Modèle d'échantillon d'une loi de Poisson $\mathrm{Pois}(\theta)$

1. l'estimation des moments est $\widehat{\theta} = \overline{x}_n$
2. l'EMV est également $\widehat{\theta} = \overline{x}_n$
3. Si on s'intéresse à la probabilité d'observer aucun événement, i.e. $g(\theta) = \mathbb{P}_\theta\{0\} = \exp(-\theta)$ alors
   (a) la méthode des moments donne $\widehat{g(\theta)}_{\mathrm{Mom}} = \frac{1}{n}\sum_i \mathbf{1}_{\{x_i=0\}}$ autrement dit on estime par la proportion de 0
   (b) Par la propriété d'invariance, l'EMV est $\widehat{g(\theta)}_{\mathrm{EMV}} = \exp(-\overline{x}_n)$. Notons qu'il s'agit d'une fonction de la statistique exhaustive $\sum_i x_i$.

- Si les deux méthodes donnent des résultats identiques pour des modèles simples, en général ils diffèrent. C'est le cas pour la probabilité de n'observer aucun événement.
- Il faut pouvoir comparer des estimateurs entre eux : définir des critères de comparaison

Pour finir sur les méthodes :

- Rappelons que dans le cadre des modèles régression linéaires, une méthode dite des moindres carrées a été utilisée (elle l'est également pour la régression non linéaire [ABC92])
- Une méthode populaire pour trouver un estimation $\widehat{\theta}$ de θ consiste à maximiser un critère de la forme

$$\theta \mapsto M_n(\theta) = \frac{1}{n}\sum_{i=1}^n m_\theta(x_i)$$

  Une solution est alors appelée une *M-estimation*. Cela inclut le cas particulier des EMV, des estimations par fonction de contrastes, estimation robuste, ... [Vaa98], [Ser80, Chap 7]

A. Antoniadis, J. Berruyer, R. Carmona.
Régression non linéaire et applications. Economica, 1992.

A. W. van der Vaart.
Asymptotic statistics. Cambridge University Press, 1998.

R. J. Serfling.
Approximation Theorems of Mathematical Statistics. Wiley, 1980.

Erreur quadratique

- Pour qu'une estimation $\widehat{\theta}$ d'un paramètre $\theta(x)$ soit de bonne qualité, une idée naturelle est de vérifier que $\widehat{\theta}(x)$ est assez proche de θ. Pour cela, une distance classique sur $\mathbb{R}^p$ comme une distance L^p peut faire l'affaire :

$$d_p(\widehat{\theta}(x), \theta) = \|\widehat{\theta}(x) - \theta\|_p = \left(\sum_i (\widehat{\theta}_i(x) - \theta_i)^p\right)^{1/p}$$

  Pour des raisons de simplicité des manipulations analytiques (et non nécessairement de pertinence, car fortement sensible au valeurs extrêmes par exemple), on choisit le plus souvent la distance euclidienne usuelle :

$$d_2(\widehat{\theta}(x), \theta) = \|\widehat{\theta}(x) - \theta\|_2 = \sqrt{\sum_i (\widehat{\theta}_i(x) - \theta_i)^2}$$

- Pour prendre en compte la variabilité issue des observations x, on s'intéresse à l'estimateur associé $\widehat{\Theta}$ à $\widehat{\theta}$ et on calcule par exemple le coût moyen sous $\mathbb{P}_\theta$ du choix de l'estimation $\widehat{\theta}$ de θ :

$$\mathrm{EQM}(\widehat{\theta}) = \mathbb{E}_\theta[d_2(\widehat{\theta}(X_1,\ldots,X_n),\theta)^2] = \mathbb{E}_\theta\left[\|\widehat{\theta}(X_1,\ldots,X_n) - \theta\|_2^2\right]$$

Définition 16 (Erreur quadratique moyenne)

Pour tout $\theta \in \mathbb{R}^p$, si $\widehat{\theta}$ est de carré intégrable sous $\mathbb{P}_\theta$ alors l'erreur quadratique moyenne ou le risque quadratique moyen est défini par

$$\mathrm{EQM}_\theta(\widehat{\theta}) = \mathbb{E}_\theta\left[\|\widehat{\theta}(X_1,\ldots,X_n) - \theta\|_2^2\right] \qquad (11)$$

Dans le cas d'un paramètre unidimensionnel, la formule se réduit à $\mathrm{EQM}(\widehat{\theta}) = \mathbb{E}_\theta\left[(\widehat{\theta}(X_1,\ldots,X_n) - \theta)^2\right]$

1. Une manière de comparer des estimations (ou estimateurs) entre elles est alors de comparer leurs EQM, et on peut vouloir rechercher la meilleure estimation $\widehat{\theta}$ possible de θ donné, c.a.d. qui réaliserait $\min_{\widehat{\theta}} \mathrm{EQM}_\theta(\widehat{\theta})$.
2. Puis déterminer une estimation optimale $\widehat{\theta}^*$ qui soit d'erreur quadratique moyenne faible pour beaucoup de θ, c'est à dire une estimation minimisant l'EQM **uniformément** en θ :

$$\mathrm{EQM}_\theta(\widehat{\theta}^*) = \min_{\widehat{\theta}} \mathrm{EQM}_\theta(\widehat{\theta}) \quad \forall \theta \in \Theta.$$

Malheureusement, cet optimum uniforme n'existe pas.
➥ Restreindre à une sous-classe d'estimations ou d'estimateurs : ici la classe des **estimateurs sans biais**

Estimateur sans biais

Définition 17 (Estimateur sans biais)

Le biais d'un estimateur $\widehat{\theta}(X_1,\ldots,X_n)$ *est défini par*

$$b_\theta(\widehat{\theta}) := \mathbb{E}_\theta[\widehat{\theta}(X_1,\ldots,X_n)] - \theta$$

Une estimation $\widehat{\theta}$ *ou l'estimateur associé est dit sans biais si* $b_\theta(\widehat{\theta}) = 0$, *et biaisé dans le cas contraire.*

Le biais mesure une erreur « systématique » d'estimation. Par exemple pour un biais négatif, $\widehat{\theta}$ aura tendance à sous-estimer le paramètre θ.

Moyenne et variance empirique

Pour la loi sous-jacente de l'échantillon, $\overline{X}_n$ est un estimateur sans biais de l'espérance mais S_n^2 est un estimateur biaisé de la variance. $\widetilde{S}_n^2$ est lui sans biais.

Modèle linéaire gaussien

Le vecteur $\widehat{a} = \left(X^\top X\right)^{-1} X^\top Y$ est un estimateur des MC sans biais des coefficients de l'hyperplan de régression.

Décomposition de l'EQM

On dispose de la décomposition de l'EQM :

$$\begin{aligned} &\mathrm{EQM}_\theta(\widehat{\theta}) \\ &= \mathbb{E}_\theta\left[\|\widehat{\theta}(X_1,\ldots,X_n) - \mathbb{E}_\theta[\widehat{\theta}(X_1,\ldots,X_n)]\|_2^{\,2}\right] + \|b_\theta(\widehat{\theta})\|_2^{\,2} \\ (12)\quad &= \mathrm{Tr}(V_\theta(\widehat{\theta})) + \|b_\theta(\widehat{\theta})\|_2^{\,2} \end{aligned}$$

Dans le cas unidimensionnel, on retrouve la décomposition dite variance/biais de l'EQM :

$$\mathrm{EQM}_\theta(\widehat{\theta}) = V_\theta(\widehat{\theta}) + b_\theta(\widehat{\theta})^2 \qquad (13)$$

- (12-13) suggère un compromis biais/variance dans la construction d'une estimateur, c'est à dire un compromis entre la précision et la variabilité d'une estimation.
- Rechercher dans la classe des estimateurs sans biais (où $\mathrm{EQM}_\theta(\widehat{\theta}) = \mathrm{Tr}(V_\theta(\widehat{\theta}))$), s'il en existe un de « variance » minimale.

Pour comparer deux estimateurs sans biais, on introduit

Définition 18 (Risque quadratique pour les ESB)

Le risque quadratique d'un estimateur sans biais $\widehat{\theta}(X_1,\ldots,X_n)$ *est définie par sa matrice de covariance*

$$V_\theta(\widehat{\theta}) = \mathbb{E}_\theta\left[(\widehat{\theta} - \mathbb{E}_\theta[\widehat{\theta}])(\widehat{\theta} - \mathbb{E}_\theta[\widehat{\theta}])^\top\right] \qquad (14)$$

et un estimateur $\widehat{\theta}_1(X_1,\ldots,X_n)$ est meilleur que $\widehat{\theta}_2(X_1,\ldots,X_n)$ pour le risque quadratique si, pour tout θ,

(15)
$V_\theta(\widehat{\theta}_1) \le V_\theta(\widehat{\theta}_2)$

i.e. la matrice $V_\theta(\widehat{\theta}_2) - V_\theta(\widehat{\theta}_1) \in \mathcal{L}^+$ i.e. symétrique semi-définie positive

Remarque 9

1. On peut vérifier que (15) implique $\mathrm{EQM}_\theta(\widehat{\theta}_1) \le \mathrm{EQM}_\theta(\widehat{\theta}_2)$
2. L'inégalité (15) ne définit qu'un ordre partiel sur $\mathcal{L}^+$

Proposition 12 (Théorème de Gauss Markov)

L'estimateur des moindres carrés des coefficients de l'hyperplan de régression est l'estimateur fonction linéaire des observations sans biais de risque quadratique (ou de variance) minimal.

À partir d'un estimateur sans biais, le résultat suivant permet d'en construire un meilleur selon le risque quadratique.

Théorème 5 (Théorème de Rao-Blackwell)

Soit s une statistique exhaustive dans un modèle de n-échantillon dominé.
Alors si $\widehat{\theta}(X_1,\ldots,X_n)$ est un estimateur sans biais de θ (ou $g(\theta)$) de variance finie,

$$\widehat{\theta}^*(X_1,\ldots,X_n) = \mathbb{E}_\theta\left[\widehat{\theta}(X_1,\ldots,X_n) \mid s(X_1,\ldots,X_n)\right]$$

est un estimateur sans biais de θ (ou $g(\theta)$) vérifiant

(16) $$\forall \theta \in \Theta, \quad V_\theta(\widehat{\theta}^*) \le V_\theta(\widehat{\theta})$$

Remarque 10

Noter l'indépendance de θ du nouvel estimateur $\widehat{\theta}^*$ car dans le cas contraire ce ne serait pas un estimateur.

Modèle de Poisson $(\mathrm{Pois}(\theta), \theta \in]0,+\infty[)$

Estimation de $g(\theta) = \exp(-\theta)$:

- l'EMV de $g(\theta)$ est $\widehat{\theta}_{\mathrm{EMV}} = \exp(-\overline{X}_n)$. Il est biaisé

$$\mathbb{E}_\theta[\widehat{\theta}_{\mathrm{EMV}}] = \exp\left(-n\theta(1-\exp(-1/n))\right)$$

- l'estimateur par la méthode des moments est $\widehat{\theta} = \frac{1}{n}\sum_i 1_{\{X_i=0\}}$ qui est sans biais. La statistique $\sum_i x_i$ est exhaustive. Il en résulte un estimateur de Rao-Blackwell associé

$$\widehat{\theta}^*(X_1,\ldots,X_n) = \left(1-\frac{1}{n}\right)^{\sum_i X_i}$$

Alors on vérifie que pour $n > 1$,

$$V_\theta(\widehat{\theta}) = \frac{\exp(-\theta)(1-\exp(-\theta))}{n} = \frac{\exp(-2\theta)(\exp(\theta)-1)}{n}$$
$$> V_\theta(\widehat{\theta}^*) = \exp(-2\theta)\left[\exp\left(\frac{\theta}{n}\right)-1\right]$$

Notons que les deux estimateurs biaisés vérifient

1. Variance empirique :

$$\mathbb{E}_\theta[S_n^2] = \frac{n-1}{n}\mathbb{E}_\theta[\tilde{S}_n^2] = \frac{n-1}{n}V_\theta(X) \xrightarrow[n\to+\infty]{} V_\theta(X)$$

2. Modèle Poisson :

$$\mathbb{E}_\theta\left[\widehat{\exp(-\theta)}\right] = \exp\left(-n\theta(1-\exp(-1/n))\right) \xrightarrow[n\to+\infty]{} \exp(-\theta)$$

Un tel estimateur est dit *asymptotiquement sans biais*. Autrement dit, le biais devient asymptotiquement nul lorsque la taille de l'échantillon devient grande.

Estimateur sans biais de variance minimale

Définition 19 (ESBVM)

Un estimateur $\widehat{\theta}^*$ de θ (ou $g(\theta)$) sans biais de variance finie est dit de variance minimale (ESBVM) s'il est de risque quadratique minimal (uniformément en θ) dans la classe $\mathcal{ESB}$ des estimateurs de θ (ou $g(\theta)$) sans biais de variance finie :

$$\forall \widehat{\theta} \in \mathcal{ESB}, \forall \theta \in \Theta, \quad V_\theta(\widehat{\theta}^*) \le V_\theta(\widehat{\theta}).$$

Remarque 11

La précédente définition n'a de sens que si la classe $\mathcal{ESB}$ est non vide. De plus, même si c'est le cas, on n'a aucune garantie qu'un ESBVM existe. Par contre, s'il existe il est unique.

Proposition 13 (Unicité de l'ESBVM)

Si on reprend les notations de la définition précédente, si $\widehat{\theta}^*$ existe alors il est unique $(\mathbb{P}_\theta, \theta \in \Theta)$-ps.

Théorème 6 (Théorème de Lehmann-Scheffé)

Avec les notations de la définition 19, pour tout ESB $\widehat{\theta}$ de θ (ou $g(\theta)$) de variance finie, si s est une statistique exhaustive complète alors

$$\widehat{\theta}^*(X_1,\ldots,X_n) = \mathbb{E}_\theta[\widehat{\theta}(X_1,\ldots,X_n) \mid s(X_1,\ldots,X_n)]$$

est $(\mathbb{P}_\theta, \theta \in \Theta)$-ps l'unique ESBVM de θ (ou $g(\theta)$).
En d'autres termes tout estimateur de $\mathcal{ESB}$ qui est fonction d'une statistique exhaustive complète est l'unique ESBVM.

Modèle Poisson

On sait que la statistique naturelle $\sum_i x_i$ est exhaustive et complète. Il en résulte que la construction de Rao-Blackwell obtenue précédemment donne $(\mathbb{P}_\theta, \theta \in \Theta)$-ps l'unique ESBVM .

Modèle binomial

- La statistique naturelle $\sum_i x_i$ est exhaustive et complète.
- $\widehat{\theta} = \overline{X}_n$ est un estimateur sans biais de θ et c'est une fonction de la statistique $\sum_i x_i$, d'où c'est l'ESBVM du paramètre θ.
- On s'intéresse au paramètre $g(\theta) = \frac{\theta(1-\theta)}{n} = V_\theta(\overline{X}_n)$:
  - Candidat estimateur (par EMV par exemple) :

$$\widehat{g(\theta)} = \frac{\widehat{\theta}(1-\widehat{\theta})}{n} = \frac{\overline{X}_n(1-\overline{X}_n)}{n}$$

  Cet estimateur de $g(\theta)$ est baisé :

$$\mathbb{E}_\theta[\widehat{g(\theta)}] = \frac{n-1}{n}\frac{\theta(1-\theta)}{n} = \frac{n-1}{n}g(\theta)$$

  - L'estimateur $\widehat{g(\theta)}_1 = \frac{n}{n-1}\widehat{g(\theta)} = \frac{\overline{X}_n(1-\overline{X}_n)}{n-1}$ est sans biais, fonction de la statistique exhaustive et complète $\sum_i x_i$, on en déduit qu'il s'agit de l'ESBVM de $g(\theta)$.

Famille exponentielle

La statistique naturelle $t(x)$ pour une famille exponentielle telle que $\alpha(\Theta)$ soit d'intérieur non vide est exhaustive complète. Il en résulte que la construction de Rao-Blackwell à partir de t et tout ESB donne $(\mathbb{P}_\theta, \theta \in \Theta)$-ps l'unique ESBVM .

Obtenir un ESBVM de $g(\theta)$ à partir d'une statistique s exhaustive et complète I

1. Si la loi de $s(X_1,\ldots,X_n)$ est connue alors rechercher une fonction h telle que $\mathbb{E}_\theta[h(s(X_1,\ldots,X_n))] = g(\theta)$

Modèle uniforme $(\mathrm{Unif}]0,\theta[, \theta \in]0,+\infty[)$

L'EMV de θ est $s(X_1,\ldots,X_n) = X_{(n)}$. Cette statistique s est exhaustive et complète. Elle a une loi admettant la densité

$$f(x) = n\frac{x^{n-1}}{\theta^n}1_{]0,\theta[}(x)$$

et alors l'ESBVM de $g(\theta)$ pour g dérivable :

$$h(X_{(n)}) = g(X_{(n)}) + \frac{1}{n}X_{(n)}g'(X_{(n)}).$$

Obtenir un ESBVM de $g(\theta)$ à partir d'une statistique s exhaustive et complète II

2. Utiliser une version de Rao-Blackwell d'un ESB. Choisir ce dernier de sorte que le calcul de son espérance conditionnelle à la stat. exhaustive et complète s soit le plus « simple » possible.

Modèle exponentiel $(\mathrm{Exp}(1/\theta), \theta \in]0, +\infty[)$

- $s(x_1, \ldots, x_n) = \sum_i x_i$ est exhaustive et complète pour $\theta \in]0, +\infty[$.
- On s'intéresse au paramètre $g(\theta) = \exp(-t/\theta) = 1 - F_\theta(t)$ (pour $t > 0$ fixé).
- $\widehat{\theta}(X_1, \ldots, X_n) = 1_{\{]t,+\infty[\}}(X_1)$ est un ESB pour $g(\theta)$ et l'ESBVM est donné par

$$\widehat{\theta}^*(X_1, \ldots, X_n) = \mathbb{E}_\theta\left[1_{\{]t,+\infty[\}}(X_1) \mid \sum_i X_i\right] = \mathbb{P}_\theta\left\{X_1 > t \mid \sum_i X_i\right\}$$

- Sachant que X_1 et $\sum_{i=2}^n X_i$ sont des v.a. indépendantes et $\sum_{i=2}^n X_i \sim \mathrm{Ga}(n-1, 1/\theta)$, on peut montrer que la loi du rapport $X_1/(\sum_i X_i)$ admet la densité par rapport à la mesure de Lebesgue

$$(n-1)(1-x)^{n-2} 1_{]0,1[}(x)$$

indépendante du paramètre θ. La statistique $X_1/(\sum_i X_i)$ est donc libre.

- On déduit du Théorème 4 que les statistiques $\sum_i X_i$ et $X_1/\sum_i X_i$ sont indépendantes. Alors

$$\mathbb{P}_\theta\{X_1 > t \mid \sum_i X_i\} = \left(\mathbb{P}_\theta\left\{X_1/\sum_i X_i > t/\sum_i x_i\right\}\right) \circ (\sum_i X_i)$$

et l'ESBVM est donné par

$$\widehat{\theta}^*(X_1, \ldots, X_n) = \left(1 - \frac{t}{\sum_i X_i}\right)^{n-1} 1_{\{\sum_i X_i > t\}}$$

Borne de Fréchet-Darmois-Rao-Cramer

- Obtenir une borne inférieure sur la matrice de covariance de tout estimateur construit sur un modèle statistique de n-échantillon
  - ↪ une borne ne dépendant que du modèle statistique et de la fonction du paramètre θ à estimer (donc calculable avant la recherche d'un estimateur explicite)
- Cas unidimensionnel : l'inégalité de Cauchy-Schwarz donne une information sur $V_\theta(\widehat{\theta}_n)$ en relation avec toute v.a. t

$$\mathrm{Cov}_\theta(\widehat{\theta}_n, t)^2 \leq V_\theta(\widehat{\theta}_n) V_\theta(t) \text{ ou encore } V_\theta(\widehat{\theta}_n) \geq \frac{\mathrm{Cov}_\theta(\widehat{\theta}_n, t)^2}{V_\theta(t)}$$

(en supposant que cette écriture ait un sens)
  - ↪ **une v.a. candidate est le score $S(\theta)$ associé au modèle** : $V_\theta(S(\theta)) = I_n(\theta)$ ne dépend que du modèle. L'objectif sera atteint si la covariance ne dépend que du modèle et de la fonction du paramètre à estimer. C'est ce que précise le prochain théorème.

Théorème 7 (Borne de Fréchet-Darmois-Rao-Cramer)

Supposons que les conditions suivantes sont satisfaites

1. *le modèle statistique de n-échantillon est homogène ;*
2. *Θ est un ouvert de $\mathbb{R}^p$;*
3. *la condition (CD1) de dérivation sous le signe somme ;*
4. *la matrice d'information de Fisher $I_n(\theta)$ existe et est définie positive pour tout $\theta \in \Theta$.*

Pour tout estimateur $\widehat{\theta}_n$ de carré intégrable pour la famille de lois associées au modèle statistique alors la matrice de covariance de l'estimateur vérifie

$$V_\theta(\widehat{\theta}_n) - \left(\frac{\partial \mathbb{E}_\theta[\widehat{\theta}_n]_i}{\partial \theta_j}\right)^\top I_n(\theta)^{-1} \left(\frac{\partial \mathbb{E}_\theta[\widehat{\theta}_n]_i}{\partial \theta_j}\right) \geq 0 \tag{17}$$

*i.e. elle est dans $\mathcal{L}^+$. La seconde matrice est appelée la **borne de FDRC***
Si $\widehat{\theta}_n$ est un ESB de θ alors la borne se simplifie en

$$V_\theta(\widehat{\theta}_n) \geq I_n(\theta)^{-1}.$$

Corollaire 1 (Paramètre unidimensionnel)

Dans le cas d'un paramètre θ unidimensionnel, sous les conditions du théorème précédent :

- *l'inégalité (17) donne la minoration suivante de la variance d'un estimateur $\widehat{\theta}_n$ de carré intégrable*

$$V_\theta(\widehat{\theta}_n) \geq \frac{\left(\frac{d\mathbb{E}_\theta[\widehat{\theta}_n]}{d\theta}\right)^2}{I_n(\theta)}; \tag{18}$$

- *en particulier, si $\widehat{\theta}_n$ est un ESB de $g(\theta)$ alors*

$$V_\theta(\widehat{\theta}_n) \geq \frac{g'(\theta)^2}{I_n(\theta)};$$

- *enfin si $\widehat{\theta}_n$ est un ESB de θ alors l'inégalité se simplifie en*

$$V_\theta(\widehat{\theta}_n) \geq \frac{1}{I_n(\theta)}.$$

n-échantillon d'une famille exponentielle de paramètre naturel θ

- Il a déjà été vérifié que $I_n(\theta)$ pour le n-échantillon d'une famille exponentielle sous forme canonique et régulière est égal à la matrice $V_\theta(\sum_i t(X_i))$.
- Si on s'intéresse au paramètre $g(\theta) = \mathbb{E}_\theta[t(X_1)]$ alors la borne de FDRC vaut $I_1(\theta)/n$. Posons $\bar{t}_n := \frac{1}{n}\sum_i t(X_i)$. C'est un ESB de $g(\theta)$ de variance qui atteint la borne de FDRC.

Remarque 12

1. La moyenne empirique de la statistique naturelle d'une famille exponentielle définit un ESB du paramètre $\mathbb{E}_\theta[t(X_1)]$ qui atteint la borne de FDRC.
   On peut montrer sous certaines conditions de régularité que la borne de FDRC est atteinte **ssi** la vraisemblance définit un modèle appartenant à une famille exponentielle (e.g. [Mon97, Chap 10,Th 4]).
2. Si l'information de Fisher dépend du paramétrage, la borne de FDRC ne dépend pas d'un reparamétrage difféomorphe.

Efficacité d'un estimateur

Définition 20 (Efficacité)

Un estimateur pour lequel la borne de Fréchet-Darmois-Rao-Cramer est atteinte est dit efficace.

famille exponentielle

On vient de voir que la moyenne empirique de la statistique naturelle d'une famille exponentielle sous forme canonique et régulière définit un estimateur efficace du paramètre $g(\theta) = \mathbb{E}_\theta[t(X_1)]$.

Modèle Poisson $\mathrm{Pois}(\theta)$

- Le modèle de Poisson appartient à la famille exponentielle et comme $\theta = \mathbb{E}_\theta[X_1]$, on sait que $\overline{X}_n$ est un ESBVM qui atteint la borne FDRC qui vaut ici $1/I_n(\theta) = \frac{\theta}{n}$
- Pour $g(\theta) = \exp(-\theta)$, on a obtenu l'ESBVM qui vaut $\widehat{g(\theta)}_n = (1 - 1/n)^{\sum_{i=1}^n X_i}$. On avait obtenu sa variance

$$V_\theta(\widehat{g(\theta)}_n) = (\exp(\theta/n) - 1)\exp(-2\theta)$$

La borne de FDRC pour $g(\theta)$ est donnée par $g'(\theta)^2/I_n(\theta) = \frac{\theta\exp(-2\theta)}{n}$, et elle n'est pas atteinte car $V_\theta(\widehat{g(\theta)}_n) = (\exp(\theta/n) - 1)\exp(-2\theta) > \frac{\theta\exp(-2\theta)}{n}$.

- On va s'intéresser aux **propriétés d'un estimateur pour des grandes tailles de l'échantillon**, i.e. à son comportement limite lorsque n tend vers l'infini. Quelques formes classiques d'asymptotique :
  1. consistance, consistance forte ou en moyenne quadratique
  2. normalité asymptotique
  3. « efficacité asymptotique »
- Un estimateur était une statistique définie sur un modèle statistique de n-échantillon

$$\left((E_X)^n, \mathscr{E}^{\otimes n}, (\mathbb{P}_\theta^{\otimes n})_{\theta\in\Theta}\right)$$

associé à un modèle de base $(E_X, \mathscr{E}, (\mathbb{P}_\theta)_{\theta\in\Theta})$. Dans cette partie, pour bien marquer la dépendance en la taille n de l'échantillon, nous placerons cette taille en indice : $\widehat{\theta}_n$.
En toute rigueur, la convergence d'un estimateur doit être évaluée sur le modèle statistique

$$\left((E_X)^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}}, (\mathbb{P}_\theta^{\otimes \mathbb{N}})_{\theta\in\Theta}\right)$$

Cependant, ici, $(\mathbb{P}_\theta)_{\theta\in\Theta}$ désignera indifféremment la famille probabilité associée à chacun des trois modèles statistiques ci-dessus.