Solutions for Exercise 2

1. (a)

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \frac{1}{\pi(1+y^2)} dy \\ &= \left[\frac{1}{\pi} \arctan y\right]_{-\infty}^{x} \\ &= \frac{1}{\pi} \arctan x - \frac{1}{\pi}\frac{-\pi}{2} \\ &= \frac{1}{\pi} \arctan x + \frac{1}{2}. \end{aligned}$$

(b)

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \frac{e^{-y}}{(1+e^{-y})^2} dy \\ &= \left[\frac{1}{1+e^{-y}}\right]_{-\infty}^{x} \\ &= \frac{1}{1+e^{-x}}. \end{aligned}$$

(c) For $x \geq 0$,

$$\begin{aligned} F_X(x) &= \int_{0}^{x} \frac{a-1}{(1+y)^a} dy \\ &= \left[-\frac{1}{(1+y)^{a-1}}\right]_{0}^{x} \\ &= 1 - \frac{1}{(1+x)^{a-1}}. \end{aligned}$$

For $x < 0$ it is obvious that $F_X(x) = 0$, so we could write the result in full as

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - \frac{1}{(1+x)^{a-1}} & x \geq 0. \end{cases}$$

(d) For $x \geq 0$,

$$\begin{aligned} F_X(x) &= \int_{0}^{x} c\tau y^{\tau-1} e^{-cy^\tau} dy \\ &= \left[-e^{-cy^\tau}\right]_{0}^{x} \\ &= 1 - e^{-cx^\tau}. \end{aligned}$$

For $x < 0$ it is obvious that $F_X(x) = 0$, so we could write the result in full as

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-cx^\tau} & x \geq 0. \end{cases}$$

2. (a) We can find the s^{th} moment about the origin, and use this result to get the mean and variance.

$$\begin{aligned}
E[X^s] = m_s &= \int_0^\infty x^s f_X(x)dx = \int_0^\infty x^s \frac{1}{(r-1)!} e^{-kx} x^{r-1} k^r dx \\
&= \int_0^\infty \frac{1}{(r-1)!} e^{-kx} x^{s+r-1} k^r dx \\
&= \frac{(s+r-1)!}{(r-1)!} \frac{1}{k^s} \int_0^\infty \frac{1}{(s+r-1)!} e^{-kx} x^{(s+r)-1} k^{s+r} dx \\
&= \frac{(s+r-1)!}{(r-1)!} \frac{1}{k^s},
\end{aligned}$$

since the integrand is a Gamma density function with shape parameter $s+r$ and scale parameter k. So

$$m_s = \frac{(s+r-1)!}{(r-1)!k^s} = \frac{r(r+1)...(r+s-1)}{k^s}$$

which is sometimes written as

$$\frac{(r+s-1)^{(s)}}{k^s}.$$

Note that this result holds also for non-integer r, using Gamma function results.

Using the result,

$$EX = m_1 = \frac{r}{k}$$

and

$$EX^2 = m_2 = \frac{r(r+1)}{k^2}.$$

The variance is

$$\begin{aligned}
m_2 - (m_1)^2 &= \frac{r(r+1)}{k^2} - \frac{r^2}{k^2} \\
&= \frac{r}{k^2}.
\end{aligned}$$

Both mean and variance increase with r increasing and decrease with k increasing.

(b) The easiest direct way here is to find the r^{th} *factorial* moment, $\mu_{(r)} = EX^{(r)} = EX(X-1)\dots(X-r+1)$. This works out very simply. Then we can convert to the mean and the variance. The critical property that makes $\mu_{(r)}$ work out easily is that for x not an integer from 1 to $r-1$

$$\frac{x^{(r)}}{x!} = \frac{x^{(r)}}{x^{(x)}} = \frac{1}{(x-r)^{(x-r)}} = \frac{1}{(x-r)!}.$$

$$\begin{aligned}
EX^{(r)} &= \sum_{x=0}^{\infty} x^{(r)} e^{-\lambda} \lambda^x / x! \\
&= \sum_{x=r}^{\infty} \frac{x^{(r)}}{x!} e^{-\lambda} \lambda^x \\
&= \sum_{x=r}^{\infty} \frac{1}{(x-r)!} e^{-\lambda} \lambda^x \\
&= \lambda^r \sum_{x=r}^{\infty} \frac{1}{(x-r)!} e^{-\lambda} \lambda^{x-r} \\
&= \lambda^r \sum_{y=0}^{\infty} \frac{1}{y!} e^{-\lambda} \lambda^y \\
&= \lambda^r.
\end{aligned}$$

The last step follows because we are adding together all the probabilities for a Poisson distribution with parameter λ.

Now it is easy to get the mean and the variance.

$$EX = EX^{(1)} = \lambda$$

and since $X(X-1) + X = X^2$,

$$m_2 = EX^{(2)} + EX = \mu_{(2)} + \mu_{(1)} = \lambda^2 + \lambda$$

and so $\mathrm{Var}X = m_2 - (m_1)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

(c) We can use the factorial moments for the negative binomial distribution too. The important identity in this case is that for x not an integer from 1 to $r-1$

$$\begin{aligned}
x^{(r)}\binom{a+x-1}{x} &= x^{(r)}\frac{(a+x-1)^{(x)}}{x!} \\
&= \frac{(a+x-1)^{(x)}}{(x-r)!} = \frac{(a+x-1)^{(x)}}{(x-r)!} \\
&= \frac{(a+x-1)^{(x-r)}(a+r-1)^{(r)}}{(x-r)!} \\
&= (a+r-1)^{(r)}\binom{a+x-1}{x-r} = (a+r-1)^{(r)}\binom{(a+r)+(x-r)-1}{x-r}
\end{aligned}$$

$$\begin{aligned}
EX^{(r)} &= \sum_{x=0}^{\infty} x^{(r)}\binom{a+x-1}{x}\left[\frac{b}{1+b}\right]^a\left[\frac{1}{1+b}\right]^x \\
&= \sum_{x=r}^{\infty} x^{(r)}\binom{a+x-1}{x}\left[\frac{b}{1+b}\right]^a\left[\frac{1}{1+b}\right]^x \\
&= \sum_{x=r}^{\infty}(a+r-1)^{(r)}\binom{(a+r)+(x-r)-1}{x-r}\left[\frac{b}{1+b}\right]^a\left[\frac{1}{1+b}\right]^x \\
&= (a+r-1)^{(r)}\left[\frac{1+b}{b}\right]^r\left[\frac{1}{1+b}\right]^r\sum_{x=r}^{\infty}\binom{(a+r)+(x-r)-1}{x-r}\left[\frac{b}{1+b}\right]^{a+r}\left[\frac{1}{1+b}\right]^{x-r} \\
&= \frac{(a+r-1)^{(r)}}{b^r}\sum_{y=0}^{\infty}\binom{(a+r)+y-1}{x-r}\left[\frac{b}{1+b}\right]^{a+r}\left[\frac{1}{1+b}\right]^y \\
&= \frac{(a+r-1)^{(r)}}{b^r} = \frac{a(a+1)(a+2)\dots(a+r-1)}{b^r}.
\end{aligned}$$

Then we quickly get

$$EX = \mu_{(1)} = \frac{a}{b}$$

and

$$\begin{aligned}
\mathrm{Var}X = \mu_{(2)} + \mu_{(1)} - (\mu_{(1)})^2 &= \frac{a(a+1)}{b^2} + \frac{a}{b} - \frac{a^2}{b^2} \\
&= \frac{a(a+1) + ab - a^2}{b^2} = \frac{a(1+b)}{b^2}.
\end{aligned}$$

(d) We can directly integrate for m_r, writing the integral as a beta function after a transformation. It is easier to work with $Y = X + 1$, noting that $EX = EY - 1$ and $\text{Var}Y = \text{Var}X$.

$$\begin{aligned} EY^r &= \int_0^\infty (x+1)^r \frac{a-1}{(1+x)^a} \\ &= \frac{a-1}{a-r-1} \int_0^\infty \frac{a-r-1}{(1+x)^{a-r}} \\ &= \frac{a-1}{a-r-1}, \end{aligned}$$

provided that $a - r > 1$ (or else the integral is not defined). So

$$EY = \frac{a-1}{a-2} \Rightarrow EX = \frac{1}{a-2}$$

for $a > 2$, and provided $a > 3$,

$$\begin{aligned} \text{Var}X = \text{Var}Y = EY^2 - (EY)^2 &= \frac{a-1}{a-3} - \left[\frac{a-1}{a-2}\right]^2 \\ &= \frac{(a-1)[(a-2)^2 - (a-1)(a-3)]}{(a-2)^2(a-3)} \\ &= \frac{a-1}{(a-2)^2(a-3)} \end{aligned}$$

3. We will do this by showing that the cumulant generating function is $K_X(t) = \mu t + \sigma^2 t^2/2$. Once we have shown this the definition of cumulants will give us our result. First evaluate the moment generating function.

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \textstyle\int_{-\infty}^{\infty} e^{tx} f_X(x)dx \\ &= \textstyle\int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} dx \\ &= \textstyle\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-[(x-\mu)^2 - 2\sigma^2 t]/(2\sigma^2)} dx \\ &= e^{\mu t + \sigma^2 t^2/2} \textstyle\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-[x-(\mu+\sigma^2 t)]^2/(2\sigma^2)} dx \\ &= e^{\mu t + \sigma^2 t^2/2} \end{aligned}$$

Taking logs now gives us the cumulant generating function and hence our result.

4. The description of Y says that $Y = |X - a|$, so

$$\begin{aligned} F_Y(y) &= P[Y \le y] \\ &= P[a - y \le X \le a + y] = P[a - y < X \le a + y] \\ &= \begin{cases} F(a+y) - F(a-y) & y \ge 0 \\ 0 & y < 0 \end{cases} \end{aligned}$$

The density function $f_Y(y)$ of Y is obtained by differentiating with respect to y, so

$$f_Y(y) = \begin{cases} f(a+y) + f(a-y) & y \ge 0 \\ 0 & y < 0. \end{cases}$$

In the case that $X \sim N(\mu, \sigma^2)$ and $a = \mu$, the density function of Y is

$$f_Y(y) = \begin{cases} \frac{2}{\sigma\sqrt{2\pi}} \exp\left[-\frac{y^2}{2\sigma^2}\right] & y \ge 0 \\ 0 & y < 0 \end{cases}$$

This is sometimes called a *half-normal distribution*. It is important for residual plots.

5. It is probably better to use a finite upper limit before integrating by parts, because some standard results for integration by parts do not allow an infinite range of integration. The question implies, however, that the mean is defined. We can use a limiting operation to get rid of the infinite range.

$$\begin{aligned}\int_0^\infty [1 - F_X(x)]dx &= \lim_{a\to\infty} \int_0^a [1 - F_X(x)]dx \\ &= \lim_{a\to\infty} \left[x[1 - F_X(x)]\big|_0^a - \int_0^a (-f_X(x))x dx \right] \\ &= \lim_{a\to\infty} a[1 - F_X(a)] + \lim_{a\to\infty} \int_0^a f_X(x) x dx \\ &= \lim_{a\to\infty} a[1 - F_X(a)] + EX.\end{aligned}$$

The result is proved if we can show that $\lim_{a\to\infty} a[1-F_X(a)] = 0$. This is not true in general (check the Pareto distributions), but *is* true provided that the mean is defined. We can see this through a simple inequality that sandwiches $a[1 - F_X(a)]$ between 0 and a quantity tending to 0. Since

$$EX = \int_0^\infty x f_x(x)dx = \lim_{a\to\infty} \int_0^a x f_x(x)dx$$

it follows that $\lim_{a\to\infty} \int_a^\infty x f_x(x)dx = 0$. Now, for $a > 0$

$$\int_a^\infty x f_x(x)dx \geq \int_a^\infty a f_x(x)dx = a[1 - F_X(a)]$$

so we see that $\lim_{a\to\infty} a[1 - F_X(a)] = 0$.

There is a similar result for random variables on the whole line

$$EX = -\int_{-\infty}^0 F_X(x)dx + \int_0^\infty [1 - F_X(x)]dx.$$

The result is also true for discrete distributions. For instance, if X is a non-negative integer-valued random variable,

$$\begin{aligned}EX = \int_0^\infty [1 - F_X(x)]dx &= \sum_{i=0}^\infty \int_i^{i+1} [1 - F_X(x)]dx \\ &= \sum_{i=0}^\infty \int_i^{i+1} [1 - F_X(i)]dx \\ &= \sum_{i=0}^\infty [1 - F_X(i)] = \sum_{i=0}^\infty \sum_{j=i+1}^\infty P[X = j].\end{aligned}$$

This is easily seen directly.

6. We have

$$f(y) = \begin{cases} 0 & y < 0 \\ y f_X(y)/\mu & y \geq 0 \end{cases}$$

so it is obvious that since $\mu > 0$, $f(y) \geq 0$. The only other property to check to verify that $f(y)$ is a density function is that it integrates to 1.

$$\int_0^\infty \frac{y f_X(y)}{\mu} dy = \frac{\mu}{\mu} = 1.$$

Now, if Y is the random variable with density function $f(y)$ we must have Var$Y \geq 0$ which is $EY^2 \geq (EY)^2$. Assuming that all the moments are defined, it is obvious that $E[Y^r] = E[X^{r+1}]/\mu$. So, from the inequality for Y,

$$E[X^3]/\mu \geq [E[X^2]/\mu]^2.$$

which is the result in the question. Many similar results are available.

7. We must check

$$P[X > x + y \mid X > x] = P[X > y].$$

This can be written in terms of the survival function of X, because for $y > 0$

$$\begin{aligned} 1 - F_X(y) = P[X > y] &= P[X > x + y \mid X > x] \\ &= \frac{P[X > x + y \cap X > x]}{P[X > x]} \\ &= \frac{P[X > x + y]}{P[X > x]} \\ &= \frac{1 - F_X(x + y)}{1 - F_X(x)}. \end{aligned}$$

If X has an exponential distribution of rate λ, then the survival function is $1 - F_X(x) = e^{-\lambda x}$. The no memory property is verified by noting that

$$e^{-\lambda y} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}}.$$

If X has a geometric distribution, then the survival function is $1 - F_X(x) = q^x$. The no memory property is verified by noting that

$$q^y = \frac{q^{x+y}}{q^x}.$$

The 'no memory' property is saying that 'old is as good as new'. If we think in terms of lifetimes, it says that you are equally likely to survive for y more years whatever your current age x may be. This is unrealistic for humans for widely different ages x, but may work as a base model in other applications.

8. We are given for $\lambda \geq 0, 0 < \rho < 1$,

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - \rho \exp(-\lambda(1-\rho)x) & x \geq 0 \end{cases}$$

which is defined as 0 for all negative x, so $F(-\infty) = 0$. Also,

$$F(\infty) = lim_{x\to\infty} F(x) = lim_{x\to\infty} 1 - \rho \exp(-\lambda(1-\rho)x) = 1.$$

All we need to check now is that $F(x)$ is non-decreasing in x, and that it is right continuous. It is clearly non-decreasing and continuous for $x < 0$, being defined as 0 there. Similarly, $1 - \rho \exp(-\lambda(1-\rho)x)$ is continuous and strictly increasing in x for $x > 0$. At $x = 0$, $F(x) = 1 - \rho$, which is the limit as $x \downarrow 0$ of $1 - \rho \exp(-\lambda(1-\rho)x)$. This shows the continuity from the right at $x = 0$. For x just less than 0, we have $F(0-) = 0 < 1 - \rho = F(0)$, so $F(x)$ increases at $x = 0$. That completes checking all the properties. The distribution has a non-zero probability of a waiting time 0, and a continuous distribution for waiting times greater than 0.

One could also simply write the given function as a mixture of two distribution functions to immediately verify that it is itself a distribution function. This can be done with

$$F(x) = (1-\rho)H_0(x) + \rho \begin{cases} (1-\exp(-\lambda(1-\rho)x)) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

which shows the distribution as a mixture with mixing probabilities $1-\rho$ and ρ of a distribution giving probability 1 to $x = 0$ and an exponential distribution with rate $\lambda(1-\rho)$.

9. First notice that the failure rate is

$$\lambda(x) = -\frac{d \ln \bar{F}(x))}{dx}$$

(a) For the Weibull distribution, $\bar{F}(x) = e^{-cx^\tau}$, so

$$\lambda(x) = -\frac{d(-cx^\tau)}{dx} = c\tau x^{\tau-1}.$$

For the Pareto distribution, $\bar{F}(x) = 1/(1+x)^{a-1}$, so

$$\lambda(x) = -\frac{d(-(a-1)\ln(1+x))}{dx} = (a-1)/(1+x).$$

Neither of these gives a very flexible shape for the failure rate if one wants to use it for modelling.

(b) We are asked to show that

$$\lambda(x) \leq \lambda(x+y)$$

when $y \geq 0$.

We are told that $\bar{F}(x+y)/\bar{F}(x)$ and so $\ln \bar{F}(x+y) - \ln \bar{F}(x)$ has a non-positive derivative with respect to x. Differentiating with respect to x,

$$-\lambda(x+y) + \lambda(x) \leq 0$$

which is the result required.