Équipe no XX

Nicholas Langevin
(111 184 631)

Alexandre Turcotte
(111 172 613)

Mathématiques actuarielles IARD 1
ACT-2005

Travail pratique 2

Travail présenté à
Andrew Luong

École d'actuariat
Université Laval
Automne 2018

Question 1

Méthode de vraisemblance avec données complètes

a) Estimation des paramètres

Le jeu de données utilisé afin de compléter ce numéro comportait les 10 données suivantes :

1500 6000 3500 3800 1800 5500 4800 4200 3900 3000

À partir de ces données et de SAS, il a été possible d'estimer les paramètres $\hat{\mu}$ et $\hat{\sigma}$ d'une loi log-normale à partir de la méthode du maximum de vraisemblance. D'ailleurs, le code SAS se trouve dans l'Annexe à la fin du rapport. Voici, les estimés des paramètres obtenus:

$\hat{\mu}$	$\hat{\sigma}$
8.1618	0.4278

Table 1: Paramètres estimés d'une loi log-normale à l'aide du maximum de vraisemblance

Il est également possible de déterminer les estimateurs de façon théorique. D'abord, il faut déterminer la fonction $g(x_i;\mu,\sigma)$ pour les dix scénarios, et ce, en fonction de la densité de X.

Alors comme $X \sim LN(\mu,\sigma)$, la densité de f(x) est la suivante :

$$f_x(x) = \frac{1}{\sqrt{2\pi}x\sigma} e^{\frac{-1}{2}\left(\frac{In(x)-\mu}{\sigma}\right)^2}$$

Ensuite, il faut appliquer la vraisemblance en tenant compte des 10 scénarios et de la densité de X.

$$\text{Ł}(\mu,\sigma) = \prod_{i=1}^{10} g(x_i;\mu,\sigma)$$

$$= f_x(1500)\cdot f_x(6000)\cdot ...\cdot f_x(3000)$$

$$= \left[\frac{1}{\sqrt{2\pi}1500\sigma} e^{\frac{-1}{2}(\frac{In(1500)-\mu}{\sigma})^2}\right]\cdot\left[\frac{1}{\sqrt{2\pi}6000\sigma} e^{\frac{-1}{2}(\frac{In(6000)-\mu}{\sigma})^2}\right]\cdot ...\cdot\left[\frac{1}{\sqrt{2\pi}3000\sigma} e^{\frac{-1}{2}(\frac{In(3000)-\mu}{\sigma})^2}\right]$$

$$= \left(\frac{1}{2\pi}\right)^5\left(\frac{1}{\sigma}\right)^{10}\left(\frac{1}{2.795153962\times 10^{35}}\right) e^{\frac{-1}{2\sigma^2}\left[(In(1500)-\mu)^2+(In(6000)-\mu)^2+...+(In(3000)-\mu)^2\right]}$$

Par la suite, il faut appliquer la log vraisemblance.

$$
\begin{aligned}
l(\mu, \sigma) = In(L(\mu, \sigma)) &= In \prod_{i=1}^{10} g(x_i; \mu, \sigma) \\
&= -5In(2\pi) - 10In(\sigma) - In(2.795153962 \times 10^{35}) \\
&- \frac{1}{2\sigma^2}\left[(In(1500) - \mu)^2 + (In(6000) - \mu)^2 + ... + (In(3000) - \mu)^2\right]
\end{aligned}
$$

Finalement, il faut dériver la log vraisemblance selon les deux paramètres et égaliser les deux équations à 0 pour pouvoir estimer les paramètres en les isolant.

Dériver selon μ

$$
\begin{aligned}
\frac{d}{d\mu} l(\mu, \sigma) &= \frac{-1}{2\sigma^2}\left[-2(In(1500) - \mu) - 2(In(6000) - \mu) - ... - 2(In(3000) - \mu)\right] \\
0 &= \frac{1}{\hat{\sigma}^2}\left[In(1500) + In(6000) + ... + In(3000) - 10\hat{\mu}\right] \\
\hat{\mu} &= \frac{1}{10}\sum_{i=1}^{10} In(x_i) \ = \ 8.161836544
\end{aligned}
$$

Dériver selon σ

$$
\begin{aligned}
\frac{d}{d\sigma} l(\mu, \sigma) &= \frac{-10}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{10}(In(x_i) - \mu)^2 \\
0 &= \frac{-10}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3}\sum_{i=1}^{10}(In(x_i) - \hat{\mu})^2 \\
\hat{\sigma}^2 &= \frac{1}{10}\sum_{i=1}^{10}(In(x_i) - 8.161836544)^2 \ = \ 0.1830137502 \\
\hat{\sigma} &= \sqrt{\hat{\sigma}^2} \ = \ 0.427801064
\end{aligned}
$$

Donc, les estimateurs de la loi log-normale, selon la méthode du maximum de vraisemblance, sont $\hat{\mu} = 8.1618$ et $\hat{\sigma} = 0.4278$.

b) Détermination des primes stop-loss

En ayant obtenu les estimés des paramètres à la question a), il est maintenant possible d'effectuer divers calculs puisque la distribution du modèle est maintenant connue. Par conséquent, il est possible de calculer la prime stop-loss $E[(X-d)_+]$, et ce, pour différentes valeurs de déductible, soit d = 2000, 2100, ..., 3000. Pour y arriver, il faut utiliser la formule suivante:

$$E\left[(X-d)_+\right] = e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}}\left(1-\Phi\left(\frac{Ind-\hat{\mu}-\hat{\sigma}^2}{\hat{\sigma}}\right)\right) - d\left(1-\Phi\left(\frac{Ind-\hat{\mu}}{\hat{\sigma}}\right)\right)$$

où $\hat{\mu} = 8.1618$, $\hat{\sigma} = 0.4278$ et d = déductible. Par conséquent, les valeurs obtenues pour la prime stop-loss sont les suivantes:

d	**Prime stop-loss**
2000	253.3579
2100	236.1653
2200	220.3997
2300	205.9213
2400	192.6061
2500	180.3436
2600	169.0355
2700	158.5939
2800	148.9402
2900	140.0041
3000	131.7223

Table 2: Valeurs de la prime stop-loss pour chacun des déductibles d donnés

À partir des valeurs du tableau 2, il est possible d'observer une baisse du niveau de la prime lorsque le déductible augmente. De plus, en regardant le graphique (A) de la figure 1, il est possible d'observer cette décroissance qui semble à première vue assez linéaire. Toutefois, en regardant le graphique (B) de la figure 1, ayant un plus grand nombre de valeurs de d, on remarque la convexité de la courbe, ce qui signifie que plus d augmente, moins la diminution est significative. Par exemple, l'écart entre 2000 et 2100 est plus important que l'écart entre 2900 et 3000. Bref, cela fait du sens puisqu'en ajoutant un déductible, la compagnie vient diminuer le risque ainsi que le montant de prestation qu'elle devra verser. Par conséquent, le niveau de la prime devrait diminuer, et ce, avec un impact de moins en moins significatif, car la valeur ajoutée d'augmenter le déductible pour la compagnie est de moins en moins importante.

c) Détermination des primes limited loss

Maintenant, toujours en utilisant les paramètres estimés en a), il faut désormais calculer l'espérance limitée, sans déductible, et ce, pour différentes limites supérieure, soit u = 3000, 3100, ..., 4000. Pour y arriver, il faut utiliser la formule suivante:

$$E\left[X\Lambda u\right] = e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}}\Phi\left(\frac{Inu-\hat{\mu}-\hat{\sigma}^2}{\hat{\sigma}}\right) + u\left(1-\Phi\left(\frac{Inu-\hat{\mu}}{\hat{\sigma}}\right)\right)$$

où $\hat{\mu} = 8.1618$, $\hat{\sigma} = 0.4278$ et u = limite. Par conséquent, les valeurs obtenues pour la prime limitée sont les suivantes:

u	**Espérance limitée**
3000	1187.0103
3100	1194.6945
3200	1201.8320
3300	1208.4691
3400	1214.6471
3500	1220.4038
3600	1225.7731
3700	1230.7858
3800	1235.4700
3900	1239.8511
4000	1243.9524

Table 3: Valeurs de l'espérance limitée pour chacune des limites u données

À partir des valeurs du tableau 3, il est possible d'observer une hausse du niveau de la prime lorsque la limite augmente. De plus, en regardant le graphique (C) de la figure 1, il est possible d'observer cette croissance. D'ailleurs, en regardant le graphique (D) de la figure 1, ayant un plus grand nombre de valeurs u, on remarque la concavité de la courbe, ce qui signifie que plus u augmente, moins l'augmentation est significative. Bref, cela fait du sens puisqu'en ayant une limite plus faible, la compagnie vient diminuer le risque ainsi que le montant de prestation qu'elle devra verser puisqu'elle ne payera pas ce qui a en excédent. Par conséquent, le niveau de la prime devrait augmenter plus la limite est grande, car elle devrait débourser de plus grande somme, et ce, avec un impact de moins en moins significatif, car, à un certain point, le fait d'augmenter la limite va affecter peu de gens, donc la hausse des sommes à débourser sera moins importante.

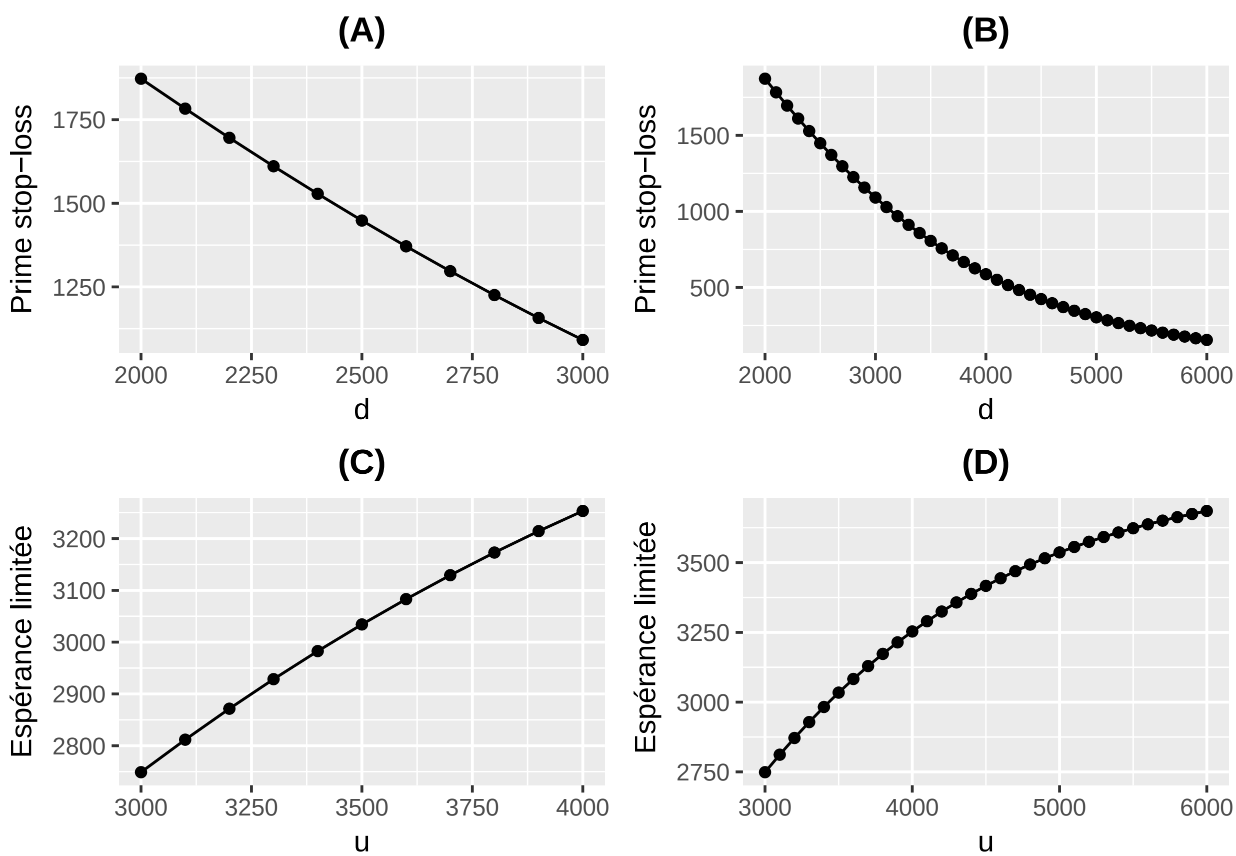

Figure 1 : (A) et (B) Représentation graphique de la prime stop-loss en fonction de d. (C) et (D)

Représentation graphique de l'espérance limitée en fonction de u. Toutefois, l'étendu des valeurs sur l'axe des abscisses est plus importante pour les graphiques (B) et (D) afin de mieux discerner la tendance.

d) Détermination de la prime stop-loss par simulation avec un déductible d = 2000

Il est également possible d'approximer la prime stop-loss par simulation. Alors, il est possible de tirer un échantillon de la loi log-normale suffisament grand pour estimer la prime stop-loss. Cela a été fait avec un m = 100 000, 500 000 et 1 000 000. Pour y arriver, il faut utiliser la formule suivante:

$$E\left[(X-2000)_+\right] = \frac{1}{m}\sum_{i=1}^{m} max(0, X-2000)$$

Par conséquent, les valeurs obtenues pour la prime stop-loss sont les suivantes:

m	approximation
100000	254.4107
500000	253.3894
1000000	253.4287

Table 4: Valeurs approximées par simulations de la prime stop-loss ayant un déductible de 2000

La valeur de la prime stop-loss calculée en b) pour un déductible de 2000 est de 253.3579. Puis, en regardant les résultats obtenus à l'aide de la simulation, on se rend compte que ceux-ci sont très près de la valeur théorique. En effet, les écarts relatifs sont de 0.42%, 0.012% et 0.028%. Par conséquent, l'approximation est très précise et donne sensiblement la même valeur que la valeur théorique.

Question 2

Méthode de vraisemblance avec données groupées

a) Estimation des paramètres

Le jeu de données utilisé afin de compléter ce numéro comportait les 10 données suivantes :

Intervalles	nombres
0-250	3
250-500	4
500-750	5
750-1000	5
1000-1250	4
1250-1500	3
1500-2000	3
2000-2500	3
2500-3000	2
3000-5000	2

Table 5: Jeux de données de la question 2

À partir de ces données et de SAS, il a été possible d'estimer les paramètres $\hat{\mu}$ et $\hat{\sigma}$ d'une loi log-normale à partir de la méthode du maximum de vraisemblance. D'ailleurs, le code SAS se trouve dans l'Annexe à la fin du rapport. Voici, les estimés des paramètres obtenus:

$\hat{\mu}$	$\hat{\sigma}$
6.8418	0.8278

Table 6: Paramètres estimés d'une loi log-normale à l'aide du maximum de vraisemblance

Donc, les estimateurs de la loi log-normale, selon la méthode du maximum de vraisemblance, sont $\hat{\mu} = 6.8418$ et $\hat{\sigma} = 0.8278$.

b) Détermination des primes stop-loss

En ayant obtenu les estimés des paramètres à la question a), il est maintenant possible d'effectuer divers calculs puisque la distribution du modèle est maintenant connue. Par conséquent, il est possible de calculer la prime stop-loss $E[(X-d)_+]$, et ce, pour différentes valeurs de déductible, soit d = 2000, 2100, ..., 3000. Pour y arriver, il faut utiliser la formule suivante:

$$E\left[(X-d)_+\right] = e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}}\left(1-\Phi\left(\frac{Ind-\hat{\mu}-\hat{\sigma}^2}{\hat{\sigma}}\right)\right) - d\left(1-\Phi\left(\frac{Ind-\hat{\mu}}{\hat{\sigma}}\right)\right)$$

où $\hat{\mu} = 6.8418$, $\hat{\sigma} = 0.8278$ et d = déductible. Par conséquent, les valeurs obtenues pour la prime stop-loss sont les suivantes:

	d	Prime
1	2000	253.3579
2	2100	236.1653
3	2200	220.3997
4	2300	205.9213
5	2400	192.6061
6	2500	180.3436
7	2600	169.0355
8	2700	158.5939
9	2800	148.9402
10	2900	140.0041
11	3000	131.7223

Table 7: Valeurs de la prime STOP-LOSS pour chacun des déductibles d donnés

c) Détermination des primes limited loss

Maintenant, toujours en utilisant les paramètres estimés en a), il faut désormais calculer l'espérance limitée, sans déductible, et ce, pour différentes limites supérieure, soit u = 3000, 3100, ..., 4000. Pour y arriver, il faut utiliser la formule suivante:

$$E\left[X \Lambda u\right] = e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}}\Phi\left(\frac{Inu-\hat{\mu}-\hat{\sigma}^2}{\hat{\sigma}}\right) + u\left(1-\Phi\left(\frac{Inu-\hat{\mu}}{\hat{\sigma}}\right)\right)$$

où $\hat{\mu} = 6.8418$, $\hat{\sigma} = 0.8278$ et u = limite. Par conséquent, les valeurs obtenues pour la prime limitée sont les suivantes:

	u	Prime
1	3000	1187.0103
2	3100	1194.6945
3	3200	1201.8320
4	3300	1208.4691
5	3400	1214.6471
6	3500	1220.4038
7	3600	1225.7731
8	3700	1230.7858
9	3800	1235.4700
10	3900	1239.8511
11	4000	1243.9524

Table 8: Valeurs de l'espérance limitée pour chacune des limites u données

Question 3

Estimateur de Bayes

a) Estimation du paramètre θ

b) Présentation graphique de la loi beta a priori

c) Détermination de la loi a posteriori ainsi que la valeur exacte de l'estimateur de Bayes

$$\begin{aligned}
f_{x_i}(x_i) &= \int_0^1 \binom{r + x_i - 1}{x_i} \theta^r (1-\theta)^{x_i} \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)} d\theta \\
&= \binom{r + x_i - 1}{x_i} \frac{B(\alpha+r, \beta+x_i)}{B(\alpha,\beta)} \int_0^1 \frac{\theta^{\alpha+r-1}(1-\theta)^{\beta+x_i-1}}{B(\alpha+r,\beta+x_i)} d\theta \\
&= \frac{\Gamma(r+x_i)}{\Gamma(r)x_i!} \frac{\Gamma(\alpha+r)\Gamma(\beta+x_i)}{\Gamma(\alpha+\beta+r+x_i)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
&= \frac{\Gamma(2+x_i)}{\Gamma(2)x_i!} \frac{\Gamma(2+2)\Gamma(2+x_i)}{\Gamma(2+2+2+x_i)} \frac{\Gamma(2+2)}{\Gamma(2)\Gamma(2)} \\
&= \frac{\Gamma(2+x_i)}{x_i!} \frac{\Gamma(4)\Gamma(2+x_i)}{\Gamma(6+x_i)} \Gamma(4) \\
&= \frac{36[\Gamma(2+x_i)]^2}{x_i!\Gamma(6+x_i)}
\end{aligned}$$

d) Approximation de l'estimateur de Bayes par simulation

Annexe

Question 1

a) Estimation des paramètres

```
# Données utilisées afin de compléter le numéro 1

data <-  c(1500, 6000, 3500, 3800, 1800, 5500, 4800, 4200, 3900, 3000)
```

```
##  [1] 1500 6000 3500 3800 1800 5500 4800 4200 3900 3000
```

```
# Code SAS permettant d'estimer les paramètres de la loi log-normale
# par la méthode du maximum de vraisemblance

data MLE;
input t cens;
datalines;
1500 1
6000 1
3500 1
3800 1
1800 1
5500 1
4800 1
4200 1
3900 1
3000 1
;
run;

proc lifereg data=MLE;
    Model t*cens(0)= /covb dist=lnormal;
run;
```

b) Détermination des primes stop-loss

```
# Les estimateurs ont été calculés à l'aide de SAS en a)
mu <- 8.1618
sigma <-  0.4278

# Valeurs des déductibles d à utiliser
d <- 100 * seq(20,30)

# Calculer les primes STOP-LOSS pour les différents déductibles d
STOP_LOSS <- function(x) {
    exp(mu + sigma^2/2) * (1 - pnorm((log(x) - mu - sigma^2)/sigma)) -
        x * (1 - pnorm((log(x) - mu)/sigma))
}
```

```
Primes_STOP_LOSS <- data.frame(d, "Prime" = sapply(d, function(x) STOP_LOSS(x)))

# Présenter les valeurs des primes STOP-LOSS en fonction de d
library(xtable)
options(xtable.comment = FALSE)
xtable(Primes_STOP_LOSS, caption = "Valeurs de la prime STOP-LOSS pour
                 chacun des déductibles d donnés",
                 align = c("c", "c", "c"),
                 digits = c(0, 0, 4))
```

c) Détermination des primes limited loss

```
# Valeurs des limites u à utiliser
u <- 100 * seq(30,40)

# Calculer l'espérance limitée pour les différentes limites u
Esp_limite <- function(x) {
    exp(mu + sigma^2/2) * pnorm((log(x) - mu - sigma^2)/sigma) +
        x * (1 - pnorm((log(x) - mu)/sigma))
}

Primes_Esp_limite <- data.frame(u, "Prime" = sapply(u, function(x) Esp_limite(x)))

# Présenter les valeurs des espérances limitées en fonction de u
library(xtable)
options(xtable.comment = FALSE)
xtable(Primes_Esp_limite, caption = "Valeurs de l'espérance limitée pour
                 chacune des limites u données",
                 align = c("c", "c", "c"),
                 digits = c(0, 0, 4))
```

d) Détermination de la prime stop-loss par simulation avec un déductible d = 2000

```
# Valeurs du déductible d à utiliser
d <- d[1]

# Nombre de simulations
m <- c(1e5, 5e5, 1e6)

# Simuler un échantillon de m simulations de la log log-normale
# avec les paramètres estimés en a)
simul <- function(x) rlnorm(x, meanlog = mu, sdlog = sigma)
simulation <- sapply(1:3, function(i) simul(m[i]))

# Approximer la prime STOP-LOSS ayant un déductible de 2000 par simulation
approximation <- sapply(1:3, function(x)
                    mean(sapply(1:m[x], function(i)
                        max(0, simulation[[x]][i] - 2000))))

approx <- data.frame(m, approximation)
```

```r
# Présenter les valeurs approximées de la prime STOP-LOSS ayant un déductible de 2000
library(xtable)
options(xtable.comment = FALSE)
xtable(approx, caption = "Valeurs approximées par simulations de la prime STOP-LOSS
                 ayant un déductible de 2000",
                 align = c("c", "c", "c"),
                 digits = c(0, 0, 4))
```

Question 2

a) Estimation des paramètres

```
# Code SAS permettant d'estimer les paramètres de la loi log-normale
# par la méthode du maximum de vraisemblance
data donnees;
input lb ub;
datalines;
0.00001 250
0.00001 250
0.00001 250
250 500
250 500
250 500
250 500
500 750
500 750
500 750
500 750
500 750
750 1000
750 1000
750 1000
750 1000
750 1000
1000 1250
1000 1250
1000 1250
1000 1250
1250 1500
1250 1500
1250 1500
1500 2000
1500 2000
1500 2000
2000 2500
2000 2500
2000 2500
2500 3000
2500 3000
3000 5000
3000 5000
```

```
;
run;

proc lifereg data=donnees;
    Model (lb,ub)= /covb dist=lnormal;
run;
```

b) Détermination des primes stop-loss

```r
# Les estimateurs ont été calculés à l'aide de SAS en a)
mu <- 6.8418
sigma <-  0.8278

# Valeurs des déductibles d à utiliser
d <- 100 * seq(20,30)

# Calculer les primes STOP-LOSS pour les différents déductibles d
STOP_LOSS <- function(x) {
    exp(mu + sigma^2/2) * (1 - pnorm((log(x) - mu - sigma^2)/sigma)) -
        x * (1 - pnorm((log(x) - mu)/sigma))
}

Primes_STOP_LOSS <- data.frame(d, "Prime" = sapply(d, function(x) STOP_LOSS(x)))

# Présenter les valeurs des primes STOP-LOSS en fonction de d
library(xtable)
options(xtable.comment = FALSE)
xtable(Primes_STOP_LOSS, caption = "Valeurs de la prime STOP-LOSS pour
                 chacun des déductibles d donnés",
                 align = c("c", "c", "c"),
                 digits = c(0, 0, 4))
```

c) Détermination des primes limited loss

```r
# Valeurs des limites u à utiliser
u <- 100 * seq(30,40)

# Calculer l'espérance limitée pour les différentes limites u
Esp_limite <- function(x) {
    exp(mu + sigma^2/2) * pnorm((log(x) - mu - sigma^2)/sigma) +
        x * (1 - pnorm((log(x) - mu)/sigma))
}

Primes_Esp_limite <- data.frame(u, "Prime" = sapply(u, function(x) Esp_limite(x)))

# Présenter les valeurs des espérances limitées en fonction de u
library(xtable)
options(xtable.comment = FALSE)
xtable(Primes_Esp_limite, caption = "Valeurs de l'espérance limitée pour
                 chacune des limites u données",
```

```
                  align = c("c", "c", "c"),
                  digits = c(0, 0, 4))
```

Question 3

a) Estimation du paramètre θ

b) Présentation graphique de la loi beta a priori

c) Détermination de la loi a posteriori ainsi que la valeur exacte de l'estimateur de Bayes

d) Approximation de l'estimateur de Bayes par simulation