Probability Homework 4

Michael Nelson

Problem 4.6

Let L be the length of time that A waits for B. Note that $\operatorname{im} L = [0,1]$. There's a $1/2$ chance that B arrives before A. In particular L is a mixture variable. Now let $\ell \in (0,1]$. Then

$$\begin{aligned} F_L(\ell) &= \frac{1}{2} + \frac{\text{area of region in plane bounded by } y = x,\ y = x + \ell,\ x = 0, \text{ and } y = 1}{\text{area of unit square}} \\ &= \frac{1}{2} + \frac{\ell - \ell^2 + \ell^2/2}{1} \\ &= \frac{1}{2} + \ell - \frac{\ell^2}{2}. \end{aligned}$$

Thus the cdf of L is given by

$$F_L(\ell) = \begin{cases} 0 & \text{if } -\infty < \ell < 0 \\ \frac{1}{2} + \ell - \frac{\ell^2}{2} & \text{if } 0 \leq \ell \leq 1 \\ 1 & \text{if } 1 < \ell < \infty \end{cases}$$

Problem 4.7

The probability that she makes it on time is given by

$$\begin{aligned} \text{P(makes it before 9:00 am)} &= \frac{\text{area of region in plane bounded by } x = 0,\ y = 5/6, y = 4/6, \text{ and } y = 1 - x}{\text{area of region in plane bounded by } x = 0,\ y = 5/6, y = 4/6, \text{ and } x = 1/2} \\ &= \frac{\frac{1}{6} \cdot \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{6} \cdot \frac{1}{6}}{\frac{1}{6} \cdot \frac{1}{2}} \\ &= \frac{1}{3} + \frac{1}{6} \\ &= \frac{1}{2}. \end{aligned}$$

Problem 4.8

Problem 4.8.a

If we are given $M = m$, then there are only two possible values which X takes, namely $X = m$ and $X = 2m$. Each value is assumed to occur with equal probability, and so therefore

$$P(X = m \mid M = m) = P(X = 2m \mid M = m) = \frac{1}{2}.$$

Also, note that

$$\begin{aligned}
P(M = x \mid X = x) &= \frac{P(M = x)P(X = x \mid M = x)}{P(X = x)} \\
&= \frac{1}{2}\frac{P(M = x)}{P(X = x)} \\
&= \frac{1}{2}\frac{\pi(x)}{P(X = x)} \\
&= \frac{1}{2}\frac{\pi(x)}{\frac{1}{2}\pi(x) + \frac{1}{2}\pi(x/2)} \\
&= \frac{\pi(x)}{\pi(x) + \pi(x/2)}.
\end{aligned}$$

A similar computation gives us

$$\begin{aligned}
P(M = x/2 \mid X = x) &= \frac{P(M = x/2)P(X = x \mid M = x/2)}{P(X = x)} \\
&= \frac{1}{2}\frac{P(M = x/2)}{P(X = x)} \\
&= \frac{1}{2}\frac{\pi(x/2)}{P(X = x)} \\
&= \frac{1}{2}\frac{\pi(x/2)}{\frac{1}{2}\pi(x) + \frac{1}{2}\pi(x/2)} \\
&= \frac{\pi(x/2)}{\pi(x) + \pi(x/2)}.
\end{aligned}$$

Problem 4.8.b

The expected winning of a trade is

$$\frac{\pi(x)}{\pi(x) + \pi(x/2)}2x + \frac{\pi(x/2)}{\pi(x) + \pi(x/2)}\frac{x}{2} = \frac{x(4\pi(x) + \pi(x/2))}{2(\pi(x) + \pi(x/2))},$$

and we have

$$\begin{aligned}
\frac{x(4\pi(x) + \pi(x/2))}{2(\pi(x) + \pi(x/2))} > x &\iff \frac{4\pi(x) + \pi(x/2)}{2(\pi(x) + \pi(x/2))} > 1 \\
&\iff 4\pi(x) + \pi(x/2) > 2\pi(x) + 2\pi(x/2) \\
&\iff 2\pi(x) > \pi(x/2).
\end{aligned}$$

Furthermore, if $\pi(x) = \lambda e^{-\lambda x}$, then

$$\begin{aligned}
2\pi(x) > \pi(x/2) &\iff 2\lambda e^{-\lambda x} > \lambda e^{-\lambda x/2} \\
&\iff 2e^{-\lambda x} > e^{-\lambda x/2} \\
&\iff 2 > e^{\lambda x - \lambda x/2} \\
&\iff 2 > e^{\lambda x/2} \\
&\iff \log 2 > \lambda x/2 \\
&\iff (2\log 2)/\lambda > x.
\end{aligned}$$

Problem 4.8.c

If we know that $X = m$, then there is only one possible value for Y, namely $Y = 2m$ (or $Y = 2x$). Similarly, if we know that $X = 2m$, then there is only one possible value for Y, namely $Y = m$ (or $Y = x/2$). Thus

$$\mathrm{P}(Y = 2x \mid X = m) = \mathrm{P}(Y = x/2 \mid X = 2m) = 1.$$

The expected winning if you keep or trade your envelope is performed doing an iterated expectation. Namely,

$$\begin{aligned}
\mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}[Y|X]] \\
&= \mathrm{E}[Y|X=m]\mathrm{P}(X=m) + \mathrm{E}[Y|X=2m]\mathrm{P}(X=2m) \\
&= \frac{1}{2}\cdot \mathrm{E}[Y|X=m] + \frac{1}{2}\cdot \mathrm{E}[Y|X=2m] \\
&= \frac{1}{2}\cdot (2m\mathrm{P}(Y=2m|X=m) + m\mathrm{P}(Y=m|X=m)) + \frac{1}{2}\cdot (2m\mathrm{P}(Y=2m|X=2m) + m\mathrm{P}(Y=m|X=2m)) \\
&= \frac{1}{2}\cdot 2m + \frac{1}{2}\cdot m \\
&= \frac{3m}{2}.
\end{aligned}$$

A similar calculation gives $\mathrm{E}(X) = 3m/2$.

Problem 4.12

We first prove the following propsition:

Proposition 0.1. *Let a, b, c be real numbers which satisfy the inequalities*

$$\begin{aligned}
0 < c < a+b \\
0 < b < a+c \\
0 < a < b+c.
\end{aligned}$$

Then there exists a triangle whose sides have lengths a, b, and c.

Remark 1. Note that if $a \leq c$ and $b \leq c$, then the only inequality that we need to show is $c < a+b$ since, in this case, we clearly have $a < b+c$ and $b < a+c$.

Proof. We may assume, without loss of generality, that $b \geq a$. Let $\mathbf{e}_1 = (1,0)^\top$ be the standard coordinate vector in $\mathbb{R}^2$ with 1 in its first entry and 0 in its second entry. In particular, the vector $a\mathbf{e}_1$ has length a. Next, let $\theta \in (0,\pi)$ and let $\mathbf{v}_\theta = (\cos\theta, \sin\theta)^\top$ be the vector in $\mathbb{R}^2$ which corresponds to a point on the unit circle. In particular, the vector $b\mathbf{v}_\theta$ has length b. Now define a function $f\colon (0,\pi) \to \mathbb{R}$ by

$$\begin{aligned}
f(\theta) &= \|b\mathbf{v}_\theta - a\mathbf{e}_1\| \\
&= \left\| \begin{pmatrix} b\cos\theta - a \\ b\sin\theta \end{pmatrix} \right\| \\
&= \sqrt{(b\cos\theta - a)^2 + b^2\sin^2\theta} \\
&= \sqrt{b^2\cos^2\theta - 2ab\cos\theta + a^2 + b^2\sin^2\theta} \\
&= \sqrt{b^2 + a^2 - 2ab\cos\theta}.
\end{aligned}$$

Observe that when $\theta = 0$, we have $f(0) = b - a$, and when $\theta = \pi$, we have $f(\pi) = b + a$. In particular, f is a continuous function whose domain is $(0,\pi)$ and whose range is $(b-a, b+a)$. Now note that the inequalities $b < a+c$ and $c < b+a$ implies $c \in (b-a, b+a)$. It follows from the intermediate value theorem that there exists a $\theta \in (0,\pi)$ such that $f(\theta) = c$. It follows that the vectors $a\mathbf{e}_1, b\mathbf{v}_\theta$, and $b\mathbf{v}_\theta - a\mathbf{e}_1$ forms a triangle with lengths a, b, and c respectively. □

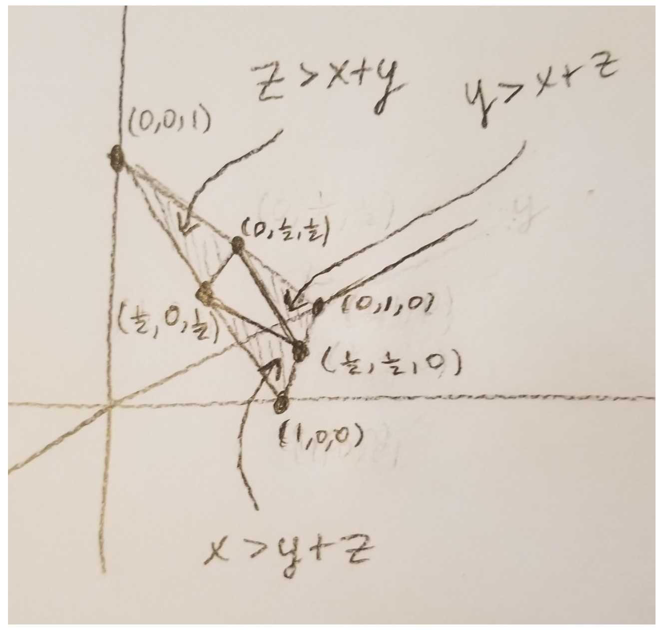
Now we can tackle problem 4.12. Assume that the length of the stick is 1. Let X, Y, and Z be random variables which correspond to the lengths of the first, second, and third segment of the stick respectively. In particular, these random variables are subject to the constraints

$$X + Y + Z = 1 \qquad \text{and} \qquad 0 \leq X, Y, Z \leq 1.$$

Thus (X, Y, Z) can be viewed as a point lying on the 2-simplex in $\mathbb{R}^3$. Now the proposition above tells us that the lengths X, Y, Z can be made into a triangle if and only if they satisfy those inequalities. The set of all points in $\mathbb{R}^3_{\geq 0}$ which satisfies

$$x + y + z = 1, \quad x < y + z, \quad y < x + z, \quad \text{and} \quad z < x + y$$

corresponds to the middle unshaded triangle in the image below



This area is easily seen to be 1/4 the area of the larger triangle. Thus the probability that these three segments can be formed into a triangle is 1/4.

Problem 4.16

Let $0 < p < 1$ and let X and Y be two independent random variables such that $X \sim \text{geom}(p) \sim Y$. We recall that this means the pmf of X and Y is given by

$$f_Y(m) = f_X(m) = \begin{cases} p(1-p)^{m-1} & m \in \mathbb{Z}_{\geq 1} \\ 0 & \text{else} \end{cases}$$

Problem 4.16.a

We first note that $\operatorname{supp} U = \mathbb{Z}_{\geq 1}$ and $\operatorname{supp} V = \mathbb{Z}$. Let $u \in \mathbb{Z}_{\geq 1}$ and let $v \in \mathbb{Z}$. Then we have

$$\begin{aligned}
f_{U,V}(u,v) &= \mathrm{P}(U = u, V = v) \\
&= \mathrm{P}(\min(X,Y) = u, X - Y = v) \\
&= \begin{cases} \mathrm{P}(X = u, Y = u - v) & \text{if } v < 0 \\ \mathrm{P}(Y = u, X = u + v) & \text{if } v \geq 0 \end{cases} \\
&= \begin{cases} f_X(u) f_Y(u - v) & \text{if } v < 0 \\ f_X(u + v) f_Y(u) & \text{if } v \geq 0 \end{cases} \\
&= \begin{cases} p(1-p)^{u-1} p(1-p)^{u-v-1} & \text{if } v < 0 \\ p(1-p)^{u+v-1} p(1-p)^{u-1} & \text{if } v \geq 0 \end{cases} \\
&= \begin{cases} p^2(1-p)^{2u-v-2} & \text{if } v < 0 \\ p^2(1-p)^{2u+v-2} & \text{if } v \geq 0 \end{cases} \\
&= p^2(1-p)^{2u+|v|-2} \\
&= p(1-p)^{2u-1} p(1-p)^{|v|-1}.
\end{aligned}$$

Thus $f_{U,V}$ can be expressed as a product of two functions, one involving only u terms and involving only v terms. It follows that U and V are independent.

Problem 4.16.b

We first note that

$$\operatorname{supp} Z = \{m/n \mid m, n \in \mathbb{Z}_{\geq 1},\ \gcd(m,n) = 1, \text{ and } n > m\}.$$

Now let $m/n \in \operatorname{supp} Z$. Then we have

$$\begin{aligned}
\mathrm{P}\left(Z = \frac{m}{n}\right) &= \sum_{k=1}^{\infty} \mathrm{P}(X = km, Y = k(n-m)) \\
&= \sum_{k=1}^{\infty} \mathrm{P}(X = km)\mathrm{P}(Y = k(n-m)) \\
&= \sum_{k=1}^{\infty} p(1-p)^{km-1} p(1-p)^{kn-km-1} \\
&= p^2(1-p)^{-2} \sum_{k=1}^{\infty} (1-p)^{kn} \\
&= \left(\frac{p}{1-p}\right)^2 \left(\frac{(1-p)^n}{1-(1-p)^n}\right).
\end{aligned}$$

In particular, note that the pmf of Z involves no m terms, so that we get the same probability for different m. The number of positive integers which are strictly less than n and also relatively prime to n is given by $\varphi(n)$ where φ is the Euler totient function. Since all probabilities must add up to 1, we arrive at the following identity

$$\begin{aligned}
1 &= \sum_{m/n \in \operatorname{supp} Z} \mathrm{P}\left(Z = \frac{m}{n}\right) \\
&= \sum_{\substack{1 \leq m < n \\ \gcd(m,n)=1}} \sum_{n=2}^{\infty} \mathrm{P}\left(Z = \frac{m}{n}\right) \\
&= \sum_{\substack{1 \leq m < n \\ \gcd(m,n)=1}} \sum_{n=2}^{\infty} \left(\frac{p}{1-p}\right)^2 \left(\frac{(1-p)^n}{1-(1-p)^n}\right) \\
&= \sum_{n=2}^{\infty} \varphi(n) \left(\frac{p}{1-p}\right)^2 \left(\frac{(1-p)^n}{1-(1-p)^n}\right).
\end{aligned}$$

Problem 4.16.c

We first note that $\operatorname{supp} X = \mathbb{Z}_{\geq 1}$ and $\operatorname{supp}(X+Y) = \mathbb{Z}_{\geq 2}$. Let $m \in \mathbb{Z}_{\geq 1}$ and let $n \in \mathbb{Z}_{\geq 2}$. Then we have

$$\begin{aligned}
f_{X,X+Y}(m,n) &= \mathrm{P}(X = m, X+Y = n) \\
&= \mathrm{P}(X = m, Y = n-m) \\
&= \mathrm{P}(X = m)\mathrm{P}(Y = n-m) \\
&= p(1-p)^{m-1} p(1-p)^{n-m-1} \\
&= p^2(1-p)^{n-2},
\end{aligned}$$

where $n > m$ (if $n \leq m$, then $\mathrm{P}(Y = n-m) = 0$ which forces $f_{X,X+Y}(m,n) = 0$). Thus

$$f_{X,X+Y}(m,n) = \begin{cases} p^2(1-p)^{n-2} & \text{if } m \in \mathbb{Z}_{\geq 1},\ n \in \mathbb{Z}_{\geq 2}, \text{ and } n > m \\ 0 & \text{else} \end{cases}$$

Problem 4.18

We have $f(x,y) \geq 0$ since $g(x,y) \geq 0$ for all $(x,y) \in \mathbb{R}^2$. Also we have

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y &= \int_{0}^{\infty}\int_{0}^{\infty} f(x,y)\mathrm{d}x\mathrm{d}y \\
&= \int_{0}^{\infty}\int_{0}^{\infty} \frac{2g(\sqrt{x^2+y^2})}{\pi\sqrt{x^2+y^2}}\mathrm{d}x\mathrm{d}y \\
&= \int_{0}^{\pi/2}\int_{0}^{\infty} \frac{2g(r)}{\pi r} r\mathrm{d}r\mathrm{d}\theta \\
&= \frac{2}{\pi}\int_{0}^{\pi/2}\int_{0}^{\infty} g(r)\mathrm{d}r\mathrm{d}\theta \\
&= \frac{2}{\pi}\int_{0}^{\pi/2} \mathrm{d}\theta \\
&= \frac{2}{\pi}\cdot\frac{\pi}{2} \\
&= 1,
\end{aligned}$$

where we changed to polar coordinates in the third line.

Problem 4.21

Let R be a random variable such that $R^2 \sim \chi^2(2)$. Thus the pdf of R is given by

$$f_R(r) = \begin{cases} re^{-r^2/2} & \text{if } 0 < r < \infty \\ 0 & \text{else} \end{cases}$$

Let Θ be a random variable such that $\theta \sim \text{uniform}[0, 2\pi)$. Thus the pdf of Θ is given by

$$f_\Theta(\theta) = \begin{cases} 1/2\pi & \text{if } 0 \leq \theta < 2\pi \\ 0 & \text{else} \end{cases}$$

Suppose that R and Θ are independent. Let $\mathcal{A}$ be the support of the bivariate random variable (R, Θ). Thus

$$\begin{aligned}
\mathcal{A} &= \{(r,\theta) \in \mathbb{R}^2 \mid f_{R,\Theta}(r,\theta) > 0\} \\
&= \{(r,\theta) \in \mathbb{R}^2 \mid f_R(r)f_\Theta(\theta) > 0\} \\
&= \mathbb{R}_{>0} \times (0, 2\pi).
\end{aligned}$$

Let $g = (g_1, g_2)\colon \mathcal{A} \to \mathbb{R}^2$ be defined by

$$g_1(r,\theta) = r\cos\theta \qquad \text{and} \qquad g_2(r,\theta) = r\sin\theta$$

for all $(r,\theta) \in \mathcal{A}$. Denote $X = g_1(R,\Theta)$, denote $Y = g_2(R,\Theta)$, and denote $\mathcal{B} = \operatorname{im} g$. Thus

$$\begin{aligned}
\mathcal{B} &= \{(x,y) \in \mathbb{R}^2 \mid x = g_1(r,\theta) \text{ and } y = g_2(r,\theta) \text{ for some } (r,\theta) \in \mathcal{A}\} \\
&= \{(x,y) \in \mathbb{R}^2 \mid x = r\cos\theta \text{ and } y = r\sin\theta \text{ for some } (r,\theta) \in \mathcal{A}\} \\
&= \mathbb{R}^2 \backslash \{(0,0)\}.
\end{aligned}$$

Then g is a one-one and onto map from $\mathcal{A}$ to $\mathcal{B}$ with inverse map $h = (h_1, h_2)\colon \mathcal{B} \to \mathcal{A}$ defined by

$$h_1(x,y) = \sqrt{x^2+y^2} \qquad \text{and} \qquad h_2(x,y) = \tan^{-1}(y/x)$$

for all $(x,y) \in \mathcal{B}$. Note that by $\tan^{-1}(y/x)$, we do *not* mean the arctangent function! The arctangent function function has range $(-\pi/2, \pi/2)$, whereas $\tan^{-1}(y/x)$ has range $[0, 2\pi)$. In more detail, we define $\tan^{-1}(y/x)$ as follows:

$$\tan^{-1}(y/x) = \begin{cases} \arctan(y/x) & \text{if } x > 0 \text{ and } y > 0 \\ \pi/2 & \text{if } x = 0 \text{ and } y > 0 \\ \pi + \arctan(y/x) & \text{if } x < 0 \\ 3\pi/2 & \text{if } x = 0 \text{ and } y < 0 \\ 2\pi + \arctan(y/x) & \text{if } x > 0 \text{ and } y < 0 \end{cases}$$

Even though $\tan^{-1}(y/x)$ and $\arctan(y/x)$ are not the same functions, they still have the same partial derivatives since they differ by a constant. Thus the absolute value of the Jacobian of h at the point $(x,y) \in \mathcal{B}$ is given by

$$\begin{aligned} \left|\mathrm{J}_{(x,y)}(h)\right| &= \left|\det\begin{pmatrix} \partial_x h_1 & \partial_y h_1 \\ \partial_x h_2 & \partial_y h_2 \end{pmatrix}\right| \\ &= \left|\det\begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{y^2+x^2} & \frac{x}{y^2+x^2} \end{pmatrix}\right| \\ &= \frac{1}{\sqrt{x^2+y^2}}. \end{aligned}$$

It follows that

$$\begin{aligned} f_{X,Y}(x,y) &= f_{R,\Theta}(h_1(x,y), h_2(x,y)) \cdot \left|\mathrm{J}_{(x,y)}(h)\right| \\ &= f_{R,\Theta}\left(\sqrt{x^2+y^2}, \tan^{-1}(y/x)\right)(x^2+y^2)^{-1/2} \\ &= f_R\left(\sqrt{x^2+y^2}\right) f_\Theta\left(\tan^{-1}(y/x)\right)(x^2+y^2)^{-1/2} \\ &= \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \end{aligned}$$

Note that the decomposition of $f_{X,Y}$ implies $X \sim \mathrm{n}(0,1) \sim Y$.

Problem 4.26

For this problem. let $\lambda, \mu > 0$ and let X and Y be independent random variables with $X \sim \text{exponential}(\lambda)$ and $Y \sim \text{exponential}(\mu)$. Thus the pdf of X is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

and the pdf of Y is given by

$$f_Y(y) = \begin{cases} \mu e^{-\mu y} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Define random variables Z and W by

$$Z = \min(X,Y) \quad \text{and} \quad W = \begin{cases} 1 & \text{if } Z = X \\ 0 & \text{if } Z = Y \end{cases}$$

Problem 4.26.a

First note that $\operatorname{supp} Z = \mathbb{R}_{\geq 0}$ and $\operatorname{supp} W = \{1, 0\}$. Now let $z \in \mathbb{R}_{\geq 0}$. First we consider the case where $W = 0$. We have

$$\begin{aligned}
\mathrm{P}(Z \leq z, W = 0) &= \mathrm{P}(\min(X, Y) \leq z, Z = Y) \\
&= \mathrm{P}(Y \leq z, Y \leq X) \\
&= \int_0^z \int_0^x f_{X,Y}(x, y) \mathrm{d}y \mathrm{d}x + \int_z^\infty \int_0^z f_{X,Y}(x, y) \mathrm{d}y \mathrm{d}x \\
&= \int_0^z \lambda e^{-\lambda x} \int_0^x \mu e^{-\mu y} \mathrm{d}y \mathrm{d}x + \int_z^\infty \lambda e^{-\lambda x} \int_0^z \mu e^{-\mu y} \mathrm{d}y \mathrm{d}x \\
&= \int_0^z \lambda e^{-\lambda x} \left(1 - e^{-\mu x}\right) \mathrm{d}x + \left(1 - e^{-\mu z}\right) e^{-\lambda z} \\
&= 1 - e^{-\lambda z} - \frac{\lambda}{\lambda + \mu} \left(1 - e^{-(\lambda+\mu)z}\right) + e^{-\lambda z} - e^{-(\lambda+\mu)z} \\
&= \frac{\mu}{\lambda + \mu} \left(1 - e^{-(\lambda+\mu)z}\right),
\end{aligned}$$

and after differentiating with respect to z, we get

$$f_{Z,W}(z, 0) = \begin{cases} \mu e^{-(\lambda+\mu)z} & \text{if } z \geq 0 \\ 0 & \text{else} \end{cases}$$

Now if $W = 1$, then arguing by symmetry, we obtain

$$\mathrm{P}(Z \leq z, W = 1) = \begin{cases} \frac{\lambda}{\lambda+\mu} \left(1 - e^{-(\lambda+\mu)z}\right) & \text{if } z \geq 0 \\ 0 & \text{else} \end{cases}$$

Similarly, we have

$$f_{Z,W}(z, 1) = \begin{cases} \lambda e^{-(\lambda+\mu)z} & \text{if } z \geq 0 \\ 0 & \text{else} \end{cases}$$

Problem 4.26.b

First note that

$$\begin{aligned}
\mathrm{P}(W = 0) &= \mathrm{P}(Y \leq X) \\
&= \int_0^\infty \int_0^x f_{X,Y}(x, y) \mathrm{d}y \mathrm{d}x \\
&= \int_0^\infty \lambda e^{-\lambda x} \left(1 - e^{-\mu x}\right) \mathrm{d}x \\
&= \lim_{z \to \infty} \left(\frac{(\lambda + \mu)(1 - e^{-\lambda z}) - \lambda(1 - e^{-(\lambda+\mu)z})}{\lambda + \mu} \right) \\
&= \frac{\mu}{\lambda + \mu}
\end{aligned}$$

We also have

$$\begin{aligned}
\mathrm{P}(Z \leq z) &= \mathrm{P}(\min\{X, Y\} \leq z) \\
&= \mathrm{P}(X \leq z) + \mathrm{P}(Y \leq z) - \mathrm{P}(X \leq z, Y \leq z) \\
&= 1 - e^{-\lambda z} + 1 - e^{-\mu z} - (1 - e^{-\lambda z})(1 - e^{-\mu z}) \\
&= 1 - e^{-(\mu+\lambda)z}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathrm{P}(Z \le z | W = 0) &= \frac{\mathrm{P}(Z \le z, W = 0)}{\mathrm{P}(W = 0)} \\
&= \frac{\mu}{\lambda + \mu}\left(1 - e^{-(\lambda+\mu)z}\right) \cdot \frac{\lambda + \mu}{\mu} \\
&= 1 - e^{-(\mu+\lambda)z} \\
&= \mathrm{P}(Z \le z).
\end{aligned}$$

Arguing by symmetry again, we also get

$$\mathrm{P}(Z \le z | W = 1) = \mathrm{P}(Z \le z).$$

It follows that Z and W are independent.

Problem 4.29

For this problem let Θ be a random variable such that $\Theta \sim \mathrm{u}(0, 2\pi)$, let R be a positive random variable, let $X = R\sin\Theta$, and let $Y = R\cos\Theta$.

Problem 4.29.a

We want to show that $Z = X/Y$ has a Cauchy distribution. First note that $\operatorname{supp} Z = \mathbb{R}$. Let $z \in \mathbb{R}$. If $z \ge 0$, then we have

$$\begin{aligned}
\mathrm{P}(Z \le z) &= \mathrm{P}\left(\frac{X}{Y} \le z\right) \\
&= \mathrm{P}\left(\tan\theta \le z\right) \\
&= \int_0^{\arctan z} \frac{1}{2\pi}\mathrm{d}t + \int_{\pi/2}^{\arctan z + \pi} \frac{1}{2\pi}\mathrm{d}t + \frac{1}{4} \\
&= \frac{1}{2\pi}\arctan z + \frac{1}{2\pi}\left(\arctan z + \pi - \frac{\pi}{2}\right) + \frac{1}{4} \\
&= \frac{1}{2\pi}\arctan z + \frac{1}{2\pi}\arctan z + \frac{1}{2} \\
&= \frac{1}{\pi}\arctan z + \frac{1}{2}.
\end{aligned}$$

If $z < 0$, then we have

$$\begin{aligned}
\mathrm{P}(Z \le z) &= \mathrm{P}\left(\frac{X}{Y} \le z\right) \\
&= \mathrm{P}\left(\tan\theta \le z\right) \\
&= \int_{\pi/2}^{\arctan z + \pi} \frac{1}{2\pi}\mathrm{d}t + \int_{3\pi/2}^{\arctan z + 2\pi} \frac{1}{2\pi}\mathrm{d}t \\
&= \frac{1}{2\pi}(2\arctan z + \pi) \\
&= \frac{1}{\pi}\arctan z + \frac{1}{2}.
\end{aligned}$$

In any case, we get

$$F_Z(z) = \begin{cases} \frac{1}{\pi}\arctan z + \frac{1}{2} & \text{if } z \in \mathbb{R} \\ 0 & \text{else} \end{cases}$$

To get the pdf, we differentiate with respect to z; we obtain

$$f_Z(z) = \begin{cases} \frac{1}{\pi}\frac{1}{1+z^2} & \text{if } z \in \mathbb{R} \\ 0 & \text{else} \end{cases}$$

Thus $Z \sim \text{Cauchy}(0,1)$.

Alternatively, we coult get the pdf as follows: let $A_0 = \{\pi/2, 3\pi/2\}$, let $A_1 = (-\pi/2, \pi/2)$, let $A_2 = (\pi/2, 3\pi/2)$, and let $A_3 = (3\pi/2, 5\pi/2)$. Then $\{A_0, A_1, A_2, A_3\}$ forms a partition of $(-\pi/2, 5\pi/2)$ which contains $(0, 2\pi)$. Also let $g_i(\theta) = \tan\theta$ for each $i = 1,2,3$. Then observe that each g_i is monotone increasing on A_i with $\operatorname{im} g_i = \mathbb{R}$ and

$$\begin{aligned} g_1^{-1}(z) &= \arctan z \\ g_2^{-1}(z) &= \arctan z + \pi \\ g_3^{-1}(z) &= \arctan z + 2\pi. \end{aligned}$$

It follows that the conditions of theorem 2.1.8 are satisfied, and thus

$$f_Z(z) = \begin{cases} \frac{1}{\pi}\frac{1}{1+z^2} & \text{if } z \in \mathbb{R} \\ 0 & \text{else} \end{cases}$$

Problem 4.29.b

We want to show that $W = 2XY/\sqrt{X^2+Y^2}$ has the same distribution as the distribution of $Y = R\sin\Theta$ (I think the book made a typo). First note that $\operatorname{supp} W = \mathbb{R}$. Let $w \in \mathbb{R}$. Then we have

$$\begin{aligned} \mathrm{P}(W \le w) &= \mathrm{P}\left(\frac{2XY}{\sqrt{X^2+Y^2}} \le w\right) \\ &= \mathrm{P}\left(2R\sin\Theta\cos\Theta \le w\right) \\ &= \mathrm{P}\left(R\sin(2\Theta) \le w\right) \\ &= \mathrm{P}\left(R\sin\Theta \le w\right) \\ &= F_Y(w) \end{aligned}$$

where in the fourth line we are using the fact that Θ is uniform and the fact that

$$\mathrm{m}\left(\{\theta \in (0,2\pi) \mid \sin(2\theta) \le w\}\right) = \mathrm{m}\left(\{\theta \in (0,2\pi) \mid \sin\theta \le w\}\right)$$

where m denotes the Lebesgue measure. In other words, these two sets have the same size, and it is equally likely that Θ lands in one over the other. In particular, if X and Y are both $\mathrm{n}(0,1)$ random variables, then $(2XY)/\sqrt{X^2+Y^2}$ is $\mathrm{n}(0,1)$ too.

Problem 4.32

We recall that the pdf of a gamma distribution is given by

$$f(x|\alpha,\beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & x \le 0 \end{cases}$$

where $\alpha > 0$ is the shape paramater of the distribution and $\beta > 0$ is the scale parameter of the distribution, and where the gamma function is defined as

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} \mathrm{d}x.$$

We also recall that a random variable X has a **Poisson distribution** (which we denote by $X \sim \text{Poisson}(\lambda)$) if the pmf of X is given by

$$f_X(k) = \begin{cases} \frac{\lambda^k e^{-\lambda}}{k!} & \text{if } k \in \mathbb{Z}_{\ge 0} \text{ (number of occurences)} \\ 0 & \text{else} \end{cases}$$

where $\lambda > 0$ is the “expected rate of occurences”.

Problem 4.32.a

We first find the marginal distribution of Y. We have

$$\begin{aligned}
f_Y(y) &= \mathrm{P}(Y = y, 0 < \Lambda < \infty) \\
&= \int_0^\infty f(y,\lambda)\mathrm{d}\lambda \\
&= \int_0^\infty f(y|\lambda)f(\lambda)\mathrm{d}\lambda \\
&= \int_0^\infty \frac{\lambda^y e^{-\lambda}}{y!}\frac{1}{\Gamma(\alpha)\beta^\alpha}\lambda^{\alpha-1}e^{-\lambda/\beta}\mathrm{d}\lambda \\
&= \frac{1}{y!\Gamma(\alpha)\beta^\alpha}\int_0^\infty \lambda^{y+\alpha-1}e^{-\lambda(1+1/\beta)}\mathrm{d}\lambda \\
&= \frac{1}{y!\Gamma(\alpha)\beta^\alpha(1+\beta^{-1})^{y+\alpha}}\int_0^\infty u^{y+\alpha-1}e^{-u}\mathrm{d}u \\
&= \frac{\Gamma(y+\alpha)}{y!\Gamma(\alpha)\beta^\alpha(1+\beta^{-1})^{y+\alpha}}.
\end{aligned}$$

where we did the change of variable with $u = \lambda(1+\beta^{-1})$. Furthermore, if $\alpha \in \mathbb{Z}$, then we have

$$\begin{aligned}
\frac{\Gamma(y+\alpha)}{y!\Gamma(\alpha)\beta^\alpha(1+\beta^{-1})^{y+\alpha}} &= \frac{(y+\alpha-1)!}{y!(\alpha-1)!\beta^\alpha(1+\beta^{-1})^{y+\alpha}} \\
&= \binom{y+\alpha-1}{y}\frac{1}{\beta^\alpha(1+\beta^{-1})^{y+\alpha}} \\
&= \binom{y+\alpha-1}{y}\frac{1}{\beta^\alpha(\frac{\beta+1}{\beta})^{y+\alpha}} \\
&= \binom{y+\alpha-1}{y}\frac{\beta^y}{(\beta+1)^{y+\alpha}} \\
&= \binom{y+\alpha-1}{y}\left(\frac{\beta}{\beta+1}\right)^y\left(\frac{1}{\beta+1}\right)^\alpha.
\end{aligned}$$

This shows us that $Y \sim \mathrm{NB}\left(\alpha, \frac{\beta}{\beta+1}\right)$. In particular, this implies

$$\mathrm{E}(Y) = \alpha\beta \qquad \text{and} \qquad \mathrm{Var}(Y) = \alpha\beta(\beta+1).$$

Problem 4.32.b

We have

$$\mathrm{P}(Y = y) = \sum_{n=0}^\infty \mathrm{P}(Y = y|N = n)\mathrm{P}(N = n)$$

$$
\begin{aligned}
f_Y(y) &= \sum_{n=0}^{\infty} f(y|n) f_N(n) \\
&= \sum_{n=0}^{\infty} f(y|n) \int_0^{\infty} f(n|\lambda) f(\lambda) \mathrm{d}\lambda \\
&= \sum_{n=y}^{\infty} \binom{n}{y} p^y (1-p)^{n-y} \int_0^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta} \mathrm{d}\lambda \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} e^{-\lambda} \lambda^{\alpha-1} e^{-\lambda/\beta} \sum_{n=y}^{\infty} \binom{n}{y} p^y (1-p)^{n-y} \frac{\lambda^n}{n!} \mathrm{d}\lambda \\
&= \frac{p^y (1-p)^{-y}}{\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} e^{-\lambda} \lambda^{\alpha-1} e^{-\lambda/\beta} \sum_{n=y}^{\infty} \binom{n}{y} (1-p)^n \frac{\lambda^n}{n!} \mathrm{d}\lambda \\
&= \frac{p^y (1-p)^{-y}}{\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} e^{-\lambda} \lambda^{\alpha-1} e^{-\lambda/\beta} \sum_{n=y}^{\infty} \frac{1}{y!(n-y)!} (1-p)^n \lambda^n \mathrm{d}\lambda \\
&= \frac{p^y (1-p)^{-y}}{\Gamma(\alpha)\beta^\alpha y!} \int_0^{\infty} e^{-\lambda} \lambda^{\alpha-1} e^{-\lambda/\beta} \sum_{n=y}^{\infty} \frac{((\lambda - p\lambda)^n}{(n-y)!} \mathrm{d}\lambda \\
&= \frac{p^y (1-p)^{-y} (1-p)^y}{\Gamma(\alpha)\beta^\alpha y!} \int_0^{\infty} \lambda^y e^{-\lambda} \lambda^{\alpha-1} e^{-\lambda/\beta} \sum_{m=0}^{\infty} \frac{(\lambda - p\lambda)^m}{m!} \mathrm{d}\lambda \\
&= \frac{p^y}{\Gamma(\alpha)\beta^\alpha y!} \int_0^{\infty} e^{-\lambda} \lambda^{y+\alpha-1} e^{-\lambda/\beta} e^{(\lambda - p\lambda)} \mathrm{d}\lambda \\
&= \frac{p^y}{\Gamma(\alpha)\beta^\alpha y!} \int_0^{\infty} \lambda^{y+\alpha-1} \frac{u}{(p+\beta^{-1})^{y+\alpha-1}} e^{(-p\lambda - \lambda/\beta)} \mathrm{d}\lambda \\
&= \frac{p^y}{\Gamma(\alpha)\beta^\alpha y! (p+\beta^{-1})(p+\beta^{-1})^{y+\alpha-1}} \int_0^{\infty} u^{y+\alpha-1} e^{-u} \mathrm{d}u \\
&= \frac{p^y \Gamma(y+\alpha)}{\Gamma(\alpha)\beta^\alpha y! (p+\beta^{-1})^{y+\alpha}}
\end{aligned}
$$

Problem 4.36

For each $1 \le i \le n$ let X_i and P_i be random variables such that $X_i|P_i \sim \text{bernoulli}(P_i)$ and $P_i \sim \text{beta}(\alpha, \beta)$. Also denote $Y = \sum_{i=1}^n X_i$.

Problem 4.36.a

We have

$$
\begin{aligned}
\mathrm{E}Y &= \mathrm{E}\left(\sum_{i=1}^n X_i\right) \\
&= \sum_{i=1}^n \mathrm{E}X_i \\
&= \sum_{i=1}^n \mathrm{E}(\mathrm{E}(X_i|P_i)) \\
&= \sum_{i=1}^n \mathrm{E}(P_i) && \text{Bernoulli mean} \\
&= \sum_{i=1}^n \frac{\alpha}{\alpha+\beta} && \text{beta mean} \\
&= \frac{n\alpha}{\alpha+\beta}
\end{aligned}
$$

Problem 4.36.b

First note that for each $1 \leq i \leq n$, we have

$$\begin{aligned} f_{X_i}(x) &= \int_0^1 f(x|p)f(p)\mathrm{d}p \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 p^{x+\alpha-1}(1-p)^{\beta-x}\mathrm{d}p. \end{aligned}$$

In particular, this implies

$$f_{X_i}(x) = \begin{cases} \frac{\beta}{\alpha+\beta} & \text{if } x = 0 \\ \frac{\alpha}{\alpha+\beta} & \text{if } x = 1 \\ 0 & \text{else} \end{cases}$$

In other words, $X_i \sim \text{Bernoulli}\left(\frac{\alpha}{\alpha+\beta}\right)$.

Also note that since X_i and X_j are independent whenever $i \neq j$, we have

$$\begin{aligned} \mathrm{Var}\left(\sum_{i=1}^n X_i\right) &= \mathrm{E}\left(\left(\sum_{i=1}^n X_i\right)^2\right) - \left(\mathrm{E}\left(\sum_{i=1}^n X_i\right)\right)^2 \\ &= \mathrm{E}\left(\sum_{i=1}^n X_i^2 + 2\sum_{1<i<j\leq n} X_i X_j\right) - \left(\sum_{i=1}^n \mathrm{E}X_i\right)^2 \\ &= \sum_{i=1}^n \mathrm{E}(X_i^2) + 2\sum_{1<i<j\leq n} \mathrm{E}(X_i)\mathrm{E}(X_j) - \left(\sum_{i=1}^n \mathrm{E}X_i\right)^2 \\ &= \sum_{i=1}^n \mathrm{Var}(X_i) + \sum_{i=1}^n \mathrm{E}(X_i)^2 + 2\sum_{1<i<j\leq n} \mathrm{E}(X_i)\mathrm{E}(X_j) - \left(\sum_{i=1}^n \mathrm{E}X_i\right)^2 \\ &= \sum_{i=1}^n \mathrm{Var}(X_i) + \left(\sum_{i=1}^n \mathrm{E}X_i\right)^2 - \left(\sum_{i=1}^n \mathrm{E}X_i\right)^2 \\ &= \sum_{i=1}^n \mathrm{Var}(X_i). \end{aligned}$$

Therefore

$$\begin{aligned} \mathrm{Var}\, Y &= \sum_{i=1}^n \mathrm{Var}\, X_i \\ &= \frac{n\alpha\beta}{(\alpha+\beta)^2}. \end{aligned}$$

Since each X_i is independent from each other, it follows that the mgf of Y is given by

$$\begin{aligned} M_Y(t) &= \prod_{i=1}^n M_{X_i}(t) \\ &= \prod_{i=1}^n \left(\frac{\beta+\alpha e^t}{\alpha+\beta}\right) \\ &= \left(\frac{\beta+\alpha e^t}{\alpha+\beta}\right)^n \\ &= \left(\left(\frac{\alpha}{\alpha+\beta}\right)e^t + \left(1 - \frac{\alpha}{\alpha+\beta}\right)\right)^n. \end{aligned}$$

This is the mgf of a binomial random variable with n trials each with probability $p = \frac{\alpha}{\alpha+\beta}$ of success. Thus $Y \sim \text{binomial}\left(n, \frac{\alpha}{\alpha+\beta}\right)$.

Problem 4.36.c

We have

$$\begin{aligned}
\mathrm{E}Y &= \mathrm{E}\left(\sum_{i=1}^{n} X_i\right) \\
&= \sum_{i=1}^{n} \mathrm{E}X_i \\
&= \sum_{i=1}^{n} \mathrm{E}(\mathrm{E}(X_i|P_i)) \\
&= \sum_{i=1}^{n} \mathrm{E}(n_i P_i) && \text{binomial mean} \\
&= \sum_{i=1}^{n} n_i \mathrm{E}(P_i) \\
&= \sum_{i=1}^{n} \frac{n_i \alpha}{\alpha+\beta} && \text{beta mean} \\
&= \frac{\alpha}{\alpha+\beta} \sum_{i=1}^{n} n_i
\end{aligned}$$

Since X_i and X_j are independent whenever $i \neq j$, we have (by the remark above)

$$\begin{aligned}
\mathrm{Var}(Y) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) \\
&= \sum_{i=1}^{n} n_i \frac{\alpha\beta(\alpha+\beta+n_i)}{(\alpha+\beta)^2(\alpha+\beta+1)},
\end{aligned}$$

where this calculation was done in the book (page 168, example 4.4.8).

Problem 4.38

Let $f(x)$ be a gamma(r, λ) pdf. Thus

$$f_X(x|r,\lambda) = \begin{cases} \frac{1}{\Gamma(r)\lambda^r} x^{r-1} e^{-x/\lambda} & x \in \mathbb{R}_{>0} \\ 0 & \text{else} \end{cases}$$

Problem 4.38.a

Suppose $r < 1$. Then

$$\begin{aligned}
\int_0^\lambda \frac{1}{\nu} e^{-x/\nu} p_\lambda(\nu) \mathrm{d}\nu &= \int_0^\lambda \frac{1}{\nu} e^{-x/\nu} p_\lambda(\nu) \mathrm{d}\nu \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^\lambda e^{-x/\nu} \nu^{r-2} (\lambda - \nu)^{-r} \mathrm{d}\nu \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_\infty^0 e^{\frac{-x(\lambda u + x)}{\lambda x}} \left(\frac{\lambda x}{\lambda u + x}\right)^{r-2} \left(\lambda - \frac{\lambda x}{\lambda u + x}\right)^{-r} (-1) \left(\frac{\lambda x}{\lambda u + x}\right)^2 \frac{1}{x} \mathrm{d}u \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^\infty e^{\frac{-\lambda u - x}{\lambda}} \left(\frac{\lambda x}{\lambda u + x}\right)^{r-2} \left(\frac{\lambda^2 u}{\lambda u + x}\right)^{-r} \left(\frac{\lambda x}{\lambda u + x}\right)^2 \frac{1}{x} \mathrm{d}u \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^\infty e^{-u - x/\lambda} \frac{\lambda^{r-2-2r+2} x^{r-2+2-1} u^{-r}}{(\lambda u + x)^{r-2-r+2}} \mathrm{d}u \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^\infty e^{-u} e^{-x/\lambda} \lambda^{-r} x^{r-1} u^{-r} \mathrm{d}u \\
&= \frac{\lambda^{-r} x^{r-1} e^{-x/\lambda}}{\Gamma(r)\Gamma(1-r)} \int_0^\infty u^{-r} e^{-u} \mathrm{d}u \\
&= \frac{\lambda^{-r} x^{r-1} e^{-x/\lambda}}{\Gamma(r)\Gamma(1-r)} \int_0^\infty u^{1-r-1} e^{-u} \mathrm{d}u \\
&= \frac{\lambda^{-r} x^{r-1} e^{-x/\lambda}}{\Gamma(r)\Gamma(1-r)} \Gamma(1-r) \\
&= \frac{\lambda^{-r} x^{r-1} e^{-x/\lambda}}{\Gamma(r)} \\
&= f_X(x|r, \lambda)
\end{aligned}$$

where we did a change of variables with $u = x/\nu - x/\lambda$ and where we needed $r < 1$ in order for the integral $\int_0^\infty u^{-r} e^{-u} \mathrm{d}u$ to converge.

Problem 4.38.b

Suppose $r < 1$. Then it's easy to see that $p_\lambda(\nu) > 0$ since $\lambda > \nu$. We also have

$$\begin{aligned}
\int_0^\lambda p_\lambda(\nu) \mathrm{d}\nu &= \int_0^\lambda \frac{1}{\Gamma(r)\Gamma(1-r)} \frac{\nu^{r-1}}{(\lambda - \nu)^r} \mathrm{d}\nu \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^\lambda \frac{\nu^{r-1}}{(\lambda - \nu)^r} \mathrm{d}\nu \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^1 \frac{(\lambda u)^{r-1}}{(\lambda - \lambda u)^r} \lambda \mathrm{d}u \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^1 \lambda^{-r+1+r-1} (1-u)^{-r} u^{r-1} \mathrm{d}u \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \int_0^1 (1-u)^{1-r-1} u^{r-1} \mathrm{d}u \\
&= \frac{1}{\Gamma(r)\Gamma(1-r)} \frac{\Gamma(r)\Gamma(1-r)}{\Gamma(r+1-r)} \\
&= 1
\end{aligned}$$

where we made the change of variable $u = \nu/\lambda$.

Problem 4.38.c

Suppose $f(x) = \int_0^\infty (e^{-x/\nu}/\nu) q_\lambda(\nu) \mathrm{d}\nu$ for some pdf $q_\lambda(\nu)$. Then observe that for all $x \in \mathbb{R}_{>0}$, we have

$$\begin{aligned}
\partial_x \log\left(\int_0^\infty (e^{-x/\nu}/\nu) q_\lambda(\nu) \mathrm{d}\nu\right) &= \frac{1}{f(x)} \int_0^\infty \partial_x (e^{-x/\nu}/\nu) q_\lambda(\nu) \mathrm{d}\nu \\
&= \frac{-1}{f(x)} \int_0^\infty (e^{-x/\nu}/\nu^2) q_\lambda(\nu) \mathrm{d}\nu \\
&< 0
\end{aligned}$$

where the last line follows from the fact that $f(x) > 0$ and $e^{-x/\nu}/\nu^2 > 0$ for all $x, \nu \in \mathbb{R}_{>0}$.

On the other hand, since $f(x)$ has a gamma distribution, we have $f'(x) > 0$ for x sufficiently small whenever $r > 1$. In particular, we have

$$\begin{aligned}
\partial_x \log(f(x)) &= \frac{1}{f(x)} \partial_x \left(\frac{1}{\Gamma(r)\lambda^r} x^{r-1} e^{-x/\lambda}\right) \\
&= \frac{1}{f(x)} \left(\frac{(r-1)}{\Gamma(r)\lambda^r} x^{r-2} e^{-x/\lambda} - \frac{1}{\Gamma(r)\lambda^{r+1}} x^{r-1} e^{-x/\lambda}\right) \\
&= \frac{\Gamma(r)\lambda^r}{x^{r-1} e^{-x/\lambda}} \left(\frac{(r-1)}{\Gamma(r)\lambda^r} x^{r-2} e^{-x/\lambda} - \frac{1}{\Gamma(r)\lambda^{r+1}} x^{r-1} e^{-x/\lambda}\right) \\
&= (r-1)x^{-1} - \lambda^{-1} \\
&= \frac{r-1}{x} - \frac{1}{\lambda} \\
&> 0 \qquad (\text{for } x < \lambda(r-1)).
\end{aligned}$$

Problem 4.40

Problem 4.40.a

We have

$$\begin{aligned}
1 &= \int_0^1 \int_0^{1-x} C x^{a-1} y^{b-1} (1-x-y)^{c-1} \mathrm{d}y \mathrm{d}x \\
&= C \int_0^1 x^{a-1} \int_0^{1-x} y^{b-1} (1-x-y)^{c-1} \mathrm{d}y \mathrm{d}x \\
&= C \int_0^1 x^{a-1} (1-x)^b \int_0^1 u^{b-1} (1-u)^{c-1} (1-x)^{c-1} \mathrm{d}u \mathrm{d}x \\
&= C \int_0^1 x^{a-1} (1-x)^{b+c-1} \int_0^1 u^{b-1} (1-u)^{c-1} \mathrm{d}u \mathrm{d}x \\
&= C \frac{\Gamma(b)\Gamma(c)}{\Gamma(b+c)} \int_0^1 x^{a-1} (1-x)^{b+c-1} \mathrm{d}x \\
&= C \frac{\Gamma(b)\Gamma(c)}{\Gamma(b+c)} \frac{\Gamma(a)\Gamma(b+c)}{\Gamma(a+b+c)} \\
&= C \frac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b+c)}.
\end{aligned}$$

It follows that $C = \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}$.

Problem 4.40.b

Let $x \in [0,1]$. To find $f_X(x)$, we integrate $f(x,y)$ as y ranges from 0 to $1-x$ since $f(x,y)$ is zero outside $[0,1-x]$. Thus we have

$$\begin{aligned}
f_X(x) &= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} x^{a-1} \int_0^{1-x} y^{b-1}(1-x-y)^{c-1} \mathrm{d}y \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} x^{a-1}(1-x)^{b+c-1} \int_0^1 u^{b-1}(1-u)^{c-1} \mathrm{d}u \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \frac{\Gamma(b)\Gamma(c)}{\Gamma(b+c)} x^{a-1}(1-x)^{b+c-1} \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b+c)} x^{a-1}(1-x)^{b+c-1}.
\end{aligned}$$

If $x \notin [0,1]$, then clearly $f_X(x) = 0$ since $f(x,y) = 0$ for all $y \in \mathbb{R}$. Thus $f_X(x) \sim \text{beta}(a, b+c)$.

Similarly, Let $y \in [0,1]$. To find $f_Y(y)$, we integrate $f(x,y)$ as x ranges from 0 to $1-y$ since $f(x,y)$ is zero outside $[0,1-y]$. Thus we have

$$\begin{aligned}
f_Y(y) &= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} y^{b-1} \int_0^{1-y} x^{a-1}(1-x-y)^{c-1} \mathrm{d}x \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} y^{b-1}(1-y)^{a+c-1} \int_0^1 u^{a-1}(1-u)^{c-1} \mathrm{d}u \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} \frac{\Gamma(a)\Gamma(c)}{\Gamma(a+c)} y^{b-1}(1-y)^{a+c-1} \\
&= \frac{\Gamma(a+b+c)}{\Gamma(b)\Gamma(a+c)} y^{b-1}(1-y)^{a+c-1}.
\end{aligned}$$

If $y \notin [0,1]$, then clearly $f_Y(y) = 0$ since $f(x,y) = 0$ for all $x \in \mathbb{R}$. Thus $f_Y(y) \sim \text{beta}(b, a+c)$.

Problem 4.40.c

Fix $x \in [0,1]$ and let $y \in \mathbb{R}$. We want to calculate the conditional distribution of $Y|X = x$. If $y > 1-x$, then we have

$$\begin{aligned}
f(y|x) &= \frac{f(x,y)}{f_Y(y)} \\
&= \frac{0}{f_Y(y)} \\
&= 0.
\end{aligned}$$

Similarly, if $y \notin [0,1]$, then $f(y|x) = 0$. Thus assume $y \in [0,1-x]$. Then we have

$$\begin{aligned}
f(y|x) &= \frac{f(x,y)}{f_Y(y)} \\
&= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)} x^{a-1} y^{b-1} (1-x-y)^{c-1} \frac{\Gamma(b)\Gamma(a+c)}{\Gamma(a+b+c)} y^{1-b}(1-y)^{1-a-c} \\
&= \frac{\Gamma(a+c)}{\Gamma(a)\Gamma(c)} x^{a-1}(1-x-y)^{c-1}(1-y)^{1-a-c}.
\end{aligned}$$

Now we want to show that $U = Y/(1-X)$ is beta(a,c). To this end, let $\mathcal{A} = \text{supp}(X,Y) = \{(x,y) \in (0,1)^2 \mid y < 1-x\}$, define $g = (g_1, g_2)\colon \mathcal{A} \to \mathbb{R}^2$ by

$$g_1(x,y) = \frac{y}{1-x} \qquad \text{and} \qquad g_2(x,y) = x$$

for all $(x,y) \in \mathcal{A}$, and denote $\mathcal{B} = \operatorname{im} g = [0,1]^2$, denote $U = g_1(X,Y)$, and denote $V = g_2(X,Y)$. Then observe that g is invertible with the inverse $h = (h_1, h_2)\colon \mathcal{B} \to \mathcal{A}$ given by

$$h_1(u,v) = v \qquad \text{and} \qquad h_2(u,v) = u(1-v)$$

Now the absolute value of the Jacobian of h at $(u, v) \in \mathcal{B}$ is given by

$$\begin{aligned} |J_{u,v}(h)| &= \left|\det \begin{pmatrix} 0 & 1 \\ 1-v & -u \end{pmatrix}\right| \\ &= 1 - v. \end{aligned}$$

It follows that

$$\begin{aligned} f_{U,V}(u, v) &= f_{X,Y}(h(u, v))\, |J_{u,v}(h)| \\ &= Cv^{a-1}u^{b-1}(1-v)^{b-1}(1 - v - u + uv)^{c-1}(1-v) \\ &= Cv^{a-1}u^{b-1}(1-v)^{b-1}(1-v)^{c-1}(1-u)^{c-1}(1-v) \\ &= Cv^{a-1}(1-v)^{b+c-1}u^{b-1}(1-u)^{c-1}. \end{aligned}$$

In particular, this implies

$$\begin{aligned} f_U(u) &= \int_0^1 Cv^{a-1}(1-v)^{b+c-1}u^{b-1}(1-u)^{c-1}\mathrm{d}v \\ &= Cu^{b-1}(1-u)^{c-1}\int_0^1 v^{a-1}(1-v)^{b+c-1}\mathrm{d}v \\ &= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}\frac{\Gamma(a)\Gamma(b+c)}{\Gamma(a+b+c)}u^{b-1}(1-u)^{c-1} \\ &= \frac{\Gamma(b+c)}{\Gamma(b)\Gamma(c)}u^{b-1}(1-u)^{c-1}. \end{aligned}$$

It follows that $U \sim \text{beta}(b, c)$.

Problem 4.40.d

First we find $\mathrm{E}(XY)$. We have

$$\begin{aligned} \mathrm{E}(XY) &= \int_0^1\int_0^{1-x} xy \cdot Cx^{a-1}y^{b-1}(1-x-y)^{c-1}\mathrm{d}y\mathrm{d}x \\ &= C\int_0^1 x^a\int_0^{1-x} y^b(1-x-y)^{c-1}\mathrm{d}y\mathrm{d}x \\ &= C\int_0^1 x^a\int_0^1 (1-x)^b u^b(1-x)^{c-1}(1-u)^{c-1}(1-x)\mathrm{d}u\mathrm{d}x \\ &= C\int_0^1 x^a(1-x)^{b+c}\int_0^1 u^b(1-u)^{c-1}\mathrm{d}u\mathrm{d}x \\ &= \frac{\Gamma(a+b+c)}{\Gamma(a)\Gamma(b)\Gamma(c)}\frac{\Gamma(b+1)\Gamma(c)}{\Gamma(b+c+1)}\frac{\Gamma(a+1)\Gamma(b+c+1)}{\Gamma(a+b+c+2)} \\ &= \frac{ab}{(a+b+c+1)(a+b+c)}. \end{aligned}$$

Now we find $\mathrm{Cov}(X, Y)$. This is given by

$$\begin{aligned} \mathrm{Cov}(X, Y) &= \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) \\ &= \frac{ab}{(a+b+c+1)(a+b+c)} - \left(\frac{a}{a+b+c}\right)\left(\frac{b}{a+b+c}\right) \\ &= \frac{ab}{(a+b+c+1)(a+b+c)} - \frac{ab}{(a+b+c)(a+b+c)} \\ &= \frac{ab(a+b+c) - (a+b+c+1)ab}{(a+b+c+1)(a+b+c)(a+b+c)} \\ &= \frac{-ab}{(a+b+c+1)(a+b+c)(a+b+c)}. \end{aligned}$$

Problem 4.45

In the table below, we provide a summary of what a bivariate normal pdf is

notation	$(X,Y) \sim \mathrm{bn}(\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2,\rho)$
pdf	$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)}$ for all $(x,y)\in\mathbb{R}^2$
parameters	$\mu_X,\mu_Y\in\mathbb{R}$ (means), $\sigma_X,\sigma_Y\in\mathbb{R}_{>0}$ (square root of variances), and $\rho\in(0,1)$ (correlation)

Problem 4.45.a

We just need to calculate the marginal distribution of X. This is because the joint distrubtion $f(x,y)$ is invariant when we swap x with y and X with Y. We have

$$
\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)}\mathrm{d}y \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)u+u^2\right)}\sigma_Y\mathrm{d}u && u\text{-substitution } u = \frac{y-\mu_Y}{\sigma_Y} \\
&= \frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(a^2-2\rho a u+u^2\right)}\mathrm{d}u && \text{denote } a = \frac{x-\mu_X}{\sigma_X} \\
&= \frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left((u-\rho a)^2+a^2(1-\rho^2)\right)}\mathrm{d}u && \text{completing the square} \\
&= \frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(v^2+a^2(1-\rho^2)\right)}\mathrm{d}v && v\text{-substitution } v = u-\rho a \\
&= \frac{e^{-\frac{1}{2}a^2}}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{v}{\sqrt{1-\rho^2}}\right)^2}\mathrm{d}v \\
&= \frac{e^{-\frac{1}{2}a^2}}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}w^2}\sqrt{1-\rho^2}\mathrm{d}w && w\text{-substitution } w = \frac{v}{\sqrt{1-\rho^2}} \\
&= \frac{e^{-\frac{1}{2}a^2}}{2\pi\sigma_X} \int_{-\infty}^{\infty} e^{-\frac{1}{2}w^2}\mathrm{d}w \\
&= \frac{e^{-\frac{1}{2}a^2}}{2\pi\sigma_X}\sqrt{2\pi} \\
&= \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} && \text{recalling that } a = \frac{x-\mu_X}{\sigma_X}
\end{aligned}
$$

It follows that $f_X \sim \mathrm{n}(\mu_X,\sigma_X^2)$, and by symmetry of the joint distribution $f(x,y)$ when swapping x with y and X with Y, we also have $f_Y \sim \mathrm{n}(\mu_Y,\sigma_Y^2)$.

Problem 4.45.b

We have

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

It follows that $f_{Y|X=x} \sim \mathrm{n}(\mu_Y + \rho(\sigma_Y/\sigma_X)(x-\mu_X), \sigma_Y^2(1-\rho^2))$.

Problem 4.45.c

We may assume that $a \neq 0$. Let $\mathcal{A} = \operatorname{supp}(X,Y) = \mathbb{R}^2$, define $g = (g_1, g_2)\colon \mathcal{A} \to \mathbb{R}^2$ by

$$
g_1(x,y) = ax + by \qquad \text{and} \qquad g_2(x,y) = y/a
$$

for all $(x,y) \in \mathcal{A}$. Denote $\mathcal{B} = \operatorname{im} g = \mathbb{R}^2$ (here is where we are using the assumption that $a \neq 0$) and denote $U = g_1(X,Y)$ and $V = g_2(X,Y)$. Then g is an intervible linear map, in particular its matrix representation with respect to the standard basis $\mathbf{e}_1 = (1,0)^\top$ and $\mathbf{e}_2 = (0,1)^\top$ is just $[g] = \left(\begin{smallmatrix} a & 0 \\ b & 1/a \end{smallmatrix}\right)$. Its inverse map $h = (h_1, h_2)\colon \mathcal{B} \to \mathcal{A}$ is given by

$$
h_1(u,v) = u/a - bv \qquad \text{and} \qquad h_2(u,v) = av
$$

for all $(u,v) \in \mathcal{B}$. The absolute value of the Jacobian of h at $(u,v) \in \mathcal{B}$ is equal to 1. It follows that

Therefore

$$
\begin{aligned}
f_U(u) &= \int_{-\infty}^{\infty} f_{U,V}(u,v)\mathrm{d}v \\
&= \int_{-\infty}^{\infty} f_{X,Y}\left(u/a - bv, av\right)\mathrm{d}v \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{u/a-bv-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{u/a-bv-\mu_X}{\sigma_X}\right)\left(\frac{av-\mu_Y}{\sigma_Y}\right)+\left(\frac{av-\mu_Y}{\sigma_Y}\right)^2\right)}\mathrm{d}v \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{u-abv-a\mu_X}{a\sigma_X}\right)^2 - 2\rho\left(\frac{u-abv-a\mu_X}{a\sigma_X}\right)\left(\frac{av-\mu_Y}{\sigma_Y}\right)+\left(\frac{av-\mu_Y}{\sigma_Y}\right)^2\right)}\mathrm{d}v \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{u}{a\sigma_X}-\frac{\mu_X}{\sigma_X}-\frac{bv}{\sigma_X}\right)^2 - 2\rho\left(\frac{u}{a\sigma_X}-\frac{\mu_X}{\sigma_X}-\frac{bv}{\sigma_X}\right)\left(\frac{av}{\sigma_Y}-\frac{\mu_Y}{\sigma_Y}\right)+\left(\frac{av}{\sigma_Y}-\frac{\mu_Y}{\sigma_Y}\right)^2\right)}\mathrm{d}v \\
&= \frac{1}{2\pi\sigma_U\sigma_V\sqrt{1-\rho_{U,V}^2}} e^{-\frac{1}{2(1-\rho_{U,V}^2)}\left(\left(\frac{u-\mu_U}{\sigma_U}\right)^2 - 2\rho_{U,V}\left(\frac{u-\mu_U}{\sigma_U}\right)\left(\frac{v-\mu_V}{\sigma_V}\right)+\left(\frac{v-\mu_V}{\sigma_V}\right)^2\right)}
\end{aligned}
$$

where $\mu_U = a\mu_X + b\mu_Y$ and $\sigma_U^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho\sigma_X\sigma_Y$.

Problem 4.46

Problem 4.46.a

We have

$$
\begin{aligned}
\mathrm{E}X &= \mathrm{E}(a_X Z_1 + b_X Z_2 + c_X) \\
&= a_X \mathrm{E}(Z_1) + b_X \mathrm{E}(Z_2) + c_X \\
&= a_X \cdot 0 + b_X \cdot 0 + c_X \\
&= c_X.
\end{aligned}
$$

A similar computation shows $\mathrm{E}Y = c_Y$. Next, we have

$$
\begin{aligned}
\operatorname{Var} X &= \mathrm{E}(X^2) - (\mathrm{E}X)^2 \\
&= \mathrm{E}(a_X^2 Z_1^2 + b_X^2 Z_2^2 + c_X^2 + 2a_X b_X Z_1 Z_2 + 2a_X c_X Z_1 + 2b_X c_X Z_2) - c_X^2 \\
&= a_X^2 \mathrm{E}(Z_1^2) + b_X^2 \mathrm{E}(Z_2^2) + c_X^2 + 2a_X b_X \mathrm{E}(Z_1)\mathrm{E}(Z_2) + 2a_X c_X \mathrm{E}(Z_1) + 2b_X c_X \mathrm{E}(Z_1) - c_X^2 \\
&= a_X^2 \mathrm{E}(Z_1^2) + b_X^2 \mathrm{E}(Z_2^2) \\
&= a_X^2 (\operatorname{Var} Z_1 + (\mathrm{E}Z_1)^2) + b_X^2 (\operatorname{Var} Z_1 + (\mathrm{E}Z_1)^2) \\
&= a_X^2 + b_X^2,
\end{aligned}
$$

where we used independence of Z_1 and Z_2 in the third step. A similar computation shows $\operatorname{Var} Y = a_Y^2 + b_Y^2$. Finally, we have

$$
\begin{aligned}
\operatorname{Cov}(X,Y) &= \mathrm{E}(XY) - \mu_X\mu_Y \\
&= \mathrm{E}((a_X Z_1 + b_X Z_2 + c_X)(a_Y Z_1 + b_Y Z_2 + c_Y)) - \mu_X\mu_Y \\
&= a_X a_Y \mathrm{E}(Z_1^2) + b_X b_Y \mathrm{E}(Z_2^2) + c_X c_Y - \mu_X\mu_Y \\
&= a_X a_Y + b_X b_Y.
\end{aligned}
$$

Problem 4.46.b

We have $\mathrm{E}X = c_X = \mu_X$ and $\mathrm{E}Y = c_Y = \mu_Y$ by the calculations above. We also have

$$\begin{aligned}\operatorname{Var} X &= a_X^2 + b_X^2 \\ &= \left(\sqrt{\frac{1+\rho}{2}}\sigma_X\right)^2 + \left(\sqrt{\frac{1-\rho}{2}}\sigma_X\right)^2 \\ &= \frac{(1+\rho)\sigma_X^2}{2} + \frac{(1-\rho)\sigma_X^2}{2} \\ &= \sigma_X^2.\end{aligned}$$

A similar computation shows $\operatorname{Var} Y = \sigma_Y^2$. Finally, we have

$$\begin{aligned}\rho_{XY} &= \frac{\operatorname{Cov}(X,Y)}{\sigma_X\sigma_Y} \\ &= \frac{a_X a_Y + b_X b_Y}{\sigma_X\sigma_Y} \\ &= \frac{\left(\sqrt{\frac{1+\rho}{2}}\sigma_X\right)\left(\sqrt{\frac{1+\rho}{2}}\sigma_Y\right) - \left(\sqrt{\frac{1-\rho}{2}}\sigma_X\right)\left(\sqrt{\frac{1-\rho}{2}}\sigma_Y\right)}{\sigma_X\sigma_Y} \\ &= \frac{1}{2}\frac{(1+\rho)\sigma_X\sigma_Y - (1-\rho)\sigma_X\sigma_Y}{\sigma_X\sigma_Y} \\ &= \frac{1}{2}\frac{(1+\rho)\sigma_X\sigma_Y - (1-\rho)\sigma_X\sigma_Y}{\sigma_X\sigma_Y} \\ &= \rho.\end{aligned}$$

Problem 4.46.c

I think we need to assume here that $\sigma_X, \sigma_Y > 0$ and $-1 < \rho < 1$. These are the parameter conditions used in the definition of **bivariate normal pdf** given in the book. You'll see why these conditions are needed in a moment.

Since Z_1 and Z_2 are independent $\mathrm{n}(0,1)$ distributions, their joint distribution is given by

$$\begin{aligned}f_{Z_1,Z_2}(z_1,z_2) &= f_{Z_1}(z_1)f_{Z_2}(z_2) \\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_1^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_1^2} \\ &= \frac{1}{2\pi}e^{-\frac{1}{2}(z_1^2+z_2^2)}\end{aligned}$$

for all $z_1, z_2 \in \mathbb{R}$. Now denote $\mathcal{A} = \operatorname{supp}(f_{Z_1,Z_2}) = \mathbb{R}^2$ and define $g = (g_1, g_2)\colon \mathcal{A} \to \mathbb{R}^2$ by

$$g_1(z_1,z_2) = a_X z_1 + b_X z_2 + c_X \qquad \text{and} \qquad g_1(z_1,z_2) = a_Y z_1 + b_Y z_2 + c_Y$$

for all $z_1, z_2 \in \mathbb{R}$ and denote $g_1(Z_1,Z_2) = X$, $g_2(Z_1,Z_2) = Y$, and $\mathcal{B} = \operatorname{im} g$. Note that $\mathcal{B} = \mathbb{R}^2$ since

$$\begin{aligned}\det\begin{pmatrix} a_X & a_Y \\ b_X & b_Y\end{pmatrix} &= a_X b_Y - a_Y b_X \\ &= -\left(\sqrt{\frac{1+\rho}{2}}\right)\sigma_X\left(\sqrt{\frac{1-\rho}{2}}\right)\sigma_Y - \left(\sqrt{\frac{1+\rho}{2}}\right)\sigma_Y\left(\sqrt{\frac{1-\rho}{2}}\right)\sigma_X \\ &= -\frac{\sqrt{1-\rho^2}}{2}\sigma_X\sigma_Y - \frac{\sqrt{1-\rho^2}}{2}\sigma_Y\sigma_X \\ &= -\sqrt{1-\rho^2}\sigma_X\sigma_Y \\ &\neq 0.\end{aligned}$$

The map g is invertible (it's just an affine transformation), with inverse $h = (h_1,h_2)\colon \mathcal{B} \to \mathcal{A}$ given by

$$h_1(x,y) = \frac{b_Y x - b_X y - (b_Y c_X - b_X c_Y)}{-\sqrt{1-\rho^2}\sigma_X\sigma_Y} \qquad \text{and} \qquad h_2(x,y) = \frac{-a_Y x + a_X y - (-a_Y c_X + a_X c_Y)}{-\sqrt{1-\rho^2}\sigma_X\sigma_Y}.$$

Now the absolute value of the Jacobian of h at $(x, y) \in \mathcal{B}$ is given by

$$
\begin{aligned}
|J_{x,y}(h)| &= \left|\det \begin{pmatrix} \partial_x h_1(x,y) & \partial_y h_1(x,y) \\ \partial_x h_1(x,y) & \partial_y h_2(x,y) \end{pmatrix}\right| \\
&= \frac{1}{(1-\rho^2)\sigma_X^2\sigma_Y^2} \left|\det \begin{pmatrix} -b_Y & b_X \\ a_Y & -a_X \end{pmatrix}\right| \\
&= \frac{1}{(1-\rho^2)\sigma_X^2\sigma_Y^2} |b_Y a_X - a_Y b_Y| \\
&= \frac{\sqrt{1-\rho^2}\sigma_X\sigma_Y}{(1-\rho^2)\sigma_X^2\sigma_Y^2} \\
&= \frac{1}{\sigma_X\sigma_Y\sqrt{1-\rho^2}}
\end{aligned}
$$

It follows that

$$
\begin{aligned}
f_{X,Y}(x,y) &= f_{Z_1,Z_2}(h(x,y))\,|J_{x,y}(h)| \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2}\left(\left(\frac{b_Y x - b_X y - (b_Y c_X - b_X c_Y)}{-\sqrt{1-\rho^2}\sigma_X\sigma_Y}\right)^2 + \left(\frac{-a_Y x + a_X y - (-a_Y c_X + a_X c_Y)}{-\sqrt{1-\rho^2}\sigma_X\sigma_Y}\right)^2\right)} \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2(1-\rho^2)}\left(\left(\frac{b_Y x - b_X y - (b_Y c_X - b_X c_Y)}{\sigma_X\sigma_Y}\right)^2 + \left(\frac{-a_Y x + a_X y - (-a_Y c_X + a_X c_Y)}{\sigma_X\sigma_Y}\right)^2\right)} \\
&= \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{\frac{-1}{2(1-\rho^2)}\left(\left(\frac{\mu_X - x}{\sigma_X}\right)^2 - 2\rho\left(\frac{\mu_X - x}{\sigma_X}\right)\left(\frac{\mu_Y - y}{\sigma_Y}\right) + \left(\frac{\mu_Y - y}{\sigma_Y}\right)^2\right)}
\end{aligned}
$$

where we obtained the last line from the third line after performing the calculations

$$
\begin{aligned}
\left(\frac{b_Y x - b_X y - (b_Y c_X - b_X c_Y)}{\sigma_X\sigma_Y}\right)^2 &= \left(\frac{-\sqrt{\frac{1-\rho}{2}}\sigma_Y x - \sqrt{\frac{1-\rho}{2}}\sigma_X y + \sqrt{\frac{1-\rho}{2}}\sigma_Y\mu_X + \sqrt{\frac{1-\rho}{2}}\sigma_X\mu_Y}{\sigma_X\sigma_Y}\right)^2 \\
&= \frac{1-\rho}{2}\left(\frac{-\sigma_Y x - \sigma_X y + \sigma_Y\mu_X + \sigma_X\mu_Y}{\sigma_X\sigma_Y}\right)^2 \\
&= \frac{1-\rho}{2}\left(\frac{(\mu_X - x)\sigma_Y + (\mu_Y - y)\sigma_X}{\sigma_X\sigma_Y}\right)^2 \\
&= \frac{1-\rho}{2}\left(\frac{\mu_X - x}{\sigma_X} + \frac{\mu_Y - y}{\sigma_Y}\right)^2 \\
&= \frac{1-\rho}{2}\left(\frac{\mu_X - x}{\sigma_X}\right)^2 + (1-\rho)\left(\frac{\mu_X - x}{\sigma_X}\right)\left(\frac{\mu_Y - y}{\sigma_Y}\right) + \frac{1-\rho}{2}\left(\frac{\mu_Y - y}{\sigma_Y}\right)^2
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\left(\frac{-a_Y x + a_X y - (-a_Y c_X + a_X c_Y)}{\sigma_X\sigma_Y}\right)^2 &= \left(\frac{-\sqrt{\frac{1+\rho}{2}}\sigma_Y x + \sqrt{\frac{1+\rho}{2}}\sigma_X y + \sqrt{\frac{1+\rho}{2}}\sigma_Y\mu_X - \sqrt{\frac{1+\rho}{2}}\sigma_X\mu_Y}{\sigma_X\sigma_Y}\right)^2 \\
&= \frac{1+\rho}{2}\left(\frac{-\sigma_Y x + \sigma_X y + \sigma_Y\mu_X - \sigma_X\mu_Y}{\sigma_X\sigma_Y}\right)^2 \\
&= \frac{1+\rho}{2}\left(\frac{\sigma_Y(\mu_X - x) - \sigma_X(\mu_Y - y)}{\sigma_X\sigma_Y}\right)^2 \\
&= \frac{1+\rho}{2}\left(\frac{\mu_X - x}{\sigma_X} - \frac{\mu_Y - y}{\sigma_Y}\right)^2 \\
&= \frac{1+\rho}{2}\left(\frac{\mu_X - x}{\sigma_X}\right)^2 - (1+\rho)\left(\frac{\mu_X - x}{\sigma_X}\right)\left(\frac{\mu_Y - y}{\sigma_Y}\right) + \frac{1+\rho}{2}\left(\frac{\mu_Y - y}{\sigma_Y}\right)^2.
\end{aligned}
$$

Thus $(X, Y) \sim \mathrm{bn}(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$.

Problem 4.46.d

This is equivalent to finding the solution set in $(x,y,z,w) \in \mathbb{R}$ to the system of equations

$$
\begin{aligned}
x^2 + y^2 &= a \\
z^2 + w^2 &= b \\
xz + yw &= c
\end{aligned}
$$

where $a, b \in \mathbb{R}_{>0}$ and $c \in (-\sqrt{ab}, \sqrt{ab})$ (here we are thinking of $a = \sigma_X^2$, $b = \sigma_Y^2$, $c = \sqrt{ab}\rho$, $x = a_X$, $y = b_X$, $z = a_Y$, and $w = b_Y$). For the first two equations to be satisfied, we need

$$
\begin{aligned}
x &= \sqrt{a}\cos\theta \\
y &= \sqrt{a}\sin\theta \\
z &= \sqrt{b}\cos\vartheta \\
w &= \sqrt{b}\sin\vartheta
\end{aligned}
$$

where $\theta, \vartheta \in [0, 2\pi)$ and where we may assume (without loss of generality) that . Note that the x and y coordinates are *uniquely* determined by the θ and ϑ coordinates. For the last equation to also be satisfied, we need

$$
\begin{aligned}
c &= (\sqrt{a}\cos\theta)(\sqrt{b}\cos\vartheta + (\sqrt{a}\sin\theta)(\sqrt{b}\sin\vartheta) \\
&= \sqrt{ab}\,(\cos\theta\cos\vartheta + \sin\theta\sin\vartheta) \\
&= \sqrt{ab}\cos(|\theta - \vartheta|) \\
&= \sqrt{ab}\cos\alpha && \text{where } \alpha = |\theta - \vartheta|
\end{aligned}
$$

Since $c/\sqrt{ab} \in (-1, 1)$ and $\alpha \in [0, 2\pi)$, there exists two solutions to this in α, namely

$$
\alpha = \arccos(c/\sqrt{ab}) \quad \text{or} \quad \alpha = \arccos(c/\sqrt{ab}) + \pi
$$

where we are using the convention that the domain of arccos is $[-1, 1]$ and the range of arccos is $[0, \pi]$. Finally, we note that there exists infinitely many $\theta, \vartheta \in [0, 2\pi)$ such that $\alpha = \theta - \vartheta$. Thus there exists infinitely many different distributions (parametrized by $\theta, \vartheta \in [0, 2\pi)$ satisfying $\theta - \vartheta = \pm\arccos(c/\sqrt{ab})$ or $\theta - \vartheta = \pm(\arccos(c/\sqrt{ab}) + \pi)$ which give rise to the same bivariate normal distribution. Putting everything back to the original notation, we find that these distributions can be described as

$$
\begin{aligned}
X - \mu_X &= \sigma_X\,(\cos\theta Z_1 + \sin\theta Y) \\
Y - \mu_Y &= \sigma_Y\,(\cos\vartheta Z_1 + \sin\vartheta Y)
\end{aligned}
$$

Problem 4.47

Problem 4.47.a

Since X and Y are independent $\mathrm{n}(0,1)$ distributions, their joint distribution is given by

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

for all $x, y \in \mathbb{R}$. Denote $\mathcal{A} = \operatorname{supp}(X,Y) = \mathbb{R}^2$ and define $g = (g_1, g_2) \colon \mathcal{A} \to \mathbb{R}^2$ by

$$
g_1(x,y) = \begin{cases} x & \text{if } xy > 0 \\ -x & \text{if } xy < 0 \\ 1 & \text{if } xy = 0 \end{cases} \qquad \text{and} \qquad g_2(x,y) = \begin{cases} -y & \text{if } xy > 0 \\ y & \text{if } xy < 0 \\ 1 & \text{if } xy = 0 \end{cases}
$$

for all $x, y \in \mathbb{R}$. Denote $\mathcal{B} = \operatorname{im} g = \mathbb{R}^2 \backslash \{(x,y) \mid xy = 0\})$ and denote $Z = g_1(X,Y)$ and $W = g_2(X,Y)$. Define the sets

$$
\begin{aligned}
A_0 &= \{(x,y) \in \mathbb{R}^2 \mid xy = 0\} \\
A_1 &= \mathbb{R}^2 \backslash A_0.
\end{aligned}
$$

Then note that $P((X,Y) \in A_0) = 0$ and g restricts to an invertible map on A_1 whose image is all of $\mathcal{B}$ and such that its inverse $h = (h_1, h_2)\colon \mathcal{B} \to \mathcal{A}$ is simply $h = -g|_{A_1}$. The absolute value of the Jacobian of h at $(z,w) \in \mathcal{B}$ is always equal to 1. Therefore

$$
\begin{aligned}
f_{Z,W}(z,w) &= f_{X,Y}(h(z,w)) \\
&= \frac{1}{2\pi} e^{-\frac{1}{2}\left(z^2+w^2\right)} \\
&= \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2}\right).
\end{aligned}
$$

where we do not need to consider cases in the second line since squaring always kills the negative sign. The factorization of $f_{Z,W}$ implies

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}.$$

Thus $Z \sim \mathrm{n}(0,1)$.

Problem 4.47.b

Assume for a contradiction that the joint distribution of Z and Y is bivariate normal. Then

$$P((Z,Y) \in \{(z,y) \in \mathbb{R}^2 \mid z > 0 \text{ and } y < 0\} \neq 0.$$

However Z and Y *always* has the same sign. Indeed, if $Y > 0$, then $Z > 0$ regardless of whether $X > 0$ or $X < 0$. Similarly, if $Y < 0$, then $Z < 0$.

Problem 4.48

Problem 4.48.a

First we calculate

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(Ax^2y^2+x^2+y^2-2Bxy-2Cx-2Dy\right)} \mathrm{d}x \\
&= e^{-\frac{1}{2}\left(y^2-2Dy\right)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(Ax^2y^2+x^2-2Bxy-2Cx\right)} \mathrm{d}x \\
&= e^{-\frac{1}{2}\left(y^2-2Dy\right)} \int_{-\infty}^{\infty} e^{-\frac{(Ay^2+1)}{2}\left(x^2-\left(\frac{2By+2C}{(Ay^2+1)}\right)x\right)} \mathrm{d}x \\
&= e^{-\frac{1}{2}\left(y^2-2Dy\right)} \int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left(x^2-2\mu x\right)} \mathrm{d}x && \text{denoting } \mu = \frac{By+C}{(Ay^2+1)} \text{ and } \sigma^2 = \frac{1}{Ay^2+1} \\
&= e^{-\frac{1}{2}\left(y^2-2Dy\right)} \int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left((x-\mu)^2+\mu^2\right)} \mathrm{d}x && \text{complete the square} \\
&= e^{-\frac{1}{2}\left(y^2-2Dy\right)} e^{-\frac{1}{2}(\mu/\sigma)^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x \\
&= \sigma\sqrt{2\pi} e^{-\frac{1}{2}\left(y^2-2Dy\right)} e^{-\frac{1}{2}(\mu/\sigma)^2}
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
f_{X|Y=y}(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(x)} \\
&\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(Ax^2y^2+x^2+y^2-2Bxy-2Cx-2Dy\right)} e^{\frac{1}{2}\left(y^2-2Dy\right)} e^{\frac{1}{2}(\mu/\sigma)^2} \\
&\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} e^{-\frac{1}{2}\left(y^2-2Dy\right)} e^{-\frac{1}{2}(\mu/\sigma)^2} e^{\frac{1}{2}\left(y^2-2Dy\right)} e^{\frac{1}{2}(\mu/\sigma)^2} \\
&= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.
\end{aligned}
$$

Similarly, we have

$$f_{Y|X=x}(y) = \frac{1}{\tau\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\lambda}{\tau}\right)^2}$$

where $\lambda = \frac{Bx+D}{(Ax^2+1)}$ and $\tau^2 = \frac{1}{Ax^2+1}$.

Problem 4.48.b

If $A = 1$, $B = 0$, and $C = D = 8$, then we have

$$f_{X,Y}(x,y) = e^{-\frac{1}{2}\left(x^2y^2+x^2+y^2-16x-16y\right)}$$

Note that a relative maximum of $f_{X,Y}$ occurs precisely when a relative minimum of $g(x,y) = x^2y^2 + x^2 + y^2 - 16x - 16y$ occurs. In fact, g is a polynomial, thus its relative minimum will be global minimum. Using some simple calculus (or using a software like wolfram alpha) we find that the global minimum of g occurs as

$$(x,y) = \left(4-\sqrt{15}, 4+\sqrt{15}\right) \qquad \text{and} \qquad (x,y) = \left(4+\sqrt{15}, \frac{1}{4+\sqrt{15}}\right)$$

Problem 4.52

Let $B_1 = (X_1, Y_1)$ be the random variable corresponding to the first bullet and let $B_2 = (X_2, Y_2)$ be the random variable corresponding to the second bullet. Set $X = X_1 - X_2$ and $Y = Y_1 - Y_2$. Then since all random variables involed are mutually independent, we have $X \sim \mathrm{n}(0,1) \sim Y$. Finally, set $R = \sqrt{X^2+Y^2}$. From problem 4.21, we know that $R^2 \sim \chi^2(2)$. In other words, the pdf of R is given by

$$f_R(r) = \begin{cases} re^{-r^2/2} & \text{if } 0 < r < \infty \\ 0 & \text{else} \end{cases}$$

This is precisely the distrubtion of the distance between the two points that we are looking for.

Problem 4.55

Let $X, Y, Z \sim \text{exponential}(\lambda)$ and let $W = \max(X,Y,Z)$. We want to find the distribution of W given that X, Y, and Z are mutually independent. In particular, the joint pdf of (X,Y,Z) is given by

$$\begin{aligned} f_{X,Y,Z}(x,y,z) &= f_X(x)f_Y(y)f_Z(z) \\ &= \lambda e^{-\lambda x}\lambda e^{-\lambda y}\lambda e^{-\lambda z} \\ &= \lambda^3 e^{-\lambda(x+y+z)} \end{aligned}$$

for all $(x,y,z) \in \mathbb{R}^3_{\geq 0}$ and is equal to zero everywhere else. Now note that the supp $W = \mathbb{R}_{\geq 0}$, so letting $w \in \mathbb{R}_{\geq 0}$, we have

$$\begin{aligned} F_W(w) &= \mathrm{P}(\max(X,Y,Z) \leq w) \\ &= \int_0^w \int_0^w \int_0^w f_{X,Y,Z}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &= \int_0^w \int_0^w \int_0^w \lambda^3 e^{-\lambda(x+y+z)}\mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &= \int_0^w \lambda e^{-\lambda x}\mathrm{d}x \int_0^w \lambda e^{-\lambda y}\mathrm{d}y \int_0^w \lambda e^{-\lambda z}\mathrm{d}z \\ &= (1-\lambda e^{-\lambda w})^3. \end{aligned}$$

Problem 4.56

Problem 4.56.a

The probability that the test for a pooled sample of k people will be positive is precisely the probability that at least one of those k people tests positive. In other words, it is the probability that all k people testing negative does not happen. This probability is given by

$$1 - q^k$$

where $q = 1 - p$.

Problem 4.56.b

We first find the pmf of X. At least m blood tests are necessary: even if all pooled blood samples test negative, we still need m tests to cover all $N = mk$ people. On the other hand, if each pooled blood sample tests positive, then we would need to perform a total of $m(k+1) = Nk + m$ tests. In general, let $0 \leq i \leq m$. Suppose that exactly i pooled blood samples test positive with the remaining pooled samples testing negative. Then the number of tests that need to be performed is

$$i(k+1) + m - i = ik + m.$$

Thus the support of X is

$$\operatorname{supp} X = \{ik + m \mid 1 \leq i \leq m\}.$$

Now the probability that exactly i pooled samples test positive while the remaining pooled samples testing negative is

$$\binom{m}{i} \left(1 - q^k\right)^i q^{k(m-i)}$$

This also gives us the probability that $ik + m$ tests are performed. Therefore the expected value of X is

$$\begin{aligned}
\mathrm{E}X &= \sum_{i=0}^{m} (ik+m) \mathrm{P}(ik + m \text{ tests are performed}) \\
&= \sum_{i=0}^{m} (ik+m) \binom{m}{i} \left(1 - q^k\right)^i q^{k(m-i)} \\
&= k \sum_{i=0}^{m} i \binom{m}{i} \left(1 - q^k\right)^i \left(q^k\right)^{m-i} + m \sum_{i=0}^{m} \binom{m}{i} \left(1 - q^k\right)^i \left(q^k\right)^{m-i} \\
&= km(1 - q^k) + m(1 - q^k + q^k)^m \\
&= km(1 - q^k) + m \\
&= (k(1 - q^k) + 1)m.
\end{aligned}$$

Problem 4.56.c

Observe that

$$\begin{aligned}
\lim_{p \to 0} \mathrm{E}X &= \lim_{p \to 0} (k(1 - q^k) + 1)m \\
&= \lim_{q \to 1} (k(1 - q^k) + 1)m \\
&= m.
\end{aligned}$$

Thus plan (ii) would be preferred if p is closed to zero, since this would imply the expected value is close to the minimal amount of tests needed!

Special Note

Now we fix m and we wish to derive an expression that identifies the optimal pool size k^* that minimizes the expected number of tests. In other words, we wish to calculate

$$\min\{(k(1-q^k)+1)m \mid 1 \le k \le N/m\} = m \min_{1\le k\le N/m}\{k(1-q^k)\} + m$$

Observe that

$$\begin{aligned} \partial_k(k(1-q^k) &= (1-q^k) + (1-q^k)k\log(1-q^k) \\ &= (1+k\log(1-q^k))(1-q^k). \end{aligned}$$

This is equal to zero if and only if $q = 1$ or

$$\begin{aligned} 0 = 1 + k\log(1-q^k) &\iff -1/k = \log(1-q^k) \\ &\iff e^{-1/k} = 1-q^k \\ &\iff 1-e^{-1/k} = q^k \\ &\iff \left(1-e^{-1/k}\right)^{1/k} = q. \end{aligned}$$

Problem 4.58

For the following problems, note that

$$\begin{aligned} \mathrm{E}[X\mathrm{E}[Y|X]] &= \int_{\mathcal{X}} x\mathrm{E}[Y|x]f_X(x)\mathrm{d}x \\ &= \int_{\mathcal{X}} x\left(\int_{-\infty}^{\infty} y f_{Y|x}(y)\mathrm{d}y\right) f_X(x)\mathrm{d}x \\ &= \int_{\mathcal{X}} x\left(\int_{-\infty}^{\infty} y \frac{f_{X,Y}(x,y)}{f_X(x)}\mathrm{d}y\right) f_X(x)\mathrm{d}x \\ &= \int_{\mathcal{X}} x\left(\int_{-\infty}^{\infty} y f_{X,Y}(x,y)\mathrm{d}y\right) \frac{1}{f_X(x)} f_X(x)\mathrm{d}x \\ &= \int_{\mathcal{X}} x \int_{-\infty}^{\infty} y f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x \\ &= \mathrm{E}[XY]. \end{aligned}$$

We shall use this identity in what follows.

Problem 4.58.a

We have

$$\begin{aligned} \mathrm{Cov}(X, \mathrm{E}[Y|X]) &= \mathrm{E}[X\mathrm{E}[Y|X]] - \mathrm{E}[X]\mathrm{E}[\mathrm{E}[Y|X]] \\ &= \mathrm{E}[X\mathrm{E}[Y|X]] - \mathrm{E}[X]\mathrm{E}[Y] \\ &= \mathrm{E}[X]\mathrm{E}[\mathrm{E}[Y|X]] - \mathrm{E}[X]\mathrm{E}[Y] \\ &= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\ &= \mathrm{Cov}(X,Y). \end{aligned}$$

Problem 4.58.b

We have

$$
\begin{aligned}
\mathrm{Cov}(X, Y - \mathrm{E}[Y|X]) &= \mathrm{E}[X(Y - \mathrm{E}[Y|X]] - \mathrm{E}[X]\mathrm{E}[Y - \mathrm{E}[Y|X]] \\
&= \mathrm{E}[XY - X\mathrm{E}[Y|X]] - \mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[X]\mathrm{E}[\mathrm{E}[Y|X]] \\
&= \mathrm{E}[XY] - \mathrm{E}[X\mathrm{E}[Y|X]]] - \mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[XY] - \mathrm{E}[X\mathrm{E}[Y|X]]] \\
&= 0.
\end{aligned}
$$

Problem 4.58.c

We have

$$
\begin{aligned}
\mathrm{Var}[Y - \mathrm{E}[Y|X]] &= \mathrm{E}[(Y - \mathrm{E}[Y|X])^2] - (\mathrm{E}[Y - \mathrm{E}[Y|X]])^2 \\
&= \mathrm{E}[Y^2] - 2\mathrm{E}[Y\mathrm{E}[Y|X]] + \mathrm{E}[\mathrm{E}[Y|X]^2] - (\mathrm{E}[Y] - \mathrm{E}[\mathrm{E}(Y|X)]])^2 \\
&= \mathrm{E}[Y^2] - 2\mathrm{E}[Y]^2 + \mathrm{E}[\mathrm{E}[Y|X]^2] - (\mathrm{E}[Y] - \mathrm{E}[\mathrm{E}(Y|X)]])^2 \\
&= \mathrm{E}[Y^2] - 2\mathrm{E}[Y]^2 + \mathrm{Var}[\mathrm{E}[Y|X]] - \mathrm{E}[Y]^2 + 2\mathrm{E}[Y]\mathrm{E}[\mathrm{E}[Y|X]] \\
&= \mathrm{E}[Y^2] - 2\mathrm{E}[Y]^2 + \mathrm{Var}[\mathrm{E}[Y|X]] - \mathrm{E}[Y]^2 + 2\mathrm{E}[Y]^2 \\
&= \mathrm{Var}(Y) - \mathrm{Var}[\mathrm{E}[Y|X]].
\end{aligned}
$$

Problem 4.59

We have

$$
\begin{aligned}
\mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{Cov}\,(\mathrm{E}[X|Z], \mathrm{E}[Y|Z]) &= \mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{E}[\mathrm{E}[X|Z]\mathrm{E}[Y|Z]] - \mathrm{E}[\mathrm{E}[X|Z]]\mathrm{E}[\mathrm{E}[Y|Z]] \\
&= \mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{E}[\mathrm{E}[X|Z]\mathrm{E}[Y|Z]] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{E}[\mathrm{E}[X|Z]Y] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{Cov}(X, Y|Z)] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{E}[X(Y|Z)] - \mathrm{E}[X]\mathrm{E}[Y|Z]] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{E}[X(Y|Z)]] - \mathrm{E}[\mathrm{E}[X]\mathrm{E}[Y|Z]] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{E}[X(Y|Z)]] - \mathrm{E}[X]\mathrm{E}[\mathrm{E}[Y|Z]] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{E}[X(Y|Z)]] - \mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[\mathrm{E}[(XY)|Z]] - \mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\
&= 0.
\end{aligned}
$$

Problem 4.60

Assume for simplicity that $X \sim \mathrm{n}(0,1)$ and $Y \sim \mathrm{n}(0,1)$. First we consider $Z_1 = Y - X$. Then we have $Z_1 \sim \mathrm{n}(0,2)$. A quick calculation shows that the joint distribution of Z_1 and Y is given by

$$
f_{Y,Z_1}(y,z) = \frac{1}{2\pi} e^{-\frac{1}{2}((y-z)^2 + y^2)}
$$

for all $(y,z) \in \mathbb{R}^2$. In particular, we obtain

$$
\begin{aligned}
f_{Y|Z_1}(y|z) &= \frac{1}{2\pi} e^{-\frac{1}{2}((y-z)^2 + y^2)} \cdot 2\pi e^{\frac{1}{4}z^2} \\
&= e^{-\frac{1}{2}((y-z)^2 + y^2) + \frac{1}{4}z^2} \\
&= e^{-y^2 + yz - \frac{1}{4}z^2}
\end{aligned}
$$

Thus

$$\begin{aligned}
\mathrm{E}[1_{Y\leq y}|Z_1=0] &= \int_{-\infty}^{\infty} 1_{Y\leq y} f_{Y|Z_1}(x|0)\mathrm{d}x \\
&= \int_{-\infty}^{y} e^{-x^2}\mathrm{d}x.
\end{aligned}$$

Now we consider $Z_2 = Y/X$. Then $Z_2 \sim \mathrm{Cauchy}(0,1)$ as shown in the book. Let us work out in detail what the joint distribution of Y and Z_2 looks like. Set $\mathcal{A} = \mathbb{R}^2 \backslash \{xy=0\}$ and define $g = (g_1, g_2) \colon \mathcal{A} \to \mathbb{R}^2$ by

$$g_1(x,y) = y \qquad \text{and} \qquad g_2(x,y) = y/x$$

for all $(x,y) \in \mathcal{A}$. Denote $\mathcal{B} = \operatorname{im} g$, $U = g_1(X,Y)$, and $V = g_2(X,Y)$ and note that g is one-one and onto with inverse $h = (h_1, h_2) \colon \mathcal{B} \to \mathcal{A}$ given by

$$h_1(u,v) = u/v \qquad \text{and} \qquad h_2(u,v) = u$$

for all $(u,v) \in \mathcal{B}$. The absolute value of the Jacobian of h at $(u,v) \in \mathcal{B}$ is given by

$$\begin{aligned}
|J_{u,v}(h)| &= \left| \det \begin{pmatrix} 1/v & -u/v^2 \\ 1 & 0 \end{pmatrix} \right| \\
&= \left| u/v^2 \right| \\
&= |u|/v^2.
\end{aligned}$$

Since $\mathrm{P}(\{xy=0\}) = 0$, it follows that

$$\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}(h(u,v))\,|J_{u,v}(h)| \\
&= \frac{|u|}{2\pi v^2} e^{-\frac{1}{2}\left(\left(\frac{u}{v}\right)^2 + u^2\right)}
\end{aligned}$$

for all $(u,v) \in \mathcal{B}$. In other words, the joint distribution of Y and Z_2 is given by

$$f_{Y,Z_2}(y,z) = \frac{|y|}{2\pi z^2} e^{-\frac{1}{2}\left(\left(\frac{y}{z}\right)^2 + y^2\right)}$$

for all $\{(y,z) \in \mathbb{R}^2 \mid yz \neq 0\}$. It follows that

$$\begin{aligned}
f_{Y|Z}(y|z) &= \frac{f_{Z_2,Y}(z,y)}{f_{Z_2}(z)} \\
&= \frac{|y|}{2\pi z^2} e^{-\frac{1}{2}\left(\left(\frac{y}{z}\right)^2 + y^2\right)} \pi(z^2+1) \\
&= \frac{|y|(z^2+1)}{2z^2} e^{-\frac{1}{2}\left(\left(\frac{y}{z}\right)^2 + y^2\right)}.
\end{aligned}$$

So if $y > 0$, then we have

$$\begin{aligned}
\mathrm{E}[1_{Y\leq y}|Z_2=1] &= \int_{-\infty}^{\infty} 1_{Y\leq y} f_{Y|Z}(x|1)\mathrm{d}x \\
&= \int_{-\infty}^{y} |x| e^{-x^2} \cdot \mathrm{d}x \\
&= \int_{-\infty}^{0} -xe^{-\frac{1}{2}\left(x^2+1\right)} \cdot \mathrm{d}x + \int_{0}^{y} xe^{-\frac{1}{2}\left(x^2+1\right)}\mathrm{d}x \\
&= 1 - \frac{1}{2}e^{-y^2}.
\end{aligned}$$

and if $y < 0$, then we have

$$\begin{aligned}
\mathrm{E}[1_{Y\leq y}|Z_2=1] &= \int_{-\infty}^{\infty} 1_{Y\leq y} f_{Y|Z}(x|1)\mathrm{d}x \\
&= \int_{-\infty}^{y} |x| e^{-x^2} \cdot \mathrm{d}x \\
&= \int_{-\infty}^{0} -xe^{-\frac{1}{2}\left(x^2+1\right)} \cdot \mathrm{d}x \\
&= \frac{1}{2}e^{-y^2}.
\end{aligned}$$

Finally we consider $Z_3 = 1_{\{Y=X\}}$. In this case, we note that the rotational symmetry of the joint distribution $f_{X,Y}$ implies

$$f_{Y|Z_3}(y|z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}.$$

Problem 4.64

Problem 4.64.a

We have

$$\begin{aligned} |a+b| &\leq ||a|+|b|| \\ &= |a|+|b|. \end{aligned}$$

Problem 4.64.b

If X is continuous, then we have

$$\begin{aligned} \mathrm{E}|X+Y| &= \int_{\mathbb{R}^2} |x+y| f_{X,Y} \mathrm{dm} \\ &\leq \int_{\mathbb{R}^2} (|x|+|y|) f_{X,Y} \mathrm{dm} \\ &= \int_{\mathbb{R}^2} |x| f_{X,Y} \mathrm{dm} + \int_{\mathbb{R}^2} |y| f_{X,Y} \mathrm{dm} \\ &= \mathrm{E}|X| + \mathrm{E}|Y|. \end{aligned}$$

A similar computation shows the same result if X is discrete.