STT 1000 - STATISTIQVES

ARTHVR CHARPENTIER

Loi de Poisson

$$f(x) = \mathbb{P}(X = x) = \exp[-\lambda]\frac{\lambda^x}{x!}$$

$$\mathbb{E}[X] = \lambda \text{ et } \mathrm{Var}[X] = \lambda$$

```
> sample_x =
```

Méthode des moments

```
> mean(sample_x)
```

Maximum de Vraisemblance (analytique)

Étant donné un échantillon $\{x_1, \cdots, x_n\}$ de loi $\mathcal{P}(\lambda)$,

$$\mathcal{L}(\lambda; x_1, \cdots, x_n) = \mathbb{P}\big[X_1 = x_1, \cdots, X_n = x_n\big] = \prod_{i=1}^{n} \mathbb{P}[X_i = x_i]$$

$$\mathcal{L}(p; x_1, \cdots, x_n) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \frac{e^{\sum x_i \log(\lambda)}}{\prod x_i!} = Ce^{-n\lambda} e^{n\overline{x}}$$

où C est une *constante*, indépendente de λ,

$$\log \mathcal{L}(\lambda; x_1, \cdots, x_n) = \log C - n\lambda + n\overline{x} \log(\lambda)$$

$$\frac{\partial \log \mathcal{L}(\lambda; x_1, \cdots, x_n)}{\partial \lambda} = -n + \frac{n\overline{x}}{\lambda}$$

de telle sorte que la dérivée s'annule si le numérateur s'annule

$$\left.\frac{\partial \log \mathcal{L}(\lambda; x_1, \cdots, x_n)}{\partial \lambda}\right|_{\lambda = \widehat{\lambda}} = 0 \iff -n + \frac{n\overline{x}}{\widehat{\lambda}} = 0$$

i.e. $\widehat{\lambda} = \overline{x}$

Maximum de Vraisemblance (numérique)

```
> logL = function(lambda){
  sum(log(dpois(sample_x,lambda=lambda)))
}
> logL(.4)
[1] -59.90693
> logL(.7)
[1] -47.44244
> vect_p = seq(.01,.99,by=.01)
> vect_L = Vectorize(logL)(vect_p)
> optim(par = .5,fn = function(p) -logL(p))
$par
[1] 0.6841797

$value
[1] 47.39777
> optimize(f =  function(p) -logL(p), interval = c(0,1))
$minimum
[1] 0.6842177

$objective
[1] 47.39777
```

Approche Bayésienne

Supposons que Λ est ici une variable aléatoire, avec

$$(X|\Lambda=\lambda)\sim\mathcal{P}(\lambda),\ f(x|\lambda)=e^{\lambda}\frac{\lambda^x}{x!},\ x\in\mathbb{N}$$

et un a-priori de loi Gamma pour Λ

$$\Lambda\sim\mathcal{G}(a,b),\ \pi(\lambda)=\frac{p^a(1-p)^b}{B(a,b)},\ p\in\mathbb{R}_+$$

Alors

$$(\Lambda|\boldsymbol{X}=\boldsymbol{x})\sim\mathcal{B}(a+n,b+n\overline{x}),\ \pi(p|\boldsymbol{x})=\frac{p^a(1-p)^b}{B(a,b)},\ p\in[0,1]$$

dont la moyenne et la variance sont

$$\mathbb{E}[\Lambda|\boldsymbol{X}=\boldsymbol{x}]=\frac{a+n\overline{x}}{b+n}\text{ et }\mathrm{Var}[\Lambda|\boldsymbol{X}=\boldsymbol{x}]=\frac{a+n\overline{x}}{(b+n)^2}$$

Comparaison Bayésien vs. Fréquentiste