Chapter 4

Multiple Random Variables

4.1 Since the distribution is uniform, the easiest way to calculate these probabilities is as the ratio of areas, the total area being 4.

a. The circle $x^2 + y^2 \le 1$ has area π, so $P(X^2 + Y^2 \le 1) = \frac{\pi}{4}$.
b. The area below the line $y = 2x$ is half of the area of the square, so $P(2X - Y > 0) = \frac{2}{4}$.
c. Clearly $P(|X + Y| < 2) = 1$.

4.2 These are all fundamental properties of integrals. The proof is the same as for Theorem 2.2.5 with bivariate integrals replacing univariate integrals.

4.3 For the experiment of tossing two fair dice, each of the points in the 36-point sample space are equally likely. So the probability of an event is (number of points in the event)/36. The given probabilities are obtained by noting the following equivalences of events.

$$\begin{aligned}
P(\{X=0, Y=0\}) &= P(\{(1,1),(2,1),(1,3),(2,3),(1,5),(2,5)\}) = \frac{6}{36} = \frac{1}{6} \\
P(\{X=0, Y=1\}) &= P(\{(1,2),(2,2),(1,4),(2,4),(1,6),(2,6)\}) = \frac{6}{36} = \frac{1}{6}
\end{aligned}$$

$$\begin{aligned}
&P(\{X=1, Y=0\}) \\
&\quad = P(\{(3,1),(4,1),(5,1),(6,1),(3,3),(4,3),(5,3),(6,3),(3,5),(4,5),(5,5),(6,5)\}) \\
&\quad = \frac{12}{36} = \frac{1}{3}
\end{aligned}$$

$$\begin{aligned}
&P(\{X=1, Y=1\}) \\
&\quad = P(\{(3,2),(4,2),(5,2),(6,2),(3,4),(4,4),(5,4),(6,4),(3,6),(4,6),(5,6),(6,6)\}) \\
&\quad = \frac{12}{36} = \frac{1}{3}
\end{aligned}$$

4.4 a. $\int_0^1 \int_0^2 C(x + 2y)\,dxdy = 4C = 1$, thus $C = \frac{1}{4}$.

b. $f_X(x) = \begin{cases} \int_0^1 \frac{1}{4}(x+2y)dy = \frac{1}{4}(x+1) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$

c. $F_{XY}(x, y) = P(X \le x, Y \le y) = \int_{-\infty}^x \int_{-\infty}^y f(v, u)\,dvdu$. The way this integral is calculated depends on the values of x and y. For example, for $0 < x < 2$ and $0 < y < 1$,

$$F_{XY}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(u,v)\,dvdu = \int_0^x \int_0^y \frac{1}{4}(u+2v)\,dvdu = \frac{x^2 y}{8} + \frac{y^2 x}{4}.$$

But for $0 < x < 2$ and $1 \le y$,

$$F_{XY}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f(u,v)\,dvdu = \int_0^x \int_0^1 \frac{1}{4}(u+2v)\,dvdu = \frac{x^2}{8} + \frac{x}{4}.$$

The complete definition of F_{XY} is

$$F_{XY}(x,y) = \begin{cases} 0 & x \le 0 \text{ or } y \le 0 \\ x^2y/8 + y^2x/4 & 0 < x < 2 \text{ and } 0 < y < 1 \\ y/2 + y^2/2 & 2 \le x \text{ and } 0 < y < 1 \\ x^2/8 + x/4 & 0 < x < 2 \text{ and } 1 \le y \\ 1 & 2 \le x \text{ and } 1 \le y \end{cases}.$$

d. The function $z = g(x) = 9/(x+1)^2$ is monotone on $0 < x < 2$, so use Theorem 2.1.5 to obtain $f_Z(z) = 9/(8z^2)$, $1 < z < 9$.

4.5 a. $P(X > \sqrt{Y}) = \int_0^1 \int_{\sqrt{y}}^1 (x+y)dxdy = \frac{7}{20}$.

b. $P(X^2 < Y < X) = \int_0^1 \int_y^{\sqrt{y}} 2xdxdy = \frac{1}{6}$.

4.6 Let A = time that A arrives and B = time that B arrives. The random variables A and B are independent uniform$(1,2)$ variables. So their joint pdf is uniform on the square $(1,2)\times(1,2)$. Let X = amount of time A waits for B. Then, $F_X(x) = P(X \le x) = 0$ for $x < 0$, and $F_X(x) = P(X \le x) = 1$ for $1 \le x$. For $x = 0$, we have

$$F_X(0) = P(X \le 0) = P(X = 0) = P(B \le A) = \int_1^2 \int_1^a 1dbda = \frac{1}{2}.$$

And for $0 < x < 1$,

$$F_X(x) = P(X \le x) = 1 - P(X > x) = 1 - P(B - A > x) = 1 - \int_1^{2-x} \int_{a+x}^2 1dbda = \frac{1}{2} + x - \frac{x^2}{2}.$$

4.7 We will measure time in minutes past 8 A.M. So $X \sim$ uniform$(0,30)$, $Y \sim$ uniform$(40,50)$ and the joint pdf is $1/300$ on the rectangle $(0,30)\times(40,50)$.

$$P(\text{arrive before 9 A.M.}) = P(X + Y < 60) = \int_{40}^{50} \int_0^{60-y} \frac{1}{300} dxdy = \frac{1}{2}.$$

4.9

$$\begin{aligned} & P(a \le X \le b, c \le Y \le d) \\ = \;& P(X \le b, c \le Y \le d) - P(X \le a, c \le Y \le d) \\ = \;& P(X \le b, Y \le d) - P(X \le b, Y \le c) - P(X \le a, Y \le d) + P(X \le a, Y \le c) \\ = \;& F(b,d) - F(b,c) - F(a,d) - F(a,c) \\ = \;& F_X(b)F_Y(d) - F_X(b)F_Y(c) - F_X(a)F_Y(d) - F_X(a)F_Y(c) \\ = \;& P(X \le b)\left[P(Y \le d) - P(Y \le c)\right] - P(X \le a)\left[P(Y \le d) - P(Y \le c)\right] \\ = \;& P(X \le b)P(c \le Y \le d) - P(X \le a)P(c \le Y \le d) \\ = \;& P(a \le X \le b)P(c \le Y \le d). \end{aligned}$$

4.10 a. The marginal distribution of X is $P(X = 1) = P(X = 3) = \frac{1}{4}$ and $P(X = 2) = \frac{1}{2}$. The marginal distribution of Y is $P(Y = 2) = P(Y = 3) = P(Y = 4) = \frac{1}{3}$. But

$$P(X = 2, Y = 3) = 0 \ne (\frac{1}{2})(\frac{1}{3}) = P(X = 2)P(Y = 3).$$

Therefore the random variables are not independent.

b. The distribution that satisfies $P(U = x, V = y) = P(U = x)P(V = y)$ where $U \sim X$ and $V \sim Y$ is

		U		
		1	2	3
	2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
V	3	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
	4	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$

4.11 The support of the distribution of (U, V) is $\{(u, v) : u = 1, 2, \ldots; v = u + 1, u + 2, \ldots\}$. This is not a cross-product set. Therefore, U and V are not independent. More simply, if we know $U = u$, then we know $V > u$.

4.12 One interpretation of "a stick is broken at random into three pieces" is this. Suppose the length of the stick is 1. Let X and Y denote the two points where the stick is broken. Let X and Y both have uniform$(0, 1)$ distributions, and assume X and Y are independent. Then the joint distribution of X and Y is uniform on the unit square. In order for the three pieces to form a triangle, the sum of the lengths of any two pieces must be greater than the length of the third. This will be true if and only if the length of each piece is less than 1/2. To calculate the probability of this, we need to identify the sample points (x, y) such that the length of each piece is less than 1/2. If $y > x$, this will be true if $x < 1/2$, $y - x < 1/2$ and $1 - y < 1/2$. These three inequalities define the triangle with vertices $(0, 1/2)$, $(1/2, 1/2)$ and $(1/2, 1)$. (Draw a graph of this set.) Because of the uniform distribution, the probability that (X, Y) falls in the triangle is the area of the triangle, which is 1/8. Similarly, if $x > y$, each piece will have length less than 1/2 if $y < 1/2$, $x - y < 1/2$ and $1 - x < 1/2$. These three inequalities define the triangle with vertices $(1/2, 0)$, $(1/2, 1/2)$ and $(1, 1/2)$. The probability that (X, Y) is in this triangle is also 1/8. So the probability that the pieces form a triangle is $1/8 + 1/8 = 1/4$.

4.13 a.

$$
\begin{aligned}
&\mathrm{E}(Y - g(X))^2 \\
&= \mathrm{E}\left((Y - \mathrm{E}(Y \mid X)) + (\mathrm{E}(Y \mid X) - g(X))\right)^2 \\
&= \mathrm{E}(Y - \mathrm{E}(Y \mid X))^2 + \mathrm{E}(\mathrm{E}(Y \mid X) - g(X))^2 + 2\mathrm{E}\left[(Y - \mathrm{E}(Y \mid X))(\mathrm{E}(Y \mid X) - g(X))\right].
\end{aligned}
$$

The cross term can be shown to be zero by iterating the expectation. Thus

$$\mathrm{E}(Y - g(X))^2 = \mathrm{E}(Y - \mathrm{E}(Y \mid X))^2 + \mathrm{E}(\mathrm{E}(Y \mid X) - g(X))^2 \geq \mathrm{E}(Y - \mathrm{E}(Y \mid X))^2, \text{ for all g}(\cdot).$$

The choice $g(X) = \mathrm{E}(Y \mid X)$ will give equality.

b. Equation (2.2.3) is the special case of a) where we take the random variable X to be a constant. Then, $g(X)$ is a constant, say b, and $\mathrm{E}(Y \mid X) = \mathrm{E}Y$.

4.15 We will find the conditional distribution of $Y|X + Y$. The derivation of the conditional distribution of $X|X + Y$ is similar. Let $U = X + Y$ and $V = Y$. In Example 4.3.1, we found the joint pmf of (U, V). Note that for fixed u, $f(u, v)$ is positive for $v = 0, \ldots, u$. Therefore the conditional pmf is

$$f(v|u) = \frac{f(u, v)}{f(u)} = \frac{\frac{\theta^{u-v}e^{-\theta}}{(u-v)!}\frac{\lambda^v e^{-\lambda}}{v!}}{\frac{(\theta+\lambda)^u e^{-(\theta+\lambda)}}{u!}} = \binom{u}{v}\left(\frac{\lambda}{\theta+\lambda}\right)^v\left(\frac{\theta}{\theta+\lambda}\right)^{u-v}, \ v = 0, \ldots, u.$$

That is $V|U \sim \text{binomial}(U, \lambda/(\theta + \lambda))$.

4.16 a. The support of the distribution of (U, V) is $\{(u, v) : u = 1, 2, \ldots; v = 0, \pm1, \pm2, \ldots\}$. If $V > 0$, then $X > Y$. So for $v = 1, 2, \ldots$, the joint pmf is

$$
\begin{aligned}
f_{U,V}(u, v) &= P(U = u, V = v) = P(Y = u, X = u + v) \\
&= p(1 - p)^{u+v-1}p(1 - p)^{u-1} = p^2(1 - p)^{2u+v-2}.
\end{aligned}
$$

If $V < 0$, then $X < Y$. So for $v = -1, -2, \ldots$, the joint pmf is

$$\begin{aligned} f_{U,V}(u,v) &= P(U=u, V=v) &= P(X=u, Y=u-v) \\ &= p(1-p)^{u-1}p(1-p)^{u-v-1} &= p^2(1-p)^{2u-v-2}. \end{aligned}$$

If $V = 0$, then $X = Y$. So for $v = 0$, the joint pmf is

$$f_{U,V}(u,0) = P(U=u, V=0) = P(X=Y=u) = p(1-p)^{u-1}p(1-p)^{u-1} = p^2(1-p)^{2u-2}.$$

In all three cases, we can write the joint pmf as

$$f_{U,V}(u,v) = p^2(1-p)^{2u+|v|-2} = \left(p^2(1-p)^{2u}\right)(1-p)^{|v|-2}, \ u = 1,2,\ldots; v = 0, \pm1, \pm2, \ldots.$$

Since the joint pmf factors into a function of u and a function of v, U and V are independent.

b. The possible values of Z are all the fractions of the form r/s, where r and s are positive integers and $r < s$. Consider one such value, r/s, where the fraction is in reduced form. That is, r and s have no common factors. We need to identify all the pairs (x,y) such that x and y are positive integers and $x/(x+y) = r/s$. All such pairs are $(ir, i(s-r))$, $i = 1, 2, \ldots$. Therefore,

$$\begin{aligned} P\left(Z = \frac{r}{s}\right) &= \sum_{i=1}^{\infty} P(X = ir, Y = i(s-r)) &= \sum_{i=1}^{\infty} p(1-p)^{ir-1}p(1-p)^{i(s-r)-1} \\ &= \frac{p^2}{(1-p)^2}\sum_{i=1}^{\infty}\left((1-p)^s\right)^i &= \frac{p^2}{(1-p)^2}\frac{(1-p)^s}{1-(1-p)^s} &= \frac{p^2(1-p)^{s-2}}{1-(1-p)^s}. \end{aligned}$$

c.

$$P(X = x, X+Y = t) = P(X = x, Y = t-x) = P(X = x)P(Y = t-x) = p^2(1-p)^{t-2}.$$

4.17 a. $P(Y = i+1) = \int_i^{i+1} e^{-x}dx = e^{-i}(1-e^{-1})$, which is geometric with $p = 1-e^{-1}$.

b. Since $Y \geq 5$ if and only if $X \geq 4$,

$$P(X-4 \leq x | Y \geq 5) = P(X-4 \leq x | X \geq 4) = P(X \leq x) = e^{-x},$$

since the exponential distribution is memoryless.

4.18 We need to show $f(x,y)$ is nonnegative and integrates to 1. $f(x,y) \geq 0$, because the numerator is nonnegative since $g(x) \geq 0$, and the denominator is positive for all $x > 0$, $y > 0$. Changing to polar coordinates, $x = r\cos\theta$ and $y = r\sin\theta$, we obtain

$$\int_0^{\infty}\int_0^{\infty} f(x,y)dxdy = \int_0^{\pi/2}\int_0^{\infty}\frac{2g(r)}{\pi r}rdrd\theta = \frac{2}{\pi}\int_0^{\pi/2}\int_0^{\infty} g(r)drd\theta = \frac{2}{\pi}\int_0^{\pi/2} 1d\theta = 1.$$

4.19 a. Since $(X_1 - X_2)\big/\sqrt{2} \sim \mathrm{n}(0,1)$, $(X_1-X_2)^2\big/2 \sim \chi^2_1$ (see Example 2.1.9).

b. Make the transformation $y_1 = \frac{x_1}{x_1+x_2}$, $y_2 = x_1 + x_2$ then $x_1 = y_1 y_2$, $x_2 = y_2(1-y_1)$ and $|J| = y_2$. Then

$$f(y_1,y_2) = \left[\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}\right]\left[\frac{1}{\Gamma(\alpha_1+\alpha_2)}y_2^{\alpha_1+\alpha_2-1}e^{-y_2}\right],$$

thus $Y_1 \sim \mathrm{beta}(\alpha_1, \alpha_2)$, $Y_2 \sim \mathrm{gamma}(\alpha_1 + \alpha_1, 1)$ and are independent.

4.20 a. This transformation is not one-to-one because you cannot determine the sign of X_2 from Y_1 and Y_2. So partition the support of (X_1, X_2) into $\mathcal{A}_0 = \{-\infty < x_1 < \infty, x_2 = 0\}$, $\mathcal{A}_1 = \{-\infty < x_1 < \infty, x_2 > 0\}$ and $\mathcal{A}_2 = \{-\infty < x_1 < \infty, x_2 < 0\}$. The support of (Y_1, Y_2) is $\mathcal{B} = \{0 < y_1 < \infty, -1 < y_2 < 1\}$. The inverse transformation from $\mathcal{B}$ to $\mathcal{A}_1$ is $x_1 = y_2\sqrt{y_1}$ and $x_2 = \sqrt{y_1 - y_1 y_2^2}$ with Jacobian

$$J_1 = \begin{vmatrix} \frac{1}{2}\frac{y_2}{\sqrt{y_1}} & \sqrt{y_1} \\ \frac{1}{2}\frac{\sqrt{1-y_2^2}}{\sqrt{y_1}} & \frac{y_2\sqrt{y_1}}{\sqrt{1-y_2^2}} \end{vmatrix} = \frac{1}{2\sqrt{1-y_2^2}}.$$

The inverse transformation from $\mathcal{B}$ to $\mathcal{A}_2$ is $x_1 = y_2\sqrt{y_1}$ and $x_2 = -\sqrt{y_1 - y_1 y_2^2}$ with $J_2 = -J_1$. From (4.3.6), $f_{Y_1,Y_2}(y_1, y_2)$ is the sum of two terms, both of which are the same in this case. Then

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= 2\left[\frac{1}{2\pi\sigma^2} e^{-y_1/(2\sigma^2)} \frac{1}{2\sqrt{1-y_2^2}}\right] \\ &= \frac{1}{2\pi\sigma^2} e^{-y_1/(2\sigma^2)} \frac{1}{\sqrt{1-y_2^2}}, \qquad 0 < y_1 < \infty, -1 < y_2 < 1. \end{aligned}$$

b. We see in the above expression that the joint pdf factors into a function of y_1 and a function of y_2. So Y_1 and Y_2 are independent. Y_1 is the square of the distance from (X_1, X_2) to the origin. Y_2 is the cosine of the angle between the positive x_1-axis and the line from (X_1, X_2) to the origin. So independence says the distance from the origin is independent of the orientation (as measured by the angle).

4.21 Since R and θ are independent, the joint pdf of $T = R^2$ and θ is

$$f_{T,\theta}(t, \theta) = \frac{1}{4\pi} e^{-t/2}, \quad 0 < t < \infty, \quad 0 < \theta < 2\pi.$$

Make the transformation $x = \sqrt{t}\cos\theta$, $y = \sqrt{t}\sin\theta$. Then $t = x^2 + y^2$, $\theta = \tan^{-1}(y/x)$, and

$$J = \begin{vmatrix} 2x & 2y \\ \frac{-y}{x^2+y^2} & \frac{-x}{x^2+y^2} \end{vmatrix} = 2.$$

Therefore

$$f_{X,Y}(x, y) = \frac{2}{4\pi} e^{-\frac{1}{2}(x^2+y^2)}, \quad 0 < x^2 + y^2 < \infty, \quad 0 < \tan^{-1} y/x < 2\pi.$$

Thus,

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}, \quad -\infty < x, y < \infty.$$

So X and Y are independent standard normals.

4.23 a. Let $y = v$, $x = u/y = u/v$ then

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{v} & -\frac{u}{v^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v}.$$

$$f_{U,V}(u, v) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha+\beta)\Gamma(\gamma)} \left(\frac{u}{v}\right)^{\alpha-1} \left(1 - \frac{u}{v}\right)^{\beta-1} v^{\alpha+\beta-1}(1-v)^{\gamma-1}\frac{1}{v}, 0 < u < v < 1.$$

Then,

$$\begin{aligned} f_U(u) &= \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} u^{\alpha-1} \int_u^1 v^{\beta-1}(1-v)^{\gamma-1}(\frac{v-u}{v})^{\beta-1} dv \\ &= \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1} \int_0^1 y^{\beta-1}(1-y)^{\gamma-1} dy \quad \left(y=\frac{v-u}{1-u}, dy=\frac{dv}{1-u}\right) \\ &= \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1} \frac{\Gamma(\beta)\Gamma(\gamma)}{\Gamma(\beta+\gamma)} \\ &= \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta+\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1}, \qquad 0<u<1. \end{aligned}$$

Thus, $U \sim \text{gamma}(\alpha, \beta+\gamma)$.

b. Let $x = \sqrt{uv}$, $y = \sqrt{\frac{u}{v}}$ then

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial x}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2}v^{1/2}u^{-1/2} & \frac{1}{2}u^{1/2}v^{-1/2} \\ \frac{1}{2}v^{-1/2}u^{-1/2} & -\frac{1}{2}u^{1/2}v^{-3/2} \end{vmatrix} = \frac{1}{2v}.$$

$$f_{U,V}(u,v) = \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} (\sqrt{uv}^{\alpha-1}(1-\sqrt{uv})^{\beta-1} \left(\sqrt{\frac{u}{v}}\right)^{\alpha+\beta-1} \left(1-\sqrt{\frac{u}{v}}\right)^{\gamma-1} \frac{1}{2v}.$$

The set $\{0<x<1, 0<y<1\}$ is mapped onto the set $\{0<u<v<\frac{1}{u}, 0<u<1\}$. Then,

$$\begin{aligned} & f_U(u) \\ &= \int_u^{1/u} f_{U,V}(u,v) dv \\ &= \underbrace{\frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1}}_{\text{Call it A}} \int_u^{1/u} \left(\frac{1-\sqrt{uv}}{1-u}\right)^{\beta-1} \left(\frac{1-\sqrt{u/v}}{1-u}\right)^{\gamma-1} \frac{(\sqrt{u/v})^{\beta}}{2v(1-u)} dv. \end{aligned}$$

To simplify, let $z = \frac{\sqrt{u/v}-u}{1-u}$. Then $v=u \Rightarrow z=1$, $v=1/u \Rightarrow z=0$ and $dz = -\frac{\sqrt{u/v}}{2(1-u)v} dv$. Thus,

$$\begin{aligned} f_U(u) &= A \int z^{\beta-1}(1-z)^{\gamma-1} dz \qquad (\text{ kernel of beta}(\beta,\gamma)) \\ &= \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1} \frac{\Gamma(\beta)\Gamma(\gamma)}{\Gamma(\beta+\gamma)} \\ &= \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta+\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1}, \qquad 0<u<1. \end{aligned}$$

That is, $U \sim \text{beta}(\alpha, \beta+\gamma)$, as in a).

4.24 Let $z_1 = x+y$, $z_2 = \frac{x}{x+y}$, then $x = z_1 z_2$, $y = z_1(1-z_2)$ and

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial z_1} & \frac{\partial x}{\partial z_2} \\ \frac{\partial y}{\partial z_1} & \frac{\partial y}{\partial z_2} \end{vmatrix} = \begin{vmatrix} z_2 & z_1 \\ 1-z_2 & -z_1 \end{vmatrix} = z_1.$$

The set $\{x>0, y>0\}$ is mapped onto the set $\{z_1>0, 0<z_2<1\}$.

$$\begin{aligned} f_{Z_1,Z_2}(z_1,z_2) &= \frac{1}{\Gamma(r)}(z_1 z_2)^{r-1} e^{-z_1 z_2} \cdot \frac{1}{\Gamma(s)}(z_1 - z_1 z_2)^{s-1} e^{-z_1+z_1 z_2} z_1 \\ &= \frac{1}{\Gamma(r+s)} z_1^{r+s-1} e^{-z_1} \cdot \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} z_2^{r-1}(1-z_2)^{s-1}, \qquad 0<z_1, 0<z_2<1. \end{aligned}$$

$f_{Z_1,Z_2}(z_1,z_2)$ can be factored into two densities. Therefore Z_1 and Z_2 are independent and $Z_1 \sim \text{gamma}(r+s,1)$, $Z_2 \sim \text{beta}(r,s)$.

4.25 For X and Z independent, and $Y = X+Z$, $f_{XY}(x,y) = f_X(x)f_Z(y-x)$. In Example 4.5.8,

$$f_{XY}(x,y) = I_{(0,1)}(x)\frac{1}{10}I_{(0,1/10)}(y-x).$$

In Example 4.5.9, $Y = X^2 + Z$ and

$$f_{XY}(x,y) = f_X(x)f_Z(y-x^2) = \frac{1}{2}I_{(-1,1)}(x)\frac{1}{10}I_{(0,1/10)}(y-x^2).$$

4.26 a.

$$\begin{aligned} P(Z\le z, W=0) &= P(\min(X,Y)\le z, Y\le X) \;=\; P(Y\le z, Y\le X) \\ &= \int_0^z\int_y^\infty \frac{1}{\lambda}e^{-x/\lambda}\frac{1}{\mu}e^{-y/\mu}dxdy \\ &= \frac{\lambda}{\mu+\lambda}\left(1-\exp\left\{-\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)z\right\}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} P(Z\le z, W=1) &= P(\min(X,Y)\le z, X\le Y) \;=\; P(X\le z, X\le Y) \\ &= \int_0^z\int_x^\infty \frac{1}{\lambda}e^{-x/\lambda}\frac{1}{\mu}e^{-y/\mu}dydx \;=\; \frac{\mu}{\mu+\lambda}\left(1-\exp\left\{-\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)z\right\}\right). \end{aligned}$$

b.

$$P(W=0) = P(Y\le X) = \int_0^\infty\int_y^\infty \frac{1}{\lambda}e^{-x/\lambda}\frac{1}{\mu}e^{-y/\mu}dxdy = \frac{\lambda}{\mu+\lambda}.$$

$$P(W=1) = 1-P(W=0) = \frac{\mu}{\mu+\lambda}.$$

$$P(Z\le z) = P(Z\le z, W=0)+P(Z\le z, W=1) = 1-\exp\left\{-\left(\frac{1}{\mu}+\frac{1}{\lambda}\right)z\right\}.$$

Therefore, $P(Z\le z, W=i) = P(Z\le z)P(W=i)$, for $i=0,1$, $z>0$. So Z and W are independent.

4.27 From Theorem 4.2.14 we know $U\sim \text{n}(\mu+\gamma, 2\sigma^2)$ and $V\sim\text{n}(\mu-\gamma,2\sigma^2)$. It remains to show that they are independent. Proceed as in Exercise 4.24.

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}\left[(x-\mu)^2+(y-\gamma)^2\right]} \qquad \text{(by independence, so} f_{XY} = f_X f_Y)$$

Let $u=x+y$, $v=x-y$, then $x=\frac{1}{2}(u+v)$, $y=\frac{1}{2}(u-v)$ and

$$|J| = \begin{vmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{vmatrix} = \frac{1}{2}.$$

The set $\{-\infty<x<\infty, -\infty<y<\infty\}$ is mapped onto the set $\{-\infty<u<\infty, -\infty<v<\infty\}$. Therefore

$$\begin{aligned} f_{UV}(u,v) &= \frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}\left[\left(\left(\frac{u+v}{2}\right)-\mu\right)^2+\left(\left(\frac{u-v}{2}\right)-\gamma\right)^2\right]}\cdot\frac{1}{2} \\ &= \frac{1}{4\pi\sigma^2}e^{-\frac{1}{2\sigma^2}\left[2\left(\frac{u}{2}\right)^2-u(\mu+\gamma)+\frac{(\mu+\gamma)^2}{2}+2\left(\frac{v}{2}\right)^2-v(\mu-\gamma)+\frac{(\mu+\gamma)^2}{2}\right]} \\ &= g(u)\frac{1}{4\pi\sigma^2}e^{-\frac{1}{2(2\sigma^2)}}\left(u-(\mu+\gamma)\right)^2\cdot h(v)e^{-\frac{1}{2(2\sigma^2)}}\left(v-(\mu-\gamma)\right)^2. \end{aligned}$$

By the factorization theorem, U and V are independent.

4.29 a. $\frac{X}{Y} = \frac{R\cos\theta}{R\sin\theta} = \cot\theta$. Let $Z = \cot\theta$. Let $A_1 = (0,\pi)$, $g_1(\theta) = \cot\theta$, $g_1^{-1}(z) = \cot^{-1} z$, $A_2 = (\pi, 2\pi)$, $g_2(\theta) = \cot\theta$, $g_2^{-1}(z) = \pi + \cot^{-1} z$. By Theorem 2.1.8

$$f_Z(z) = \frac{1}{2\pi}\left|\frac{-1}{1+z^2}\right| + \frac{1}{2\pi}\left|\frac{-1}{1+z^2}\right| = \frac{1}{\pi}\frac{1}{1+z^2}, \quad -\infty < z < \infty.$$

b. $XY = R^2\cos\theta\sin\theta$ then $2XY = R^2 2\cos\theta\sin\theta = R^2\sin 2\theta$. Therefore $\frac{2XY}{R} = R\sin 2\theta$. Since $R = \sqrt{X^2+Y^2}$ then $\frac{2XY}{\sqrt{X^2+Y^2}} = R\sin 2\theta$. Thus $\frac{2XY}{\sqrt{X^2+Y^2}}$ is distributed as $\sin 2\theta$ which is distributed as $\sin\theta$. To see this let $\sin\theta \sim f_{\sin\theta}$. For the function $\sin 2\theta$ the values of the function $\sin\theta$ are repeated over each of the 2 intervals $(0,\pi)$ and $(\pi, 2\pi)$. Therefore the distribution in each of these intervals is the distribution of $\sin\theta$. The probability of choosing between each one of these intervals is $\frac{1}{2}$. Thus $f_{2\sin\theta} = \frac{1}{2}f_{\sin\theta} + \frac{1}{2}f_{\sin\theta} = f_{\sin\theta}$. Therefore $\frac{2XY}{\sqrt{X^2+Y^2}}$ has the same distribution as $Y = \sin\theta$. In addition, $\frac{2XY}{\sqrt{X^2+Y^2}}$ has the same distribution as $X = \cos\theta$ since $\sin\theta$ has the same distribution as $\cos\theta$. To see this let consider the distribution of $W = \cos\theta$ and $V = \sin\theta$ where $\theta \sim$ uniform$(0, 2\pi)$. To derive the distribution of $W = \cos\theta$ let $A_1 = (0,\pi)$, $g_1(\theta) = \cos\theta$, $g_1^{-1}(w) = \cos^{-1} w$, $A_2 = (\pi, 2\pi)$, $g_2(\theta) = \cos\theta$, $g_2^{-1}(w) = 2\pi - \cos^{-1} w$. By Theorem 2.1.8

$$f_W(w) = \frac{1}{2\pi}\left|\frac{-1}{\sqrt{1-w^2}}\right| + \frac{1}{2\pi}\left|\frac{1}{\sqrt{1-w^2}}\right| = \frac{1}{\pi}\frac{1}{\sqrt{1-w^2}}, -1 \le w \le 1.$$

To derive the distribution of $V = \sin\theta$, first consider the interval $(\frac{\pi}{2}, \frac{3\pi}{2})$. Let $g_1(\theta) = \sin\theta$, $4g_1^{-1}(v) = \pi - \sin^{-1} v$, then

$$f_V(v) = \frac{1}{\pi}\frac{1}{\sqrt{1-v^2}}, \quad -1 \le v \le 1.$$

Second, consider the set $\{(0, \frac{\pi}{2}) \cup (\frac{3\pi}{2}, 2\pi)\}$, for which the function $\sin\theta$ has the same values as it does in the interval $(\frac{-\pi}{2}, \frac{\pi}{2})$. Therefore the distribution of V in $\{(0, \frac{\pi}{2}) \cup (\frac{3\pi}{2}, 2\pi)\}$ is the same as the distribution of V in $(\frac{-\pi}{2}, \frac{\pi}{2})$ which is $\frac{1}{\pi}\frac{1}{\sqrt{1-v^2}}$, $-1 \le v \le 1$. On $(0, 2\pi)$ each of the sets $(\frac{\pi}{2}, \frac{3\pi}{2})$, $\{(0, \frac{\pi}{2}) \cup (\frac{3\pi}{2}, 2\pi)\}$ has probability $\frac{1}{2}$ of being chosen. Therefore

$$f_V(v) = \frac{1}{2}\frac{1}{\pi}\frac{1}{\sqrt{1-v^2}} + \frac{1}{2}\frac{1}{\pi}\frac{1}{\sqrt{1-v^2}} = \frac{1}{\pi}\frac{1}{\sqrt{1-v^2}}, \quad -1 \le v \le 1.$$

Thus W and V has the same distribution.

Let X and Y be iid n$(0,1)$. Then $X^2 + Y^2 \sim \chi^2_2$ is a positive random variable. Therefore with $X = R\cos\theta$ and $Y = R\sin\theta$, $R = \sqrt{X^2+Y^2}$ is a positive random variable and $\theta = \tan^{-1}(\frac{Y}{X}) \sim$ uniform$(0,1)$. Thus $\frac{2XY}{\sqrt{X^2+Y^2}} \sim X \sim$ n$(0,1)$.

4.30 a.

$$\begin{aligned}
\mathrm{E}Y &= \mathrm{E}\left\{\mathrm{E}(Y|X)\right\} = \mathrm{E}X = \frac{1}{2}. \\
\mathrm{Var}Y &= \mathrm{Var}\left(\mathrm{E}(Y|X)\right) + \mathrm{E}\left(\mathrm{Var}(Y|X)\right) = \mathrm{Var}X + \mathrm{E}X^2 = \frac{1}{12} + \frac{1}{3} = \frac{5}{12}. \\
\mathrm{E}XY &= \mathrm{E}[\mathrm{E}(XY|X)] = \mathrm{E}[X\mathrm{E}(Y|X)] = \mathrm{E}X^2 = \frac{1}{3} \\
\mathrm{Cov}(X,Y) &= \mathrm{E}XY - \mathrm{E}X\mathrm{E}Y = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}.
\end{aligned}$$

b. The quick proof is to note that the distribution of $Y|X = x$ is n$(1,1)$, hence is independent of X. The bivariate transformation $t = y/x$, $u = x$ will also show that the joint density factors.

4.31 a.

$$\mathrm{E}Y = \mathrm{E}\{\mathrm{E}(Y|X)\} = \mathrm{E}nX = \frac{n}{2}.$$

$$\mathrm{Var}Y = \mathrm{Var}\left(\mathrm{E}(Y|X)\right) + \mathrm{E}\left(\mathrm{Var}(Y|X)\right) = \mathrm{Var}(nX) + \mathrm{E}nX(1-X) = \frac{n^2}{12} + \frac{n}{6}.$$

b.

$$P(Y=y, X \le x) = \binom{n}{y} x^y (1-x)^{n-y}, \quad y = 0,1,\ldots,n, \quad 0 < x < 1.$$

c.

$$P(y=y) = \binom{n}{y} \frac{\Gamma(y+1)\Gamma(n-y+1)}{\Gamma(n+2)}.$$

4.32 a. The pmf of Y, for $y = 0,1,\ldots$, is

$$\begin{aligned} f_Y(y) &= \int_0^\infty f_Y(y|\lambda) f_\Lambda(\lambda) d\lambda = \int_0^\infty \frac{\lambda^y e^{-\lambda}}{y!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} d\lambda \\ &= \frac{1}{y!\Gamma(\alpha)\beta^\alpha} \int_0^\infty \lambda^{(y+\alpha)-1} \exp\left\{\frac{-\lambda}{\left(\frac{\beta}{1+\beta}\right)}\right\} d\lambda \\ &= \frac{1}{y!\Gamma(\alpha)\beta^\alpha} \Gamma(y+\alpha) \left(\frac{\beta}{1+\beta}\right)^{y+\alpha}. \end{aligned}$$

If α is a positive integer,

$$f_Y(y) = \binom{y+\alpha-1}{y} \left(\frac{\beta}{1+\beta}\right)^y \left(\frac{1}{1+\beta}\right)^\alpha,$$

the negative binomial$(\alpha, 1/(1+\beta))$ pmf. Then

$$\begin{aligned} \mathrm{E}Y &= \mathrm{E}(\mathrm{E}(Y|\Lambda)) = \mathrm{E}\Lambda = \alpha\beta \\ \mathrm{Var}Y &= \mathrm{Var}(\mathrm{E}(Y|\Lambda)) + \mathrm{E}(\mathrm{Var}(Y|\Lambda)) = \mathrm{Var}\Lambda + \mathrm{E}\Lambda = \alpha\beta^2 + \alpha\beta = \alpha\beta(\beta+1). \end{aligned}$$

b. For $y = 0,1,\ldots$, we have

$$\begin{aligned} P(Y=y|\lambda) &= \sum_{n=y}^\infty P(Y=y|N=n,\lambda) P(N=n|\lambda) \\ &= \sum_{n=y}^\infty \binom{n}{y} p^y (1-p)^{n-y} \frac{e^{-\lambda}\lambda^n}{n!} \\ &= \sum_{n=y}^\infty \frac{1}{y!(n-y)!} \left(\frac{p}{1-p}\right)^y [(1-p)\lambda]^n e^{-\lambda} \\ &= e^{-\lambda} \sum_{m=0}^\infty \frac{1}{y!m!} \left(\frac{p}{1-p}\right)^y [(1-p)\lambda]^{m+y} \qquad (\text{let } m = n-y) \\ &= \frac{e^{-\lambda}}{y!} \left(\frac{p}{1-p}\right)^y [(1-p)\lambda]^y \left[\sum_{m=0}^\infty \frac{[(1-p)\lambda]^m}{m!}\right] \\ &= e^{-\lambda} (p\lambda)^y e^{(1-p)\lambda} \\ &= \frac{(p\lambda)^y e^{-p\lambda}}{y!}, \end{aligned}$$

the Poisson($p\lambda$) pmf. Thus $Y|\Lambda \sim$ Poisson($p\lambda$). Now calculations like those in a) yield the pmf of Y, for $y = 0, 1, \ldots$, is

$$f_Y(y) = \frac{1}{\Gamma(\alpha)y!(p\beta)^\alpha}\Gamma(y+\alpha)\left(\frac{p\beta}{1+p\beta}\right)^{y+\alpha}.$$

Again, if α is a positive integer, $Y \sim$ negative binomial($\alpha, 1/(1+p\beta)$).

4.33 We can show that H has a negative binomial distribution by computing the mgf of H.

$$\mathrm{E}e^{Ht} = \mathrm{E}\mathrm{E}\left(e^{Ht}\middle| N\right) = \mathrm{E}\mathrm{E}\left(e^{(X_1+\cdots+X_N)t}\middle| N\right) = \mathrm{E}\left\{\left[\mathrm{E}\left(e^{X_1t}\middle| N\right)\right]^N\right\},$$

because, by Theorem 4.6.7, the mgf of a sum of independent random variables is equal to the product of the individual mgfs. Now,

$$\mathrm{E}e^{X_1t} = \sum_{x_1=1}^{\infty} e^{x_1t}\frac{-1}{\log p}\frac{(1-p)^{x_1}}{x_1} = \frac{-1}{\log p}\sum_{x_1=1}^{\infty}\frac{(e^t(1-p))^{x_1}}{x_1} = \frac{-1}{\log p}\left(-\log\left\{1-e^t(1-p)\right\}\right).$$

Then

$$\begin{aligned}
\mathrm{E}\left(\frac{\log\{1-e^t(1-p)\}}{\log p}\right)^N &= \sum_{n=0}^{\infty}\left(\frac{\log\{1-e^t(1-p)\}}{\log p}\right)^n\frac{e^{-\lambda}\lambda^n}{n!} \qquad \text{(since } N \sim \text{Poisson)}\\
&= e^{-\lambda}e^{\frac{\lambda\log(1-e^t(1-p))}{\log p}}\sum_{n=0}^{\infty}\frac{e^{\frac{-\lambda\log(1-e^t(1-p))}{\log p}}\left(\frac{\lambda\log(1-e^t(1-p))}{\log p}\right)^n}{n!}.
\end{aligned}$$

The sum equals 1. It is the sum of a Poisson$\Big([\lambda\log(1-e^t(1-p))]/[\log p]\Big)$ pmf. Therefore,

$$\begin{aligned}
\mathrm{E}(e^{Ht}) &= e^{-\lambda}\left[e^{\log(1-e^t(1-p))}\right]^{\lambda/\log p} = \left(e^{\log p}\right)^{-\lambda/\log p}\left(\frac{1}{1-e^t(1-p)}\right)^{-\lambda/\log p}\\
&= \left(\frac{p}{1-e^t(1-p)}\right)^{-\lambda/\log p}.
\end{aligned}$$

This is the mgf of a negative binomial(r, p), with $r = -\lambda/\log p$, if r is an integer.

4.34 a.

$$\begin{aligned}
P(Y=y) &= \int_0^1 P(Y=y|p)f_p(p)dp\\
&= \int_0^1\binom{n}{y}p^y(1-p)^{n-y}\frac{1}{B(\alpha,\beta)}p^{\alpha-1}(1-p)^{\beta-1}dp\\
&= \binom{n}{y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 p^{y+\alpha-1}(1-p)^{n+\beta-y-1}dp\\
&= \binom{n}{y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(y+\alpha)\Gamma(n+\beta-y)}{\Gamma(\alpha+n+\beta)}, \qquad y = 0, 1, \ldots, n.
\end{aligned}$$

b.

$$\begin{aligned}
P(X=x) &= \int_0^1 P(X=x|p)f_P(p)dp\\
&= \int_0^1\binom{r+x-1}{x}p^r(1-p)^x\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}dp
\end{aligned}$$

$$\begin{aligned}
&= \binom{r+x-1}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 p^{(r+\alpha)-1}(1-p)^{(x+\beta)-1}dp \\
&= \binom{r+x-1}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(r+\alpha)\Gamma(x+\beta)}{\Gamma(r+x+\alpha+\beta)} \qquad x=0,1,\ldots
\end{aligned}$$

Therefore,

$$\mathrm{E}X = \mathrm{E}[\mathrm{E}(X|P)] = \mathrm{E}\left[\frac{r(1-P)}{P}\right] = \frac{r\beta}{\alpha-1},$$

since

$$\begin{aligned}
\mathrm{E}\left[\frac{1-P}{P}\right] &= \int_0^1\left(\frac{1-P}{P}\right)\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{\alpha-1}(1-p)^{\beta-1}dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 p^{(\alpha-1)-1}(1-p)^{(\beta+1)-1}dp \quad = \quad \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha-1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta)} \\
&= \frac{\beta}{\alpha-1}.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(\mathrm{Var}(X|P)) + \mathrm{Var}(\mathrm{E}(X|P)) \quad = \quad \mathrm{E}\left[\frac{r(1-P)}{P^2}\right] + \mathrm{Var}\left(\frac{r(1-P)}{P}\right) \\
&= r\frac{(\beta+1)(\alpha+\beta)}{\alpha(\alpha-1)} + r^2\frac{\beta(\alpha+\beta-1)}{(\alpha-1)^2(\alpha-2)},
\end{aligned}$$

since

$$\begin{aligned}
\mathrm{E}\left[\frac{1-P}{P^2}\right] &= \int_0^1\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{(\alpha-2)-1}(1-p)^{(\beta+1)-1}dp \quad = \quad \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha-2)\Gamma(\beta+1)}{\Gamma(\alpha+\beta-1)} \\
&= \frac{(\beta+1)(\alpha+\beta)}{\alpha(\alpha-1)}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}\left(\frac{1-P}{P}\right) &= \mathrm{E}\left[\left(\frac{1-P}{P}\right)^2\right] - \left(\mathrm{E}\left[\frac{1-P}{P}\right]\right)^2 \quad = \quad \frac{\beta(\beta+1)}{(\alpha-2)(\alpha-1)} - (\frac{\beta}{\alpha-1})^2 \\
&= \frac{\beta(\alpha+\beta-1)}{(\alpha-1)^2(\alpha-2)},
\end{aligned}$$

where

$$\begin{aligned}
\mathrm{E}\left[\left(\frac{1-P}{P}\right)^2\right] &= \int_0^1\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}p^{(\alpha-2)-1}(1-p)^{(\beta+2)-1}dp \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha-2)\Gamma(\beta+2)}{\Gamma(\alpha-2+\beta+2)} \quad = \quad \frac{\beta(\beta+1)}{(\alpha-2)(\alpha-1)}.
\end{aligned}$$

4.35 a. $\mathrm{Var}(X) = \mathrm{E}(\mathrm{Var}(X|P)) + \mathrm{Var}(\mathrm{E}(X|P))$. Therefore,

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}[nP(1-P)] + \mathrm{Var}(nP) \\
&= n\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)} + n^2\mathrm{Var}P \\
&= n\frac{\alpha\beta(\alpha+\beta+1-1)}{(\alpha+\beta^2)(\alpha+\beta+1)} + n^2\mathrm{Var}P
\end{aligned}$$

$$\begin{aligned}
&= \frac{n\alpha\beta(\alpha+\beta+1)}{(\alpha+\beta^2)(\alpha+\beta+1)} - \frac{n\alpha\beta}{(\alpha+\beta^2)(\alpha+\beta+1)} + n^2\mathrm{Var}P \\
&= n\frac{\alpha}{\alpha+\beta}\frac{\beta}{\alpha+\beta} - n\mathrm{Var}P + n^2\mathrm{Var}P \\
&= n\mathrm{E}P(1-\mathrm{E}P) + n(n-1)\mathrm{Var}P.
\end{aligned}$$

b. $\mathrm{Var}(Y) = \mathrm{E}(\mathrm{Var}(Y|\Lambda)) + \mathrm{Var}(\mathrm{E}(Y|\Lambda)) = \mathrm{E}\Lambda + \mathrm{Var}(\Lambda) = \mu + \frac{1}{\alpha}\mu^2$ since $\mathrm{E}\Lambda = \mu = \alpha\beta$ and $\mathrm{Var}(\Lambda) = \alpha\beta^2 = \frac{(\alpha\beta)^2}{\alpha} = \frac{\mu^2}{\alpha}$. The "extra-Poisson" variation is $\frac{1}{\alpha}\mu^2$.

4.37 a. Let $Y = \sum X_i$.

$$\begin{aligned}
P(Y=k) &= P(Y=k, \frac{1}{2} < c = \frac{1}{2}(1+p) < 1) \\
&= \int_0^1 (Y=k|c=\frac{1}{2}(1+p))P(P=p)dp \\
&= \int_0^1 \binom{n}{k}[\frac{1}{2}(1+p)]^k[1-\frac{1}{2}(1+p)]^{n-k}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}dp \\
&= \int_0^1 \binom{n}{k}\frac{(1+p)^k}{2^k}\frac{(1-p)^{n-k}}{2^{n-k}}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1}dp \\
&= \binom{n}{k}\frac{\Gamma(a+b)}{2^n\Gamma(a)\Gamma(b)}\sum_{j=0}^{k}\int_0^1 p^{k+a-1}(1-p)^{n-k+b-1}dp \\
&= \binom{n}{k}\frac{\Gamma(a+b)}{2^n\Gamma(a)\Gamma(b)}\sum_{j=0}^{k}\binom{k}{j}\frac{\Gamma(k+a)\Gamma(n-k+b)}{\Gamma(n+a+b)} \\
&= \sum_{j=0}^{k}\left[\left(\frac{\binom{k}{j}}{2^n}\right)\left(\binom{n}{k}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(k+a)\Gamma(n-k+b)}{\Gamma(n+a+b)}\right)\right].
\end{aligned}$$

A mixture of beta-binomial.

b.

$$\mathrm{E}Y = \mathrm{E}(\mathrm{E}(Y|c)) = \mathrm{E}[nc] = \mathrm{E}\left[n\left(\frac{1}{2}(1+p)\right)\right] = \frac{n}{2}\left(1+\frac{a}{a+b}\right).$$

Using the results in Exercise 4.35(a),

$$\mathrm{Var}(Y) = n\mathrm{E}C(1-\mathrm{E}C) + n(n-1)\mathrm{Var}C.$$

Therefore,

$$\begin{aligned}
\mathrm{Var}(Y) &= n\mathrm{E}\left[\frac{1}{2}(1+P)\right]\left(1-\mathrm{E}\left[\frac{1}{2}(1+P)\right]\right) + n(n-1)\mathrm{Var}\left(\frac{1}{2}(1+P)\right) \\
&= \frac{n}{4}(1+\mathrm{E}P)(1-\mathrm{E}P) + \frac{n(n-1)}{4}\mathrm{Var}P \\
&= \frac{n}{4}\left(1-\left(\frac{a}{a+b}\right)^2\right) + \frac{n(n-1)}{4}\frac{ab}{(a+b)^2(a+b+1)}.
\end{aligned}$$

4.38 a. Make the transformation $u = \frac{x}{\nu} - \frac{x}{\lambda}$, $du = \frac{-x}{\nu^2}d\nu$, $\frac{\nu}{\lambda-\nu} = \frac{x}{\lambda u}$. Then

$$\int_0^\lambda \frac{1}{\nu}e^{-x/\nu}\frac{1}{\Gamma(r)\Gamma(1-r)}\frac{\nu^{r-1}}{(\lambda-\nu)^r}d\nu$$

$$= \frac{1}{\Gamma(r)\Gamma(1-r)}\int_0^\infty \frac{1}{x}\left(\frac{x}{\lambda u}\right)^r e^{-(u+x/\lambda)}du$$
$$= \frac{x^{r-1}e^{-x/\lambda}}{\lambda^r\Gamma(r)\Gamma(1-r)}\int_0^\infty \left(\frac{1}{u}\right)^r e^{-u}du = \frac{x^{r-1}e^{-x/\lambda}}{\Gamma(r)\lambda^r},$$

since the integral is equal to $\Gamma(1-r)$ if $r<1$.

b. Use the transformation $t=\nu/\lambda$ to get

$$\int_0^\lambda p_\lambda(\nu)d\nu = \frac{1}{\Gamma(r)\Gamma(1-r)}\int_0^\lambda \nu^{r-1}(\lambda-\nu)^{-r}d\nu = \frac{1}{\Gamma(r)\Gamma(1-r)}\int_0^1 t^{r-1}(1-t)^{-r}dt = 1,$$

since this is a beta$(r,1-r)$.

c.

$$\frac{d}{dx}\log f(x) = \frac{d}{dx}\left[\log\frac{1}{\Gamma(r)\lambda^r}+(r-1)\log x - x/\lambda\right] = \frac{r-1}{x}-\frac{1}{\lambda} > 0$$

for some x, if $r>1$. But,

$$\frac{d}{dx}\left[\log\int_0^\infty \frac{e^{-x/\nu}}{\nu}q_\lambda(\nu)d\nu\right] = \frac{-\int_0^\infty \frac{1}{\nu^2}e^{-x/\nu}q_\lambda(\nu)d\nu}{\int_0^\infty \frac{1}{\nu}e^{-x/\nu}q_\lambda(\nu)d\nu} < 0 \quad \forall x.$$

4.39 a. Without loss of generality lets assume that $i<j$. From the discussion in the text we have that

$$f(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n|x_j)$$
$$= \frac{(m-x_j)!}{x_1!\cdot\cdots\cdot x_{j-1}!\cdot x_{j+1}!\cdot\cdots\cdot x_n!}$$
$$\times\left(\frac{p_1}{1-p_j}\right)^{x_1}\cdot\cdots\cdot\left(\frac{p_{j-1}}{1-p_j}\right)^{x_{j-1}}\left(\frac{p_{j+1}}{1-p_j}\right)^{x_{j+1}}\cdot\cdots\cdot\left(\frac{p_n}{1-p_j}\right)^{x_n}.$$

Then,

$$f(x_i|x_j)$$
$$= \sum_{(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_{j-1},x_{j+1},\ldots,x_n)} f(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n|x_j)$$
$$= \sum_{(x_k\neq x_i,x_j)}\frac{(m-x_j)!}{x_1!\cdot\cdots\cdot x_{j-1}!\cdot x_{j+1}!\cdot\cdots\cdot x_n!}$$
$$\times\left(\frac{p_1}{1-p_j}\right)^{x_1}\cdot\cdots\cdot\left(\frac{p_{j-1}}{1-p_j}\right)^{x_{j-1}}\left(\frac{p_{j+1}}{1-p_j}\right)^{x_{j+1}}\cdot\cdots\cdot\left(\frac{p_n}{1-p_j}\right)^{x_n}$$
$$\times\frac{(m-x_i-x_j)!\left(1-\frac{p_i}{1-p_j}\right)^{m-x_i-x_j}}{(m-x_i-x_j)!\left(1-\frac{p_i}{1-p_j}\right)^{m-x_i-x_j}}$$
$$= \frac{(m-x_j)!}{x_i!(m-x_i-x_j)!}\left(\frac{p_i}{1-p_j}\right)^{x_i}\left(1-\frac{p_i}{1-p_j}\right)^{m-x_i-x_j}$$
$$\times\sum_{(x_k\neq x_i,x_j)}\frac{(m-x_i-x_j)!}{x_1!\cdot\cdots\cdot x_{i-1}!,x_{i+1}!\cdot\cdots\cdot x_{j-1}!,x_{j+1}!\cdot\cdots\cdot x_n!}$$
$$\times\left(\frac{p_1}{1-p_j-p_i}\right)^{x_1}\cdot\cdots\cdot\left(\frac{p_{i-1}}{1-p_j-p_i}\right)^{x_{i-1}}\left(\frac{p_{i+1}}{1-p_j-p_i}\right)^{x_{i+1}}$$

$$\begin{aligned} &\times (\frac{p_{j-1}}{1-p_j-p_i})^{x_{j-1}}(\frac{p_{j+1}}{1-p_j-p_i})^{x_{j+1}}\cdot\dots\cdot(\frac{p_n}{1-p_j-p_i})^{x_n} \\ = \quad & \frac{(m-x_j)!}{x_i!(m-x_i-x_j)!}(\frac{p_i}{1-p_j})^{x_i}\left(1-\frac{p_i}{1-p_j}\right)^{m-x_i-x_j}. \end{aligned}$$

Thus $X_i|X_j = x_j \sim \text{binomial}(m-x_j, \frac{p_i}{1-p_j})$.

b.

$$f(x_i, x_j) = f(x_i|x_j)f(x_j) = \frac{m!}{x_i!x_j!(m-x_j-x_i)!}p_i^{x_i}p_j^{x_j}(1-p_j-p_i)^{m-x_j-x_i}.$$

Using this result it can be shown that $X_i + X_j \sim \text{binomial}(m, p_i + p_j)$. Therefore,

$$\text{Var}(X_i + X_j) = m(p_i + p_j)(1 - p_i - p_j).$$

By Theorem 4.5.6 $\text{Var}(X_i + X_j) = \text{Var}(X_i) + \text{Var}(X_j) + 2\text{Cov}(X_i, X_j)$. Therefore,

$$\text{Cov}(X_i, X_j) = \frac{1}{2}[m(p_i+p_j)(1-p_i-p_j) - mp_i(1-p_i) - mp_i(1-p_i)] = \frac{1}{2}(-2mp_ip_j) = -mp_ip_j.$$

4.41 Let a be a constant. $\text{Cov}(a, X) = \text{E}(aX) - \text{E}a\text{E}X = a\text{E}X - a\text{E}X = 0$.

4.42

$$\rho_{XY,Y} = \frac{\text{Cov}(XY, Y)}{\sigma_{XY}\sigma_Y} = \frac{\text{E}(XY^2) - \mu_{XY}\mu_Y}{\sigma_{XY}\sigma_Y} = \frac{\text{E}X\text{E}Y^2 - \mu_X\mu_Y\mu_Y}{\sigma_{XY}\sigma_Y},$$

where the last step follows from the independence of X and Y. Now compute

$$\begin{aligned} \sigma_{XY}^2 &= \text{E}(XY)^2 - [\text{E}(XY)]^2 \quad = \quad \text{E}X^2\text{E}Y^2 - (\text{E}X)^2(\text{E}Y)^2 \\ &= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2 \quad = \quad \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2. \end{aligned}$$

Therefore,

$$\rho_{XY,Y} = \frac{\mu_X(\sigma_Y^2+\mu_Y^2) - \mu_X\mu_Y^2}{(\sigma_X^2\sigma_Y^2+\sigma_X^2\mu_Y^2+\sigma_Y^2\mu_X^2)^{1/2}\sigma_Y} = \frac{\mu_X\sigma_Y}{(\mu_X^2\sigma_Y^2+\mu_Y^2\sigma_X^2+\sigma_X^2\sigma_Y^2)^{1/2}}.$$

4.43

$$\begin{aligned} \text{Cov}(X_1 + X_2, X_2 + X_3) &= \text{E}(X_1 + X_2)(X_2 + X_3) - \text{E}(X_1 + X_2)\text{E}(X_2 + X_3) \\ &= (4\mu^2 + \sigma^2) - 4\mu^2 \quad = \quad \sigma^2 \\ \text{Cov}(X_1+X_2)(X_1-X_2) &= \text{E}(X_1 + X_2)(X_1 - X_2) \quad = \quad \text{E}X_1^2 - X_2^2 \quad = \quad 0. \end{aligned}$$

4.44 Let $\mu_i = \text{E}(X_i)$. Then

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n X_i\right) &= \text{Var}\left(X_1 + X_2 + \cdots + X_n\right) \\ &= \text{E}\left[(X_1 + X_2 + \cdots + X_n) - (\mu_1 + \mu_2 + \cdots + \mu_n)\right]^2 \\ &= \text{E}\left[(X_1 - \mu_1) + (X_2 - \mu_2) + \cdots + (X_n - \mu_n)\right]^2 \\ &= \sum_{i=1}^n \text{E}(X_i - \mu_i)^2 + 2\sum_{1\le i<j\le n}\text{E}(X_i - \mu_i)(X_j - \mu_j) \\ &= \sum_{i=1}^n \text{Var}X_i + 2\sum_{1\le i<j\le n}\text{Cov}(X_i, X_j). \end{aligned}$$

4.45 a. We will compute the marginal of X. The calculation for Y is similar. Start with

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right\}\right]$$

and compute

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\omega^2-2\rho\omega z+z^2\right)}\sigma_Y dz,$$

where we make the substitution $z = \frac{y-\mu_Y}{\sigma_Y}$, $dy = \sigma_Y dz$, $\omega = \frac{x-\mu_X}{\sigma_X}$. Now the part of the exponent involving ω^2 can be removed from the integral, and we complete the square in z to get

$$f_X(x) = \frac{e^{-\frac{\omega^2}{2(1-\rho^2)}}}{2\pi\sigma_X\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[(z^2-2\rho\omega z+\rho^2\omega^2)-\rho^2\omega^2\right]}dz$$
$$= \frac{e^{-\omega^2/2(1-\rho^2)}e^{\rho^2\omega^2/2(1-\rho^2)}}{2\pi\sigma_X\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z-\rho\omega)^2}dz.$$

The integrand is the kernel of normal pdf with $\sigma^2 = (1-\rho^2)$, and $\mu = \rho\omega$, so it integrates to $\sqrt{2\pi}\sqrt{1-\rho^2}$. Also note that $e^{-\omega^2/2(1-\rho^2)}e^{\rho^2\omega^2/2(1-\rho^2)} = e^{-\omega^2/2}$. Thus,

$$f_X(x) = \frac{e^{-\omega^2/2}}{2\pi\sigma_X\sqrt{1-\rho^2}}\sqrt{2\pi}\sqrt{1-\rho^2} = \frac{1}{\sqrt{2\pi}\sigma_X}e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2},$$

the pdf of $\mathrm{n}(\mu_X, \sigma_X^2)$.

b.

$$f_{Y|X}(y|x)$$
$$= \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}}{\frac{1}{\sqrt{2\pi}\sigma_X}e^{-\frac{1}{2\sigma_X^2}(x-\mu_X)^2}}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2-(1-\rho^2)\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2-2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}}e^{-\frac{1}{2\sigma_Y^2\sqrt{(1-\rho^2}}\left[(y-\mu_Y)-\left(\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)\right]^2},$$

which is the pdf of $\mathrm{n}\Big((\mu_Y - \rho(\sigma_Y/\sigma_X)(x-\mu_X), \sigma_Y\sqrt{1-\rho^2}\Big)$.

c. The mean is easy to check,

$$\mathrm{E}(aX+bY) = a\mathrm{E}X + b\mathrm{E}Y = a\mu_X + b\mu_Y,$$

as is the variance,

$$\mathrm{Var}(aX+bY) = a^2\mathrm{Var}X + b^2\mathrm{Var}Y + 2ab\mathrm{Cov}(X,Y) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y.$$

To show that $aX+bY$ is normal we have to do a bivariate transform. One possibility is $U = aX+bY$, $V = Y$, then get $f_{U,V}(u,v)$ and show that $f_U(u)$ is normal. We will do this in the standard case. Make the indicated transformation and write $x = \frac{1}{a}(u-bv)$, $y = v$ and obtain

$$|J| = \begin{vmatrix} 1/a & -b/a \\ 0 & 1 \end{vmatrix} = \frac{1}{a}.$$

Then

$$f_{UV}(u,v) = \frac{1}{2\pi a\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left[\frac{1}{a}(u-bv)\right]^2 - 2\frac{\rho}{a}(u-bv)+v^2\right]}.$$

Now factor the exponent to get a square in u. The result is

$$-\frac{1}{2(1-\rho^2)}\left[\frac{b^2+2\rho ab+a^2}{a^2}\right]\left[\frac{u^2}{b^2+2\rho ab+a^2} - 2\left(\frac{b+a\rho}{b^2+2\rho ab+a^2}\right)uv + v^2\right].$$

Note that this is joint bivariate normal form since $\mu_U = \mu_V = 0$, $\sigma_v^2 = 1$, $\sigma_u^2 = a^2+b^2+2ab\rho$ and

$$\rho^* = \frac{\mathrm{Cov}(U,V)}{\sigma_U\sigma_V} = \frac{\mathrm{E}(aXY+bY^2)}{\sigma_U\sigma_V} = \frac{a\rho+b}{\sqrt{a^2+b^2+ab\rho}},$$

thus

$$(1-\rho^{*2}) = 1 - \frac{a^2\rho^2+ab\rho+b^2}{a^2+b^2+2ab\rho} = \frac{(1-\rho^2)a^2}{a^2+b^2+2ab\rho} = \frac{(1-\rho^2)a^2}{\sigma_u^2}$$

where $a\sqrt{1-\rho^2} = \sigma_U\sqrt{1-\rho^{*2}}$. We can then write

$$f_{UV}(u,v) = \frac{1}{2\pi\sigma_U\sigma_V\sqrt{1-\rho^{*2}}}\exp\left[-\frac{1}{2\sqrt{1-\rho^{*2}}}\left(\frac{u^2}{\sigma_U^2} - 2\rho\frac{uv}{\sigma_U\sigma_V} + \frac{v^2}{\sigma_V^2}\right)\right],$$

which is in the exact form of a bivariate normal distribution. Thus, by part a), U is normal.

4.46 a.

$$\begin{aligned}
\mathrm{E}X &= a_X\mathrm{E}Z_1 + b_X\mathrm{E}Z_2 + \mathrm{E}c_X = a_X 0 + b_X 0 + c_X = c_X \\
\mathrm{Var}X &= a_X^2\mathrm{Var}Z_1 + b_X^2\mathrm{Var}Z_2 + \mathrm{Var}c_X = a_X^2 + b_X^2 \\
\mathrm{E}Y &= a_Y 0 + b_Y 0 + c_Y = c_Y \\
\mathrm{Var}Y &= a_Y^2\mathrm{Var}Z_1 + b_Y^2\mathrm{Var}Z_2 + \mathrm{Var}c_Y = a_Y^2 + b_Y^2 \\
\mathrm{Cov}(X,Y) &= \mathrm{E}XY - \mathrm{E}X\cdot\mathrm{E}Y \\
&= \mathrm{E}[(a_Xa_YZ_1^2 + b_Xb_YZ_2^2 + c_Xc_Y + a_Xb_YZ_1Z_2 + a_Xc_YZ_1 + b_Xa_YZ_2Z_1 \\
&\qquad + b_Xc_YZ_2 + c_Xa_YZ_1 + c_Xb_YZ_2) - c_Xc_Y] \\
&= a_Xa_Y + b_Xb_Y,
\end{aligned}$$

since $\mathrm{E}Z_1^2 = \mathrm{E}Z_2^2 = 1$, and expectations of other terms are all zero.

b. Simply plug the expressions for a_X, b_X, etc. into the equalities in a) and simplify.

c. Let $D = a_Xb_Y - a_Yb_X = -\sqrt{1-\rho^2}\sigma_X\sigma_Y$ and solve for Z_1 and Z_2,

$$\begin{aligned}
Z_1 &= \frac{b_Y(X-c_X) - b_X(Y-c_Y)}{D} = \frac{\sigma_Y(X-\mu_X)+\sigma_X(Y-\mu_Y)}{\sqrt{2(1+\rho)}\sigma_X\sigma_Y} \\
Z_2 &= \frac{\sigma_Y(X-\mu_X)+\sigma_X(Y-\mu_Y)}{\sqrt{2(1-\rho)}\sigma_X\sigma_Y}.
\end{aligned}$$

Then the Jacobian is

$$J = \begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial y} \\ \frac{\partial z_2}{\partial x} & \frac{\partial z_2}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{b_Y}{D} & \frac{-b_X}{D} \\ \frac{-a_Y}{D} & \frac{a_X}{D} \end{pmatrix} = \frac{a_X b_Y}{D^2} - \frac{a_Y b_X}{D^2} = \frac{1}{D} = \frac{1}{-\sqrt{1-\rho^2}\sigma_X\sigma_Y},$$

and we have that

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\sigma_Y(x-\mu_X)+\sigma_X(y-\mu_Y))^2}{2(1+\rho)\sigma_X^2\sigma_Y^2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\sigma_Y(x-\mu_X)+\sigma_X(y-\mu_Y))^2}{2(1-\rho)\sigma_X^2\sigma_Y^2}} \frac{1}{\sqrt{1-\rho^2}\sigma_X\sigma_Y} \\ &= (2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})^{-1} \exp\left(-\frac{1}{2(1-\rho^2)} \left(\frac{x-\mu_X}{\sigma_X}\right)^2 \right) \\ &- 2\rho\frac{x-\mu_X}{\sigma_X}\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2, \quad -\infty < x < \infty,\ -\infty < y < \infty, \end{aligned}$$

a bivariate normal pdf.

d. Another solution is

$$\begin{aligned} a_X &= \rho\sigma_X b_X = \sqrt{(1-\rho^2)}\sigma_X \\ a_Y &= \sigma_Y b_Y = 0 \\ c_X &= \mu_X \\ c_Y &= \mu_Y. \end{aligned}$$

There are an infinite number of solutions. Write $b_X = \pm\sqrt{\sigma_X^2 - a_X^2}$,$b_Y = \pm\sqrt{\sigma_Y^2 - a_Y^2}$, and substitute b_X,b_Y into $a_X a_Y = \rho\sigma_X\sigma_Y$. We get

$$a_X a_Y + \left(\pm\sqrt{\sigma_X^2 - a_X^2}\right)\left(\pm\sqrt{\sigma_Y^2 - a_Y^2}\right) = \rho\sigma_X\sigma_Y.$$

Square both sides and simplify to get

$$(1-\rho^2)\sigma_X^2\sigma_Y^2 = \sigma_X^2 a_Y^2 - 2\rho\sigma_X\sigma_Y a_X a_Y + \sigma_Y^2 a_X^2.$$

This is an ellipse for $\rho \neq \pm 1$, a line for $\rho = \pm 1$. In either case there are an infinite number of points satisfying the equations.

4.47 a. By definition of Z, for $z < 0$,

$$\begin{aligned} P(Z \le z) &= P(X \le z \text{ and } XY > 0) + P(-X \le z \text{ and } XY < 0) \\ &= P(X \le z \text{ and } Y < 0) + P(X \ge -z \text{ and } Y < 0) \qquad \text{(since } z < 0) \\ &= P(X \le z)P(Y < 0) + P(X \ge -z)P(Y < 0) \qquad \text{(independence)} \\ &= P(X \le z)P(Y < 0) + P(X \le z)P(Y > 0) \qquad \text{(symmetry of } X \text{and } Y) \\ &= P(X \le z)(P(Y < 0) + P(Y > 0)) \\ &= P(X \le z). \end{aligned}$$

By a similar argument, for $z > 0$, we get $P(Z > z) = P(X > z)$, and hence, $P(Z \le z) = P(X \le z)$. Thus, $Z \sim X \sim \mathrm{n}(0, 1)$.

b. By definition of Z, $Z > 0 \Leftrightarrow$ either (i)$X < 0$ and $Y > 0$ or (ii)$X > 0$ and $Y > 0$. So Z and Y always have the same sign, hence they cannot be bivariate normal.

4.49 a.

$$\begin{aligned} f_X(x) &= \int (af_1(x)g_1(y) + (1-a)f_2(x)g_2(y))dy \\ &= af_1(x)\int g_1(y)dy + (1-a)f_2(x)\int g_2(y)dy \\ &= af_1(x) + (1-a)f_2(x). \\ f_Y(y) &= \int (af_1(x)g_1(y) + (1-a)f_2(x)g_2(y))dx \\ &= ag_1(y)\int f_1(x)dx + (1-a)g_2(y)\int f_2(x)dx \\ &= ag_1(y) + (1-a)g_2(y). \end{aligned}$$

b. ($\Rightarrow$) If X and Y are independent then $f(x,y) = f_X(x)f_Y(y)$. Then,

$$\begin{aligned} &f(x,y) - f_X(x)f_Y(y) \\ &\quad= af_1(x)g_1(y) + (1-a)f_2(x)g_2(y) \\ &\qquad - [af_1(x) + (1-a)f_2(x)][ag_1(y) + (1-a)g_2(y)] \\ &\quad= a(1-a)[f_1(x)g_1(y) - f_1(x)g_2(y) - f_2(x)g_1(y) + f_2(x)g_2(y)] \\ &\quad= a(1-a)[f_1(x) - f_2(x)][g_1(y) - g_2(y)] \\ &\quad= 0. \end{aligned}$$

Thus $[f_1(x) - f_2(x)][g_1(y) - g_2(y)] = 0$ since $0 < a < 1$.
($\Leftarrow$) if $[f_1(x) - f_2(x)][g_1(y) - g_2(y)] = 0$ then

$$f_1(x)g_1(y) + f_2(x)g_2(y) = f_1(x)g_2(y) + f_2(x)g_1(y).$$

Therefore

$$\begin{aligned} &f_X(x)f_Y(y) \\ &\quad= a^2f_1(x)g_1(y) + a(1-a)f_1(x)g_2(y) + a(1-a)f_2(x)g_1(y) + (1-a)^2f_2(x)g_2(y) \\ &\quad= a^2f_1(x)g_1(y) + a(1-a)[f_1(x)g_2(y) + f_2(x)g_1(y)] + (1-a)^2f_2(x)g_2(y) \\ &\quad= a^2f_1(x)g_1(y) + a(1-a)[f_1(x)g_1(y) + f_2(x)g_2(y)] + (1-a)^2f_2(x)g_2(y) \\ &\quad= af_1(x)g_1(y) + (1-a)f_2(x)g_2(y) = f(x,y). \end{aligned}$$

Thus X and Y are independent.

c.

$$\begin{aligned} \mathrm{Cov}(X,Y) &= a\mu_1\xi_1 + (1-a)\mu_2\xi_2 - [a\mu_1 + (1-a)\mu_2][a\xi_1 + (1-a)\xi_2] \\ &= a(1-a)[\mu_1\xi_1 - \mu_1\xi_2 - \mu_2\xi_1 + \mu_2\xi_2] \\ &= a(1-a)[\mu_1 - \mu_2][\xi_1 - \xi_2]. \end{aligned}$$

To construct dependent uncorrelated random variables let $(X,Y) \sim af_1(x)g_1(y) + (1-a)f_2(x)g_2(y)$ where f_1, f_2, g_1, g_2 are such that $f_1 - f_2 \neq 0$ and $g_1 - g_2 \neq 0$ with $\mu_1 = \mu_2$ or $\xi_1 = \xi_2$.

d. (i) $f_1 \sim$ binomial(n,p), $f_2 \sim$ binomial(n,p), $g_1 \sim$ binomial(n,p), $g_2 \sim$ binomial$(n, 1-p)$.
(ii) $f_1 \sim$ binomial(n,p_1), $f_2 \sim$ binomial(n,p_2), $g_1 \sim$ binomial(n,p_1), $g_2 \sim$ binomial(n,p_2).
(iii) $f_1 \sim$ binomial$(n_1, \frac{p}{n_1})$, $f_2 \sim$ binomial$(n_2, \frac{p}{n_2})$, $g_1 \sim$ binomial(n_1,p), $g_2 \sim$ binomial(n_2,p).

4.51 a.

$$\begin{aligned} P(X/Y \le t) &= \begin{cases} \frac{1}{2}t & t > 1 \\ \frac{1}{2} + (1-t) & t \le 1 \end{cases} \\ P(XY \le t) &= t - t\log t \qquad 0 < t < 1. \end{aligned}$$

b.

$$\begin{aligned} P(XY/Z \le t) &= \int_0^1 P(XY \le zt)dz \\ &= \begin{cases} \int_0^1 \left[\frac{zt}{2} + (1-zt)\right] dz & \text{if } t \le 1 \\ \int_0^{\frac{1}{t}} \left[\frac{zt}{2} + (1-zt)\right] dz + \int_{\frac{1}{t}}^1 \frac{1}{2zt}dz & \text{if } t \le 1 \end{cases} \\ &= \begin{cases} 1 - t/4 & \text{if } t \le 1 \\ t - \frac{1}{4t} + \frac{1}{2t}\log t & \text{if } t > 1 \end{cases}. \end{aligned}$$

4.53

$$\begin{aligned} P(\text{Real Roots}) &= P(B^2 > 4AC) \\ &= P(2\log B > \log 4 + \log A + \log C) \\ &= P(-2\log B \le -\log 4 - \log A - \log C) \\ &= P\left(-2\log B \le -\log 4 + (-\log A - \log C)\right). \end{aligned}$$

Let $X = -2\log B$, $Y = -\log A - \log C$. Then $X \sim$ exponential(2), $Y \sim$ gamma(2,1), independent, and

$$\begin{aligned} P(\text{Real Roots}) &= P(X < -\log 4 + Y) \\ &= \int_{\log 4}^{\infty} P(X < -\log 4 + y) f_Y(y)dy \\ &= \int_{\log 4}^{\infty} \int_0^{-\log 4 + y} \frac{1}{2} e^{-x/2} dx y e^{-y} dy \\ &= \int_{\log 4}^{\infty} \left(1 - e^{-\frac{1}{2}\log 4} e^{-y/2}\right) y e^{-y} dy. \end{aligned}$$

Integration-by-parts will show that $\int_a^\infty ye^{-y/b} = b(a+b)e^{-a/b}$ and hence

$$P(\text{Real Roots}) = \frac{1}{4}(1 + \log 4) - \frac{1}{24}\left(\frac{2}{3} + \log 4\right) = .511.$$

4.54 Let $Y = \prod_{i=1}^n X_i$. Then $P(Y \le y) = P(\prod_{i=1}^n X_i \le y) = P(\sum_{i=1}^n -\log X_i \ge -\log y)$. Now, $-\log X_i \sim$ exponential(1) = gamma(1,1). By Example 4.6.8, $\sum_{i=1}^n -\log X_i \sim$ gamma$(n,1)$. Therefore,

$$P(Y \le y) = \int_{-\log y}^{\infty} \frac{1}{\Gamma(n)} z^{n-1} e^{-z} dz,$$

and

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} \int_{-\log y}^{\infty} \frac{1}{\Gamma(n)} z^{n-1} e^{-z} dz \\ &= -\frac{1}{\Gamma(n)} (-\log y)^{n-1} e^{-(-\log y)} \frac{d}{dy}(-\log y) \\ &= \frac{1}{\Gamma(n)} (-\log y)^{n-1}, \qquad 0 < y < 1. \end{aligned}$$

4.55 Let X_1, X_2, X_3 be independent exponential(λ) random variables, and let $Y = \max(X_1, X_2, X_3)$, the lifetime of the system. Then

$$\begin{aligned} P(Y \le y) &= P\left(\max(X_1, X_2, X_3) \le y\right) \\ &= P\left(X_1 \le y \text{ and } X_2 \le y \text{ and } X_3 \le y\right) \\ &= P(X_1 \le y)P(X_2 \le y)P(X_3 \le y). \end{aligned}$$

by the independence of X_1, X_2 and X_3. Now each probability is $P(X_1 \le y) = \int_0^y \frac{1}{\lambda} e^{-x/\lambda} dx = 1 - e^{-y/\lambda}$, so

$$P(Y \le y) = \left(1 - e^{-y/\lambda}\right)^3, \qquad 0 < y < \infty,$$

and the pdf is

$$f_Y(y) = \begin{cases} 3\left(1 - e^{-y/\lambda}\right)^2 e^{-y/\lambda} & y > 0 \\ 0 & y \le 0. \end{cases}$$

4.57 a.

$$\begin{aligned} A_1 &= [\frac{1}{n}\sum_{x=1}^{n} x_i^1]^{\frac{1}{1}} = \frac{1}{n}\sum_{x=1}^{n} x_i, \quad \text{the arithmetic mean.} \\ A_{-1} &= [\frac{1}{n}\sum_{x=1}^{n} x_i^{-1}]^{-1} = \frac{1}{\frac{1}{n}(\frac{1}{x_1} + \cdots + \frac{1}{x_n})}, \quad \text{the harmonic mean.} \end{aligned}$$

$$\begin{aligned} \lim_{r\to 0} \log A_r &= \lim_{r\to 0} \log[\frac{1}{n}\sum_{x=1}^{n} x_i^r]^{\frac{1}{r}} = \lim_{r\to 0} \frac{1}{r}\log[\frac{1}{n}\sum_{x=1}^{n} x_i^r] = \lim_{r\to 0} \frac{\frac{1}{n}\sum_{i=1}^n r x_i^{r-1}}{\frac{1}{n}\sum_{i=1}^n x_i^r} \\ &= \lim_{r\to 0} \frac{\frac{1}{n}\sum_{i=1}^n x_i^r \log x_i}{\frac{1}{n}\sum_{i=1}^n x_i^r} = \frac{1}{n}\sum_{i=1}^{n} \log x_i = \frac{1}{n}\log(\prod_{i=1}^{n} x_i). \end{aligned}$$

Thus $A_0 = \lim_{r\to 0} A_r = \exp(\frac{1}{n}\log(\prod_{i=1}^n x_i)) = (\prod_{i=1}^n x_i)^{\frac{1}{n}}$, the geometric mean. The term $r x_i^{r-1} = x_i^r \log x_i$ since $r x_i^{r-1} = \frac{d}{dr} x_i^r = \frac{d}{dr}\exp(r \log x_i) = \exp(r\log x_i)\log x_i = x_i^r \log x_i$.

b. (i) if $\log A_r$ is nondecreasing then for $r \le r'$ $\log A_r \le \log A_{r'}$, then $e^{\log A_r} \le e^{\log A_{r'}}$. Therefore $A_r \le A_{r'}$. Thus A_r is nondecreasing in r.

(ii) $\frac{d}{dr}\log A_r = \frac{-1}{r^2}\log(\frac{1}{n}\sum_{x=1}^n x_i^r) + \frac{1}{r}\frac{\frac{1}{n}\sum_{i=1}^n r x_i^{r-1}}{\frac{1}{n}\sum_{i=1}^n x_i^r} = \frac{1}{r^2}\left[\frac{r\sum_{i=1}^n x_i^r \log x_i}{\sum_{x=1}^n x_i^r} - \log(\frac{1}{n}\sum_{x=1}^n x_i^r)\right]$, where we use the identity for $r x_i^{r-1}$ showed in a).

(iii)

$$\begin{aligned} &\frac{r\sum_{i=1}^n x_i^r \log x_i}{\sum_{x=1}^n x_i^r} - \log(\frac{1}{n}\sum_{x=1}^{n} x_i^r) \\ &= \log(n) + \frac{r\sum_{i=1}^n x_i^r \log x_i}{\sum_{x=1}^n x_i^r} - \log(\sum_{x=1}^{n} x_i^r) \\ &= \log(n) + \sum_{i=1}^{n}\left[\frac{x_i^r}{\sum_{i=1}^n x_i^r} r\log x_i - \frac{x_i^r}{\sum_{i=1}^n x_i^r}\log(\sum_{x=1}^{n} x_i^r)\right] \\ &= \log(n) + \sum_{i=1}^{n}\left[\frac{x_i^r}{\sum_{i=1}^n x_i^r}(r\log x_i - \log(\sum_{x=1}^{n} x_i^r))\right] \\ &= \log(n) - \sum_{i=1}^{n}\frac{x_i^r}{\sum_{i=1}^n x_i^r}\log(\frac{\sum_{x=1}^n x_i^r}{x_i^r}) = \log(n) - \sum_{i=1}^{n} a_i \log(\frac{1}{a_i}). \end{aligned}$$

We need to prove that $\log(n) \geq \sum_{i=1}^n a_i \log(\frac{1}{a_i})$. Using Jensen inequality we have that $\mathrm{E}\log(\frac{1}{a}) = \sum_{i=1}^n a_i \log(\frac{1}{a_i}) \leq \log(\mathrm{E}\frac{1}{a}) = \log(\sum_{i=1}^n a_i \frac{1}{a_i}) = \log(n)$ which establish the result.

4.59 Assume that $\mathrm{E}X = 0$, $\mathrm{E}Y = 0$, and $\mathrm{E}Z = 0$. This can be done without loss of generality because we could work with the quantities $X - \mathrm{E}X$, etc. By iterating the expectation we have

$$\mathrm{Cov}(X,Y) = \mathrm{E}XY = \mathrm{E}[\mathrm{E}(XY|Z)].$$

Adding and subtracting $\mathrm{E}(X|Z)\mathrm{E}(Y|Z)$ gives

$$\mathrm{Cov}(X,Y) = \mathrm{E}[\mathrm{E}(XY|Z) - \mathrm{E}(X|Z)\mathrm{E}(Y|Z)] + \mathrm{E}[\mathrm{E}(X|Z)\mathrm{E}(Y|Z)].$$

Since $\mathrm{E}[\mathrm{E}(X|Z)] = \mathrm{E}X = 0$, the second term above is $\mathrm{Cov}[\mathrm{E}(X|Z)\mathrm{E}(Y|Z)]$. For the first term write

$$\mathrm{E}[\mathrm{E}(XY|Z) - \mathrm{E}(X|Z)\mathrm{E}(Y|Z)] = \mathrm{E}\left[\mathrm{E}\left\{XY - \mathrm{E}(X|Z)\mathrm{E}(Y|Z)\right| Z\}\right]$$

where we have brought $\mathrm{E}(X|Z)$ and $\mathrm{E}(Y|Z)$ inside the conditional expectation. This can now be recognized as $\mathrm{E}\mathrm{Cov}(X,Y|Z)$, establishing the identity.

4.61 a. To find the distribution of $f(X_1|Z)$, let $U = \frac{X_2-1}{X_1}$ and $V = X_1$. Then $x_2 = h_1(u,v) = uv+1$, $x_1 = h_2(u,v) = v$. Therefore

$$f_{U,V}(u,v) = f_{X,Y}(h_1(u,v), h_2(u,v))|J| = e^{-(uv+1)}e^{-v}v,$$

and

$$f_U(u) = \int_0^\infty ve^{-(uv+1)}e^{-v}dv = \frac{e^{-1}}{(u+1)^2}.$$

Thus $V|U = 0$ has distribution ve^{-v}. The distribution of $X_1|X_2$ is e^{-x_1} since X_1 and X_2 are independent.

b. The following Mathematica code will draw the picture; the solid lines are B_1 and the dashed lines are B_2. Note that the solid lines increase with $x1$, while the dashed lines are constant. Thus B_1 is informative, as the range of X_2 changes.

```
e = 1/10;
Plot[{-e*x1 + 1, e*x1 + 1, 1 - e, 1 + e}, {x1, 0, 5},
PlotStyle -> {Dashing[{}], Dashing[{}],Dashing[{0.15, 0.05}],
  Dashing[{0.15, 0.05}]}]
```

c.

$$\begin{aligned} P(X_1 \leq x|B_1) &= P(V \leq v^*| - \epsilon < U < \epsilon) = \frac{\int_0^{v^*}\int_{-\epsilon}^{\epsilon} ve^{-(uv+1)}e^{-v}dudv}{\int_0^{\infty}\int_{-\epsilon}^{\epsilon} ve^{-(uv+1)}e^{-v}dudv} \\ &= \frac{e^{-1}\left[\frac{e^{-v^*(1+\epsilon)}}{1+\epsilon} - \frac{1}{1+\epsilon} - \frac{e^{-v^*(1-\epsilon)}}{1-\epsilon} + \frac{1}{1-\epsilon}\right]}{e^{-1}\left[-\frac{1}{1+\epsilon} + \frac{1}{1-\epsilon}\right]}. \end{aligned}$$

Thus $\lim_{\epsilon\to 0} P(X_1 \leq x|B_1) = 1 - e^{-v^*} - v^*e^{-v^*} = \int_0^{v^*} ve^{-v}dv = P(V \leq v^*|U = 0)$.

$$P(X_1 \leq x|B_2) = \frac{\int_0^x \int_0^{1+\epsilon} e^{-(x_1+x_2)}dx_2dx_1}{\int_0^{1+\epsilon} e^{-x_2}dx_2} = \frac{e^{-(x+1+\epsilon)} - e^{-(1+\epsilon)} - e^{-x} + 1}{1 - e^{-(1+\epsilon)}}.$$

Thus $\lim_{\epsilon\to 0} P(X_1 \leq x|B_2) = 1 - e^x = \int_0^x e^{x_1}dx_1 = P(X_1 \leq x|X_2 = 1)$.

4.63 Since $X = e^Z$ and $g(z) = e^z$ is convex, by Jensen's Inequality $\mathrm{E}X = \mathrm{E}g(Z) \geq g(\mathrm{E}Z) = e^0 = 1$. In fact, there is equality in Jensen's Inequality if and only if there is an interval I with $P(Z \in I) = 1$ and $g(z)$ is linear on I. But e^z is linear on an interval only if the interval is a single point. So $\mathrm{E}X > 1$, unless $P(Z = \mathrm{E}Z = 0) = 1$.

4.64 a. Let a and b be real numbers. Then,

$$|a+b|^2 = (a+b)(a+b) = a^2 + 2ab + b^2 \leq |a|^2 + 2|ab| + |b|^2 = (|a| + |b|)^2.$$

Take the square root of both sides to get $|a+b| \leq |a| + |b|$.

b. $|X+Y| \leq |X| + |Y| \Rightarrow \mathrm{E}|X+Y| \leq \mathrm{E}(|X| + |Y|) = \mathrm{E}|X| + \mathrm{E}|Y|$.

4.65 Without loss of generality let us assume that $\mathrm{E}g(X) = \mathrm{E}h(X) = 0$. For part (a)

$$\begin{aligned}
\mathrm{E}(g(X)h(X)) &= \int_{-\infty}^{\infty} g(x)h(x)f_X(x)dx \\
&= \int_{\{x:h(x)\leq 0\}} g(x)h(x)f_X(x)dx + \int_{\{x:h(x)\geq 0\}} g(x)h(x)f_X(x)dx \\
&\leq g(x_0)\int_{\{x:h(x)\leq 0\}} h(x)f_X(x)dx + g(x_0)\int_{\{x:h(x)\geq 0\}} h(x)f_X(x)dx \\
&= \int_{-\infty}^{\infty} h(x)f_X(x)dx \\
&= g(x_0)\mathrm{E}h(X) = 0.
\end{aligned}$$

where x_0 is the number such that $h(x_0) = 0$. Note that $g(x_0)$ is a maximum in $\{x : h(x) \leq 0\}$ and a minimum in $\{x : h(x) \geq 0\}$ since $g(x)$ is nondecreasing. For part (b) where $g(x)$ and $h(x)$ are both nondecreasing

$$\begin{aligned}
\mathrm{E}(g(X)h(X)) &= \int_{-\infty}^{\infty} g(x)h(x)f_X(x)dx \\
&= \int_{\{x:h(x)\leq 0\}} g(x)h(x)f_X(x)dx + \int_{\{x:h(x)\geq 0\}} g(x)h(x)f_X(x)dx \\
&\geq g(x_0)\int_{\{x:h(x)\leq 0\}} h(x)f_X(x)dx + g(x_0)\int_{\{x:h(x)\geq 0\}} h(x)f_X(x)dx \\
&= \int_{-\infty}^{\infty} h(x)f_X(x)dx \\
&= g(x_0)\mathrm{E}h(X) = 0.
\end{aligned}$$

The case when $g(x)$ and $h(x)$ are both nonincreasing can be proved similarly.