ACSC/STAT 3703, Actuarial Models I

WINTER 2024
Toby Kenney

Homework Sheet 4

Model Solutions

Basic Questions

1. *A distribution has survival function*

$$S(x) = e^{-\log(x)^{1.2}}$$

for $x \geqslant 0$. How does the tail weight of this distribution compare to that of a log-normal distribution with $\mu = 0$ and $\sigma^2 = 1$, when tail-weight is assessed by

(a) Asymptotic behaviour of hazard rate.

We differentiate $S(x)$ to get

$$f(x) = 1.2\frac{\log(x)^{0.2}}{x}e^{-\log(x)^{1.2}}$$

so

$$\lambda(x) = \frac{f(x)}{S(x)} = 1.2\frac{\log(x)^{0.2}}{x}$$

We see that $\lambda(x) \to 0$ as $x \to \infty$.

For the log-normal distribution, we have $S(x) = \Phi(\log(x))$ and $f(x) = \frac{1}{\sqrt{2\pi}x}e^{-\frac{\log(x)^2}{2}}$, so $\lambda(x) = \frac{e^{-\frac{\log(x)^2}{2}}}{\sqrt{2\pi}x\Phi(\log(x))}$. Asymptotically, this can be approximated by

$$\lambda(x) \approx -\frac{f'(x)}{f(x)} = -\frac{d}{dx}\log(f(x)) = \frac{d}{dx}\left(\frac{\log(x)^2}{2} + \frac{\log(2\pi)}{2} + \log(x)\right) = \frac{\log(x)}{x} + \frac{1}{x}$$

We see that this is asymptotically larger, so the distribution has a heavier tail than the log-normal distribution in terms of the hazard rate function.

(b) Existence of moments.

The kth raw moment of the log-normal distribution is $\mu_k = M_Z(k) = e^{k\mu + \frac{\sigma k^2}{2}}$. In particular, all moments exist. However, the moment generating function is undefined for any $t > 0$.

For the given distribution, the kth moment is

$$\int_0^\infty kx^{k-1}S(x)\,dx = \int_0^\infty kx^{k-1}e^{-\log(x)^{1.2}}\,dx$$

We see that $e^{-\log(x)^{1.2}} = x^{-\log(x)^{0.2}} < e^{m^6}x^{-m}$ for any m and x, since if $x < e^{m^5}$, $e^{m^6}x^{-m} > 1$, while if $x \geqslant m^5$, we have $\log(x)^{0.2} > m$, so $x^{-\log(x)^{0.2}} < e^{m^6}x^{-m}$. This means that

$$\int_0^\infty kx^{k-1}e^{-\log(x)^{1.2}}\,dx < \int_0^\infty kx^{k-1}e^{(k+2)^6}x^{-(k+2)}\,dx$$

which clearly converges, so all finite moments exist.

On the other hand,

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_0^\infty \frac{e^{tx}}{t}S_X(x)\,dx = \int_0^\infty \frac{e^{tx-\log(x)^{1.2}}}{t}\,dx$$

and since $tx - \log(x)^{1.2} > 0$ for sufficiently large x, this does not converge for any $t > 0$, so the moment generating function is undefined for all $t > 0$. Thus we cannot compare the tails in terms of existence of moments.

2. *Which coherence properties are satisfied by the following measure of risk?*

$$\rho(X) = \frac{\mathbb{E}(X) + \sqrt[3]{\mathbb{E}(X^3)}}{2}$$

Give a proof or a counterexample for each property.

Sub-additivity For random variables X and Y, we have $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ and

$$\begin{aligned}\mathbb{E}((X+Y)^3) &= \mathbb{E}(X^3 + Y^3 + 3X^2Y + 3XY^2)\\ &= \mathbb{E}(X^3) + \mathbb{E}(Y^3) + 3\mathbb{E}(X^2Y) + 3\mathbb{E}(XY^2)\\ &\leqslant \mathbb{E}(X^3) + \mathbb{E}(Y^3) + 3\left(\mathbb{E}(X^3)\right)^{\frac{2}{3}}\left(\mathbb{E}(Y^3)\right)^{\frac{1}{3}} + 3\left(\mathbb{E}(X^3)\right)^{\frac{1}{3}}\left(\mathbb{E}(Y^3)\right)^{\frac{2}{3}}\\ &= \left(\sqrt[3]{\mathbb{E}(X^3)} + \sqrt{\mathbb{E}(Y^3)}\right)^3\end{aligned}$$

Monotonicity X^3 is a monotone function, so if $P(X > Y) = 1$ then $P(X^3 > Y^3) = 1$, so $\mathbb{E}(X) > \mathbb{E}(Y)$ and $\mathbb{E}(X^3) > \mathbb{E}(Y^3)$, from which monotonicity follows easily.

Positive homogeneity For any $c > 0$, we have $\mathbb{E}(cX) = c\mathbb{E}(X)$ and $\mathbb{E}((cX)^3) = c^3\mathbb{E}(X^3)$, so $\sqrt[3]{\mathbb{E}((cX)^3)} = c\sqrt[3]{\mathbb{E}(X^3)}$. Adding these together gives positive homogeneity.

Translation invariance Let $X = 1$ with probability $\frac{1}{8}$, and $X = 0$ with probability $\frac{7}{8}$. We now have $\mathbb{E}(X) = \frac{1}{8} = \mathbb{E}(X^3)$, so

$$\rho(X) = \frac{1}{2}\left(\frac{1}{8} + \frac{1}{2}\right) = \frac{5}{16}$$

On the other hand, $\mathbb{E}(X+1) = \frac{9}{8}$ and $\mathbb{E}((X+1)^3) = \frac{15}{8}$, so

$$\rho(X+1) = \frac{1}{2}\left(\frac{9}{8} + \frac{\sqrt[3]{15}}{2}\right) \neq \frac{21}{16}$$

3. *Calculate the TVaR at the 95% level of a distribution with survival function $S_X(x) = e^{\sqrt{3}-\sqrt{x+3}}$ for $x > 0$.*

The VaR at the 95% level is the solution to $S_X(x) = 0.05$, which is

$$\begin{aligned}
e^{\sqrt{3}-\sqrt{x+3}} &= 0.05 \\
\sqrt{x+3} &= \log(20) + \sqrt{3} \\
x &= \left(\log(20) + \sqrt{3}\right)^2 - 3 \\
&= 19.3519328621
\end{aligned}$$

The TVaR is therefore

$$\begin{aligned}
19.3519328621 + \frac{1}{0.05}\int_{19.3519328621}^{\infty} S(x)\,dx &= 19.3519328621 + 20\int_{19.3519328621}^{\infty} e^{\sqrt{3}-\sqrt{x+3}}\,dx \\
&= 19.3519328621 + 20e^{\sqrt{3}}\int_{\sqrt{22.3519328621}}^{\infty} 2ue^{-u}\,du \\
&= 19.3519328621 + 40e^{\sqrt{3}}\left(\left[-ue^{-u}\right]_{\sqrt{22.3519328621}}^{\infty} + \int_{\sqrt{22.3519328621}}^{\infty} e^{-u}\,du\right) \\
&= 19.3519328621 + 40e^{\sqrt{3}}\left(\sqrt{22.3519328621}e^{-\sqrt{22.3519328621}} + e^{-\sqrt{22.3519328621}}\right) \\
&= 19.3519328621 + 40e^{\sqrt{3}}\left(\sqrt{22.3519328621} + 1\right)e^{-\sqrt{22.3519328621}} \\
&= 30.8074990244
\end{aligned}$$

4. *Which of the following density functions with parameters α, β and γ are scale distributions? Which have scale parameters?*

   (i) $f(x) = Ce^{-\frac{x}{\beta}-\frac{x^\alpha}{\gamma}}\left(\frac{x^{\alpha+2}}{\gamma\beta^2}\right)$

   (ii) $f(x) = C\left(\frac{\beta^\alpha}{(\beta+x)^\alpha} + \frac{\beta^\gamma}{\beta^\gamma+x^\gamma}\right)$

(iii) $f(x)=C(x+\alpha)^{-3}(x+\beta)^{-5}\left(x^2+\alpha\right)^{-2}$

[In each case C is a normalising constant that may depend on α, β and γ, but not on x.]

(i) is a scale distribution since

$$f_{cX}(x)=c^{-1}f_X\left(\frac{x}{c}\right)=c^{-1}Ce^{-\frac{x}{c\beta}-\frac{x^\alpha}{c^\alpha\gamma}}\left(\frac{x^{\alpha+2}}{c^{\alpha+2}\gamma\beta^2}\right)$$

Which is the density of $f_X(x)$ with β replaced by $c\beta$ and γ replaced by $c^\alpha\gamma$.

(ii) We can rewrite

$$f(x)=C\left(\frac{1}{\left(1+\frac{x}{\beta}\right)^\alpha}+\frac{1}{1+\left(\frac{x}{\beta}\right)^\gamma}\right)$$

from which it is clear that this is a scale distribution and β is a scale parameter.

(iii) is not a scale distribution, since for example

$$\begin{aligned}f_{2X}(x)&=\frac{1}{2}f_X\left(\frac{x}{c}\right)\\&=512C\left(x+2\alpha\right)^{-3}\left(x+2\beta\right)^{-5}\left(x^2+4\alpha\right)^{-2}\end{aligned}$$

which is not the same distribution form.

5. *An insurance company observes the following sample of claims (in thousands):*

   0.8 1.7 2.6 3.6 5.5 7.1 11.4 20.6

   They use a kernel density model with uniform kernel with bandwidth 2.

   What is the TVaR at the 95% level of the fitted distribution?

   There are 8 sample points, and the kernels about the smallest 7 all have support contained in $(-\infty,13.4]$, so $S(18.6)=\frac{1}{8}$, and for $18.6<x<22.6$, we have

$$S(x)=\frac{1}{32}(22.6-x)$$

   Therefore, the VaR at the 95% level is the solution to

$$\begin{aligned}\frac{1}{32}(22.6-x)&=0.05\\22.6-x&=1.6\\x&=21\end{aligned}$$

The TVaR is therefore

$$\begin{aligned}
&21 + \frac{1}{0.05}\int_{21}^{22.6} \frac{22.6 - x}{32}\,dx\\
=&21 + 20\int_{0}^{1.6} \frac{x}{32}\,dx\\
=&21 + 20\left[\frac{x^2}{64}\right]_0^{1.6}\\
=&21 + 20\frac{1.6^2}{64}\\
=&21.8
\end{aligned}$$

Standard Questions

6. *An inverse gamma distribution with α and $\theta = 1$ has mean $\frac{1}{\alpha-1}$ and variance $\frac{1}{(\alpha-1)^2(\alpha-2)}$. You can simulate n random variables following this inverse gamma distribution with the command*

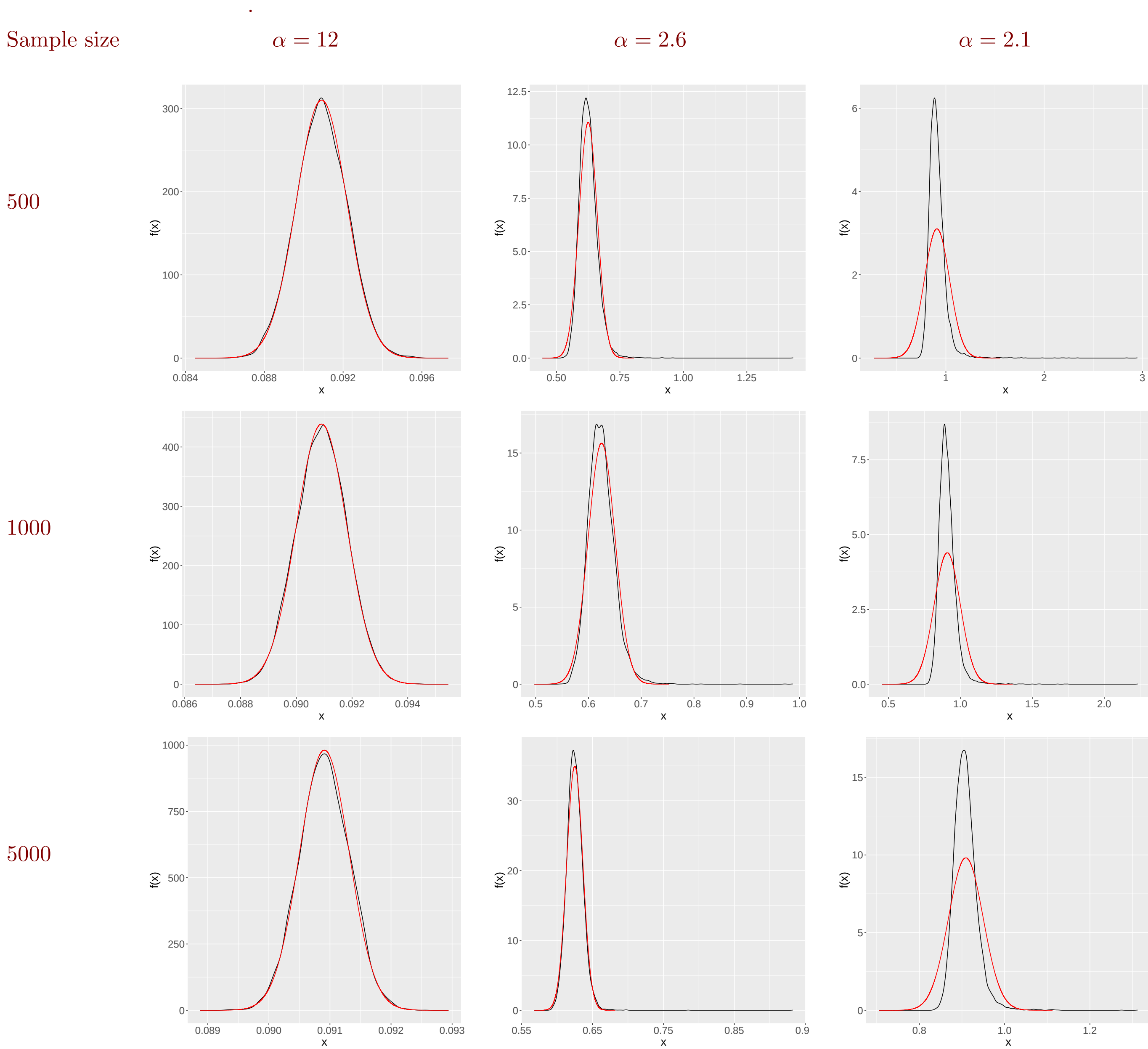
```
sim=1/gamma(n,shape=alpha)
```

[This is simulating a gamma distribution then taking the inverse.]

Based on the central limit theorem, if we take the average of a sample of n inverse gamma random variables, this should approximately follow a normal distribution with mean $\frac{1}{\alpha-1}$ and variance $\frac{1}{n(\alpha-1)^2(\alpha-2)}$. Plot the distribution of this sample average for $\alpha = 12$, $\alpha = 2.6$ and $\alpha = 2.1$, for sample sizes 500, 1000, and 5000, and compare it with the normal distribution.

We run the simulations using the following code

```
library(ggplot2)

InvGammaCLTplot<-function(alpha,n,nsamp){
### alpha is the inverse gamma shape parameter
### n is the sample size
### m is the number of samples

    samp<-1/rgamma(n*nsamp,shape=alpha)
    ## simulate Inverse Gamma Random Varibles

    samples<-matrix(samp,n,nsamp)
    means<-colMeans(samples)
    ## arranging into a matrix and using the column means function is
    ## an efficient way to calculate the sample means. You could also
    ## use a loop.

    dm<-1/(alpha-1)
    dv<-1/(alpha-1)^2/(alpha-2)

    x<-seq_len(100000)*0.0001*sqrt(dv/n)+dm-5*sqrt(dv/n)
    ## x covers 5 standard deviations either side of the mean

    return(
        ggplot(data=data.frame(x=means),mapping=aes(x=x))+
        geom_density()+
        geom_line(data=data.frame(x=x,y=dnorm(x-dm,sd=sqrt(dv/n))),
                  mapping=aes(x=x,y=y),
                  colour="red")+
        scale_y_continuous(name="f(x)")+
        theme(axis.title=element_text(size=18),
            axis.text=element_text(size=16),
            plot.title=element_text(size=18,hjust=0.5))
    )

}


for(alpha in c(12,2.6,2.1)){
    for(ss in c(500,1000,5000)){
        pdf(paste("alpha",alpha,"ssize",ss,".pdf",sep=""))
        print(InvGammaCLTplot(alpha,ss,10000))
        dev.off()
    }
}
```


Sample size
$\alpha = 12$
$\alpha = 2.6$
$\alpha = 2.1$
500
1000
5000
f(x)
x