Solutions for Exercise 4

1.

$$\begin{aligned}
P(K = k) &= \sum_{r=k}^{\infty} P[K = k|R = r]P(R = r) \\
&= \sum_{r=k}^{\infty} \binom{r}{k} p^k (1-p)^{r-k} e^{-\lambda}\lambda^r/r! \\
&= \frac{e^{-\lambda}\lambda^k p^k}{k!} \sum_{r=k}^{\infty} \frac{[\lambda(1-p)]^{r-k}}{(r-k)!} \\
&= \frac{e^{-\lambda}\lambda^k p^k}{k!} e^{\lambda(1-p)} \\
&= e^{-\lambda p}\frac{(\lambda p)^k}{k!}.
\end{aligned}$$

So K has a Poisson distribution with mean λp, which is what one might expect for a random thinning of random events.

2. Suppose that the density function for Λ is, for $\lambda > 0$,

$$f_\Lambda(\lambda) = e^{-k\lambda}k^r\lambda^{r-1}/\Gamma(r),$$

where k and r are positive. Then the marginal distribution of X is given by

$$\begin{aligned}
P(X = x) &= \int_0^\infty P[X = x|\Lambda = \lambda]f_\Lambda(\lambda)d\lambda \\
&= \int_0^\infty e^{-\lambda}\frac{\lambda^x}{x!}e^{-k\lambda}\frac{k^r\lambda^{r-1}}{\Gamma(r)}d\lambda \\
&= \frac{\Gamma(r+x)k^r}{(1+k)^{x+r}x!\Gamma(r)} \int_0^\infty e^{-\lambda(1+k)}\frac{\lambda^{r+x-1}(1+k)^{r+x}}{\Gamma(r+x)}d\lambda \\
&= \frac{\Gamma(r+x)k^r}{(1+k)^{x+r}x!\Gamma(r)} \\
&= \binom{r+x-1}{x}\left[\frac{1}{1+k}\right]^x\left[\frac{k}{1+k}\right]^r.
\end{aligned}$$

This is a natural way to produce a negative binomial distribution for which there is no requirement that r be integer.

3. The marginal density function for X is, for $x > 0$,

$$\begin{aligned}
f_X(x) &= \int_0^\infty f_{X|\Theta}(x|\Theta = \theta)f_\Theta(\theta)d\theta \\
&= \int_0^\infty \theta e^{-x\theta}e^{-\lambda\theta}\frac{\lambda^\alpha\theta^{\alpha-1}}{\Gamma(\alpha)}d\theta \\
&= \frac{\lambda^\alpha\Gamma(\alpha+1)}{(x+\lambda)^{(\alpha+1)}\Gamma(\alpha)} \int_0^\infty e^{-\theta(x+\lambda)}\frac{\theta^{(\alpha+1)-1}(x+\lambda)^{(\alpha+1)}}{\Gamma(\alpha+1)}d\theta \\
&= \frac{\lambda^\alpha\Gamma(\alpha+1)}{(x+\lambda)^{(\alpha+1)}\Gamma(\alpha)} \\
&= \frac{\alpha}{\lambda\left(1+\frac{x}{\lambda}\right)^{\alpha+1}}.
\end{aligned}$$

This is a Pareto density function (with scale parameter λ).

4.

$$\begin{aligned}\text{Cov}[X,Y] &= E[XY] - E[X]E[Y] \\ &= E[XE[Y|X]] - E[X]E[E[Y|X]] \\ &= E[X(1+X)/2] - E[X]E[(1+X)/2] \\ &= E[X]/2 + E[X^2]/2 - E[X]/2 - (E[X])^2/2 \\ &= [\text{Var}X]/2.\end{aligned}$$

For $0 < x < 1$,

$$f_X(x) = k(1-x)$$

for some constant k. One can easily see that $k = 2$, so for $0 < x < 1$

$$f_X(x) = 2(1-x).$$

The mean of X is $1/3$, and

$$E[X^2] = \int_0^1 x^2 2(1-x)dx = 2/3 - 1/2 = 1/6.$$

The variance of X is $1/6 - 1/9 = 1/18$, and the $\text{Cov}(X,Y) = 1/36$.

Notice that $E[Y|X]$, which is often called the regression of Y on X, is linear in X, since $E[Y|X] = (1+X)/2$. The regression slope coefficient is $1/2$. When the regression is linear we know that the slope coefficient is

$$\frac{\text{Cov}(X,Y)}{\text{Var}X}$$

in concordance with the results used above.

5. For $0 < x < 1$ the support for the conditional distribution of Y given $X = x$ is the strip going from $(x,0)$ to (x,x). Outside this strip the conditional density is 0. To get the marginal density of X, we must integrate along strips at each different value of x. For $0 < x < 1$,

$$\begin{aligned}f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy &= \int_0^x 8xydy \\ &= 8x\int_0^x ydy = 4x^3.\end{aligned}$$

Obviously, the marginal density function for X at values x outside the interval $(0,1)$ must be 0. The conditional density function of Y given $X = x$ is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

which for $0 < x < 1$ and $0 < y < x$ is

$$f_{Y|X}(y|x) = \frac{8xy}{4x^3} = \frac{2y}{x^2}.$$

The conditional mean and variance are now easy to find. For $0 < x < 1$,

$$\begin{aligned} E[Y^r|X = x] &= \int_0^x y^r \frac{2y}{x^2} dy \\ &= \int_0^x \frac{2y^{r+1}}{x^2} dy \\ &= \frac{2x^{r+2}}{(r+2)x^2} \\ &= \frac{2x^r}{r+2}. \end{aligned}$$

For other values x, the conditional mean values are not defined. We get $E[Y|X = x] = 2x/3$, $E[Y^2|X = x] = 2x^2/4 = x^2/2$, $\mathrm{Var}[Y|X = x] = x^2/2 - 4x^2/9 = x^2/18$.

$E[XY|X = x] = xE[Y|X = x] = 2x^2/3$, so

$$\begin{aligned} E[XY] = E[E[XY|X]] &= E[2X^2/3] \\ &= \int_0^1 \frac{2x^2}{3} 4x^3 dx \\ &= \int_0^1 \frac{8x^5}{3} dx \\ &= 4/9. \end{aligned}$$

To get the covariance of X, Y we need also $E[X] = \int_0^1 x4x^3 dx = 4/5$, and $E[Y] = E[E[Y|X]] = E[2X/3] = 8/15$. The covariance of X, Y is

$$\frac{4}{9} - \frac{4}{5}\frac{8}{15} = 4/225.$$

Notice that $E[Y|X]$, which is often called the regression of Y on X, is linear in X, since $E[Y|X] = 2X/3$. The regression slope coefficient is $2/3$. When the regression is linear we know that the slope coefficient is

$$\frac{\mathrm{Cov}(X, Y)}{\mathrm{Var}X}$$

so we could say from this that $\mathrm{Cov}(X, Y) = \frac{2}{3}\mathrm{Var}X$, as may be checked directly.

6. It is always true that if $X \perp Y$ then $\mathrm{Cov}(X, Y) = 0$ and so the correlation is 0. It is necessary to prove that for this family of distributions the reverse implication is also true.

   If $a = 0$ then the joint distribution is uniform over the unit square, and so $X \perp Y$. It is enough therefore to show that when $\mathrm{Cov}(X, Y) = 0$, $a = 0$.

   The marginal density function of X for $0 < x < 1$ is

$$f_X(x) = \int_0^1 [1 - a(1 - 2x)(1 - 2y)]dy = 1.$$

So X, and by symmetry Y have $U(0,1)$ distributions. It follows that $EX = EY = 1/2$.

$$
\begin{aligned}
E(XY) &= \int_0^1 \int_0^1 xy[1 - a(1-2x)(1-2y)]dydx \\
&= \frac{x^2}{2}\Big|_0^1 \frac{y^2}{2}\Big|_0^1 - a\int_0^1 x(1-2x)dx \int_0^1 y(1-2y)dy \\
&= \frac{1}{4} - a\left[\frac{x^2}{2} - \frac{2x^3}{3}\right]\Big|_0^1 \left[\frac{y^2}{2} - \frac{2y^3}{3}\right]\Big|_0^1 \\
&= 1/4 - a[1/2 - 2/3]^2 \\
&= 1/4 - a/36.
\end{aligned}
$$

So $\mathrm{Cov}(X,Y) = 1/4 - a/36 - (1/2)^2 = -a/36$. If the the correlation of X, Y is 0, then so is the covariance and so $a = 0$. This implies that $X \perp Y$.

Note. This is another bivariate family with regressions that are linear. It is easy to see that

$$E[Y|X] = 1/2 - a(1-2x)/6$$

so that the slope coefficient is a/3. We can verify directly that

$$\mathrm{Cov}(X,Y) = \frac{a}{3}\mathrm{Var}X = \frac{a}{3}\frac{1}{12}.$$

7. Consider exponential distributions $X \sim \mathrm{Exp}(\lambda)$ and $Y \sim \mathrm{Exp}(\mu)$. If the means are equal the $\lambda = \mu$. Let $Z = X + Y$ then

$$
\begin{aligned}
f_Z(z) &= \textstyle\int_{-\infty}^{\infty} f_X(x) f_Y(z-x)dx \\
&= \textstyle\int_0^z \lambda e^{-\lambda x}\lambda e^{-\lambda(z-x)} \\
&= \lambda^2 \textstyle\int_0^z e^{-\lambda z}dx \\
&= \lambda^2 z e^{-\lambda z} \text{ for } z \geq 0
\end{aligned}
$$

For the general case

$$
\begin{aligned}
f_Z(z) &= \textstyle\int_0^z \lambda e^{-\lambda x}\mu e^{-\mu(z-x)}dx \\
&= \lambda\mu e^{-\mu z}\textstyle\int_0^z e^{-x(\lambda-\mu)}dx \\
&= \lambda\mu e^{-\mu z}[-e^{-x(\lambda-\mu)}/(\lambda-\mu)]_0^z \\
&= [\lambda\mu/(\lambda-\mu)](e^{-\mu z} - e^{-\lambda z}) \text{ for } z \geq 0
\end{aligned}
$$

8. We will interpret $\sum_{i=1}^N X_i$ as 0 for $N = 0$, and $\prod_{i=1}^N$ operations as giving 1 when $N = 0$. Then

$$
\begin{aligned}
M_Y(t) &= E[e^{Yt}] \\
&= E[E[e^{Yt}|N]] = E[E[e^{t\sum_{i=1}^N X_i}|N]] \\
&= E[\prod_{i=1}^N E[e^{tX_i}]] \\
&= E[\prod_{i=1}^N M_X(t)] \\
&= E[M_X(t)^N] = E[e^{N\ln M_X(t)}] \\
&= E[e^{NK_X(t)}] = M_N(K_X(t)).
\end{aligned}
$$

Taking logs on both sides gives the result, though the form above may often be more useful.

The example has a geometric distribution for N.

$$\begin{aligned} M_N(t) &= \sum_{n=1}^{\infty} e^{tn} p^{n-1}(1-p) \\ &= e^t(1-p)\sum_{n=1}^{\infty}(pe^t)^{n-1} \\ &= \frac{e^t(1-p)}{1-pe^t}. \end{aligned}$$

X_i has a Gamma$(\alpha, 2)$ distribution with MGF

$$M_X(t) = \frac{\alpha^2}{(\alpha-t)^2}.$$

Using the result we obtained above,

$$\begin{aligned} M_Y(t) = M_N(K_X(t)) &= \frac{e^{K_X(t)}(1-p)}{1-pe^{K_X(t)}} \\ &= \frac{M_X(t)(1-p)}{1-pM_X(t)} \\ &= \frac{(1-p)\frac{\alpha^2}{(\alpha-t)^2}}{1-p\frac{\alpha^2}{(\alpha-t)^2}} \\ &= \frac{(1-p)\alpha^2}{(\alpha-t)^2-p\alpha^2} \\ &= \frac{(1-p)\alpha^2}{[(\alpha-t)-\sqrt{p}\alpha][(\alpha-t)+\sqrt{p}\alpha]} \\ &= \frac{(1-p)\alpha^2}{[\alpha(1-\sqrt{p})-t][\alpha(1+\sqrt{p})-t]} \\ &= \frac{(1-\sqrt{p})(1+\sqrt{p})\alpha}{2\sqrt{p}}\frac{1}{\alpha(1-\sqrt{p})-t} - \frac{(1-\sqrt{p})(1+\sqrt{p})\alpha}{2\sqrt{p}}\frac{1}{\alpha(1+\sqrt{p})-t} \\ &= \frac{1+\sqrt{p}}{2\sqrt{p}}\frac{\alpha(1-\sqrt{p})}{\alpha(1-\sqrt{p})-t} - \frac{1-\sqrt{p}}{2\sqrt{p}}\frac{\alpha(1+\sqrt{p})}{\alpha(1+\sqrt{p})-t} \end{aligned}$$

We can identify the moment generating function as a mixture of two exponential distributions with scale parameters $\alpha(1-\sqrt{p})$ and $\alpha(1+\sqrt{p})$ and mixing weights $\frac{1+\sqrt{p}}{2\sqrt{p}}$ and $-\frac{1-\sqrt{p}}{2\sqrt{p}}$. One of the weights is negative. The corresponding density function is

$$\begin{aligned} f_Y(y) &= \frac{1+\sqrt{p}}{2\sqrt{p}}\alpha(1-\sqrt{p})e^{-\alpha(1-\sqrt{p})y} - \frac{1-\sqrt{p}}{2\sqrt{p}}\alpha(1+\sqrt{p})e^{-\alpha(1+\sqrt{p})y} \\ &= \frac{\alpha(1-p)}{2\sqrt{p}}\left[e^{-\alpha(1-\sqrt{p})y} - e^{-\alpha(1+\sqrt{p})y}\right]. \end{aligned}$$

9. Notice that because the variances of X, Y are both 1, $\kappa_{11} = \rho$. We need to work out

$$\text{Cov}(X^2, Y^2) = E[X^2Y^2] - E[X^2]E[Y^2] = \mu_{22} - \mu_{20}\mu_{02}$$

because we have zero means. So

$$\begin{aligned} \text{Cov}(X^2, Y^2) &= (\kappa_{22} + \kappa_{20}\kappa_{02} + 2\kappa_{11}^2) - \kappa_{20}\kappa_{02} \\ &= \kappa_{22} + 2\kappa_{11}^2 = 2\kappa_{11}^2. \end{aligned}$$

Also

$$\begin{aligned}\mathrm{Var}(X^2) &= E[X^4] - [E(X^2)]^2 = \mu_{40} - \mu_{20}^2 \\ &= [\kappa_{40} + 3\kappa_{20}^2] - \kappa_{20}^2 \\ &= \kappa_{40} + 2\kappa_{20}^2 = 0 + 2 \times 1^2 = 2.\end{aligned}$$

The variance of Y^2 is the same as the variance of X^2, so the correlation coefficient is

$$\frac{2\kappa_{11}^2}{2} = \rho^2.$$

This is less than ρ unless $\rho^2 = 1$: squaring reduces the strength of the linear association.

10. Once you know X you also know X^2, so the joint distribution of $X, Y = X^2$ is a degenerate distribution with all its probability in the (X, Y) plane concentrated on the curve $Y = X^2$. Nevertheless, we can use the usual definitions even for a case like this.

$$M_{X,X^2}(s,t) = E[e^{sX+tX^2}] = \int_{-\infty}^{\infty} e^{sx+tx^2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Now we can make the integrand like a normal density function.

$$sx + tx^2 - x^2/2 = -x^2(1-2t)/2 + sx = -(1-2t)(x - s/(1-2t))^2/2 + s^2/[2(1-2t)]$$

So provided that $1 - 2t > 0$, we manufacture a normal density with mean $\frac{s}{1-2t}$ and variance $\frac{1}{1-2t}$.

$$\begin{aligned}M_{X,X^2}(s,t) &= \frac{e^{s^2/[2(1-2t)]}}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi/(1-2t)}} e^{-(1-2t)(x-s/(1-2t))^2/2} dx \\ &= \frac{e^{s^2/[2(1-2t)]}}{\sqrt{1-2t}}.\end{aligned}$$

Putting $t = 0$ gives $e^{s^2/2}$, which is the mgf for X, and putting $s = 0$ gives $\frac{1}{\sqrt{1-2t}}$ which is the $\chi^2_{(1)}$ mgf for X^2.

11. Suppose that $W = XY$ then

$$M_W(u) = E(M_{W|X}(u|X)) = E(E(e^{uXY}|X)) = E(M_Y(uX))$$

Since Y is a standard normal. $M_Y(t) = e^{t^2/2}$. Thus

$$M_W(u) = E(e^{u^2X^2/2}) = M_{X^2}(u^2/2)$$

But X is also standard normal so X^2 is χ^2_1 and $M_{X^2}(t) = (1-2t)^{-1/2}$. Thus

$$M_W(u) = \frac{1}{\sqrt{1-u^2}}$$

This MGF does not correspond to any 'standard' distributions. However, if we take $Z = UV + XY$. Then by independence the MGF of the sum is the product of the MGFs so

$$M_Z(u) = \frac{1}{1-u^2}$$

which is the MGF of a Laplace distribution.

12. The transformation here is $u = x/y$ and $v = x + y$ so the inverse transformation is $x = uv/(1+u)$ and $y = v/(1+u)$. So

$$\begin{array}{ll} \frac{\partial x}{\partial u} = v/(1+u)^2 & \frac{\partial x}{\partial v} = u/(1+u) \\ \frac{\partial y}{\partial u} = -v/(1+u)^2 & \frac{\partial y}{\partial v} = 1/(1+u) \end{array}$$

and the Jacobian of the transformation is $v/(1+u)^2$. The joint density is

$$f_{U,V}(u,v) = f_{X,Y}(x(u,v), y(u,v))v/(1+u)^2 = \lambda^2 v e^{-\lambda v}/(1+u^2) \text{ for } u, v > 0$$

13. In this case we will work out the Jacobian for the transformation and then invert. Here $x = r\cos\theta$ and $y = r\sin\theta$ so

$$\begin{array}{ll} \frac{\partial x}{\partial r} = \cos\theta & \frac{\partial x}{\partial \theta} = -r\sin\theta \\ \frac{\partial y}{\partial r} = \sin\theta & \frac{\partial y}{\partial \theta} = r\cos\theta \end{array}$$

So the Jacobian is just r. Thus

$$f_{X,Y}(x,y) = f_{R,\Theta}(r(x,y), \theta(x,y))/r$$

But from the question R and Θ are independent and Θ is uniformly distributed over $[0, 2\pi]$. Thus

$$f_{X,Y}(x,y) = f_R(r)/2\pi r = \frac{1}{2\pi\sqrt{x^2+y^2}} f_R(\sqrt{x^2+y^2})$$

This is a bivariate density that has circular contours of equal density.

14. (a) Let's use the result for the mean from the solutions for Exercise 4.

$$E[X] = -\int_{-\infty}^{0} F_X(z)dz + \int_0^\infty [1 - F_X(z)]dz$$

$$E[Y] = -\int_{-\infty}^{0} F_Y(z)dz + \int_0^\infty [1 - F_Y(z)]dz$$

Taking these two together,

$$E[Y] - E[X] = \int_{-\infty}^{0} [F_X(z) - F_Y(z)dz + \int_0^\infty [F_X(z) - F_Y(z)]dz$$

so that since the integrands are positive for all z

$$E[Y] - E[X] > 0.$$

(b) This is a very weak result, so easy to prove.

$$F_X(z) > F_Y(z)$$

so

$$F_{X,Y}(z,\infty) > F_{X,Y}(\infty, z)$$
$$F_{X,Y}(z,\infty) - F_{X,Y}(z,z) > F_{X,Y}(\infty,z) - F_{X,Y}(z,z).$$

So

$$F_{X,Y}(z,\infty) - F_{X,Y}(z,z) > 0$$

and since $P(Y > X) \geq F_{X,Y}(z,\infty) - F_{X,Y}(z,z)$, the result follows.

(c) Here we can just calculate directly:

$$\begin{aligned} P(Y > X) &= \int_{-\infty}^{\infty} [\int_{x}^{\infty} f_X(x) f_Y(y) dy] dx \\ &= \int_{-\infty}^{\infty} [1 - F_Y(x)] f_X(x) dx \\ &> \int_{-\infty}^{\infty} [1 - F_X(x)] f_X(x) dx \\ &= -[1 - F_X(x)]^2/2 \Big|_{-\infty}^{\infty} \\ &= 1/2. \end{aligned}$$

