STT 1000 - STATISTIQVES

ARTHVR CHARPENTIER

Loi Gamma (shape/scale - scale/rate)

- écriture shape/scale

$$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}} \text{ pour } x \geq 0,$$

avec $k > 0$ (shape) et $\theta > 0$ (scale).

$$\mathbb{E}[X] = k\theta \text{ et } \text{Var}[X] = k\theta^2$$

- écriture shape/rate

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \text{ pour } x \geq 0,$$

avec $\alpha > 0$ (shape) et $\beta > 0$ (rate).

$$\mathbb{E}[X] = \frac{\alpha}{\beta} \text{ et } \text{Var}[X] = \frac{\alpha}{\beta^2}$$

Loi Gamma

La log-vraisemblance s'écrit

$$\log L(k,\theta) = (k-1)\sum_{i=1}^{N}\ln(x_i) - \sum_{i=1}^{N}\frac{x_i}{\theta} - Nk\ln(\theta) - N\ln(\Gamma(k))$$

- k fixé

$$\hat{\theta} = \frac{1}{kN}\sum_{i=1}^{N}x_i$$

Loi Gamma

Il existe des estimateurs

$$\hat{k} = \frac{N\sum_{i=1}^N x_i}{N\sum_{i=1}^N x_i \ln(x_i) - \sum_{i=1}^N \ln(x_i)\sum_{i=1}^N x_i}$$

et

$$\hat{\theta} = \frac{1}{N^2}\left(N\sum_{i=1}^N x_i \ln(x_i) - \sum_{i=1}^N \ln(x_i)\sum_{i=1}^N x_i\right)$$

Loi Gamma vs Loi Exponentielle

$$\mu(t) = \mathbb{E}[X - t | X > t] = \frac{1}{\overline{F}(t)} \int_t^\infty \overline{F}(x) dx$$

donc

$$\mu(t) = \frac{\alpha}{\beta} \frac{\overline{F}_{\alpha+1,\beta}(x)}{\overline{F}_{\alpha,\beta}(x)} - x \sim \frac{1}{\beta}$$

pour x grand COMPLETER
https://freakonometrics.hypotheses.org/tag/imrl

Loi Gamma

Soit $x_1, \cdots, x_n$ un échantillon i.i.d. tiré suivant une loi $\mathcal{G}(2, \theta)$, avec $\theta > 0$ inconnu. Soit $\widehat{\theta} = 2\overline{x}^{-1}$

$\widehat{\theta}$ n'est pas un estimateur sans biais de θ : $\mathbb{E}\left[\widehat{\theta}\right] = \dfrac{2n}{2n-1}\theta \neq \theta$

En effet,

$$S_n = X_1 + X_2 + \cdots + X_n \sim \mathcal{G}(2n, \theta)$$

de densité $g(s) = \dfrac{\theta^{2n}}{(2n-1!)} s^{2n-1} \exp[-\theta s]$ sur $\mathbb{R}_+$. Alors

$$\mathbb{E}\left[\frac{2n}{S}\right] = 2n \int_0^\infty \frac{\theta^{2n}}{(2n-1)!} \frac{s^{2n-1}}{s} \exp[-\theta s] ds$$

$$\mathbb{E}\left[\frac{2n}{S}\right] = \frac{2n\theta}{2n-1} \int_0^\infty \frac{\theta^{2n-1}}{(2n-2)!} s^{2n-2} \exp[-\theta s] ds = \frac{2n\theta}{2n-1}$$

$\widehat{\theta}$ est un estimateur asymp. Gaussien : $\sqrt{n}\left(\widehat{\theta}-\theta\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\theta^2}{2}\right)$

Si $X \sim \mathcal{G}(2,\theta)$, alors $\mathbb{E}[X] = 2\theta^{-1}$ et $\text{Var}[X] = 2\theta^{-2}$, donc par le théorème central limite,

$$\sqrt{n}\left(\overline{X}-\tau\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,\tau^2/2) \text{ où } \tau = 2\theta^{-2}$$

Posons $g(x) = 2/x$, alors d'après la Δ-methode,

$$\sqrt{n}\left(\widehat{\theta}-\theta\right) = \sqrt{n}\left(g(\overline{X})-g(\tau)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \left[g'(\tau)\right]^2\tau^2/2\right)$$

Loi Gamma: maximum de vraisemblance

$$\log \mathcal{L}(\alpha,\beta) = n\alpha\log(\beta) - n\log[\Gamma(\alpha)] + (\alpha-1)\sum_{i=1}^n \log(x_i) - \beta\sum_{i=1}^n x_i$$

$$\nabla \log \mathcal{L}(\alpha,\beta) = \begin{pmatrix} n\log(\beta) - n[\log\Gamma(\alpha)]' + \displaystyle\sum_{i=1}^n \log(x_i) \\ \dfrac{n\alpha}{\beta} - \displaystyle\sum_{i=1}^n x_i \end{pmatrix}$$

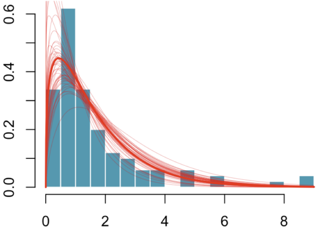
$$H\log \mathcal{L}(\alpha,\beta) = \frac{1}{n(1-\alpha[\log\Gamma(\alpha)]'')}\begin{pmatrix} \alpha & \beta \\ \beta & \beta^2[\log(\Gamma(\alpha))]'' \end{pmatrix}$$

Loi Gamma: vraisemblance profilée

Une loi Gamma est a un paramètre vectoriel, $\boldsymbol{\theta} = (\alpha, \beta)$

```
> set.seed(123)
> x = exp(rnorm(100))
> library(MASS)
> (F = fitdistr(x,"gamma"))
     shape        rate
  1.3562877    0.8207035
 (0.1732663)  (0.1263275)

> log_lik = function(theta){
+    a = theta[1]
+    b = theta[2]
+    logL = sum(log(dgamma(x,a,b)))
+    return(-logL)
+ }
> optim(c(1,1),log_lik)
$par
[1] 1.3558113 0.8206505
```



$$\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}, \widehat{\beta}) = \underset{\boldsymbol{\theta} \in \mathbb{R}_+^2}{\text{argmax}} \Big\{ \log \mathcal{L}(\boldsymbol{\theta}) \Big\}$$

est l'estimateur du maximum de vraisemblance

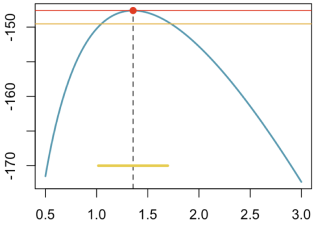
Loi Gamma: vraisemblance profilée

On peut aussi considérer

$$\widehat{\alpha} = \underset{\alpha\in\mathbb{R}_+}{\text{argmax}}\Big\{ \underbrace{\max_{\beta\in\mathbb{R}_+}\Big\{\log\mathcal{L}(\alpha,\beta)\Big\}}_{\text{fonction de } \alpha}\Big\}$$

$\alpha \mapsto \displaystyle\max_{\beta\in\mathbb{R}_+}\Big\{\log\mathcal{L}(\alpha,\beta)\Big\}$ est appelée vraisemblance profilée

```
> prof_log_lik = function(a){
+   b = (optim(1,function(z) -sum(log(dgamma(x,a,z))))$par
+   return(-sum(log(dgamma(x,a,b))))
+ }
> optim(1,prof_log_lik)
$par
[1] 1.356445
```



Loi du Chi-Deux

$$f_\nu(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-y^2/2} y^{\nu-1} \text{ pour } x > 0.$$

$$\mathbb{E}[X] = \nu,\ \mathbb{E}[X^2] = \nu(\nu+2) \text{ et } \text{Var}[X] = 2\nu.$$