

Retour sur le maximum de vraisemblance

- Soient x_1, \dots, x_n des données supposées être une réalisation d'un échantillon aléatoire $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} f(x; \theta)$ (densité / masse). La **vraisemblance** pour θ est la fonction

$$L(\theta) = f(x_1; \theta) \times f(x_2; \theta) \times \dots \times f(x_n; \theta) = \prod_{i=1}^n f(x_i; \theta).$$

- $\hat{\theta}_{\text{ML}}$ satisfait

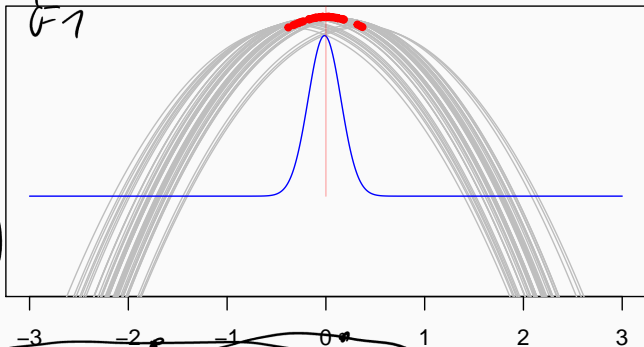
$$L(\hat{\theta}_{\text{ML}}) \geq L(\theta) \quad \text{pour tout } \theta$$

- Interprétation** : dans le cas discret, on maximise $\Pr_{\theta}(X_1 = x_1, \dots, X_n = x_n)$
- Il est plus facile de maximiser $\ell(\theta) := \log L(\theta)$, souvent en résolvant $d\ell(\theta)/d\theta = 0$, et vérifiant qu'il s'agit bien d'un maximum

La vraisemblance est une fonction aléatoire

$$l(\theta) = \sum_{i=1}^{40} \log f(x_i, \theta)$$

$$X_i \sim \mathcal{N}(\theta, 1)$$



Fonctions de vraisemblances correspondantes à 50 échantillons de taille $n = 40$. La vraie valeur est $\theta = 0$ et les 50 maxima sont en rouge.

Distribution asymptotique/approximative de $\hat{\theta}_{ML}$

- Dans des modèles "réguliers" (normal, exponentiel, Poisson, ...) on a $\ell'(\hat{\theta}_{ML}) = 0$
- Comme $\ell(\theta)$ est une somme de variables aléatoires iid, le théorème central limite s'applique
- Grâce à la méthode delta $\hat{\theta}_{ML} \stackrel{\text{app}}{\sim} \mathcal{N}(\theta, 1/I_n(\theta))$ avec **l'information de Fisher**

$$g(y) = (\ell')^{-1}(y)$$

$$I_n(\theta) = -\mathbb{E}_\theta(\ell''(\theta)) = \mathbb{E}_\theta(J_n(\theta)), \quad J_n(\theta) = -\ell''(\theta)$$

$1/I_n(\theta)$ est la **variance asymptotique** de $\hat{\theta}_{ML}$, $I_n(\theta)$ est la courbure

- **L'information observée** est $J_n(\hat{\theta}_{ML})$
- **Attention !** Certains modèles, tel que $U[0, \theta]$, ne sont pas réguliers

Exemple x_1, \dots, x_n réalisations d'une loi $\exp(\lambda)$ avec densité $\lambda e^{-\lambda x} I(x \geq 0)$, $\lambda > 0$

Exemple

Exemple x_1, \dots, x_n réalisations d'une loi $\text{exp}(\lambda)$ avec densité $\lambda e^{-\lambda x} I(x \geq 0)$, $\lambda > 0$

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

information
observée

$$L(\lambda) = n \log \lambda - \lambda n \bar{x}_n$$

$$L'(\lambda) = \frac{n}{\lambda} - n \bar{x}_n, \quad L'(\lambda) = 0 \Leftrightarrow \lambda = \frac{1}{\bar{x}_n}$$

$$L''(\lambda) = -\frac{n}{\lambda^2} < 0 \Rightarrow \hat{\lambda}_{ML} = \frac{1}{\bar{x}_n}$$

information
de Fisher

$$I_n(\lambda) = \mathbb{E}_\lambda \left[\frac{n}{\lambda^2} \right] = \frac{n}{\lambda^2}$$

$$J_n(\lambda) = -L''(\lambda) = \frac{n}{\lambda^2}$$

$\hat{\lambda}_{ML} \overset{\text{app}}{\sim} N\left(\lambda, \frac{\lambda^2}{n}\right)$

$$J_n(\hat{\lambda}_{ML}) = \frac{n}{\hat{\lambda}_{ML}^2} = n \bar{x}_n^2$$

3.2 Estimation par intervalle

Les intervalles de confiance

Un élément clé de la statistique est de donner une idée de l'incertitude d'un constat

Soit θ un paramètre inconnu, et soit $\tilde{\theta}_n = 1$ une estimation de θ basée sur y_1, \dots, y_n :

- si $n = 10^5$ on est beaucoup plus sûr que $\theta \approx \tilde{\theta}_n$ que si $n = 10$
- pour exprimer ceci on aimerait donner un intervalle qui serait plus large quand $n = 10$ que quand $n = 10^5$, pour expliciter l'incertitude liée à $\tilde{\theta}_n$

Définition: Soient $Y \equiv Y_1, \dots, Y_n$ des données issues d'une loi F de paramètre $\theta \in \mathbb{R}$.

Un **intervalle de confiance** (IC, 'confidence interval' en anglais) (L_n, U_n) pour θ est une statistique sous forme d'intervalle qui contient θ avec une probabilité spécifiée $1 - \alpha$

$$\theta \in [L_n, U_n]$$

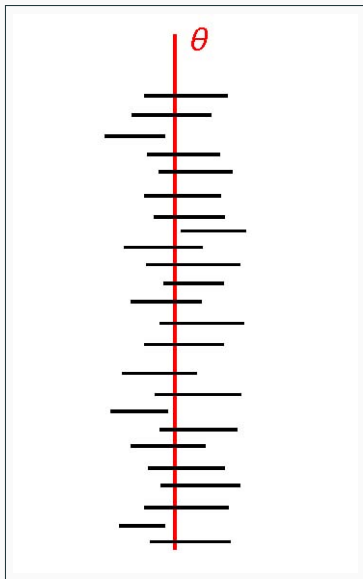
$$\Pr_{\theta}(L_n \leq \theta \leq U_n) = 1 - \alpha \quad \forall \theta$$

souvent
 $\alpha = 0,05$

- $1 - \alpha \in (0, 1)$ est le **niveau**, souvent $\alpha \in \{0.05, 0.01, 0.1\}$
- Les bornes $L_n = L_n(Y)$, $U_n = U_n(Y)$ sont des statistiques et non pas des inconnus
- Un IC **bilatéral**, de la forme (L_n, U_n) , est le plus souvent utilisé
- Un IC **unilatéral à droite** est $(-\infty, U_n]$ tel que $\Pr_{\theta}(U_n \geq \theta) = 1 - \alpha$
- Un IC **unilatéral à gauche** est $[L_n, \infty)$ tel que $\Pr_{\theta}(L_n \leq \theta) = 1 - \alpha$

Interprétation d'un intervalle de confiance

- (L_n, U_n) est un intervalle aléatoire qui contient θ avec probabilité $(1 - \alpha)$
- Si on répète l'expérience avec d'autres données, on aura un autre intervalle de confiance
- Nous ne savons pas si notre IC contient θ , mais cette procédure nous fournit une garantie statistique que cet événement a une probabilité $(1 - \alpha)$



Construire un IC dans le cas normal

- Soient $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$, et

$$\bar{Y}_n = \frac{1}{n} \sum_{j=1}^n Y_j$$

la moyenne de l'échantillon. On a $\bar{Y}_n \sim \mathcal{N}(\mu, 1/n)$ et donc

$$Z_n = n^{1/2}(\bar{Y}_n - \mu) \sim \mathcal{N}(0, 1)$$

a une distribution qui ne dépend pas de μ . On obtient

$$\Pr_{\mu} \left(\bar{Y}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \leq \mu \leq \bar{Y}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \right) = 1 - \alpha$$

avec $z_{\beta} = \Phi^{-1}(\beta)$ les quantiles de la loi $\mathcal{N}(0, 1)$

- Si $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, et **on connaît** σ^2 , alors

$$\frac{Z_n}{\sigma} = \frac{n^{1/2}(\bar{Y}_n - \mu)}{\sigma} \sim \mathcal{N}(0, 1)$$

et l'intervalle de confiance pour μ est

$$\left[\bar{Y}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}}\sigma, \bar{Y}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}}\sigma \right]$$

- La largeur \downarrow avec n , \uparrow avec σ et avec le niveau $(1 - \alpha)$, car $\Pr_{\theta}(L_n \leq \theta \leq U_n) \uparrow$

Construire un IC dans le cas normal

- Soient $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, 1)$, et

$$\bar{Y}_n = \frac{1}{n} \sum_{j=1}^n Y_j = \Pr(-b \leq -Z \leq a)$$

la moyenne de l'échantillon. On a $\bar{Y}_n \sim \mathcal{N}(\mu, 1/n)$ et donc

$$Z_n = n^{1/2}(\bar{Y}_n - \mu) \sim \mathcal{N}(0, 1)$$

a une distribution qui ne dépend pas de μ . On obtient

$$\Pr_{\mu} \left(\bar{Y}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \leq \mu \leq \bar{Y}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \right) = 1 - \alpha$$

avec $z_{\beta} = \Phi^{-1}(\beta)$ les quantiles de la loi $\mathcal{N}(0, 1)$

- Si $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, et **on connaît** σ^2 , alors

$$\frac{Z_n}{\sigma} = \frac{n^{1/2}(\bar{Y}_n - \mu)}{\sigma} \sim \mathcal{N}(0, 1)$$

et l'intervalle de confiance pour μ est

$$\left[\bar{Y}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma, \bar{Y}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma \right]$$

- La largeur \downarrow avec n , \uparrow avec σ et avec le niveau $(1 - \alpha)$, car $\Pr_{\theta}(L_n \leq \theta \leq U_n) \uparrow$

$$Z_n = \sqrt{n}(\bar{Y}_n - \mu) \sim N(0, 1)$$

$$\forall \alpha \quad 1 - \alpha = P_{\mu} \left(z_{\frac{\alpha}{2}} \leq Z_n \leq z_{1-\frac{\alpha}{2}} \right) \phi\left(\frac{\alpha}{2}\right) = z_{\frac{\alpha}{2}} \quad z_{1-\frac{\alpha}{2}}$$

Is der μ

$-z_{1-\frac{\alpha}{2}}$

$$= P_{\mu} \left(-z_{1-\frac{\alpha}{2}} \leq \sqrt{n}(\mu - \bar{Y}_n) \leq -z_{\frac{\alpha}{2}} \right)$$

$$= P_{\mu} \left(\bar{Y}_n - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}} \leq \mu \leq \bar{Y}_n + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}} \right)$$

Exemple

Exemple On suppose que la résistance Y d'un certain type d'équipements électriques est distribuée selon une loi normale avec $\sigma = 0.12$ ohm.

Un échantillon de taille $n = 64$ a donné comme moyenne la valeur $\bar{y}_n = 5.34$ ohm.

Trouver un intervalle de confiance pour μ au niveau 95%.

$$\alpha = 0.05$$
$$1 - \frac{\alpha}{2} = 0.975$$

$$z_{1-\frac{\alpha}{2}} \approx 1.96$$

(tableau)

$$\left[\bar{y}_n - \frac{z_{1-\frac{\alpha}{2}} \sigma}{\sqrt{n}}, \bar{y}_n + \frac{z_{1-\frac{\alpha}{2}} \sigma}{\sqrt{n}} \right]$$
$$\approx \left[5.34 - \frac{1.96}{8} 0.12, 5.34 + \frac{1.96}{8} 0.12 \right]$$
$$\approx [5.31, 5.37]$$

Intervalle de Student

- Soient $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ avec σ^2 **inconnue**
- $\left[\bar{Y}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma, \bar{Y}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma \right]$ n'est plus un intervalle de confiance
- On suppose $n > 1$ et estime σ^2 par

$$S_n^2 = \frac{1}{n-1} \sum_{j=1}^n (Y_j - \bar{Y}_n)^2 \quad \mathbb{E}_{\mu, \sigma^2} [S_n^2] = \sigma^2$$

- On remplace σ par S_n : soit

$$T_n = \frac{Z_n}{S_n} = n^{1/2} \frac{\bar{Y}_n - \mu}{S_n} = \frac{Z_n/\sigma}{\sqrt{S_n^2/\sigma^2}}$$

- Nominateur et dénominateur indép, leurs lois ne dépendent pas de (μ, σ^2)
- T_n suit une **loi de Student avec $n - 1$ degrés de liberté**, $T_n \sim t_{n-1}$, qui ne dépend de $\theta = (\mu, \sigma^2)$. C'est une loi symétrique, on dénote les quantiles par $t_{n-1, \beta}$
- L'intervalle de confiance qui en résulte est

$$\left[\bar{Y}_n - \frac{t_{n-1, 1-\alpha/2}}{\sqrt{n}} S_n, \bar{Y}_n + \frac{t_{n-1, 1-\alpha/2}}{\sqrt{n}} S_n \right]$$

- La largeur \downarrow avec n , \uparrow avec S_n et avec le niveau $(1 - \alpha)$

Exemple

Exemple Pour déterminer le point de fusion μ d'un certain alliage, on a procédé à $n = 9$ observations qui ont donné une moyenne $\bar{y}_n = 1040^\circ$ avec $s_n = 16^\circ$. On suppose que les données suivent une loi normale.

Trouver un intervalle de confiance pour μ à niveau 95%.

$\alpha = 0.05$

≈ 2.306

$t_{8, 0.975}$

$$\left[1040 - \frac{2.306}{\sqrt{9}} 16, 1040 + \frac{2.306}{\sqrt{9}} 16 \right]^\circ$$
$$\approx [1028, 1052]^\circ$$

Intervalles de confiance pour μ dans le cas normal : résumé

Soient $Y_1, \dots, Y_n \stackrel{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$ avec μ inconnu

- Pour σ **connue** $n^{1/2} \frac{\bar{Y}_n - \mu}{\sigma} \sim \mathcal{N}(0, 1)$ et on a les IC de niveau $1 - \alpha$

bilatéral

$$\left[\bar{Y}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma, \bar{Y}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \sigma \right]$$

Unilatéral

$$\left[\bar{Y}_n - \frac{z_{1-\alpha}}{\sqrt{n}} \sigma, \infty \right] \quad \text{ou} \quad \left[-\infty, \bar{Y}_n + \frac{z_{1-\alpha}}{\sqrt{n}} \sigma \right]$$

avec $z_\beta = \Phi^{-1}(\beta)$ les quantiles de la loi $\mathcal{N}(0, 1)$

- Pour σ **inconnue** (et $n > 1$) $n^{1/2} \frac{\bar{Y}_n - \mu}{S_n} \sim t_{n-1}$ et on a les IC de niveau $1 - \alpha$

$$\left[\bar{Y}_n - \frac{t_{n-1, 1-\alpha/2}}{\sqrt{n}} S_n, \bar{Y}_n + \frac{t_{n-1, 1-\alpha/2}}{\sqrt{n}} S_n \right]$$

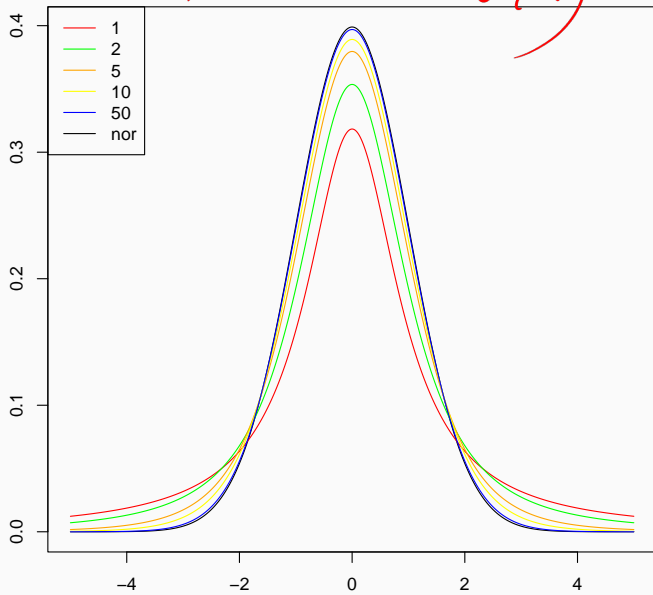
$$\left[\bar{Y}_n - \frac{t_{n-1, 1-\alpha}}{\sqrt{n}} S_n, \infty \right] \quad \text{ou} \quad \left[-\infty, \bar{Y}_n + \frac{t_{n-1, 1-\alpha}}{\sqrt{n}} S_n \right]$$

avec t_{n-1} la **loi de Student avec $n - 1$ degrés de liberté**, $t_{n-1, \beta}$ sont les quantiles de cette distribution, S_n définie à la diapositive [182](#).

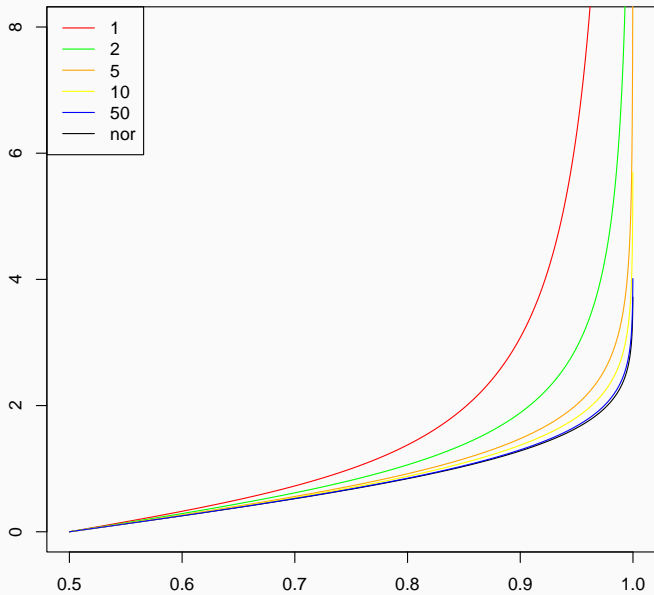
t_{n-1} symétrique comme $\mathcal{N}(0, 1)$, mais quantiles plus grands : $|t_{n-1, \beta}| > |z_\beta|$, $\beta \neq \frac{1}{2}$ 184

Densités de t_k et $\mathcal{N}(0, 1)$

$k \rightarrow \infty \Rightarrow \text{loi } T_k \approx \mathcal{N}(0, 1)$



Quantiles de t_k et $\mathcal{N}(0,1)$



Pivots

- Une fonction $T(Y_1, \dots, Y_n, \theta)$ dont la loi est connue et ne dépend pas de θ s'appelle un **pivot**

$$Y_2, Y_n \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu, 1)$$

- Attention !** Ce n'est pas une statistique car c'est une fonction de θ

- Si $a \leq b$,

$$Z_n = \sqrt{n}(\bar{Y}_n - \mu) \sim \mathcal{N}(0, 1) \text{ pivot}$$

$$\alpha_1 = \Pr(T < a), \quad \alpha_2 = \Pr(T > b), \quad \alpha_1 + \alpha_2 = \alpha \in [0, 1]$$

(souvent $\alpha_1 = \alpha_2 = \alpha/2$) alors

$$\Pr_{\theta}(a \leq T \leq b) = \Pr_{\theta}(T \leq b) - \Pr_{\theta}(T < a) = (1 - \alpha_2) - \alpha_1 = 1 - \alpha$$

- Si on peut isoler θ , on peut trouver des variables aléatoires L et U telles que

$$\Pr_{\theta}(L \leq \theta \leq U) = 1 - \alpha$$

Exemples T_n (diapositive [182](#)), Z_n (diapositive [179](#)).