Modern Statistical Methods Assignment

James Aley

February 2016

1 Question 1

1.1 (a)

The Pareto distribution has probability density function $f(x)$ given by:

$$f(x) = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}} \quad x \geq 0$$

To calculate the cumulative distribution function, we integrate this function over the support $x \geq 0$:

$$\begin{aligned}
F(u) &= \int_0^u f(x)\,\mathrm{d}x \\
&= \int_0^u \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}\,\mathrm{d}x \\
&= \left[-\beta^\alpha(x+\beta)^{-\alpha}\right]_0^u \\
&= 1 - \frac{\beta^\alpha}{(u+\beta)^\alpha} \\
&= 1 - \left(\frac{\beta}{u+\beta}\right)^\alpha
\end{aligned}$$

Now, to calculate the expected value, using integration by parts:

$$
\begin{aligned}
\mathbf{E}\left[X\right] &= \int_0^\infty x f(x) \mathrm{d}x \\
&= \int_0^\infty \frac{x\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}\,\mathrm{d}x \\
&= \left[-\beta^\alpha x(x+\beta)^{-\alpha} + \int \beta^\alpha (x+\beta)^{-\alpha}\right]_0^\infty \\
&= \left[-\beta^\alpha x(x+\beta)^{-\alpha} + \frac{\beta^\alpha}{1-\alpha}(x+\beta)^{1-\alpha}\right]_0^\infty
\end{aligned}
$$

We can see by taking limits on the upper bounds, that we require $\alpha > 1$ in order for this integral to converge.

$$
\lim_{x\to\infty} \left[-\beta^\alpha x(x+\beta)^{-\alpha} + \frac{\beta^\alpha}{1-\alpha}(x+\beta)^{1-\alpha}\right]_0^\infty = 0 \quad (\alpha > 1)
$$

Therefore, substituting in $x = 0$ and subtracting from 0 gives us:

$$
\begin{aligned}
\mathbf{E}\left[X\right] &= 0 - \frac{\beta^\alpha}{\beta^\alpha}\left(\frac{\beta}{1-\alpha}\right) \\
&= \frac{\beta}{\alpha-1}
\end{aligned}
$$

For the median, we need to find the value m that puts equal probability density on either side of it. That is to say that $F(x) = \frac{1}{2}$

$$
\begin{aligned}
F(m) &= \frac{1}{2} \\
1 - \left(\frac{\beta}{m+\beta}\right)^\alpha &= \frac{1}{2} \\
\frac{\beta}{m+\beta} &= \frac{1}{2}^{\frac{1}{\alpha}} \\
m &= \beta\left(\frac{1}{2^{\frac{1}{\alpha}}} - 1\right)
\end{aligned}
$$

To calculate the variance, $\mathbf{Var}\left[X\right]$, we shall make use of the formula:

$$
\mathbf{Var}\left[X\right] = \mathbf{E}\left[X^2\right] - \mathbf{E}\left[X\right]^2
$$

We already have the first moment available from from calculating the mean earlier, so now we need to calculate the second moment, $\mathbf{E}\left[X^2\right]$ as follows. This time, we'll need to apply integration by parts twice successively.

$$
\begin{aligned}
\mathbf{E}\left[X^2\right] &= \int_0^\infty x^2 f(x)\,\mathrm{d}x \\
&= \int_0^\infty \frac{x^2 \alpha \beta^\alpha}{(x+\beta)^{\alpha+1}}\,\mathrm{d}x \\
&= \alpha\beta^\alpha \left[\frac{-x^2(x+\beta)^{-\alpha}}{\alpha} - \int \frac{-2x(x+\beta)^{-\alpha}}{\alpha}\right]_0^\infty \\
&= \beta^\alpha \left[-x^2(x+\beta)^{-\alpha} + \left(\frac{2x(x+\beta)^{1-\alpha}}{(1-\alpha)} - \int \frac{2(x+\beta)^{1-\alpha}}{(1-\alpha)}\right)\right]_0^\infty \\
&= \beta^\alpha \left[-x^2(x+\beta)^{-\alpha} + \frac{2x(x+\beta)^{1-\alpha}}{(1-\alpha)} + \frac{2(x+\beta)^{2-\alpha}}{(1-\alpha)(2-\alpha)}\right]_0^\infty
\end{aligned}
$$

This integral will only converge if we require that $\alpha > 2$, in which case we take limits for x as before:

$$
\lim_{x\to\infty} \beta^\alpha \left[-x^2(x+\beta)^{-\alpha} + \frac{2x(x+\beta)^{1-\alpha}}{(1-\alpha)} + \frac{2(x+\beta)^{2-\alpha}}{(1-\alpha)(2-\alpha)}\right] = 0
$$

Note that the requirement that $\alpha > 2$ is important both so that the x^2 component above grows more slowly than the denominator, and so that the final quotient is finite. Again, we substitute in $x = 0$ and subtract from the 0 obtained from this limit to get an expression for the second moment:

$$
\begin{aligned}
\mathbf{E}\left[X^2\right] &= 0 - \beta^\alpha \left(\frac{2\beta^{2-\alpha}}{(1-\alpha)(2-\alpha)}\right) \\
&= \frac{2\beta^2}{(\alpha-1)(\alpha-2)}
\end{aligned}
$$

Returning to the expression for $\mathbf{Var}[X]$, we have:

$$
\begin{aligned}
\mathbf{Var}[X] &= \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2 \\
&= \frac{2\beta^2}{(\alpha-1)(\alpha-2)} - \left(\frac{\beta}{\alpha-1}\right)^2 \\
&= \frac{2\beta^2\alpha(\alpha-1) - \beta^2(\alpha-2)}{(\alpha-1)^2(\alpha-2)} \\
&= \frac{\beta^2\alpha}{(\alpha-1)^2(\alpha-2)}
\end{aligned}
$$

1.2 (b)

To generate samples from the Pareto distribution, using uniform variables, we can invert the CDF obtained in the previous section as follows.

$$
\begin{aligned}
F(x) &= 1 - \left(\frac{\beta}{x+\beta}\right)^{\alpha} \\
\frac{\beta}{x+\beta} &= (1 - F(x))^{\frac{1}{\alpha}} \\
x &= \beta\left[(1 - F(x))^{-\frac{1}{\alpha}} - 1\right]
\end{aligned}
$$

This gives us an expression for $F^{-1}(x)$, the inverse cumulative distribution function:

$$F^{-1}(u) = \beta\left[(1 - u)^{-\frac{1}{\alpha}} - 1\right]$$

We may now use values for u, which should be realisations of $U \sim \text{Uniform}(0, 1)$. These may be generated using a congruential generator, or any standard method for generating pseudo-random uniform numbers. The `runif()` function can be used to achieve this in R.