ECE 2 - Année 2021-2022
ENC Bessières
Mathématiques - F. Gaunard
http://frederic.gaunard.com

Approfondissement

Autour de la loi Gamma

Partie I : Une intégrale à deux paramètres

Soit $t > 0$ un réel fixé.
Pour tout couple (p, q) d'entiers naturels, on pose

$$I(p,q) = \int_0^t x^p (t-x)^q \, \mathrm{d}x.$$

(1) Montrer que , pour tout couple $(p,q) \in \mathbb{N} \times \mathbb{N}$

$$I(p,q) = \frac{q}{p+1} I(p+1, q-1).$$

(2) En déduire que, pour tout couple $(p,q) \in \mathbb{N} \times \mathbb{N}$

$$I(p,q) = \frac{p!q!}{(p+q)!} I(p+q, 0).$$

(3) Déterminer$I(p+q,0)$ et montrer finalement que

$$I(p,q) = \frac{p!q!}{(p+q+1)!}.$$

Partie II : Loi Gamma

Pour tout $k \in \mathbb{N}$, on pose

$$J_k = \int_0^{+\infty} x^k e^{-x} \mathrm{d}x.$$

(4) (a) Justifier que, pour tout $k \in \mathbb{N}$, l'intégrale J_k est bien convergente. Calculer J_0.
(b) Montrer par récurrence que $J_{k+1} = (k+1)J_k$. En déduire que $J_k = k!$.

Soit $\lambda > 0$ un réel fixé.
Pour $k \in \mathbb{N}^*$, on introduit alors la fonction g_k définie sur $\mathbb{R}$ par

$$g_k(x) = \begin{cases} 0, & \text{si } x < 0 \\ \dfrac{\lambda}{(k-1)!} e^{-\lambda x} (\lambda x)^{k-1}, & \text{si } x \geq 0 \end{cases}$$

(5) Montrer que g_k peut être considérée comme une densité de probabilité.
On note X une variable de densité g_k. On dit que X suit la *loi gamma* de paramètres k et λ et on note $X \hookrightarrow \Gamma(k, \lambda)$.

(6) Reconnaître la loi $\Gamma(1, \lambda)$.

(7) Montrer que X admet une espérance et une variance et que

$$E(X) = \frac{k}{\lambda}, \quad \text{et} \quad V(X) = \frac{k}{\lambda^2}.$$

(8) On pose $Y = 1/X$. Montrer que Y admet une espérance et une variance et que

$$E(Y) = \frac{\lambda}{k-1}, \quad \text{et} \quad V(Y) = \frac{\lambda^2}{(k-1)^2(k-2)}.$$

Partie III : Somme(s) de variables aléatoires indépendantes

Dans toute cette partie, on considère encore un paramètre réel $\lambda > 0$.

Si f et g sont deux fonctions définies sur $\mathbb{R}$, le *produit de convolution* de f et de g est la fonction $f \star g$ définie par

$$(f \star g)(t) = \int_{-\infty}^{+\infty} f(x)g(t-x)\mathrm{d}x$$

définie pour tout réel x où l'intégrale ci-dessus converge.

On **admet** le résultat suivant. Si X et Y sont deux variables aléatoires de densités respectives f et g alors $X + Y$ est encore une variable aléatoire dont une densité est donnée par $f \star g$.

(9) Soient Y_1, Y_2 deux variables aléatoires indépendantes de même loi $\mathcal{E}(\lambda)$.
  (a) Rappeler l'expression d'une densité f de Y_1 (aussi densité de Y_2).
  (b) Soient $x, t \in \mathbb{R}$. Montrer que

$$f(x)f(t-x) \neq 0 \Longleftrightarrow 0 \leq x \leq t$$

  (c) Montrer alors que $Y_1 + Y_2 \hookrightarrow \Gamma(2, \lambda)$.

(10) (a) Soient $k, n \in \mathbb{N}^*$. Montrer, à l'aide de la Question (3) de la Partie I, que $(g_k \star g_n) = g_{k+n}$ où g_i représente une densité de la loi $\Gamma(i, \lambda)$ introduite dans la Partie II.
  (b) En déduire que si X et Y sont deux variables aléatoires indépendantes de lois respectives $\Gamma(k, \lambda)$ et $\Gamma(n, \lambda)$, alors $X + Y \hookrightarrow \Gamma(k+n, \lambda)$.

(11) (a) Déduire des Questions (9) et (10) que si $(Y_1, ..., Y_n)$ est un $n-$échantillon de la loi $\mathcal{E}(\lambda)$, alors

$$S_n = Y_1 + Y_2 + ... + Y_n \hookrightarrow \Gamma(n, \lambda).$$

  (b) Retrouver les résultats de la Question (7) de la Partie II.
  (c) Soit (S_n) une suite de variables aléatoires telle que $S_n \hookrightarrow \Gamma(n, \lambda)$. Montrer que

$$\frac{\lambda}{\sqrt{n}} S_n - \sqrt{n} \xrightarrow{\mathcal{L}} Z, \qquad Z \hookrightarrow \mathcal{N}(0, 1).$$

Partie IV - Estimation du paramètre d'une loi exponentielle par le maximum de vraisemblance

On suppose que le paramètre λ est inconnu que l'on souhaite estimer à partir d'un $n-$échantillon $Y_1, Y_2, .., Y_n$ de la loi exponentielle $\mathcal{E}(\lambda)$ (avec $n \geq 2$). On note f une densité de Y_i.
On utilise la méthode du *maximum de vraisemblance*.

(12) On considère des réels $x_1, x_2, .., x_n$ strictement positifs, ainsi que la fonction L, à valeurs dans $\mathbb{R}$, définie sur $]0;+\infty[$ par

$$L(\lambda) = \prod_{k=1}^{n} f(x_k), \qquad \lambda > 0.$$

(13) Exprimer $L(\lambda)$, puis $\ln(L(\lambda))$ en fonction de λ, $x_1, ..., x_n$.

(14) On considère la fonction φ, définie pour tout réel $\lambda > 0$ par

$$\varphi(\lambda) = n\ln(\lambda) - \lambda \sum_{=1}^{n} x_k.$$

Montrer que la fonction φ admet un maximum, atteint en un seul réel que l'on notera z et que l'on exprimera en fonction de $x_1, x_2, ..., x_n$. Que peut-on dire de z pour la fonction L?

On pose dorénavant, toujours avec $n \geq 2$,

$$Z_n = \frac{n}{Y_1 + Y_2 + ... + Y_n}.$$

L'estimateur Z_n est appelée *estimateur du maximum de vraisemblance* pour λ.

(15) À l'aide des résultats des Parties II et III, montrer que

$$E(Z_n) = \frac{n}{n-1}\lambda, \quad \text{et} \quad V(Z_n) = \frac{n^2}{(n-1)^2(n-2)}\lambda^2.$$

et déduire que Z_n est biaisé mais asymptotiquement sans biais pour λ.

(16) En déduire un estimateur $\tilde{Z}_n$ non biaisé pour λ. Quel est son risque quadratique?

(17) Soit $\alpha \in]0;1[$. On note $t_\alpha = \Phi^{-1}(1-\alpha/2)$ où Φ désigne la fonction de répartition de la loi $\mathcal{N}(0,1)$. À l'aide de la Question (11c) de la Partie III, montrer que

$$\left[Z_n\left(1 - \frac{t_\alpha}{\sqrt{n}}\right) ; \ Z_n\left(1 + \frac{t_\alpha}{\sqrt{n}}\right) \right]$$

est un intervalle de confiance asymptotique au seuil $1-\alpha$ pour λ.

(18) Compléter le programme `SciLab` suivant sachant qu'il a permis d'afficher le résultat ci-contre

```
N=1000
alpha=0.05
Y=grand(1,N, 'exp', ........)
Z=length(Y)/sum(Y)
t=cdfnor("X",0,1, 1-alpha/2, alpha/2)
A=Z*(1-t/sqrt(N))
B=Z*(1+t/sqrt(N))
c=(N-1)/N
disp(c*Z)
disp([A,B])
```

```
2.0405910

1.9160322   2.1692351
```