Travail pratique 2 ACT-2001

Olivier Bourret

Numéro 1

a) Comparer les queues de distribution

Tout d'abord, on doit valider que l'espérence des distribtions est la même.

$$E[X] = \frac{1}{\lambda} = \frac{1}{0.02} = 50 = E[Y]$$

Donc,

$$\begin{aligned}
\lim_{x\to\infty} \frac{f_X(x)}{f_Y(x)} &= \lim_{x\to\infty} 0.02e^{-0.02x} \cdot \frac{\sqrt{2\pi}}{e^{-1/2(x-50)^2}} \\
&= \lim_{x\to\infty} \frac{0.02\sqrt{2\pi}e^{-0.02x}}{e^{-1/2x^2}e^{50x}e^-1250} \\
&= \lim_{x\to\infty} 0.02\sqrt{2\pi}e^{0.5(x^2-100.04x+2500)} \\
&= \lim_{x\to\infty} 0.02\sqrt{2\pi}e^{0.5((x-50.02)^2-2.0004} \quad \longrightarrow \text{Completion de carré} \\
&= 0.02\sqrt{2\pi}e^{\infty} \\
&= \infty
\end{aligned}$$

Alors, puisque $\lim_{x\to\infty} \frac{f_X(x)}{f_Y(x)} = \infty$, X a une queue plus lourde que Y.

b) X est light tail ou heavy tail?

Une des premières façon de vérifier, c'est de regarder la f.g.m. de la loi de Burr. Puisqu'elle n'exite pas, alors X est *"Heavy tail"*.

Une deuxième est de regarder avec la méthode des moments si elle existe toujours ou non. On sait que $E[X^k] = \frac{\lambda^{1/\tau}}{\Gamma(\alpha)}\Gamma(1+\frac{1}{\tau})\Gamma(\alpha-\frac{k}{\tau}),\ -\tau < k < \alpha\tau$. Puisque k est limité dans l'intervalle $(-\tau, \alpha\tau)$, alors $E[X^k]$ existe jusqu'à un maximum de $k \leq \alpha\tau$. Nous pouvons donc dire par la méthode des moments que X est *"Heavy tail"*

Une troisième manière de déterminer la queue de distribition est de la comparer à une loi exponentielle. Voici la manière que j'ai utilisée.

$$\begin{aligned}
\lim_{x\to\infty} e^{rx}\overline{F}_X(x) &= \lim_{x\to\infty} e^{rx}\frac{\lambda^\alpha}{(\lambda+x^\tau)^\alpha} \\
&= \lim_{x\to\infty} \frac{\lambda^\alpha e^{rx}}{x^{\tau\alpha}\left(\frac{\lambda}{x^\tau}+1\right)^\alpha} \quad \longrightarrow \text{Exponentielle croit plus vite qu'un polynôme} \\
&= \infty \forall r > 0
\end{aligned}$$

Puisque $\lim_{x\to\infty} e^{rx}\overline{F}_X(x) = \infty,\ \forall\ r > 0$, on peut dire que X est *"Heavy tail"*.

Numéro 2

a) Trouver la distribution de S et calculer $E[S]$, $Var(S)$ et la $Pr(S > 1000)$

Pour des X_i indépendants,

$$\begin{aligned} M_S(t) = E[e^{St}] &= E[e^{(X_1+...+X_{20})t}] \\ &= E[e^{X_1 t}] \cdot ... \cdot E[e^{X_{20} t}] \\ &= E[e^{Xt}]^{\sum_{i=1}^{20} n_i} \\ &= \left(\frac{\beta}{\beta + t}\right)^{\sum_{i=1}^{20} n_i} \\ &= \left(\frac{0.05}{0.05 + t}\right)^{210} \Rightarrow S \sim Erlang(n = 210, \beta = 0.05) \end{aligned}$$

Pour l'espérance:

$$E[S] = \frac{\alpha}{\beta} = \frac{210}{0.05} = 4\,200$$

Pour la variance:

$$Var(S) = \frac{\alpha}{\beta^2} = \frac{210}{0.05^2} = 84\,000$$

Pour $Pr(S > 1000)$:

$$\begin{aligned} Pr(S > 1000) = \overline{F}_S(1000) &= \sum_{j=0}^{209} \frac{(0.05(1000))^j}{j!} \\ &= 1 \quad \longrightarrow \text{Calculé avec R} \end{aligned}$$

Pour le calcul de la fonction de répartition, j'ai utilisé la fonction r $qgamma(1000, 210, 0.05)$, mais cela génère de NaN, pusiqu'on obitent une division par l'infini dans le calcul. Alors, j'ai vérifié la $Pr(S* > 1000)$ pour $S* = X_1 + ... + X_{18}$ de manière similaire pour S. Puisque la probabilité donne 1, en ajoutant X_{19} et X_{20}, la probabilité est encore 1.

b) Utiliser l'approximation normale pour calculer $E[S]$, $Var(S)$ et la $Pr(S > 1000)$

$$\begin{aligned} E[S] = \sum_{i=1}^{20} \mu_{X_i} &= \frac{n_1}{\beta} + \frac{n_2}{\beta} + ... + \frac{n_{20}}{\beta} \\ &= \frac{1}{0.05} + \frac{2}{0.05} + ... + \frac{20}{0.05} \\ &= \frac{210}{0.05} = 4\,200 \end{aligned}$$

$$
\begin{aligned}
Var(S) = \sum_{i=1}^{20} \sigma {X_i}^2 &= \frac{n_1}{\beta^2} + \frac{n_2}{\beta^2} + ... + \frac{n_{20}}{\beta^2} \\
&= \frac{1}{0.05^2} + \frac{2}{0.05^2} + ... + \frac{20}{0.05^2} \\
&= \frac{210}{0.05^2} = 84\,000
\end{aligned}
$$

$$
\begin{aligned}
Pr(S > 1\,000) &= Pr\left(\frac{S - E[S]}{\sqrt{Var(S)}} > \frac{1\,000 - 4\,200}{\sqrt{84\,000}}\right) \\
&= Pr(Z > -11.041) \\
&= \Phi(11.041) = 1
\end{aligned}
$$

Numéro 3

a) Trouver λ et la fonction de répartition de C

$$\lambda_S = \lambda_1 + \lambda_2 + \lambda_3 = (4-1) + (4-2) + (4-3) = 6$$

$$\begin{aligned}
F_C(x) &= \frac{\lambda_1}{\lambda_S} \cdot F_{B_1}(x) + \frac{\lambda_2}{\lambda_S} \cdot F_{B_2}(x) + \frac{\lambda_3}{\lambda_S} \cdot F_{B_3}(x) \\
&= \frac{3}{6}(1 - e^{-0.05x}) + \frac{2}{6}(1 - e^{-0.05x}[1 + 0.05x]) + \frac{1}{6}(1 - e^{-0.05x}[1 + 0.05x + \frac{(0.05x)^2}{2}]) \\
&= 1 - e^{-0.05x} - 0.025xe^{-0.05x} - \frac{0.0025}{12}x^2e^{-0.05x}
\end{aligned}$$

b) Calculer la $Pr(B_{2,1} + B_{2,2} \leq 40)$

En utilisé la fgm, on a déjà vu qu'une somme de loi Erlang donne aussi une loi Erlang, alors $B_{2,1} + B_{2,2} \sim Erlang(n = 2 + 2 = 4, \beta = 0.05)$. Alors,

$$Pr((B_{2,1} + B_{2,2} \leq 40) = 1 - e^{0.05(40)} \left[1 + 0.05(40) + \frac{(0.05(40))^2}{2} + \frac{(0.05(40))^3}{3!}\right] = 0.142\,876\,54$$

c) Calculer la $Pr(C_1 + C_2 \leq 120)$

Pour C_1 et C_2 i.i.d. on peut trouver le résultat rechercher de la manière suivante.

$$\begin{aligned}
Pr(C_1 + C_2 \leq 120) &= Pr(C_2 \leq 120 - C_1) \\
&= \int_0^{120} \int_0^{120-c_1} f_{C_1,C_2}(c_1, c_2) dc_2\, dc_1 \\
&= \int_0^{120} \int_0^{120-c_1} f_C(c_1) \cdot f_C(c_2) dc_2\, dc_1 \\
&= \int_0^{120} f_C(c_1) \cdot F_C(120 - c_1) dc_1
\end{aligned}$$

Trouvons quelques résutats pour nous aider à faire cette intégrale.

$$\begin{aligned}
F_C(120 - c_1) &= 1 - e^{-0.05(120-c_1)} - 0.025(120 - c_1)e^{-0.05(120-c_1)} - \frac{0.0025}{12}(120 - c_1)^2e^{-0.05(120-c_1)} \\
&= 1 - 7e^{-6}e^{0.05c_1} + 0.075c_1e^{-6}e^{0.05c_1} - \frac{0.0025}{12}c_1^2e^{-6}e^{0.05c_1} \\
&= \frac{1}{4\,800}\left(4\,800 - 33\,600e^{-6}e^{0.05c_1} + 360c_1e^{-6}e^{0.05c_1} - c_1^2e^{-6}e^{0.05c_1}\right)
\end{aligned}$$

$$
\begin{aligned}
f_C(x) &= \frac{d}{dx}\left[F_C(x)\right] \\
&= \frac{d}{dx}\left[1 - e^{-0.05x} - 0.025xe^{-0.05x} - \frac{0.0025}{12}x^2e^{-0.05x}\right] \\
&= -.05e^{-0.05x} - 0.025e^{-0.05x} + 0.00125xe^{-0.05x} - \frac{0.005}{12}xe^{-0.05x} + \frac{0.000125}{12}x^2e^{-0.05x} \\
&= \frac{1}{96\,000}\left(2\,400e^{-0.05x} + 80xe^{-0.05x} + x^2e^{-0.05x}\right)
\end{aligned}
$$

$$
\begin{aligned}
f_C(c1)\cdot F_C(120-c_1) &= \frac{1}{96\,000}\left(2\,400e^{-0.05c_1} + 80c_1e^{-0.05c_1} + c_1^2e^{-0.05c_1}\right) \\
&\quad \cdot \frac{1}{4\,800}\left(4\,800 - 33\,600e^{-6}e^{0.05c_1} + 360c_1e^{-6}e^{0.05c_1} - c_1^2e^{-6}e^{0.05c_1}\right) \\
&= \frac{1}{460\,800\,000}(11\,520\,000e^{-0.05c_1} + 384\,000c_1e^{-0.05c_1} + 4\,800c_1^2e^{-0.05c_1} \\
&\quad - 80\,640\,000e^{-6} - 2\,688\,000e^{-6}c_1 - 33\,600e^{-6}c_1^2 \\
&\quad + 864\,000e^{-6}c_1 + 28\,800e^{-6}c_1^2 + 360e^{-6}c_1^3 \\
&\quad - 2\,400e^{-6}c_1^2 - 80e^{-6}c_1^3 - e^{-6}c_1^4) \\
&= \frac{1}{460\,800\,000}(11\,520\,000e^{-0.05c_1} + 384\,000c_1e^{-0.05c_1} + 4\,800c_1^2e^{-0.05c_1} \\
&\quad - 80\,640\,000e^{-6} - 1\,824\,000e^{-6}c_1 - 7\,200e^{-6}c_1^2 + 280e^{-6}c_1^3 - e^{-6}c_1^4)
\end{aligned}
$$

Alors, on peut faire l'intégrale suivante que l'on séparera de la manière suivante pour alléger sa lecture

$$
\begin{aligned}
\int_0^{120} f_C(c_1)\cdot F_C(120-c_1) = \int_0^{120} &\frac{1}{460\,800\,000}(11\,520\,000e^{-0.05c_1} + 384\,000c_1e^{-0.05c_1} + 4\,800c_1^2e^{-0.05c_1} \\
&- 80\,640\,000e^{-6} - 1\,824\,000e^{-6}c_1 - 7\,200e^{-6}c_1^2 + 280e^{-6}c_1^3 - e^{-6}c_1^4)dc_1
\end{aligned}
$$

$$
\begin{aligned}
= \frac{1}{460\,800\,000}\Bigg[&\int_0^{120} 11\,520\,000e^{-0.05c_1}dc_1 && \textbf{(1)} \\
&+ \int_0^{120} +384\,000c_1e^{-0.05c_1}dc_1 && \textbf{(2)} \\
&+ \int_0^{120} 4\,800c_1^2e^{-0.05c_1}dc_1 && \textbf{(3)} \\
&- \int_0^{120} 80\,640\,000e^{-6}dc_1 && \textbf{(4)} \\
&- \int_0^{120} 1\,824\,000e^{-6}c_1dc_1 && \textbf{(5)} \\
&- \int_0^{120} 7\,200e^{-6}c_1^2dc_1 && \textbf{(6)} \\
&+ \int_0^{120} 280e^{-6}c_1^3dc_1 && \textbf{(7)} \\
&- \int_0^{120} e^{-6}c_1^4)dc_1\Bigg] && \textbf{(8)}
\end{aligned}
$$

Donc, on a pour chaque intégrale.

$$\textbf{(1)} \qquad 11\,520\,000\int_0^{120} e^{-0.05c_1}dc_1 = 11\,520\,000\frac{e^{-0.05c_1}}{-0.05}\Bigg|_{c_1=0}^{120}$$
$$= 229\,828\,895.5$$

$$\textbf{(2)} \qquad 384\,000\int_0^{120} c_1e^{-0.05c_1}dc_1 = 384\,000\left[\frac{xe^{-0.05c_1}}{-0.05}\Bigg|_{c_1=0}^{120} - \int_0^{120}\frac{e^{-0.05c_1}}{-0.05}dc_1\right]$$
$$= 384\,000\left[\frac{xe^{-0.05c_1}}{-0.05}\Bigg|_{c_1=0}^{120} - \frac{e^{-0.05c_1}}{(0.05)^2}\Bigg|_{c_1=0}^{120}\right]$$
$$= 384\,000 \cdot 393.059\,494$$
$$= 150\,934\,845.7$$

$$\textbf{(3)} \qquad 4\,800\int_0^{120} c_1^2e^{-0.05c_1}dc_1 = 4\,800\left[\frac{c_1^2e^{-0.05c_1}}{-0.05}\Bigg|_{c_1=0}^{120} + 40\int_0^{120} c_1e^{-0.05c_1}dc_1\right]$$

$*$Le résultat de $\int_0^{120} c_1e^{-0.05c_1}dc_1$ a été calculé en (2)

$$= 4\,800 \cdot 15\,008.499$$
$$= 72\,040\,795.82$$

$$\textbf{(4)} \qquad -\,80\,840\,000\int_0^{120} e^{-6}dc_1 = -80\,840\,0000e^{-6}c_1\big|_{c_1=0}^{120}$$
$$= -23\,986\,389.06$$

$$\textbf{(5)} \qquad -\,1\,824\,000\int_0^{120} e^{-6}c_1dc_1 = -1\,824\,000e^{-6}\frac{c_1^2}{2}\Bigg|_{c_1=0}^{120}$$
$$= -32\,552\,956.59$$

$$\textbf{(6)} \qquad 7\,200\int_0^{120} e^{-6}c_1^2dc_1 = 7\,200e^{-6}\frac{c_1^3}{3}\Bigg|_{c_1=0}^{120}$$
$$= -10\,279\,881.03$$

$$\textbf{(7)} \qquad 280\int_0^{120} e^{-6}c_1^3dc_1 = 280e^{-6}\frac{c_1^4}{4}\Bigg|_{c_1=0}^{120}$$
$$= 35\,979\,583.59$$

$$\int_0^{120} -e^{-6} c_1^4 dc_1 = -e^{-6} \frac{c_1^5}{5}\bigg|_{c_1=0}^{120} = -12\,335\,857.23 \tag{8}$$

Ainsi, je peux calculer le résultat finale en additionnant toutes les valeurs calculées.

$$Pr(C_1 + C_2 \leq 120) = \frac{409\,629\,036.7}{460\,800\,000} = 0.888\,951\,903$$

Numéro 4

a) Estimer $E[X_i]$ et la $Var(X_i)$ pour $i = \{1, 2, 3\}$

Grâce au paquetage actuar, il est simple de simuler le 100 000 données avec les fonctions rcompound et rcomppois. Voici ce que j'ai fait pour simuler mes réalisations et calculer la moyenne et la variance.

```r
library(actuar)

# Simulation des 100000 réalisations pour les X_i
BinComp_sim <- rcompound(100000, rbinom(2,0.5), rpareto(3,600))
PoiComp_sim <- rcomppois(100000, 2, rgamma(3,0.01))
BinNegComp_sim <- rcompound(100000, rnbinom(2,0.5), rlnorm(4,2))


# Calcul de E[X_i]
Esp_X1 <- mean(BinComp_sim)
Esp_X2 <- mean(PoiComp_sim)
Esp_X3 <- mean(BinNegComp_sim)


# Calcul de la Var(X_i)
Var_X1 <- var(BinComp_sim)
Var_X2 <- var(PoiComp_sim)
Var_X3 <- var(BinNegComp_sim)
```

Voici les résultats obtenus

Table 1: Espérance et variance

X_i	$E[X_i]$	$Var(X_i)$
1	301.98	286995
2	599.55	239645
3	815.71	30578221

b) Estimer la fonction stop-loss avec $d = 200$ ($\pi_{X_i}(200)$

Avec les valeurs simulées précédemment, j'ai pu calculer la valeur de la fonction stop-loss avec un déductible de 200. Voici le code informatique utilisé.

```r
# Calcule de la fonction stop-loss
EsSL_X1 <- mean(pmax(BinComp_sim-200,0))
EsSL_X2 <- mean(pmax(PoiComp_sim-200,0))
EsSL_X3 <- mean(pmax(BinNegComp_sim-200,0))
```

J'obtiens les résultats suivants:

Table 2: Fonction stop-loss

π_{X_1}	π_{X_2}	π_{X_3}
190.37	432.68	710.26

c) Estimer la $VaR_{90}(X_i)$ et la $TVaR_{90}(X_i)$

En réutilisant les simulations en a) et en triant en ordre croissant les données, on peut trouver la $VaR_p(X_i)$ et $TVaR_p(X_i)$ ainsi:

```r
# VaR(X_i)
VaR90_X1 <- sort(BinComp_sim)[90000]
VaR90_X2 <- sort(PoiComp_sim)[90000]
VaR90_X3 <- sort(BinNegComp_sim)[90000]

# TVaR(X_i)
TVaR90_X1 <- mean(sort(BinComp_sim)[90001:100000])
TVaR90_X2 <- mean(sort(PoiComp_sim)[90001:100000])
TVaR90_X3 <- mean(sort(BinNegComp_sim)[90001:100000])
```

Voici les résultats obtenus:

Table 3: VaR et TVaR

X_i	$VaR_p(X_i)$	$TVaR_p(X_i)$
1	778.20	1461.49
2	1269.93	1623.98
3	1753.27	5842.19

d) Estimer $\pi_S(600)$, la $VaR_{90}(S)$ et la $TVaR_{90}(S)$

Supposons que X_1, X_2 et X_3 sont indépendants. Alors, on additionne les valeurs que nous avons simulés en a) (les résultats non triés) et les calcules restent les mêmes que ceux précédents. J'obtiens:

```r
S <- BinComp_sim + PoiComp_sim + BinNegComp_sim

EsSL_S <- mean(pmax(S-600,0))
VaR90_S <- sort(S)[90000]
TVaR90_S <- mean(sort(S)[90001:100000])
```

Table 4: Stop-loss, VaR et TVaR de S

π_S	$VaR_p(S)$	$TVaR_p(S)$
1178.66	3047.79	7044.66

Numéro 5

Dans ce numéro, il est demandé de déterminer si la mesure de risque ρ est convexe. Pour être convexe, elle doit remplir 3 critères. Commençons (de manière presque pas aléatoire) avec le 2^e critère.

$$\begin{aligned}
\rho(X + c) &= 1.2E[X + c] \\
&= 1.2(E[X] + E[c]) \\
&= 1.2E[X] + 1.2c \\
&= \rho(X) + 1.2c \neq \rho(X) + c
\end{aligned}$$

Donc, puisque $\rho(X)$ n'est pas invariante par translation, alors, elle n'est pas une mesure de risque convexe.