1. Using the definition of conditional probability for $P(A|B \cap C)$ and $P(A|B \cap \overline{C})$, the RHS can be written as

$$\frac{P(A \cap B \cap C)}{P(B \cap C)}\, P(C) + \frac{P(A \cap B \cap \overline{C})}{P(B \cap \overline{C})}\, P(\overline{C}).$$

Because B and C are independent by assumption, $P(B \cap C) = P(B)P(C)$ and $P(B \cap \overline{C}) = P(B)P(\overline{C})$. Therefore, we have

$$\frac{P(A \cap B \cap C)}{P(B)P(C)}\, P(C) + \frac{P(A \cap B \cap \overline{C})}{P(B)P(\overline{C})}\, P(\overline{C}) = \frac{P(A \cap B \cap C) + P(A \cap B \cap \overline{C})}{P(B)}.$$

Now write $P(A \cap B \cap C) = P(A \cap B|C)P(C)$ and $P(A \cap B \cap \overline{C}) = P(A \cap B|\overline{C})P(\overline{C})$ by the multiplication rule. The last expression becomes

$$\frac{P(A \cap B|C)P(C) + P(A \cap B|\overline{C})P(\overline{C})}{P(B)} = \frac{P(A \cap B)}{P(B)} = P(A|B).$$

The penultimate step follows by noting $P(A \cap B|C)P(C) + P(A \cap B|\overline{C})P(\overline{C})$ is the LOTP expansion of $P(A \cap B)$, as C and $\overline{C}$ partition the sample space S. The last step follows from the definition of $P(A|B)$.

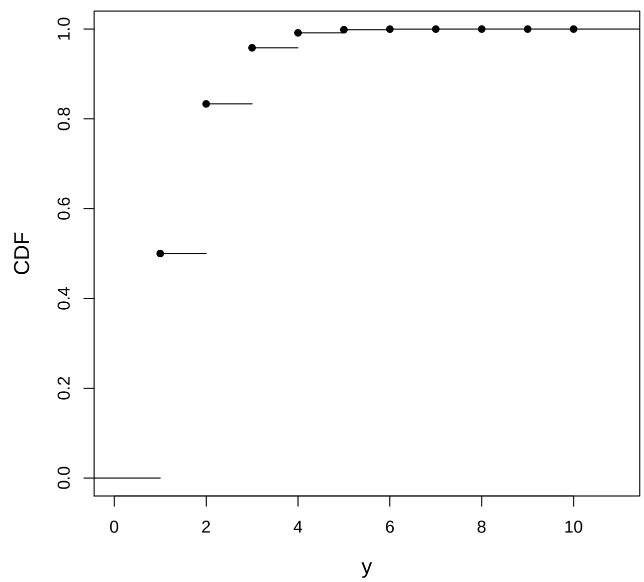
2. It is helpful to calculate the first few probabilities given by the pmf $p_Y(y) = P(Y = y)$:

$$\begin{aligned}
p_Y(1) = P(Y=1) &= \frac{1}{(1+1)!} = \frac{1}{2} \\
p_Y(2) = P(Y=2) &= \frac{2}{(2+1)!} = \frac{2}{6} \\
p_Y(3) = P(Y=3) &= \frac{3}{(3+1)!} = \frac{3}{24} \\
p_Y(4) = P(Y=4) &= \frac{4}{(4+1)!} = \frac{4}{120} \\
p_Y(5) = P(Y=5) &= \frac{5}{(5+1)!} = \frac{5}{720} \\
&\vdots
\end{aligned}$$

For these values of y, the cdf $F_Y(y) = P(Y \le y)$ is

$$\begin{aligned}
F_Y(1) = P(Y \le 1) &= \frac{1}{2} \\
F_Y(2) = P(Y \le 2) &= \frac{1}{2} + \frac{2}{6} = \frac{5}{6} \\
F_Y(3) = P(Y \le 3) &= \frac{1}{2} + \frac{2}{6} + \frac{3}{24} = \frac{23}{24} \\
F_Y(4) = P(Y \le 4) &= \frac{1}{2} + \frac{2}{6} + \frac{3}{24} + \frac{4}{120} = \frac{119}{120} \\
F_Y(5) = P(Y \le 5) &= \frac{1}{2} + \frac{2}{6} + \frac{3}{24} + \frac{4}{120} + \frac{5}{720} = \frac{719}{720} \\
&\vdots
\end{aligned}$$

This pattern continues. Because Y is a discrete random variable, the cdf $F_Y(y) = P(Y \le y)$ is a step function; the heights of the steps at each y correspond to the probability Y assumes



that value; i.e., $p_Y(y) = P(Y = y)$. A graph of this function is shown above. I constructed this graph in R, so I asked R to take the calculations out to $y = 10$.

3. (a) Using the definition of mathematical expectation, we have

$$E(Y^k) = \sum_{y=0}^{\infty} y^k p_Y(y) = \sum_{y=0}^{\infty} y^k \, \frac{\lambda^y e^{-\lambda}}{y!}.$$

In the last sum, the $y = 0$ term is 0, so we can write

$$E(Y^k) = \sum_{y=1}^{\infty} y^k \, \frac{\lambda^y e^{-\lambda}}{y!}$$

instead. Now, re-index the last sum by letting $x = y - 1$ (so that $y = x + 1$). We have

$$\begin{aligned} E(Y^k) = \sum_{x=0}^{\infty} (x+1)^k \, \frac{\lambda^{x+1} e^{-\lambda}}{(x+1)!} &= \sum_{x=0}^{\infty} (x+1)^{k-1}(x+1) \, \frac{\lambda^x \lambda e^{-\lambda}}{(x+1)x!} \\ &= \lambda \sum_{x=0}^{\infty} (x+1)^{k-1} \, \frac{\lambda^x e^{-\lambda}}{x!}. \end{aligned}$$

Because $\lambda^x e^{-\lambda}/x!$ is the Poisson(λ) pmf, the last sum equals $E[(X + 1)^{k-1}]$ where $X \sim$ Poisson(λ). However, X and Y have the same distribution, so this expectation also equals $E[(Y + 1)^{k-1}]$. Thus, the result.

(b) Take $k = 3$. From part (a), we have

$$E(Y^3) = \lambda E[(Y + 1)^2].$$

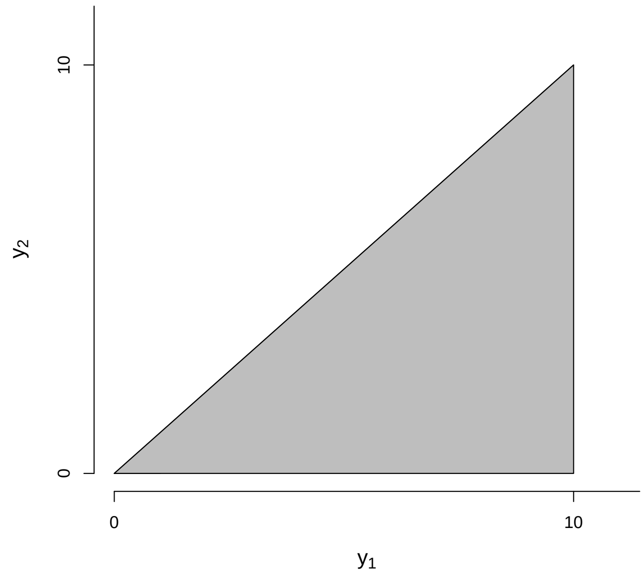
Write $(Y + 1)^2 = Y^2 + 2Y + 1$ so that

$$E[(Y + 1)^2] = E(Y^2 + 2Y + 1) = E(Y^2) + 2E(Y) + 1.$$

We know $E(Y) = \lambda$ and
$$E(Y^2) = V(Y) + [E(Y)]^2 = \lambda + \lambda^2.$$
Thus,
$$E[(Y+1)^2] = \lambda + \lambda^2 + 2\lambda + 1 = \lambda^2 + 3\lambda + 1.$$
Finally,
$$E(Y^3) = \lambda(\lambda^2 + 3\lambda + 1) = \lambda^3 + 3\lambda^2 + \lambda,$$
as claimed. You could also get $E(Y^3)$ by differentiating the mgf $m_Y(t) = \exp\{\lambda(e^t - 1)\}$ three times and evaluating the third derivative at $t = 0$. This would probably take a little longer.

4. The 0.25 quantile $\phi_{0.25}$ solves
$$\int_0^{\phi_{0.25}} \frac{1}{20}(3+y)dy = 0.25.$$
We have
$$\begin{aligned} \frac{1}{20}\left(3y + \frac{y^2}{2}\right)\Bigg|_0^{\phi_{0.25}} - 0.25 = 0 &\implies \frac{1}{20}\left(3\phi_{0.25} + \frac{\phi_{0.25}^2}{2}\right) - 0.25 = 0 \\ &\implies 0.025\phi_{0.25}^2 + 0.15\phi_{0.25} - 0.25 = 0. \end{aligned}$$
From the quadratic formula, we have
$$\begin{aligned} \phi_{0.25} = \frac{-0.15 \pm \sqrt{(0.15)^2 - 4(0.025)(-0.25)}}{2(0.025)} &= \frac{-0.15 \pm \sqrt{(0.15)^2 + 0.025}}{0.05} \\ &\implies \phi_{0.25} \approx 1.36 \quad \text{or} \quad \phi_{0.25} \approx -7.36. \end{aligned}$$
We discard the negative solution as it falls outside the support. Therefore, $\phi_{0.25} \approx 1.36$.

5. A graph of the bivariate support $R = \{(y_1, y_2) : 0 < y_2 < y_1 < 10\}$ is shown at the top of the next page. When integrating the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$ to determine both marginals, the limits of integration come from this picture. For $0 < y_1 < 10$, the marginal pdf of Y_1 is
$$f_{Y_1}(y_1) = \int_{y_2=0}^{y_1} \frac{1}{10y_1} dy_2 = \frac{1}{10y_1}\int_{y_2=0}^{y_1} 1 dy_2 = \frac{1}{10y_1}(y_1 - 0) = \frac{1}{10}.$$
For $0 < y_2 < 10$, the marginal pdf of Y_2 is
$$f_{Y_2}(y_2) = \int_{y_1=y_2}^{10} \frac{1}{10y_1} dy_1 = \frac{1}{10}\int_{y_1=y_2}^{10} \frac{1}{y_1} dy_1 = \frac{1}{10}\left(\ln y_1 \Big|_{y_1=y_2}^{10}\right) = \frac{1}{10}(\ln 10 - \ln y_2).$$
Summarizing,
$$f_{Y_1}(y_1) = \begin{cases} \frac{1}{10}, & 0 < y_1 < 10 \\ 0, & \text{otherwise} \end{cases}$$
and
$$f_{Y_2}(y_2) = \begin{cases} \frac{1}{10}(\ln 10 - \ln y_2), & 0 < y_2 < 10 \\ 0, & \text{otherwise.} \end{cases}$$



6. Regarding Y_2 as a function of Y_1 and Y_2, we can calculate $E(Y_2)$ using the joint pdf. A graph of the bivariate support $R = \{(y_1, y_2) : 0 < y_1 < y_2 < \infty\}$ is shown at the top of the next page. The limits on the double integral below come from this picture. We have

$$\begin{aligned}
E(Y_2) &= \int_{y_2=0}^{\infty}\int_{y_1=0}^{y_2} y_2 \times \underbrace{15e^{-2y_1}e^{-3y_2}}_{f_{Y_1,Y_2}(y_1,y_2)}\, dy_1 dy_2 \\
&= 15\int_{y_2=0}^{\infty} y_2 e^{-3y_2}\left(\int_{y_1=0}^{y_2} e^{-2y_1}dy_1\right)dy_2 \\
&= 15\int_{y_2=0}^{\infty} y_2 e^{-3y_2}\left(-\frac{1}{2}e^{-2y_1}\Big|_{y_1=0}^{y_2}\right)dy_2 \\
&= \frac{15}{2}\int_{y_2=0}^{\infty} y_2 e^{-3y_2}\left(1 - e^{-2y_2}\right)dy_2 \\
&= \frac{15}{2}\int_{y_2=0}^{\infty}\left(y_2 e^{-3y_2} - y_2 e^{-5y_2}\right)dy_2 \;=\; \frac{15}{2}\left(\int_{y_2=0}^{\infty} y_2 e^{-3y_2}dy_2 - \int_{y_2=0}^{\infty} y_2 e^{-5y_2}dy_2\right).
\end{aligned}$$

The integrand in the first integral is

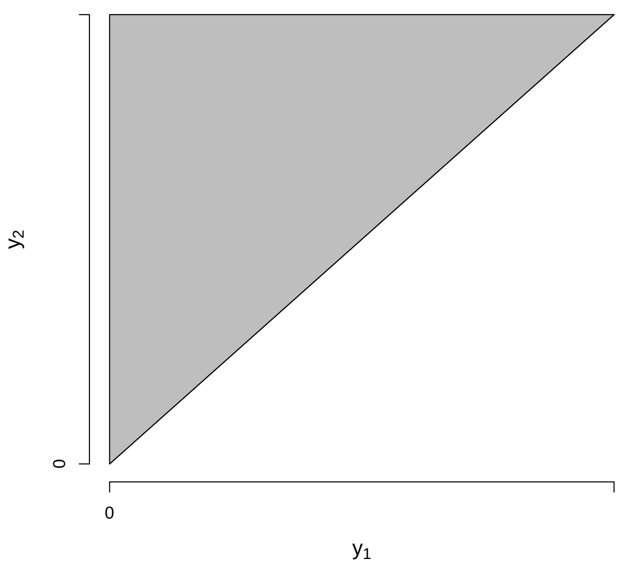
$$y_2 e^{-3y_2} = y_2^{2-1}e^{-y_2/(\frac{1}{3})},$$

a gamma kernel with shape $\alpha = 2$ and scale $\beta = 1/3$. Therefore, because integration is over $(0, \infty)$,

$$\int_{y_2=0}^{\infty} y_2 e^{-3y_2}dy_2 = \Gamma(2)\left(\frac{1}{3}\right)^2 = \frac{1}{9}.$$

An analogous argument shows

$$\int_{y_2=0}^{\infty} y_2 e^{-5y_2}dy_2 = \Gamma(2)\left(\frac{1}{5}\right)^2 = \frac{1}{25}.$$



Therefore,

$$E(Y_2) = \frac{15}{2}\left(\frac{1}{9} - \frac{1}{25}\right) \approx 0.53.$$

You could also derive the marginal pdf of Y_2 first and then calculate $E(Y_2)$ from it. For $0 < y_2 < \infty$, we have

$$f_{Y_2}(y_2) = \int_{y_1=0}^{y_2} 15e^{-2y_1}e^{-3y_2}dy_1 = \frac{15}{2}e^{-3y_2}(1 - e^{-2y_2})$$

and

$$E(Y_2) = \int_0^\infty y_2 \times \underbrace{\frac{15}{2}e^{-3y_2}(1 - e^{-2y_2})}_{f_{Y_2}(y_2)} \approx 0.53$$

as in the solution above.

7. The cdf for $Y \sim \mathcal{U}(a, b)$ is

$$F_Y(y) = \begin{cases} 0, & y \leq a \\ \dfrac{y-a}{b-a}, & a < y < b \\ 1, & y \geq b. \end{cases}$$

Therefore,

$$\int_0^\infty F_Y(y)[1 - F_Y(y)]dy = \int_a^b \frac{y-a}{b-a}\left(1 - \frac{y-a}{b-a}\right)dy.$$

In the last integral, use a u-substitution with

$$u = \frac{y-a}{b-a} \implies du = \frac{1}{b-a}\,dy.$$

The limits change under this transformation; as $y : a \to b$, we have $u : 0 \to 1$. Therefore,

$$\int_0^\infty F_Y(y)[1 - F_Y(y)]dy = (b-a)\int_0^1 u(1-u)du = (b-a) \times \frac{\Gamma(2)\Gamma(2)}{\Gamma(4)} = \frac{b-a}{6}.$$

Note the integrand

$$u(1-u) = u^{2-1}(1-u)^{2-1}$$

is a beta$(2,2)$ kernel and we are integrating over $(0,1)$. Finally, recall the mean of Y is $\mu = (a+b)/2$, the midpoint of a and b. Therefore, we have

$$g = \frac{1}{\mu}\int_0^\infty F_Y(y)[1 - F_Y(y)]dy = \frac{1}{\frac{a+b}{2}} \times \frac{b-a}{6} = \frac{b-a}{3(a+b)}.$$

8. It is easiest to conceptualize the sample points (outcomes) as follows:

(__ __ __ __ __),

where the first position says which hotel the first person checks into, the second position says which hotel the second person checks into, and so on. For example, the outcome

(1 1 1 1 1)

means that each person checks into the first hotel. How many outcomes are there in the sample space? For the first person, there are 8 hotels possible (1, 2, 3, 4, 5, 6, 7, and 8). For the second person, there are 8 hotels possible, and so on. By the basic rule of counting, there are

$$N = 8 \times 8 \times 8 \times 8 \times 8 = 8^5 = 32768$$

different outcomes in the sample space; i.e., different ways 5 people could check into 8 hotels. Define the event

$$A = \{\text{each person checks into a different hotel}\}.$$

What is n_a, the number of outcomes are in A? For example, we are looking for outcomes like

(1 2 3 4 5);

i.e., outcomes where each number is different. Combinatorially, n_a equals the number of ways to permute 5 unique integers selected from 8; i.e.,

$$n_a = P_5^8 = \frac{8!}{(8-5)!} = \frac{40320}{6} = 6720.$$

Assuming equally likely outcomes, we have

$$P(A) = \frac{n_a}{N} = \frac{6720}{32768} \approx 0.205.$$

9. The multiplicative constant

$$c = \frac{1}{\Gamma(\alpha)\beta^\alpha}$$

is free of y, so we can disregard it henceforth and just work with the gamma kernel. Using the product rule for derivatives, we have

$$\frac{d}{dy}\, y^{\alpha-1}e^{-y/\beta} = (\alpha-1)y^{\alpha-2}e^{-y/\beta} + y^{\alpha-1}\left(-\frac{1}{\beta}\right)e^{-y/\beta} = y^{\alpha-2}e^{-y/\beta}\left[(\alpha-1) - \frac{y}{\beta}\right].$$

For $y > 0$, the only way this derivative can be zero is if

$$(\alpha - 1) - \frac{y}{\beta} = 0 \implies \alpha - 1 = \frac{y}{\beta} \implies y = (\alpha - 1)\beta,$$

as claimed. Note that I did carry out the Second Derivative Test here to verify $y = (\alpha - 1)\beta$ maximizes $f_Y(y)$; i.e., I calculated

$$\frac{d^2}{dy^2}\, y^{\alpha-1}e^{-y/\beta}\Big|_{y=(\alpha-1)\beta} = -(\alpha-1)y^{\alpha-2}e^{-y/\beta} < 0,$$

provided that $\alpha > 1$. The second derivative is messy, so I didn't show it above.

10. Let's define the events

$$\begin{aligned} A &= \{\text{Allen hits the target}\} \\ B &= \{\text{Ben hits the target}\} \\ C &= \{\text{Chris hits the target}\}. \end{aligned}$$

Furthermore, define

$$\begin{aligned} D &= \{\text{exactly two shooters hit the target}\} \\ G &= \{\text{at least one shooter hits the target}\}. \end{aligned}$$

We want calculate $P(D|G)$. From the definition of conditional probability, we have

$$P(D|G) = \frac{P(D \cap G)}{P(G)}.$$

However, note that $D \subset G$ so $D \cap G = D$. Therefore, we want to calculate $P(D)/P(G)$. To find $P(D)$, note that we can write

$$D = (A \cap B \cap \overline{C}) \cup (A \cap \overline{B} \cap C) \cup (\overline{A} \cap B \cap C),$$

the union of three pairwise mutually exclusive events (i.e., the intersection of any two events is the null set). From Axiom 3, we have

$$P(D) = P(A \cap B \cap \overline{C}) + P(A \cap \overline{B} \cap C) + P(\overline{A} \cap B \cap C).$$

If we make the additional assumption that A, B, and C are mutually independent, then

$$\begin{aligned} P(D) &= P(A)P(B)P(\overline{C}) + P(A)P(\overline{B})P(C) + P(\overline{A})P(B)P(C) \\ &= p_1p_2(1-p_3) + p_1(1-p_2)p_3 + (1-p_1)p_2p_3. \end{aligned}$$

Instead of calculating $P(G)$, it is easier to calculate $P(\overline{G})$, where the complement event

$$\overline{G} = \{\text{no shooter hits the target}\}.$$

Assuming A, B, and C are mutually independent, we have

$$P(\overline{G}) = P(\overline{A} \cap \overline{B} \cap \overline{C}) = P(\overline{A})P(\overline{B})P(\overline{C}) = (1-p_1)(1-p_2)(1-p_3).$$

Therefore,

$$P(D|G) = \frac{P(D)}{P(G)} = \frac{P(D)}{1 - P(\overline{G})} = \frac{p_1p_2(1-p_3) + p_1(1-p_2)p_3 + (1-p_1)p_2p_3}{1 - (1-p_1)(1-p_2)(1-p_3)}.$$

11. The random variable Y follows a hypergeometric distribution with $r = 4$ (number of MS students/Class 1 objects), $N = 32$ (population size), and $n = 2$ (sample size). We have

$$\begin{aligned}
P(Y=0) &= \frac{\binom{4}{0}\binom{28}{2}}{\binom{32}{2}} = \frac{1 \times 378}{496} = \frac{378}{496} \\
P(Y=1) &= \frac{\binom{4}{1}\binom{28}{1}}{\binom{32}{2}} = \frac{4 \times 28}{496} = \frac{112}{496} \\
P(Y=2) &= \frac{\binom{4}{2}\binom{28}{0}}{\binom{32}{2}} = \frac{6 \times 1}{496} = \frac{6}{496}.
\end{aligned}$$

We can write the pmf of Y as follows:

y	0	1	2
$p_Y(y)$	378/496	112/496	6/496

The mgf of Y is

$$\begin{aligned}
m_Y(t) = E(e^{tY}) &= \sum_{y=0}^{2} e^{ty} p_Y(y) \\
&= e^{t(0)}\frac{378}{496} + e^{t(1)}\frac{112}{496} + e^{t(2)}\frac{6}{496} = \frac{378}{496} + \left(\frac{112}{496}\right)e^{t} + \left(\frac{6}{496}\right)e^{2t}.
\end{aligned}$$

12. Regard each income tax form as a "trial" with our usual Bernoulli trial assumptions; i.e.,

1. Each form either contains "serious errors" (success) or it doesn't (failure).
2. The probability $p = 0.05$ a form contains "serious errors" is the same for all forms.
3. The statuses of all forms (i.e., whether they contain "serious errors" or not) are mutually independent.

Let Y denote the number of forms the auditor will review until he finds the first form with "serious errors." Under the Bernoulli trial assumptions above, $Y \sim \text{geom}(p = 0.05)$. We have

$$\begin{aligned}
P(Y \geq 25) = 1 - P(Y \leq 24) &= 1 - \sum_{y=1}^{24} (0.95)^{y-1}(0.05) \\
&= 1 - (0.05)\left[\frac{1 - (0.95)^{25-1}}{1 - 0.95}\right] = (0.95)^{24} \approx 0.292.
\end{aligned}$$

Note that $\sum_{y=1}^{24}(0.95)^{y-1}$ is a finite geometric sum with common ratio $r = 0.95$.

13. The second moment of Y is

$$E(Y^2) = \int_{\mathbb{R}} y^2 f_Y(y)dy = \int_0^1 y^2 \left(-\ln y\right) dy.$$

This integral can be computed using integration by parts. Let

$$\begin{array}{ll} u = -\ln y & du = -\dfrac{1}{y}dy \\ dv = y^2\ dy & v = \dfrac{y^3}{3}. \end{array}$$

With these selections, we have

$$\begin{aligned} E(Y^2) = \int_0^1 y^2 \left(-\ln y\right) dy = -\frac{y^3}{3}\ln y\bigg|_0^1 - \int_0^1 \frac{y^3}{3}\left(-\frac{1}{y}\right) dy &= (0+0) + \frac{1}{3}\int_0^1 y^2 dy \\ &= \frac{1}{3}\left(\frac{y^3}{3}\bigg|_0^1\right) = \frac{1}{9}. \end{aligned}$$

14. First, the covariance of X and Y is

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = 3 - (1)(5) = -2.$$

The covariance of U_1 and U_2 is

$$\begin{aligned} \text{Cov}(U_1,U_2) &= \text{Cov}(X+3Y, 2X-Y) \\ &= \text{Cov}(X,2X) + \text{Cov}(X,-Y) + \text{Cov}(3Y,2X) + \text{Cov}(3Y,-Y) \\ &= 2\text{Cov}(X,X) - \text{Cov}(X,Y) + 6\text{Cov}(Y,X) - 3\text{Cov}(Y,Y). \end{aligned}$$

Recall the covariance of any random variable with itself is the random variable's variance. In addition, $\text{Cov}(X,Y) = \text{Cov}(Y,X)$. Therefore,

$$\begin{aligned} \text{Cov}(U_1,U_2) &= 2V(X) - \text{Cov}(X,Y) + 6\text{Cov}(X,Y) - 3V(Y) \\ &= 2V(X) + 5\text{Cov}(X,Y) - 3V(Y) = 2(4) + 5(-2) - 3(9) = -29. \end{aligned}$$

Now,

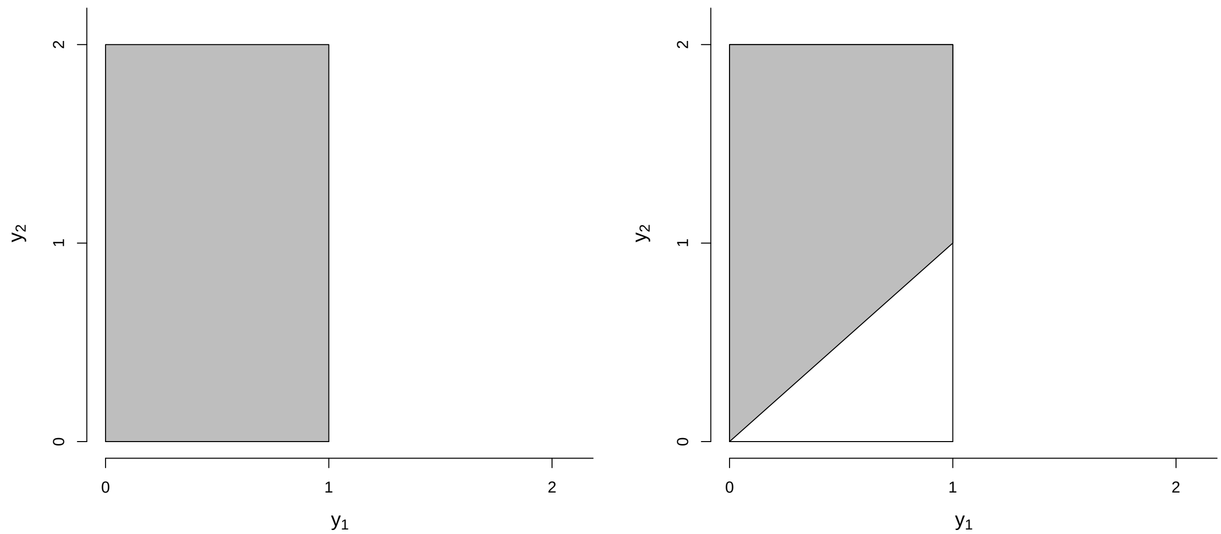
$$\begin{aligned} V(U_1) = V(X+3Y) &= V(X) + V(3Y) + 2\text{Cov}(X,3Y) \\ &= V(X) + 9V(Y) + 6\text{Cov}(X,Y) = 4 + 9(9) + 6(-2) = 73 \end{aligned}$$

and

$$\begin{aligned} V(U_2) = V(2X-Y) &= V(2X) + V(-Y) + 2\text{Cov}(2X,-Y) \\ &= 4V(X) + V(Y) - 4\text{Cov}(X,Y) = 4(4) + 9 - 4(-2) = 33. \end{aligned}$$

Finally, the correlation of U_1 and U_2 is

$$\rho = \frac{\text{Cov}(U_1,U_2)}{\sqrt{V(U_1)}\sqrt{V(U_2)}} = \frac{-29}{\sqrt{73}\sqrt{33}} \approx -0.591.$$



15. The bivariate support $R = \{(y_1, y_2) : 0 < y_1 < 1,\ 0 < y_2 < 2\}$ is shown above (left). We want to calculate $P(Y_2 > Y_1) = P(Y_2 - Y_1 > 0)$. This probability is the volume under the joint pdf over the set

$$B = \{(y_1, y_2) : 0 < y_1 < 1,\ 0 < y_2 < 2,\ y_2 - y_1 > 0\}.$$

The boundary of this set is $y_2 - y_1 = 0 \Longrightarrow y_2 = y_1$, a linear function of y_1 with intercept 0 and slope 1. The set B is the trapezoidal region shown above (right). The limits on the double integral below come from this picture. It is easier to integrate in the y_2 direction first; we have

$$\begin{aligned}
P(Y_2 - Y_1 > 0) &= \int_{y_1=0}^{1}\int_{y_2=y_1}^{2} \frac{1}{5}(y_1 + 2y_2)\ dy_2 dy_1 \\
&= \frac{1}{5}\int_{y_1=0}^{1} \left[(y_1 y_2 + y_2^2)\Big|_{y_2=y_1}^{2} \right] dy_1 \\
&= \frac{1}{5}\int_{y_1=0}^{1} \left[(2y_1 + 4) - (y_1^2 + y_1^2) \right] dy_1 \\
&= \frac{1}{5}\int_{y_1=0}^{1} (-2y_1^2 + 2y_1 + 4) dy_1 \\
&= \frac{1}{5}\left(-\frac{2y_1^3}{3} + y_1^2 + 4y_1 \right)\Big|_0^1 = \frac{1}{5}\left(-\frac{2}{3} + 1 + 4 \right) = \frac{1}{5}\left(\frac{13}{3}\right) = \frac{13}{15}.
\end{aligned}$$