum_{i=1}^{n} (x_i - \mu)$$

Loi normale – Matrice d'information de Fisher

- $\mathbb{E}\left[\sum_{i=1}^{n}(X_i - \mu)^2\right] = n\sigma^2$ et $\mathbb{E}\left[\sum_{i=1}^{n}(X_i - \mu)\right] = 0$
- Matrice d'information de Fisher :

$$\boxed{I(\mu, \sigma^2) = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}}$$

- La matrice est diagonale : μ et σ^2 sont orthogonaux.
- Les estimateurs $\hat{\mu}$ et $\hat{\sigma}^2$ sont asymptotiquement indépendants (et même exactement indépendants pour la loi normale).

Loi normale – Bornes de Cramér-Rao

- Pour μ :
$$\mathbb{V}[\hat{\mu}] \geq \frac{1}{I_{\mu\mu}} = \frac{\sigma^2}{n}$$
Variance exacte de $\hat{\mu}$: $\dfrac{\sigma^2}{n} \Rightarrow$ efficace à distance finie.

- Pour σ^2 :
$$\mathbb{V}[\hat{\sigma}^2] \geq \frac{1}{I_{\sigma^2\sigma^2}} = \frac{2\sigma^4}{n}$$
Variance exacte de $\hat{\sigma}^2_{MV}$: $\dfrac{2(n-1)\sigma^4}{n^2} = \dfrac{2\sigma^4}{n}\left(1 - \dfrac{1}{n}\right)$

- L'estimateur MV de σ^2 est asymptotiquement efficace (mais pas exactement efficace à distance finie).

Loi normale – Erreur quadratique moyenne

- L'erreur quadratique moyenne (EQM) combine biais et variance :

$$\text{EQM}(\hat{\theta}) = \mathbb{V}[\hat{\theta}] + \text{Biais}(\hat{\theta})^2$$

- Comparaison pour l'estimation de σ^2 :

Estimateur	Biais	Variance	EQM
$\hat{\sigma}^2_{MV}$	$-\sigma^2/n$	$\frac{2(n-1)\sigma^4}{n^2}$	$\frac{(2n-1)\sigma^4}{n^2}$
S^2	0	$\frac{2\sigma^4}{n-1}$	$\frac{2\sigma^4}{n-1}$

- Pour $n \geq 2$: EQM$(\hat{\sigma}^2_{MV}) <$ EQM(S^2) (l'estimateur MV biaisé a une EQM plus faible).

Loi normale – Résumé

- Estimateurs du maximum de vraisemblance :

$$\hat{\mu}_{MV} = \bar{X}, \qquad \hat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- Propriétés :

	$\hat{\mu}$	$\hat{\sigma}^2_{MV}$
Biais	0	$-\sigma^2/n$
Variance	σ^2/n	$2(n-1)\sigma^4/n^2$
Efficace	Oui (exact)	Asymptotiquement
Distribution	$\mathcal{N}(\mu, \sigma^2/n)$	$\frac{\sigma^2}{n}\chi^2_{n-1}$

- Les estimateurs $\hat{\mu}$ et $\hat{\sigma}^2$ sont indépendants (théorème de Cochran).

Plan

Loi de Bernoulli – Présentation

- La loi de Bernoulli modélise une épreuve à deux issues (succès/échec) :
  - Pile ou face
  - Réussite ou échec d'un traitement médical
  - Défectueux ou conforme dans un contrôle qualité
  - Clic ou non sur une publicité en ligne
  - Vote pour ou contre une proposition
- C'est la brique de base pour construire :
  - Loi binomiale (somme de n Bernoulli i.i.d.)
  - Loi géométrique (nombre d'essais jusqu'au premier succès)

Loi de Bernoulli – Modèle

- $X_i \sim \text{Bernoulli}(p)$ avec $p \in (0,1)$:

$$P(X = x) = p^x(1-p)^{1-x}, \quad x \in \{0,1\}$$

  Ou : $P(X = 1) = p$, $P(X = 0) = 1 - p$

- Moments :

$$\mathbb{E}[X] = p, \qquad \mathbb{V}[X] = p(1-p)$$

- La variance est maximale pour $p = 1/2$ (incertitude maximale).

Loi de Bernoulli – Log-vraisemblance

- Vraisemblance :

$$L(p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

- En notant $S = \sum_{i=1}^{n} x_i$ le nombre de succès : $L(p) = p^S(1-p)^{n-S}$

- Log-vraisemblance :

$$\ell(p) = S \log p + (n-S)\log(1-p)$$

Loi de Bernoulli – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial p} = \frac{S}{p} - \frac{n-S}{1-p} = 0 \quad \Rightarrow \quad S = pn$$

- Solution :

$$\boxed{\hat{p}_{MV} = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}}$$

  C'est la proportion empirique de succès.

- Vérification : $\dfrac{\partial^2 \ell}{\partial p^2} = -\dfrac{S}{p^2} - \dfrac{n-S}{(1-p)^2} < 0 \checkmark$

Loi de Bernoulli – Variance exacte de l'estimateur

- La somme $S = \sum_{i=1}^{n} X_i$ suit une loi binomiale $\mathrm{Bin}(n, p)$:

$$\mathbb{E}[S] = np, \qquad \mathbb{V}[S] = np(1-p)$$

- Puisque $\hat{p} = S/n$: $\mathbb{E}[\hat{p}] = p$ (estimateur sans biais).

- Variance exacte :

$$\boxed{\mathbb{V}[\hat{p}] = \frac{\mathbb{V}[S]}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}}$$

Loi de Bernoulli – Information de Fisher

- Dérivée seconde de la log-vraisemblance :

$$\frac{\partial^2 \ell}{\partial p^2} = -\frac{S}{p^2} - \frac{n-S}{(1-p)^2}$$

- Espérance (avec $\mathbb{E}[S] = np$) :

$$\mathbb{E}\left[\frac{\partial^2 \ell}{\partial p^2}\right] = -\frac{np}{p^2} - \frac{n-np}{(1-p)^2} = -\frac{n}{p} - \frac{n}{1-p}$$

- Information de Fisher :

$$\boxed{I(p) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial p^2}\right] = \frac{n}{p} + \frac{n}{1-p} = \frac{n}{p(1-p)}}$$

Loi de Bernoulli – Efficacité de l'estimateur

- Borne de Cramér-Rao : $\mathbb{V}[\hat{p}] \geq \frac{1}{I(p)} = \frac{p(1-p)}{n}$
- Variance exacte de $\hat{p}$: $\frac{p(1-p)}{n}$

Résultat remarquable

L'estimateur $\hat{p} = \bar{X}$ atteint exactement la borne de Cramér-Rao.
C'est un estimateur efficace (pas seulement asymptotiquement).

Loi de Bernoulli – Distribution de l'estimateur

- Distribution exacte (puisque $S = n\hat{p} \sim \text{Bin}(n, p)$) :

$$P(\hat{p} = k/n) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

- Distribution asymptotique (par le TCL) :

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, p(1-p))$$

soit $\hat{p} \overset{\text{approx}}{\sim} \mathcal{N}\left(p, \frac{p(1-p)}{n}\right)$

- Cette approximation est bonne pour $np \geq 5$ et $n(1-p) \geq 5$.

Plan

Loi de Poisson – Présentation

- La loi de Poisson modélise le nombre d'événements rares dans un intervalle :
  - Nombre d'appels reçus par heure dans un centre
  - Nombre de défauts par mètre de tissu
  - Nombre d'accidents par jour sur une autoroute
  - Nombre de mutations par génome
  - Nombre de clients arrivant dans une file d'attente
- Limite de la loi binomiale Bin(n, p) quand $n \to \infty$ et $p \to 0$ avec $np \to \lambda$.

Loi de Poisson – Modèle

- $X_i \sim \text{Poisson}(\lambda)$ avec $\lambda > 0$:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

- Moments :

$$\mathbb{E}[X] = \lambda, \qquad \mathbb{V}[X] = \lambda$$

- L'égalité espérance = variance est caractéristique de la loi de Poisson.

Loi de Poisson – Calcul des moments

- Espérance :

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \boxed{\lambda}$$

- Calcul de $\mathbb{E}[X(X-1)]$:

$$\mathbb{E}[X(X-1)] = \sum_{k=2}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-2)!} = \lambda^2 e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda^2$$

- Donc $\mathbb{E}[X^2] = \lambda^2 + \lambda$, et :

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \boxed{\lambda}$$

Loi de Poisson – Log-vraisemblance

- Vraisemblance :

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!}$$

- En notant $S = \sum_{i=1}^n x_i$: $L(\lambda) = \dfrac{\lambda^S e^{-n\lambda}}{\prod_{i=1}^n x_i!}$

- Log-vraisemblance :

$$\ell(\lambda) = S \log \lambda - n\lambda - \sum_{i=1}^{n} \log(x_i!)$$

Loi de Poisson – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial \lambda} = \frac{S}{\lambda} - n = 0$$

- Solution :

$$\boxed{\hat{\lambda}_{MV} = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}}$$

  L'estimateur est la moyenne empirique.

- Vérification : $\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{S}{\lambda^2} < 0 \checkmark$

Loi de Poisson – Variance exacte de l'estimateur

- La somme $S = \sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$, donc :

$$\mathbb{E}[S] = n\lambda, \qquad \mathbb{V}[S] = n\lambda$$

- L'estimateur est sans biais : $\mathbb{E}[\hat{\lambda}] = \lambda$
- Variance exacte :

$$\boxed{\mathbb{V}[\hat{\lambda}] = \frac{\mathbb{V}[S]}{n^2} = \frac{n\lambda}{n^2} = \frac{\lambda}{n}}$$

Loi de Poisson – Information de Fisher

- Dérivée seconde de la log-vraisemblance :

$$\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{S}{\lambda^2}$$

- Espérance (avec $\mathbb{E}[S] = n\lambda$) :

$$\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \lambda^2}\right] = -\frac{n\lambda}{\lambda^2} = -\frac{n}{\lambda}$$

- Information de Fisher :

$$\boxed{I(\lambda) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \lambda^2}\right] = \frac{n}{\lambda}}$$

Loi de Poisson – Efficacité de l'estimateur

- Borne de Cramér-Rao : $\mathbb{V}[\hat{\lambda}] \geq \dfrac{1}{I(\lambda)} = \dfrac{\lambda}{n}$
- Variance exacte de $\hat{\lambda}$: $\dfrac{\lambda}{n}$

Résultat remarquable

L'estimateur $\hat{\lambda} = \bar{X}$ atteint exactement la borne de Cramér-Rao.
C'est un estimateur efficace (pas seulement asymptotiquement).

Loi de Poisson – Distribution de l'estimateur

- Distribution exacte (puisque $n\hat{\lambda} = S \sim \text{Poisson}(n\lambda)$) :

$$P(\hat{\lambda} = k/n) = \frac{(n\lambda)^k e^{-n\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

- Distribution asymptotique (par le TCL) :

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{d} \mathcal{N}(0, \lambda)$$

soit $\hat{\lambda} \overset{\text{approx}}{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$

Plan

Loi exponentielle – Présentation

- La loi exponentielle modélise la durée de vie ou le temps d'attente :
  - Temps entre deux appels dans un centre téléphonique
  - Durée de vie d'un composant électronique
  - Temps d'attente à un guichet
  - Temps entre deux tremblements de terre
- Propriété sans mémoire (seule loi continue la vérifiant) :

$$P(X > s + t \mid X > s) = P(X > t)$$

Loi exponentielle – Modèle

- $X_1, \ldots, X_n \overset{iid}{\sim} \mathsf{Exp}(\lambda)$ avec $\lambda > 0$:

$$f(x; \lambda) = \lambda e^{-\lambda x}, \quad x \geq 0$$

- Moments :

$$\mathbb{E}[X] = \frac{1}{\lambda}, \qquad \mathbb{V}[X] = \frac{1}{\lambda^2}$$

- La somme $S = \sum_{i=1}^n X_i$ est une statistique suffisante.

Loi exponentielle – Log-vraisemblance

- Vraisemblance :

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right)$$

- Log-vraisemblance :

$$\ell(\lambda) = \log L(\lambda) = n \log \lambda - \lambda \sum_{i=1}^{n} x_i$$

Loi exponentielle – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

- Solution :

$$\boxed{\hat{\lambda}_{MV} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}}}$$

- Vérification : $\dfrac{\partial^2 \ell}{\partial \lambda^2} = -\dfrac{n}{\lambda^2} < 0$ ✓

Loi exponentielle – Variance exacte de l'estimateur

- On a $\hat{\lambda} = n/S$ où $S = \sum_{i=1}^{n} X_i$.
- La somme S a pour densité (loi Gamma) :

$$f_S(s) = \frac{\lambda^n s^{n-1} e^{-\lambda s}}{(n-1)!}, \quad s > 0$$

- Calcul de $\mathbb{E}[1/S]$ (pour $n > 1$) :

$$\mathbb{E}\left[\frac{1}{S}\right] = \frac{\lambda^n}{(n-1)!} \int_0^{\infty} s^{n-2} e^{-\lambda s} \, \mathrm{d}s$$

Loi exponentielle – Variance exacte (suite)

- En utilisant $\int_0^\infty s^k e^{-\lambda s}\,\mathrm{d}s = \frac{k!}{\lambda^{k+1}}$:

$$\mathbb{E}\left[\frac{1}{S}\right] = \frac{\lambda^n}{(n-1)!} \cdot \frac{(n-2)!}{\lambda^{n-1}} = \frac{\lambda}{n-1}$$

- Donc $\mathbb{E}[\hat{\lambda}] = n \cdot \frac{\lambda}{n-1} = \frac{n\lambda}{n-1}$ (estimateur biaisé à distance finie).
- Calcul de $\mathbb{E}[1/S^2]$ (pour $n > 2$) :

$$\mathbb{E}\left[\frac{1}{S^2}\right] = \frac{\lambda^n}{(n-1)!} \cdot \frac{(n-3)!}{\lambda^{n-2}} = \frac{\lambda^2}{(n-1)(n-2)}$$

Loi exponentielle – Variance exacte (fin)

- $\mathbb{E}[\hat{\lambda}^2] = n^2 \cdot \mathbb{E}\left[\frac{1}{S^2}\right] = \dfrac{n^2\lambda^2}{(n-1)(n-2)}$
- Variance :

$$\mathbb{V}[\hat{\lambda}] = \frac{n^2\lambda^2}{(n-1)(n-2)} - \frac{n^2\lambda^2}{(n-1)^2} = \frac{n^2\lambda^2}{(n-1)^2} \cdot \frac{1}{n-2}$$

$$\boxed{\mathbb{V}[\hat{\lambda}] = \frac{n^2\lambda^2}{(n-1)^2(n-2)}}$$

- Pour n grand : $\mathbb{V}[\hat{\lambda}] \approx \dfrac{\lambda^2}{n}$

Loi exponentielle – Information de Fisher

- Information de Fisher :

$$I(\lambda) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \lambda^2}\right] = -\mathbb{E}\left[-\frac{n}{\lambda^2}\right] = \frac{n}{\lambda^2}$$

- Borne de Cramér-Rao : $\mathbb{V}[\hat{\lambda}] \geq \dfrac{1}{I(\lambda)} = \dfrac{\lambda^2}{n}$
- Comparaison :
  - Variance exacte : $\dfrac{n^2\lambda^2}{(n-1)^2(n-2)} \approx \dfrac{\lambda^2}{n}$ pour n grand
  - L'estimateur atteint la borne de Cramér-Rao seulement asymptotiquement

Plan

Loi géométrique – Présentation

- La loi géométrique modélise le nombre d'essais jusqu'au premier succès :
  - Nombre de lancers de dé jusqu'à obtenir un 6
  - Nombre de clients démarchés jusqu'à une vente
  - Nombre de tentatives jusqu'à la réussite d'un examen
  - Position du premier défaut dans une chaîne de production
- C'est la version discrète de la loi exponentielle, elle vérifie aussi la propriété sans mémoire.

Loi géométrique – Modèle

- $X_i \sim \mathsf{Geom}(p)$ avec $p \in (0,1)$:

$$P(X=k) = (1-p)^{k-1}p, \quad k = 1,2,3,\dots$$

- Moments :

$$\mathbb{E}[X] = \frac{1}{p}, \qquad \mathbb{V}[X] = \frac{1-p}{p^2}$$

- La somme $S = \sum_{i=1}^{n} X_i$ suit une loi binomiale négative $\mathsf{NegBin}(n,p)$.

Loi géométrique – Séries utiles

- Série géométrique (pour $|q| < 1$) : $\displaystyle\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$
- Dérivée première par rapport à q :

$$\sum_{k=1}^{\infty} k q^{k-1} = \frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{1}{1-q}\right) = \frac{1}{(1-q)^2}$$

- Dérivée seconde par rapport à q :

$$\sum_{k=2}^{\infty} k(k-1) q^{k-2} = \frac{\mathrm{d}^2}{\mathrm{d}q^2}\left(\frac{1}{1-q}\right) = \frac{2}{(1-q)^3}$$

- Avec $q = 1 - p$, on a $1 - q = p$, donc :

$$\boxed{\sum_{k=1}^{\infty} k q^{k-1} = \frac{1}{p^2}, \qquad \sum_{k=2}^{\infty} k(k-1) q^{k-2} = \frac{2}{p^3}}$$

Loi géométrique – Série du logarithme

- De la série géométrique $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$ pour $|x| < 1$, en intégrant de 0 à q :

$$\int_0^q \frac{\mathrm{d}x}{1-x} = \sum_{k=0}^{\infty} \int_0^q x^k \,\mathrm{d}x$$

- À gauche : $-\log(1-q)$. À droite : $\sum_{j=1}^{\infty} \frac{q^j}{j}$

- Conclusion : $\boxed{\displaystyle\sum_{k=1}^{\infty} \frac{q^k}{k} = -\log(1-q)}$

- Avec $q = 1-p$: $\displaystyle\sum_{k=1}^{\infty} \frac{(1-p)^k}{k} = -\log(p)$

Loi géométrique – Calcul des moments

- Espérance (avec $q = 1 - p$) :

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k \cdot q^{k-1} p = p \sum_{k=1}^{\infty} k q^{k-1} = p \cdot \frac{1}{p^2} = \boxed{\frac{1}{p}}$$

- Calcul de $\mathbb{E}[X(X-1)]$:

$$\mathbb{E}[X(X-1)] = \sum_{k=2}^{\infty} k(k-1) q^{k-1} p = pq \sum_{k=2}^{\infty} k(k-1) q^{k-2} = pq \cdot \frac{2}{p^3} = \frac{2q}{p^2}$$

- Donc $\mathbb{E}[X^2] = \frac{2q}{p^2} + \frac{1}{p} = \frac{2q+p}{p^2} = \frac{1+q}{p^2}$

- Variance :

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2} = \boxed{\frac{1-p}{p^2}}$$

Loi géométrique – Log-vraisemblance

- Vraisemblance :

$$L(p) = \prod_{i=1}^{n} (1-p)^{x_i - 1} p = p^n (1-p)^{\sum_{i=1}^n x_i - n}$$

- Log-vraisemblance (avec $S = \sum_{i=1}^n x_i$) :

$$\ell(p) = n \log p + (S - n) \log(1-p)$$

Loi géométrique – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial p} = \frac{n}{p} - \frac{S-n}{1-p} = 0 \quad \Rightarrow \quad n = pS$$

- Solution :

$$\boxed{\hat{p}_{MV} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}}}$$

Loi géométrique – Biais de l'estimateur

- Par l'inégalité de Jensen (puisque $g(x) = 1/x$ est convexe) :

$$\mathbb{E}[\hat{p}] = n \cdot \mathbb{E}\left[\frac{1}{S}\right] > n \cdot \frac{1}{\mathbb{E}[S]} = p$$

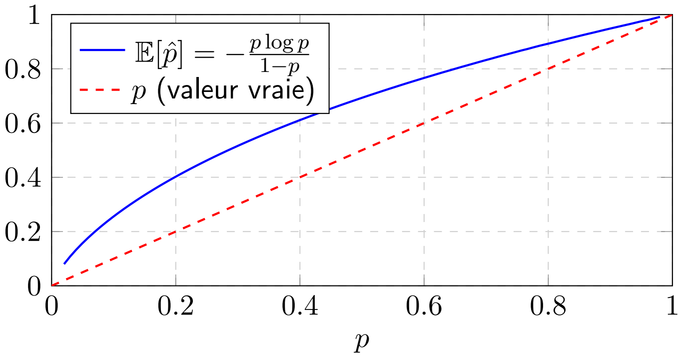
  L'estimateur surestime p en moyenne.

- Calcul exact pour $n = 1$ (avec $\sum_{k=1}^{\infty} \frac{q^k}{k} = -\log(1-q)$) :

$$\mathbb{E}\left[\frac{1}{X}\right] = \frac{p}{q} \sum_{k=1}^{\infty} \frac{q^k}{k} = -\frac{p \log p}{1-p}$$

- Exemple : pour $p = 0.5$, $\mathbb{E}[\hat{p}] = \log 2 \approx 0.693$, soit un biais de $+0.193$.

Loi géométrique – Biais ($n = 1$) : représentation graphique



Le biais (écart entre les courbes) est toujours positif : l'estimateur surestime p.

Loi géométrique – Biais : formule exacte (1/5)

- Pour $S \sim \mathrm{NegBin}(n, p)$ avec $P(S = k) = \binom{k-1}{n-1} p^n q^{k-n}$:

$$\mathbb{E}\left[\frac{1}{S}\right] = \sum_{k=n}^{\infty} \frac{1}{k} \binom{k-1}{n-1} p^n q^{k-n} \qquad (q = 1 - p)$$

- En notant que : $\dfrac{1}{k} = \displaystyle\int_0^1 t^{k-1}\,\mathrm{d}t$, il vient :

$$\mathbb{E}\left[\frac{1}{S}\right] = \sum_{k=n}^{\infty} \int_0^1 t^{k-1} \binom{k-1}{n-1} p^n q^{k-n}\,\mathrm{d}t$$

$$\Leftrightarrow \int_0^1 p^n t^{n-1} \sum_{j=0}^{\infty} \binom{j+n-1}{n-1} (tq)^j\,\mathrm{d}t$$

Loi géométrique – Biais : formule exacte (2/5)

- Sachant que $\sum_{j=0}^{\infty} \binom{j+n-1}{n-1} x^j = (1-x)^{-n}$, il vient :

$$\mathbb{E}\left[\frac{1}{S}\right] = p^n \int_0^1 \frac{t^{n-1}}{(1-tq)^n}\,\mathrm{d}t$$

- En posant $u = 1 - tq$, on a $\mathrm{d}t = -\frac{\mathrm{d}u}{q}$, et :

$$\mathbb{E}\left[\frac{1}{S}\right] = \left(\frac{p}{q}\right)^n \int_{1-q}^1 \frac{(1-u)^{n-1}}{u^n}\,\mathrm{d}u$$

- En posant $v = (1-u)/u$, on a $\mathrm{d}u = -\frac{1}{(1+v)^2}\,\mathrm{d}v$, et :

$$\mathbb{E}\left[\frac{1}{S}\right] = \left(\frac{p}{q}\right)^n \int_0^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v$$

Loi géométrique – Biais : formule exacte (3/5)

Vérification pour $n = 1$:

$$\mathbb{E}\left[\frac{1}{X}\right] = \frac{p}{q}\int_0^{q/p} \frac{1}{1+v}\,\mathrm{d}v = \frac{p}{q}\log\left(1+\frac{q}{p}\right) = \frac{p}{q}\log\left(\frac{1}{p}\right) = -\frac{p\log p}{1-p} \quad \checkmark$$

Loi géométrique – Biais : formule exacte (4/5)

- Pour $v \in [0, q/p]$, on a :
  - $v \geq 0 \Rightarrow 1 + v \geq 1 \Rightarrow \frac{1}{1+v} \leq 1$
  - $v \leq \frac{q}{p} \Rightarrow 1 + v \leq \frac{p+q}{p} = \frac{1}{p} \Rightarrow \frac{1}{1+v} \geq p$
- Ainsi, pour tout $v \in [0, q/p]$: $p \leq \frac{1}{1+v} \leq 1$
- En multipliant par $v^{n-1} \geq 0$, puis en intégrant sur $[0, q/p]$:

$$\frac{p}{n}\left(\frac{q}{p}\right)^n \leq \int_0^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v \leq \frac{1}{n}\left(\frac{q}{p}\right)^n$$

Loi géométrique – Biais : formule exacte (5/5)

Finalement, en multipliant par $n\left(\frac{p}{q}\right)^n$:

$$p \leq n\left(\frac{p}{q}\right)^n \int_0^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v \leq 1$$

$$\Leftrightarrow p \leq \mathbb{E}[\hat{p}] \leq 1$$

L'estimateur $\hat{p}$ surestime p en moyenne, mais reste borné par 1 (c'est heureux).

Loi géométrique – Convergence : concentration de v^{n-1}

- Pour n grand, v^{n-1} se concentre près de la borne supérieure $v = q/p$.

- Considérons $v \in [0, q/p]$ avec $q/p > 0$:
  - Si $v < q/p$, alors $\frac{v}{q/p} < 1$, donc $\left(\frac{v}{q/p}\right)^{n-1} \to 0$ quand $n \to \infty$
  - Plus v est petit par rapport à q/p, plus v^{n-1} décroît vite

- Découpons l'intégrale :

$$\int_0^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v = \underbrace{\int_0^{q/p-\varepsilon} \frac{v^{n-1}}{1+v}\,\mathrm{d}v}_{I_1} + \underbrace{\int_{q/p-\varepsilon}^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v}_{I_2}$$

avec $\varepsilon > 0$ petit.

Loi géométrique – Convergence : la première intégrale I_1

- Sur $[0, q/p - \varepsilon]$, on a $v \leq q/p - \varepsilon$ et $\frac{1}{1+v} \leq 1$, donc :

$$I_1 = \int_0^{q/p-\varepsilon} \frac{v^{n-1}}{1+v}\,\mathrm{d}v \leq \int_0^{q/p-\varepsilon} v^{n-1}\,\mathrm{d}v = \frac{(q/p-\varepsilon)^n}{n}$$

- En comparant avec l'intégrale totale :

$$\frac{I_1}{\int_0^{q/p} v^{n-1}\,\mathrm{d}v} \leq \frac{(q/p-\varepsilon)^n/n}{(q/p)^n/n} = \left(1 - \frac{\varepsilon p}{q}\right)^n \xrightarrow[n\to\infty]{} 0$$

- La contribution de I_1 décroît **exponentiellement** vers 0.

Loi géométrique – Convergence : la seconde intégrale I_2

- En $v = q/p$, on a $\frac{1}{1+v} = \frac{1}{1+q/p} = \frac{p}{p+q} = p$.
- Pour v proche de q/p, on peut approximer $\frac{1}{1+v}$ par sa valeur en q/p.
- Ainsi, avec cette approximation :

$$\begin{aligned} I_2 = \int_{q/p-\varepsilon}^{q/p} \frac{v^{n-1}}{1+v} \,\mathrm{d}v &\approx p \int_{q/p-\varepsilon}^{q/p} v^{n-1} \,\mathrm{d}v \\ &\approx p \cdot \frac{1}{n} \left[(q/p)^n - (q/p-\varepsilon)^n \right] \\ &\approx \frac{p}{n} (q/p)^n \left[1 - \left(1 - \frac{\varepsilon p}{q} \right)^n \right] \end{aligned}$$

- Pour n grand : $\left(1 - \frac{\varepsilon p}{q}\right)^n \to 0$, donc : $I_2 \approx \frac{p}{n} \left(\frac{q}{p}\right)^n$

Loi géométrique – Convergence : conclusion

- Puisque $I_1/I \to 0$ et $I_2 \approx \frac{p}{n}(q/p)^n$:

$$\int_0^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v \approx \frac{p}{n}\left(\frac{q}{p}\right)^n \quad \text{pour } n \text{ grand}$$

- On a donc :

$$\mathbb{E}[\hat{p}] = n\left(\frac{p}{q}\right)^n \int_0^{q/p} \frac{v^{n-1}}{1+v}\,\mathrm{d}v \approx n\left(\frac{p}{q}\right)^n \cdot \frac{p}{n}\left(\frac{q}{p}\right)^n = p$$

et

$$\boxed{\text{Biais}(\hat{p}) = \mathbb{E}[\hat{p}] - p \xrightarrow[n\to\infty]{} 0}$$

L'estimateur est **asymptotiquement sans biais**.

Loi géométrique – Variance : approche exacte

- Variance exacte :

$$\mathbb{V}[\hat{p}] = \mathbb{E}[\hat{p}^2] - \mathbb{E}[\hat{p}]^2 = n^2\mathbb{E}\left[\frac{1}{S^2}\right] - n^2\mathbb{E}\left[\frac{1}{S}\right]^2$$

- On connaît $\mathbb{E}[1/S]$. Pour $\mathbb{E}[1/S^2]$, on utilise :

$$\frac{1}{k^2} = \int_0^1 \frac{-\log t}{1}\, t^{k-1}\,\mathrm{d}t$$

- Il faut alors calculer l'intégrale suivante :

$$\mathbb{E}\left[\frac{1}{S^2}\right] = p^n \int_0^1 \frac{(-\log t)\, t^{n-1}}{(1-tq)^n}\,\mathrm{d}t$$

- Cette intégrale n'a pas de forme simple, mais on peut montrer que pour n grand :

$$\mathbb{V}[\hat{p}] \sim \frac{p^2(1-p)}{n}$$

Loi géométrique – La méthode delta

- Problème : on connaît $\mathbb{E}[S]$ et $\mathbb{V}[S]$, mais on veut $\mathbb{V}[g(S)]$ pour $g(x) = n/x$.

- Méthode delta : si $\sqrt{n}(S_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$, alors pour g dérivable en μ :

$$\sqrt{n}(g(S_n) - g(\mu)) \xrightarrow{d} \mathcal{N}(0, [g'(\mu)]^2\sigma^2)$$

- Conséquence pratique pour n grand :

$$\boxed{\mathbb{V}[g(S)] \approx [g'(\mathbb{E}[S])]^2 \cdot \mathbb{V}[S]}$$

- Justification par développement de Taylor de g autour de $\mathbb{E}[S]$:

$$g(S) \approx g(\mathbb{E}[S]) + g'(\mathbb{E}[S])(S - \mathbb{E}[S])$$

d'où $\mathbb{V}[g(S)] \approx [g'(\mathbb{E}[S])]^2\mathbb{V}[S]$.

Loi géométrique – Variance : approximation par méthode delta

- On a $\hat{p} = n/S = g(S)$ avec $g(x) = n/x$.
- Application : $g'(x) = -n/x^2$, donc $g'(\mathbb{E}[S]) = -n/(n/p)^2 = -p^2/n$.

$$\mathbb{V}[\hat{p}] \approx \left(\frac{p^2}{n}\right)^2 \cdot \mathbb{V}[S] = \frac{p^4}{n^2} \cdot \frac{n(1-p)}{p^2}$$

$$\boxed{\mathbb{V}[\hat{p}] \approx \frac{p^2(1-p)}{n}}$$

- Cette approximation est asymptotique : elle devient exacte quand $n \to \infty$.

Loi géométrique – Information de Fisher

- Dérivée seconde :

$$\frac{\partial^2 \ell}{\partial p^2} = -\frac{n}{p^2} - \frac{S-n}{(1-p)^2}$$

- Sachant que $\mathbb{E}[S] = n\mathbb{E}[X] = n/p$:

$$I(p) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial p^2}\right] = \frac{n}{p^2} + \frac{n/p - n}{(1-p)^2} = \frac{n}{p^2} + \frac{n(1-p)}{p(1-p)^2}$$

$$\boxed{I(p) = \frac{n}{p^2(1-p)}}$$

- Borne de Cramér-Rao (cas sans biais) : $\mathbb{V}[\hat{p}] \geq \dfrac{1}{I(p)} = \dfrac{p^2(1-p)}{n}$

Loi géométrique – Efficacité : MSE vs borne de Cramér-Rao

- La variance peut être calculée avec de l'intégration numérique.
- On peut aussi calculer l'erreur quadratique moyenne de l'estimateur :

$$\mathsf{MSE}[\hat{p}] = \mathbb{V}[\hat{p}] + \mathsf{Biais}[\hat{p}]^2$$

- Exemple pour $p = 0.5$:

n	Biais^2	Variance	CR	MSE	CR/Var
1	0.037	0.102	0.125	0.139	123%
5	0.002	0.028	0.025	0.031	89%
10	0.001	0.014	0.0125	0.014	92%
50	0.000	0.003	0.0025	0.003	98%

- Le biais domine en petit échantillon, puis devient négligeable.

Loi géométrique – Efficacité selon p

- L'efficacité dépend fortement de p (ici pour $n = 5$) :

p	$\mathbb{E}[\hat{p}]$	Biais	Variance	CR	MSE	CR/Var
0.2	0.236	+0.036	0.011	0.0064	0.012	58%
0.5	0.549	+0.049	0.028	0.025	0.031	89%
0.8	0.828	+0.028	0.021	0.0256	0.022	122%

- Pour p petit : Variance $>$ CR (estimateur à variance élevée)
- Pour p grand : Variance $<$ CR (possible car l'estimateur est biaisé)
- Explication : pour p petit, les X_i sont grands et variables, amplifiant le biais de $\hat{p} = 1/\bar{X}$.

Loi géométrique – Efficacité asymptotique

- Résumé :
  - L'estimateur $\hat{p} = n/S$ est biaisé pour n fini
  - Le biais est positif : $\mathbb{E}[\hat{p}] > p$
  - Le biais tend vers 0 quand $n \to \infty$: estimateur asymptotiquement sans biais
- Quand $n \to \infty$:
$$\mathbb{V}[\hat{p}] \sim \frac{p^2(1-p)}{n} = \frac{1}{I(p)}$$
- L'EMV $\hat{p}$ est asymptotiquement efficace, comme attendu.

Loi géométrique – Un estimateur sans biais

- À partir de l'expression du biais, on peut construire un estimateur sans biais. Pour $n \geq 2$, considérons l'estimateur :

$$\boxed{\tilde{p} = \frac{n-1}{S-1}}$$

- On utilise l'identité : $\dfrac{1}{k-1}\binom{k-1}{n-1} = \dfrac{1}{n-1}\binom{k-2}{n-2}$
- Calcul de l'espérance :

$$\begin{aligned}
\mathbb{E}\left[\frac{n-1}{S-1}\right] &= (n-1)\sum_{k=n}^{\infty} \frac{1}{k-1}\binom{k-1}{n-1} p^n q^{k-n} \\
&= \sum_{k=n}^{\infty} \binom{k-2}{n-2} p^n q^{k-n} = p^n \sum_{j=0}^{\infty} \binom{j+n-2}{n-2} q^j \\
&= p^n \cdot \frac{1}{(1-q)^{n-1}} = p^n \cdot \frac{1}{p^{n-1}} = \boxed{p}
\end{aligned}$$

Loi géométrique – Variance de l'estimateur sans biais

- Pour $n = 2$, on a $\tilde{p} = \frac{1}{S-1}$ avec $S \sim \mathrm{NegBin}(2, p)$.

- En utilisant des techniques similaires aux calculs précédents :

$$\mathbb{E}\left[\frac{1}{(S-1)^2}\right] = \frac{p}{2}(-\log p) = -\frac{p \log p}{2}$$

- Variance exacte pour $n = 2$:

$$\mathbb{V}[\tilde{p}] = \mathbb{E}\left[\frac{1}{(S-1)^2}\right] - p^2 = -\frac{p \log p}{2} - p^2$$

Loi géométrique – La borne de Cramér-Rao est-elle atteignable ?

- Comparaison pour $n = 2$, $p = 0.5$:

Estimateur	Variance	MSE	Biais
$\tilde{p} = (n-1)/(S-1)$ (sans biais)	0.097	0.097	0
$\hat{p} = n/S$ (EMV)	0.067	0.080	0.114
Borne de Cramér-Rao	0.0625	–	–

- On a $\mathbb{V}[\tilde{p}] > \mathbb{V}[\hat{p}] > 0.0625 = \text{BCR}$: aucun estimateur n'est efficace à distance finie.

- Cependant, puisque :

$$\tilde{p} = \frac{n-1}{S-1} = \hat{p} \cdot \frac{n-1}{n}$$

L'EMV $\hat{p}$ étant asymptotiquement efficient, l'estimateur sans biais $\tilde{p}$ l'est aussi.

Plan

Loi Gamma – Présentation

- La loi Gamma généralise la loi exponentielle et modélise :
  - Temps d'attente jusqu'au α-ième événement (processus de Poisson)
  - Précipitations, débits de rivières (hydrologie)
  - Temps de service dans les files d'attente
  - Distribution des revenus, taille des sinistres (assurance)
- Cas particuliers :
  - $\alpha = 1$: loi exponentielle
  - $\alpha = n/2$, $\beta = 1/2$: loi du chi-deux à n degrés de liberté

Loi Gamma – Modèle (forme connue)

- $X_i \sim \mathrm{Gamma}(\alpha, \beta)$ avec $\alpha > 0$ connu et $\beta > 0$ inconnu :

$$f(x; \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0$$

- Moments :

$$\mathbb{E}[X] = \frac{\alpha}{\beta}, \qquad \mathbb{V}[X] = \frac{\alpha}{\beta^2}$$

- La somme $S = \sum_{i=1}^n X_i \sim \mathrm{Gamma}(n\alpha, \beta)$.

Loi Gamma – Log-vraisemblance

- Log-vraisemblance :

$$\ell(\beta) = n\alpha \log \beta - n \log \Gamma(\alpha) + (\alpha - 1) \sum_{i=1}^{n} \log x_i - \beta \sum_{i=1}^{n} x_i$$

Loi Gamma – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial \beta} = \frac{n\alpha}{\beta} - \sum_{i=1}^{n} x_i = 0$$

- Solution :

$$\boxed{\hat{\beta}_{MV} = \frac{n\alpha}{\sum_{i=1}^{n} X_i} = \frac{\alpha}{\bar{X}}}$$

- Si α était aussi inconnu, pas de forme explicite pour $\hat{\alpha}$.

Loi Gamma – Variance exacte de l'estimateur

- On a $\hat{\beta} = n\alpha/S$ où $S \sim \text{Gamma}(n\alpha, \beta)$.
- Pour $n\alpha > 2$:

$$\mathbb{E}[\hat{\beta}] = \frac{n\alpha\beta}{n\alpha - 1}, \quad \mathbb{E}[\hat{\beta}^2] = \frac{(n\alpha)^2\beta^2}{(n\alpha-1)(n\alpha-2)}$$

- Variance exacte :

$$\boxed{\mathbb{V}[\hat{\beta}] = \frac{(n\alpha)^2\beta^2}{(n\alpha-1)^2(n\alpha-2)}}$$

- Pour n grand : $\mathbb{V}[\hat{\beta}] \approx \dfrac{\beta^2}{n\alpha}$

Loi Gamma – Information de Fisher

- Dérivée seconde : $\dfrac{\partial^2 \ell}{\partial \beta^2} = -\dfrac{n\alpha}{\beta^2}$
- Information de Fisher :

$$\boxed{I(\beta) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \beta^2}\right] = \frac{n\alpha}{\beta^2}}$$

- Borne de Cramér-Rao : $\mathbb{V}[\hat{\beta}] \geq \dfrac{\beta^2}{n\alpha}$
- Plus α est grand, plus l'estimation est précise.

Plan

Loi log-normale – Présentation

- La loi log-normale modélise des phénomènes multiplicatifs :
  - Prix des actions, rendements financiers
  - Taille des particules, diamètre des gouttelettes
  - Distribution des revenus, des richesses
  - Durée de survie en médecine
  - Concentrations chimiques
- Caractéristique : X suit une loi log-normale $\Leftrightarrow \log X$ suit une loi normale.

Loi log-normale – Modèle

- $X_i \sim \text{LogNormal}(\mu, \sigma^2)$ si $\log X_i \sim \mathcal{N}(\mu, \sigma^2)$:

$$f(x; \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad x > 0$$

- Moments de X :

$$\mathbb{E}[X] = e^{\mu+\sigma^2/2}, \qquad \mathbb{V}[X] = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$$

Loi log-normale – Log-vraisemblance

- En posant $Y_i = \log X_i \sim \mathcal{N}(\mu, \sigma^2)$:

$$\ell(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \sum_{i=1}^{n}\log x_i$$
$$-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log x_i - \mu)^2$$

- C'est la log-vraisemblance d'un échantillon normal sur les $Y_i = \log X_i$.

Loi log-normale – Estimateurs MV

- Conditions du premier ordre :

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\log x_i - \mu) = 0, \qquad \frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(\log x_i - \mu)^2 = 0$$

- Solutions :

$$\boxed{\hat{\mu}_{MV} = \frac{1}{n}\sum_{i=1}^{n}\log X_i = \overline{\log X}}, \qquad \boxed{\hat{\sigma}^2_{MV} = \frac{1}{n}\sum_{i=1}^{n}(\log X_i - \hat{\mu})^2}$$

Loi log-normale – Variance exacte des estimateurs

- Les $Y_i = \log X_i$ sont i.i.d. $\mathcal{N}(\mu, \sigma^2)$.
- Pour $\hat{\mu}$: $\hat{\mu} = \bar{Y} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$, donc $\boxed{\mathbb{V}[\hat{\mu}] = \frac{\sigma^2}{n}}$
- Pour $\hat{\sigma}^2$: $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1}$, donc $\boxed{\mathbb{V}[\hat{\sigma}^2] = \frac{2(n-1)\sigma^4}{n^2}}$

Loi log-normale – Matrice d'information de Fisher

- Dérivées secondes :

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{n}{\sigma^2}, \quad \frac{\partial^2 \ell}{\partial \mu \partial \sigma^2} = -\frac{1}{\sigma^4} \sum_{i=1}^{n} (\log x_i - \mu)$$

- Matrice d'information de Fisher :

$$\boxed{I(\mu, \sigma^2) = n \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}}$$

- Les paramètres sont orthogonaux : $\hat{\mu}$ et $\hat{\sigma}^2$ sont asymptotiquement indépendants.

Plan

Loi de Pareto – Présentation

- La loi de Pareto modélise des phénomènes avec des valeurs extrêmes fréquentes :
  - Distribution des revenus et des richesses (loi des 80-20)
  - Taille des villes, des entreprises
  - Popularité des sites web, des mots dans un texte
  - Montants des sinistres en assurance
  - Magnitude des tremblements de terre
- Propriété : décroissance en loi de puissance (queue épaisse) : $P(X > x) \propto x^{-\alpha}$

Loi de Pareto – Modèle

- $X_i \sim \text{Pareto}(\alpha, x_m)$ avec $x_m > 0$ connu et $\alpha > 0$ inconnu :

$$f(x;\alpha) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}, \quad x \geq x_m$$

- Moments (pour $\alpha > 2$) :

$$\mathbb{E}[X] = \frac{\alpha x_m}{\alpha - 1}, \qquad \mathbb{V}[X] = \frac{\alpha x_m^2}{(\alpha-1)^2(\alpha-2)}$$

- Attention : la variance n'existe que si $\alpha > 2$.

Loi de Pareto – Log-vraisemblance

- Log-vraisemblance :

$$\ell(\alpha) = n \log \alpha + n\alpha \log x_m - (\alpha + 1) \sum_{i=1}^{n} \log x_i$$

Loi de Pareto – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + n \log x_m - \sum_{i=1}^{n} \log x_i = 0$$

- Solution :

$$\boxed{\hat{\alpha}_{MV} = \frac{n}{\sum_{i=1}^{n} \log X_i - n \log x_m} = \frac{n}{\sum_{i=1}^{n} \log(X_i / x_m)}}$$

C'est l'inverse de la moyenne des log-rapports.

Loi de Pareto – Variance exacte de l'estimateur

- Posons $Y_i = \log(X_i/x_m)$. Si $X_i \sim \text{Pareto}(\alpha, x_m)$, alors $Y_i \sim \text{Exp}(\alpha)$.
- Donc $S = \sum_{i=1}^{n} Y_i \sim \text{Gamma}(n, \alpha)$ et $\hat{\alpha} = n/S$.
- Pour $n > 2$: $\mathbb{E}[\hat{\alpha}] = \dfrac{n\alpha}{n-1}$

$$\boxed{\mathbb{V}[\hat{\alpha}] = \frac{n^2\alpha^2}{(n-1)^2(n-2)}}$$

- Asymptotiquement : $\mathbb{V}[\hat{\alpha}] \approx \dfrac{\alpha^2}{n}$

Loi de Pareto – Information de Fisher

- Dérivée seconde : $\dfrac{\partial^2 \ell}{\partial \alpha^2} = -\dfrac{n}{\alpha^2}$ (ne dépend pas des observations !)
- Information de Fisher : $\boxed{I(\alpha) = \dfrac{n}{\alpha^2}}$
- Borne de Cramér-Rao : $\mathbb{V}[\hat{\alpha}] \geq \dfrac{\alpha^2}{n}$
- Même forme que pour la loi exponentielle (lien via la transformation log).

Plan

Loi binomiale négative – Présentation

- La loi binomiale négative modélise le nombre d'échecs avant r succès :
  - Nombre de lancers ratés avant de gagner r fois
  - Nombre de clients non intéressés avant r ventes
  - Données de comptage avec surdispersion (alternative à Poisson)
  - Nombre d'accidents par conducteur (hétérogénéité)
- Relations :
  - $r = 1$: loi géométrique
  - Mélange Poisson-Gamma : modélise la surdispersion

Loi binomiale négative – Modèle

- $X_i \sim \text{NegBin}(r, p)$ avec $r > 0$ connu et $p \in (0, 1)$ inconnu :

$$P(X = k) = \binom{k + r - 1}{k} (1 - p)^k p^r, \quad k = 0, 1, 2, \ldots$$

(X = nombre d'échecs avant r succès)

- Moments :

$$\mathbb{E}[X] = \frac{r(1 - p)}{p}, \qquad \mathbb{V}[X] = \frac{r(1 - p)}{p^2}$$

Loi binomiale négative – Log-vraisemblance

- Log-vraisemblance (avec $S = \sum_{i=1}^{n} x_i$) :

$$\ell(p) = \text{const} + S \log(1 - p) + nr \log p$$

Loi binomiale négative – Estimateur MV

- Condition du premier ordre :

$$\frac{\partial \ell}{\partial p} = -\frac{S}{1-p} + \frac{nr}{p} = 0 \quad \Rightarrow \quad nrp = S(1-p)$$

- Solution :

$$\boxed{\hat{p}_{MV} = \frac{nr}{nr+S} = \frac{r}{r+\bar{X}}}$$

Loi binomiale négative – Variance de l'estimateur

- Méthode delta avec $g(x) = r/(r+x)$ et $g'(x) = -r/(r+x)^2$.
- Avec $\mathbb{E}[\bar{X}] = r(1-p)/p$: $r + \mathbb{E}[\bar{X}] = r/p$, donc $g'(\mathbb{E}[\bar{X}]) = -p^2/r$.
- Variance approchée :

$$\mathbb{V}[\hat{p}] \approx \frac{p^4}{r^2} \cdot \frac{r(1-p)}{np^2} = \boxed{\frac{p^2(1-p)}{nr}}$$

Loi binomiale négative – Information de Fisher

- Dérivée seconde : $\dfrac{\partial^2 \ell}{\partial p^2} = -\dfrac{S}{(1-p)^2} - \dfrac{nr}{p^2}$
- Espérance (avec $\mathbb{E}[S] = nr(1-p)/p$) :

$$I(p) = \frac{nr(1-p)/p}{(1-p)^2} + \frac{nr}{p^2} = \boxed{\frac{nr}{p^2(1-p)}}$$

- Borne de Cramér-Rao : $\mathbb{V}[\hat{p}] \geq \dfrac{p^2(1-p)}{nr}$ ✓

Plan

Loi de Cauchy – Présentation

- La loi de Cauchy apparaît dans plusieurs contextes :
  - Rapport de deux variables normales indépendantes centrées
  - Distribution de certains phénomènes physiques (résonance)
  - Modélisation de données avec valeurs extrêmes fréquentes
  - Finance : modèles à queues épaisses
- Exemple : si $X, Y \overset{iid}{\sim} \mathcal{N}(0,1)$, alors $X/Y \sim \text{Cauchy}(0,1)$.

Loi de Cauchy – Modèle

- $X_i \sim \text{Cauchy}(\mu, \sigma)$ avec $\mu \in \mathbb{R}$ (position) et $\sigma > 0$ (échelle) :

$$f(x; \mu, \sigma) = \frac{1}{\pi\sigma\left(1 + \left(\frac{x-\mu}{\sigma}\right)^2\right)}$$

- Fonction de répartition :

$$F(x; \mu, \sigma) = \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\sigma}\right) + \frac{1}{2}$$

- Propriétés de symétrie :
  - μ est la médiane et le mode
  - Distribution symétrique autour de μ

Loi de Cauchy – Absence de moments

Propriété fondamentale

La loi de Cauchy n'a ni espérance ni variance !

- Pour que $\mathbb{E}[X]$ existe, il faut que l'intégrale soit absolument convergente, i.e. $\mathbb{E}[|X|] < \infty$.
- Vérification pour $\mu = 0$, $\sigma = 1$:

$$\mathbb{E}[|X|] = \int_{-\infty}^{\infty} \frac{|x|}{\pi(1+x^2)}\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\infty} \frac{x}{1+x^2}\,\mathrm{d}x = \frac{2}{\pi}\left[\frac{1}{2}\log(1+x^2)\right]_0^{\infty} = +\infty$$

- L'intégrale diverge logarithmiquement : $\mathbb{E}[X]$ n'existe pas.

Loi de Cauchy – Conséquences sur les estimateurs classiques

- Si $X_1, \ldots, X_n \overset{iid}{\sim}$ Cauchy(μ, σ) :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \sim \text{Cauchy}(\mu, \sigma)$$

La moyenne suit la même loi quelle que soit la taille de l'échantillon !

  - Fonction caractéristique d'une v.a. de Cauchy $\varphi(t) = e^{i\mu t - \sigma |t|}$
  - Pour la somme : $\varphi_{\sum X_i}(t) = e^{in\mu t - n\sigma |t|}$
  - Pour la moyenne : $\varphi_{\bar{X}}(t) = \varphi_{\sum X_i}(t/n) = e^{i\mu t - \sigma |t|}$

Loi de Cauchy – Pourquoi l'EMV fonctionne quand même

- Les conditions de Cramér-Rao sont satisfaites
- Ces conditions ne requièrent pas l'existence des moments de X !
- L'information de Fisher fait intervenir $\mathbb{E}\left[\left(\frac{\partial \log f}{\partial \theta}\right)^2\right]$, pas $\mathbb{E}[X]$.

Loi de Cauchy – Log-vraisemblance

- Cas général (μ et σ inconnus) :

$$\ell(\mu, \sigma) = -n \log \pi - n \log \sigma - \sum_{i=1}^{n} \log \left(1 + \left(\frac{x_i - \mu}{\sigma} \right)^2 \right)$$

- Cas simplifié ($\sigma = 1$ connu, estimation de μ seul) :

$$\ell(\mu) = -n \log \pi - \sum_{i=1}^{n} \log \left(1 + (x_i - \mu)^2 \right)$$

Loi de Cauchy – Équations du maximum de vraisemblance

- Dérivée par rapport à μ (avec $\sigma = 1$) :

$$\frac{\partial \ell}{\partial \mu} = \sum_{i=1}^{n} \frac{2(x_i - \mu)}{1 + (x_i - \mu)^2} = 0$$

- En posant $u_i = x_i - \mu$, on cherche μ tel que :

$$\sum_{i=1}^{n} \frac{u_i}{1 + u_i^2} = 0$$

- C'est une somme pondérée où les observations extrêmes reçoivent un poids faible (la fonction $u \mapsto \frac{u}{1+u^2}$ est bornée par $\pm 1/2$).

Loi de Cauchy – Robustesse de l'EMV

- La fonction $\psi(u) = \frac{2u}{1+u^2}$ qui apparaît dans l'équation de l'EMV :
  - Est bornée : $|\psi(u)| \leq 1$
  - Tend vers 0 quand $|u| \to \infty$
- L'EMV de μ est naturellement robuste aux valeurs extrêmes, contrairement à la moyenne empirique.
- Une observation très éloignée a une contribution bornée à l'équation d'estimation.

Loi de Cauchy – Résolution numérique

- L'équation $\sum_{i=1}^{n} \frac{x_i-\mu}{1+(x_i-\mu)^2} = 0$ n'a pas de forme fermée.
- Méthodes de résolution :
  - Newton-Raphson avec $\frac{\partial^2 \ell}{\partial \mu^2} = \sum_{i=1}^{n} \frac{2[(x_i-\mu)^2-1]}{[1+(x_i-\mu)^2]^2}$
  - Point fixe : $\mu^{(k+1)} = \frac{\sum_i w_i(\mu^{(k)})x_i}{\sum_i w_i(\mu^{(k)})}$ avec $w_i(\mu) = \frac{1}{1+(x_i-\mu)^2}$
  - Optimisation directe de $\ell(\mu)$
- La médiane empirique est un bon point de départ.

Loi de Cauchy – Information de Fisher pour μ

- Score (avec $\sigma = 1$) :
$$\frac{\partial \log f}{\partial \mu} = \frac{2(x-\mu)}{1+(x-\mu)^2}$$
- Information de Fisher pour une observation :
$$I_1(\mu) = \mathbb{E}\left[\left(\frac{\partial \log f}{\partial \mu}\right)^2\right] = \int_{-\infty}^{\infty} \frac{4(x-\mu)^2}{[1+(x-\mu)^2]^2} \cdot \frac{1}{\pi[1+(x-\mu)^2]} \, \mathrm{d}x$$
- En posant $u = x - \mu$: $I_1(\mu) = \frac{4}{\pi} \int_{-\infty}^{\infty} \frac{u^2}{(1+u^2)^3} \, \mathrm{d}u$

Loi de Cauchy – Calcul de l'information de Fisher

- On utilise $\int_{-\infty}^{\infty} \frac{u^2}{(1+u^2)^3}\,\mathrm{d}u = \frac{\pi}{8}$ (par résidus, ou par parties).
- Résultat : $I_1(\mu) = \frac{4}{\pi} \cdot \frac{\pi}{8} = \frac{1}{2}$
- Information de Fisher pour n observations :

$$\boxed{I_n(\mu) = \frac{n}{2}}$$

Loi de Cauchy – Variance asymptotique de l'EMV

- Borne de Cramér-Rao : $\mathbb{V}[\hat{\mu}_{MV}] \geq \frac{1}{I_n(\mu)} = \frac{2}{n}$
- L'EMV est asymptotiquement efficace :

$$\sqrt{n}(\hat{\mu}_{MV} - \mu) \xrightarrow{d} \mathcal{N}(0, 2)$$

- Variance asymptotique :

$$\boxed{\mathbb{V}[\hat{\mu}_{MV}] \approx \frac{2}{n}}$$

- Malgré l'absence de moments de X, l'EMV a une variance qui décroît en $1/n$!

Loi de Cauchy – Cas général avec σ inconnu

- Équations de vraisemblance :

$$\frac{\partial \ell}{\partial \mu} = \frac{2}{\sigma^2} \sum_{i=1}^{n} \frac{x_i - \mu}{1 + \left(\frac{x_i-\mu}{\sigma}\right)^2} = 0$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{2}{\sigma^3} \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{1 + \left(\frac{x_i-\mu}{\sigma}\right)^2} = 0$$

- Système non linéaire à résoudre numériquement (Newton-Raphson multidimensionnel).

Loi de Cauchy – Matrice d'information de Fisher

- Pour le modèle complet (μ, σ) :

$$I_1(\mu, \sigma) = \frac{1}{2\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

  Les paramètres sont orthogonaux.

- Pour n observations :

$$\boxed{I_n(\mu, \sigma) = \frac{n}{2\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}$$

- Variances asymptotiques :

$$\mathbb{V}[\hat{\mu}_{MV}] \approx \frac{2\sigma^2}{n}, \qquad \mathbb{V}[\hat{\sigma}_{MV}] \approx \frac{2\sigma^2}{n}$$

Loi de Cauchy – Comparaison des estimateurs de μ

- La médiane $\tilde{\mu}$ est un estimateur consistant avec :

$$\mathbb{V}[\tilde{\mu}] \approx \frac{1}{4n[f(\mu)]^2} = \frac{\pi^2\sigma^2}{4n}$$

Pour $\sigma = 1$: $\mathbb{V}[\tilde{\mu}] \approx \dfrac{2.47}{n}$

- Tableau comparatif (pour $\sigma = 1$) :

Estimateur	Consistant ?	Var. asympt.	Efficacité
Moyenne $\bar{X}$	Non	$\nexists$	–
Médiane $\tilde{\mu}$	Oui	$\pi^2/(4n) \approx 2.47/n$	81%
EMV $\hat{\mu}_{MV}$	Oui	$2/n$	100%

Loi de Cauchy – Efficacité relative

- Rapport des variances :

$$e(\tilde{\mu}, \hat{\mu}_{MV}) = \frac{\mathbb{V}[\hat{\mu}_{MV}]}{\mathbb{V}[\tilde{\mu}]} = \frac{2/n}{\pi^2/(4n)} = \frac{8}{\pi^2} \approx 0.81$$

- Pour atteindre la même précision que l'EMV avec n observations, la médiane nécessite environ $n/0.81 \approx 1.23n$ observations.

- Compromis pratique :
  - La médiane est simple à calculer et robuste
  - L'EMV est plus efficace mais nécessite une optimisation numérique
  - La médiane est un excellent point de départ pour l'algorithme de l'EMV

Loi de Cauchy – Résumé

- La loi de Cauchy n'a pas de moments (espérance, variance).
- La moyenne empirique ne converge pas (suit toujours une loi de Cauchy).
- Malgré cela, l'EMV est consistant et asymptotiquement normal.
- L'information de Fisher est bien définie : $I_n(\mu) = n/2$ (pour $\sigma = 1$).
- Variance asymptotique de l'EMV : $2\sigma^2/n$.
- L'EMV est plus efficace que la médiane (rapport $\approx 81\%$).
- L'EMV est naturellement robuste aux valeurs extrêmes.

Plan

Loi uniforme – Présentation

- La loi uniforme modélise l'incertitude totale dans un intervalle :
  - Erreurs d'arrondi
  - Position d'un point sur un segment
  - Génération de nombres pseudo-aléatoires
  - Modèle de base en théorie des probabilités

Attention : cas non-régulier !

La loi uniforme est un exemple où les conditions de régularité de Cramér-Rao ne sont pas satisfaites.

Loi uniforme – Modèle

- $X_i \sim U(0, \theta)$ avec $\theta > 0$ inconnu :

$$f(x;\theta) = \begin{cases} \frac{1}{\theta} & \text{si } 0 \leq x \leq \theta \\ 0 & \text{sinon} \end{cases}$$

- Moments : $\mathbb{E}[X] = \frac{\theta}{2}$, $\mathbb{V}[X] = \frac{\theta^2}{12}$

Le problème

Le support $[0, \theta]$ dépend du paramètre θ !

Loi uniforme – Vraisemblance

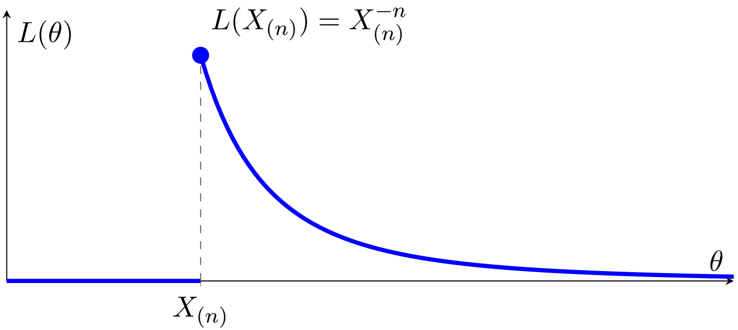
- Vraisemblance :

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \begin{cases} \frac{1}{\theta^n} & \text{si } \theta \geq \max_i x_i \\ 0 & \text{sinon} \end{cases}$$

  - $L(\theta) = 0$ si $\theta < X_{(n)} = \max(X_1, \dots, X_n)$
  - $L(\theta) = \theta^{-n}$ est décroissante pour $\theta \geq X_{(n)}$
  - Maximum atteint en $\theta = X_{(n)}$

- Estimateur MV : $\boxed{\hat{\theta}_{MV} = X_{(n)} = \max(X_1, \dots, X_n)}$

Loi uniforme – Vraisemblance (représentation graphique)



- La vraisemblance est nulle pour $\theta < X_{(n)}$, puis décroît en θ^{-n}.
- Le maximum est atteint en $\theta = X_{(n)}$ (discontinuité).

Loi uniforme – Pourquoi c'est un cas singulier ?

- Les conditions de Cramér-Rao supposent que :
  1. Le support de $f(x;\theta)$ ne dépend pas de θ VIOLÉ
  2. On peut intervertir dérivation et intégration
  3. La log-vraisemblance est deux fois dérivable en θ
- Conséquences :
  - L'information de Fisher classique ne s'applique pas
  - La borne de Cramér-Rao n'est pas valide
  - L'estimateur peut converger plus vite que $1/\sqrt{n}$!

Loi uniforme – Distribution de l'estimateur

- Fonction de répartition du maximum :

$$F_{X_{(n)}}(x) = P(X_{(n)} \leq x) = P(\text{tous les } X_i \leq x) = \prod_{i=1}^{n} P(X_i \leq x)$$

Pour $0 \leq x \leq \theta$: $F_{X_{(n)}}(x) = \left(\frac{x}{\theta}\right)^n$

- Densité du maximum :

$$\boxed{f_{X_{(n)}}(x) = \frac{nx^{n-1}}{\theta^n}, \quad 0 \leq x \leq \theta}$$

- C'est une loi Beta$(n, 1)$ mise à l'échelle sur $[0, \theta]$.

Loi uniforme – Calcul de l'espérance

- Calcul direct :

$$\mathbb{E}[X_{(n)}] = \int_0^\theta x \cdot \frac{nx^{n-1}}{\theta^n}\,\mathrm{d}x = \frac{n}{\theta^n}\int_0^\theta x^n\,\mathrm{d}x = \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1}$$

$$\boxed{\mathbb{E}[X_{(n)}] = \frac{n}{n+1}\theta}$$

- Biais : $\mathrm{Biais}(\hat{\theta}) = \mathbb{E}[X_{(n)}] - \theta = -\frac{\theta}{n+1}$
- L'estimateur MV est biaisé (il sous-estime systématiquement θ).

Loi uniforme – Estimateur sans biais

- Puisque $\mathbb{E}[X_{(n)}] = \dfrac{n}{n+1}\theta$, on peut définir :

$$\boxed{\tilde{\theta} = \frac{n+1}{n} X_{(n)}}$$

  Alors $\mathbb{E}[\tilde{\theta}] = \theta$: c'est un estimateur sans biais.

- Autre estimateur sans biais : la moyenne $\bar{X}$ vérifie $\mathbb{E}[\bar{X}] = \theta/2$, donc $\tilde{\theta}_2 = 2\bar{X}$ est aussi sans biais, mais il est moins efficace que $\tilde{\theta}$.

Loi uniforme – Calcul de la variance

- Moment d'ordre 2 :

$$\mathbb{E}[X_{(n)}^2] = \int_0^\theta x^2 \cdot \frac{nx^{n-1}}{\theta^n}\,\mathrm{d}x = \frac{n}{\theta^n} \cdot \frac{\theta^{n+2}}{n+2} = \frac{n}{n+2}\theta^2$$

- Variance :

$$\mathbb{V}[X_{(n)}] = \mathbb{E}[X_{(n)}^2] - \mathbb{E}[X_{(n)}]^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\right)^2 \theta^2$$

$$\boxed{\mathbb{V}[X_{(n)}] = \frac{n\theta^2}{(n+1)^2(n+2)}}$$

Loi uniforme – Convergence exceptionnelle

- Développement asymptotique :

$$\mathbb{V}[X_{(n)}] = \frac{n\theta^2}{(n+1)^2(n+2)} \sim \frac{\theta^2}{n^2} \quad \text{quand } n \to \infty$$

- Comparaison des vitesses de convergence :

Cas	Variance	Vitesse
Cas régulier (Cramér-Rao)	$\approx c/n$	$O(1/n)$
Loi uniforme	$\approx \theta^2/n^2$	$O(1/n^2)$

Résultat remarquable

L'estimateur du maximum converge deux fois plus vite qu'habituellement !

Loi uniforme – Comparaison des estimateurs

- Variance de $\tilde{\theta} = \frac{n+1}{n} X_{(n)}$:

$$\mathbb{V}[\tilde{\theta}] = \left(\frac{n+1}{n}\right)^2 \mathbb{V}[X_{(n)}] = \frac{\theta^2}{n(n+2)}$$

- Variance de $\tilde{\theta}_2 = 2\bar{X}$:

$$\mathbb{V}[\tilde{\theta}_2] = 4 \cdot \mathbb{V}[\bar{X}] = 4 \cdot \frac{\theta^2/12}{n} = \frac{\theta^2}{3n}$$

- Rapport d'efficacité :

$$\frac{\mathbb{V}[\tilde{\theta}]}{\mathbb{V}[\tilde{\theta}_2]} = \frac{\theta^2/[n(n+2)]}{\theta^2/(3n)} = \frac{3}{n+2} \to 0$$

L'estimateur basé sur le maximum est beaucoup plus efficace.

Loi uniforme – Distribution asymptotique

- Posons $Y_n = n(\theta - X_{(n)})$. Pour $y \geq 0$:

$$P(Y_n > y) = P\left(X_{(n)} < \theta - \frac{y}{n}\right) = \left(1 - \frac{y}{n\theta}\right)^n$$

- Quand $n \to \infty$: $\left(1 - \frac{y}{n\theta}\right)^n \to e^{-y/\theta}$

Distribution asymptotique

$$\boxed{n(\theta - X_{(n)}) \xrightarrow{d} \text{Exp}(1/\theta)}$$

Convergence vers une loi exponentielle, pas une loi normale !

Loi uniforme – Résumé des originalités

1. Estimateur MV : $\hat{\theta} = X_{(n)} = \max(X_1, \ldots, X_n)$
2. Distribution exacte : Beta$(n, 1)$ sur $[0, \theta]$
3. Biais : $-\theta/(n+1)$ (sous-estimation systématique)
4. Variance exacte : $\dfrac{n\theta^2}{(n+1)^2(n+2)} = O(1/n^2)$
5. Distribution asymptotique : exponentielle, pas normale
6. Information de Fisher : non applicable (conditions violées)

Plan

Tableau récapitulatif (1/2)

Loi	Param.	EMV	Variance
Normale	μ	$\bar{X}$	$\dfrac{\sigma^2}{n}$ (exacte, efficace)
Normale	σ^2	$\frac{1}{n}\sum(X_i - \bar{X})^2$	$\dfrac{2(n-1)\sigma^4}{n^2}$
Bernoulli	p	$\bar{X}$	$\dfrac{p(1-p)}{n}$ (exacte, efficace)
Poisson	λ	$\bar{X}$	$\dfrac{\lambda}{n}$ (exacte, efficace)
Exponentielle	λ	$1/\bar{X}$	$\dfrac{n^2\lambda^2}{(n-1)^2(n-2)}$
Géométrique	p	$1/\bar{X}$	$\approx \dfrac{p^2(1-p)}{n}$

Tableau récapitulatif (2/2)

Loi	Param.	EMV	Variance
Gamma (α connu)	β	$\alpha/\bar{X}$	$\dfrac{(n\alpha)^2\beta^2}{(n\alpha-1)^2(n\alpha-2)}$
Log-normale	μ	$\overline{\log X}$	$\dfrac{\sigma^2}{n}$ (exacte)
Pareto (x_m connu)	α	$n/\sum \log(X_i/x_m)$	$\dfrac{n^2\alpha^2}{(n-1)^2(n-2)}$
Binomiale nég.	p	$r/(r+\bar{X})$	$\approx \dfrac{p^2(1-p)}{nr}$
Cauchy	μ	numérique	$\approx \dfrac{2\sigma^2}{n}$
Uniforme	θ	$X_{(n)}$	$\dfrac{n\theta^2}{(n+1)^2(n+2)}$

Légende : Bleu = pas de moments, EMV fonctionne ; Rouge = conditions violées, $O(1/n^2)$

Conclusion

- L'EMV est généralement asymptotiquement efficace (atteint la borne de Cramér-Rao).
- La matrice d'information de Fisher permet de quantifier la précision de l'estimation.
- Les conditions de régularité sont essentielles : leur violation (loi uniforme) peut conduire à des comportements atypiques.
- L'absence de moments (loi de Cauchy) n'empêche pas l'EMV de fonctionner si les conditions de régularité sont satisfaites.
- La distribution asymptotique est généralement normale, sauf dans les cas non-réguliers (loi uniforme $\to$ exponentielle).
- Quand l'estimateur est une fonction non linéaire de statistiques simples, la méthode delta permet d'approximer la variance.