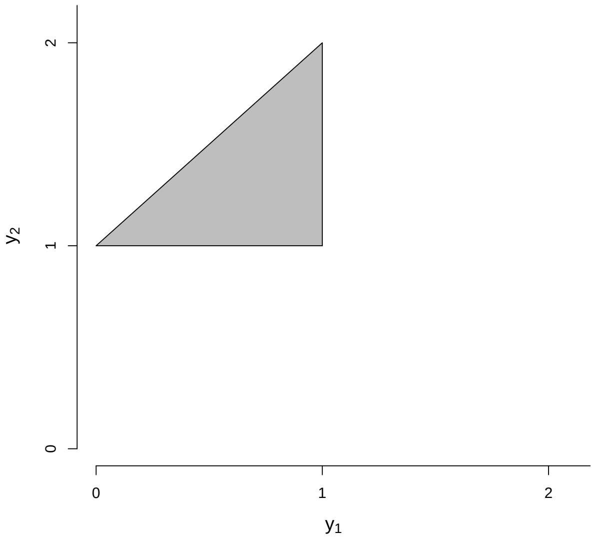
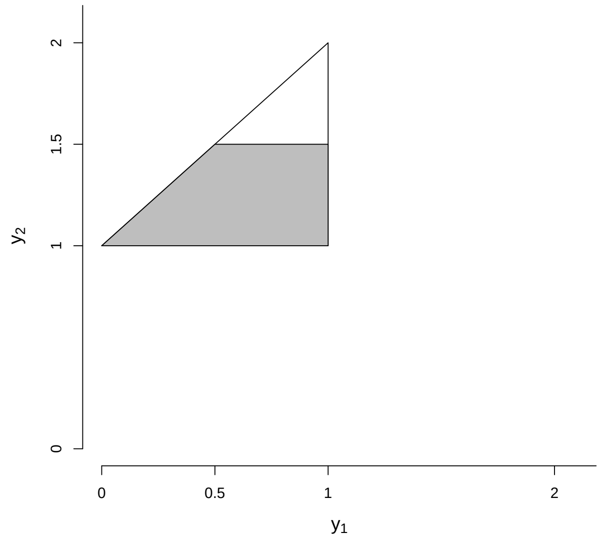
1. (a) A graph of the support $R = \{(y_1, y_2) : 0 \leq y_1 \leq 1,\ 1 \leq y_2 \leq y_1 + 1\}$ is shown below (left).





The equation of the diagonal line (boundary of the support) is $y_2 = y_1 + 1$. We have

$$\begin{aligned}
\iint_{\mathbb{R}^2} f_{Y_1,Y_2}(y_1, y_2)dy_1dy_2 &= \int_{y_1=0}^{1}\int_{y_2=1}^{y_1+1} \frac{20}{7}y_1^2y_2\ dy_2dy_1 \\
&= \frac{10}{7}\int_{y_1=0}^{1} y_1^2 \left(\frac{y_2^2}{2}\bigg|_{y_2=1}^{y_1+1} \right) dy_1 \\
&= \frac{10}{7}\int_{y_1=0}^{1} y_1^2 \left[(y_1+1)^2 - 1^2\right] dy_1 \\
&= \frac{10}{7}\int_{y_1=0}^{1} y_1^2(y_1^2 + 2y_1 + 1 - 1)dy_1 \\
&= \frac{10}{7}\int_{y_1=0}^{1} (y_1^4 + 2y_1^3)dy_1 \ = \ \frac{10}{7}\left(\frac{y_1^5}{5} + \frac{2y_1^4}{4} \right)\bigg|_{y_1=0}^{1} = \frac{10}{7}\left(\frac{1}{5} + \frac{1}{2} \right) = 1.
\end{aligned}$$

Therefore, $f_{Y_1,Y_2}(y_1, y_2)$ is a valid pdf. We can also check this in R:

```
library(pracma)
joint.pdf <- function(y1,y2) (20/7)*y1^2*y2
y2max <- function(y1) y1+1
integral2(joint.pdf,0,1,1,y2max)
$`Q`
[1] 1
$error
[1] 1.387779e-17
```

Note that you could also carry out the integration above by integrating with respect to y_1 first; i.e.,

$$\iint_{\mathbb{R}^2} f_{Y_1,Y_2}(y_1, y_2)dy_1dy_2 = \int_{y_2=1}^{2}\int_{y_1=y_2-1}^{1} \frac{20}{7}y_1^2y_2\ dy_1dy_2.$$

(b) Because the support R involves a (linear) constraint between y_1 and y_2; i.e., $y_2 \leq y_1+1$, the random variables Y_1 and Y_2 are dependent. You could also show this by deriving the marginal pdfs. For $0 < y_1 < 1$, we have

$$f_{Y_1}(y_1) = \int_{y_2=1}^{y_1+1} \frac{20}{7} y_1^2 y_2 \; dy_2 \;\; = \;\; \frac{10}{7}(y_1^4 + 2y_1^3).$$

For $1 < y_2 < 2$, we have

$$f_{Y_2}(y_2) = \int_{y_1=y_2-1}^{1} \frac{20}{7} y_1^2 y_2 \; dy_1 \;\; = \;\; \frac{20}{21}(2y_2 - 3y_2^2 + 3y_2^3 - y_2^4).$$

It is easy to see

$$f_{Y_1,Y_2}(y_1,y_2) \neq f_{Y_1}(y_1) f_{Y_2}(y_2)$$

and hence Y_1 and Y_2 are dependent.
(c) To find $P(Y_2 < 1.5)$, we can integrate the joint pdf $f_{Y_1,Y_2}(y_1,y_2)$ over the trapezoidal region on the last page (top, right). It is easiest to integrate in the y_1 direction first. We have

$$\begin{aligned}
P(Y_2 < 1.5) &= \int_{y_2=1}^{1.5} \int_{y_1=y_2-1}^{1} \frac{20}{7} y_1^2 y_2 \; dy_1 dy_2 \\
&= \frac{20}{7} \int_{y_2=1}^{1.5} y_2 \left(\frac{y_1^3}{3} \right) \Bigg|_{y_1=y_2-1}^{1} dy_2 \\
&= \frac{20}{21} \int_{y_2=1}^{1.5} y_2 \left[1 - (y_2-1)^3 \right] dy_2 \\
&= \frac{20}{21} \int_{y_2=1}^{1.5} (2y_2 - 3y_2^2 + 3y_2^3 - y_2^4) dy_2 \qquad (*) \\
&= \frac{20}{21} \left(y_2^2 - y_2^3 + \frac{3y_2^4}{4} - \frac{y_2^5}{5} \right) \Bigg|_1^{1.5} \\
&= \frac{20}{21} \left[\left((1.5)^2 - (1.5)^3 + \frac{3(1.5)^4}{4} - \frac{(1.5)^5}{5} \right) - \left(1 - 1 + \frac{3}{4} - \frac{1}{5} \right) \right] \approx 0.574.
\end{aligned}$$

If you derived the marginal pdf of Y_2 in part (b), then you could simply calculate

$$P(Y_2 < 1.5) = \int_1^{1.5} \underbrace{\frac{20}{21}(2y_2 - 3y_2^2 + 3y_2^3 - y_2^4)}_{= f_{Y_2}(y_2)} dy_2 \approx 0.574.$$

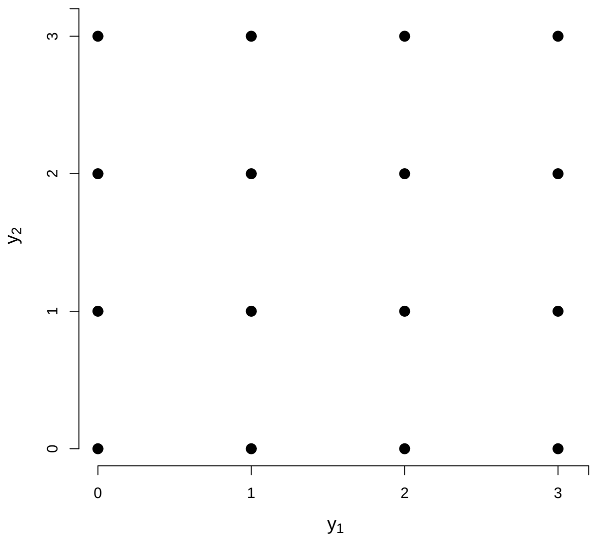
Essentially, this is what one is doing in **(*)** above.

```
> integrand <- function(y){(20/21)*(2*y-3*y^2+3*y^3-y^4)}
> integrate(integrand,lower=1,upper=1.5)
0.5744048 with absolute error < 6.4e-15
```

2. (a) The two-dimensional support of Y_1 and Y_2 is the set of the 16 ordered pairs

$$R = \{(0,0),(1,0),(2,0),(3,0),(0,1),(1,1),(2,1),(3,1),(0,2),(1,2),(2,2),(3,2),(0,3),(1,3),(2,3),(3,3)\}.$$

This two-dimensional support set is shown at the top of the next page.



(b) The marginal pmfs of Y_1 and Y_2 are obtained by summing $p_{Y_1,Y_2}(y_1, y_2)$ over the values of y_2 and y_1, respectively. These functions are shown in the margins below:

$p_{Y_1,Y_2}(y_1, y_2)$	$y_2 = 0$	$y_2 = 1$	$y_2 = 2$	$y_2 = 3$	$p_{Y_1}(y_1)$
$y_1 = 0$	0.01	0.02	0.01	0.01	0.05
$y_1 = 1$	0.01	0.06	0.12	0.06	0.25
$y_1 = 2$	0.02	0.08	0.30	0.10	0.50
$y_1 = 3$	0.01	0.04	0.05	0.10	0.20
$p_{Y_2}(y_2)$	0.05	0.20	0.48	0.27	

That is, the marginal pmf of Y_1 is

y_1	0	1	2	3
$p_{Y_1}(y_1)$	0.05	0.25	0.50	0.20

and the marginal pmf of Y_2 is

y_2	0	1	2	3
$p_{Y_2}(y_2)$	0.05	0.20	0.48	0.27

For independence to hold, we would need

$$p_{Y_1,Y_2}(y_1, y_2) = p_{Y_1}(y_1)p_{Y_2}(y_2)$$

to hold for all (y_1, y_2) in the support R. However, this condition does not even hold for the first value $(0, 0)$; i.e.,

$$0.01 = p_{Y_1,Y_2}(0, 0) \neq p_{Y_1}(0)p_{Y_2}(0) = 0.05(0.05) = 0.0025.$$

Therefore, Y_1 and Y_2 are not independent.

(c) We want the conditional pmf $p_{Y_1|Y_2}(y_1|y_2 = 1)$. This is a univariate pmf with four possible values of Y_1, namely, 0, 1, 2, and 3. These conditional probabilities are calculated below:

$$\begin{aligned}
p_{Y_1|Y_2}(y_1 = 0|y_2 = 1) &= \frac{p_{Y_1,Y_2}(0,1)}{p_{Y_2}(1)} = \frac{0.02}{0.20} = 0.10 \\
p_{Y_1|Y_2}(y_1 = 1|y_2 = 1) &= \frac{p_{Y_1,Y_2}(1,1)}{p_{Y_2}(1)} = \frac{0.06}{0.20} = 0.30 \\
p_{Y_1|Y_2}(y_1 = 2|y_2 = 1) &= \frac{p_{Y_1,Y_2}(2,1)}{p_{Y_2}(1)} = \frac{0.08}{0.20} = 0.40 \\
p_{Y_1|Y_2}(y_1 = 3|y_2 = 1) &= \frac{p_{Y_1,Y_2}(3,1)}{p_{Y_2}(1)} = \frac{0.04}{0.20} = 0.20.
\end{aligned}$$

We can display the conditional pmf of Y_1 given $Y_2 = 1$ in the following table:

y_1	0	1	2	3
$p_{Y_1\|Y_2}(y_1\|y_2 = 1)$	0.10	0.30	0.40	0.20

(d) We can get $E(Y_1 - 2Y_2)$ using the joint pmf of Y_1 and Y_2; i.e.,

$$\begin{aligned}
E(Y_1 - 2Y_2) &= \sum\sum_{(y_1,y_2)\in R} (y_1 - 2y_2)p_{Y_1,Y_2}(y_1, y_2) \\
&= [0 - 2(0)](0.01) + [0 - 2(1)](0.02) + [0 - 2(2)](0.01) + [0 - 2(3)](0.01) \\
&\quad + [1 - 2(0)](0.01) + [1 - 2(1)](0.06) + [1 - 2(2)](0.12) + [1 - 2(3)](0.06) \\
&\quad + [2 - 2(0)](0.02) + [2 - 2(1)](0.08) + [2 - 2(2)](0.30) + [2 - 2(3)](0.10) \\
&\quad + [3 - 2(0)](0.01) + [3 - 2(1)](0.04) + [3 - 2(2)](0.05) + [3 - 2(3)](0.10) = -2.09.
\end{aligned}$$

We could also write $E(Y_1 - 2Y_2) = E(Y_1) - 2E(Y_2)$ using the linearity properties of expectation and then use the marginal pmfs to get the marginal means. That is,

$$E(Y_1) = \sum_{y_1=0}^{3} y_1 p_{Y_1}(y_1) = 0(0.05) + 1(0.25) + 2(0.50) + 3(0.20) = 1.85$$

and

$$E(Y_2) = \sum_{y_2=0}^{3} y_2 p_{Y_2}(y_2) = 0(0.05) + 1(0.20) + 2(0.48) + 3(0.27) = 1.97.$$

Thus,

$$E(Y_1 - 2Y_2) = E(Y_1) - 2E(Y_2) = 1.85 - 2(1.97) = -2.09.$$

To get $V(Y_1 - 2Y_2)$, we also have options.

Option 1: Use the variance computing formula with the random variable "$Y_1 - 2Y_2$," that is,

$$\begin{aligned}
V(Y_1 - 2Y_2) &= E[(Y_1 - 2Y_2)^2] - [E(Y_1 - 2Y_2)]^2 \\
&= E[(Y_1 - 2Y_2)^2] - (-2.09)^2.
\end{aligned}$$

Therefore, all we have to get is $E[(Y_1 - 2Y_2)^2]$, the second moment of $Y_1 - 2Y_2$. Using the joint pmf, we have

$$\begin{aligned}
E[(Y_1 - 2Y_2)^2] &= \sum\sum_{(y_1,y_2)\in R} (y_1 - 2y_2)^2 p_{Y_1,Y_2}(y_1, y_2) \\
&= [0 - 2(0)]^2(0.01) + [0 - 2(1)]^2(0.02) + [0 - 2(2)]^2(0.01) + [0 - 2(3)]^2(0.01) \\
&\quad + [1 - 2(0)]^2(0.01) + [1 - 2(1)]^2(0.06) + [1 - 2(2)]^2(0.12) + [1 - 2(3)]^2(0.06) \\
&\quad + [2 - 2(0)]^2(0.02) + [2 - 2(1)]^2(0.08) + [2 - 2(2)]^2(0.30) + [2 - 2(3)]^2(0.10) \\
&\quad + [3 - 2(0)]^2(0.01) + [3 - 2(1)]^2(0.04) + [3 - 2(2)]^2(0.05) + [3 - 2(3)]^2(0.10) \\
&= 7.21.
\end{aligned}$$

Therefore,

$$V(Y_1 - 2Y_2) = 7.21 - (-2.09)^2 \approx 2.84.$$

Option 2: We can get $V(Y_1 - 2Y_2)$ using what we learned about variances of linear combinations; i.e., $Y_1 - 2Y_2$ is a linear combination of Y_1 and Y_2 and

$$\begin{aligned}
V(Y_1 - 2Y_2) &= V(Y_1) + V(-2Y_2) + 2\text{Cov}(Y_1, -2Y_2) \\
&= V(Y_1) + (-2)^2V(Y_2) + 2(1)(-2)\text{Cov}(Y_1, Y_2) \\
&= V(Y_1) + 4V(Y_2) - 4\text{Cov}(Y_1, Y_2).
\end{aligned}$$

We can get $V(Y_1)$ and $V(Y_2)$ from the marginal pmfs. The marginal second moments are

$$\begin{aligned}
E(Y_1^2) &= \sum_{y_1=0}^{3} y_1^2 p_{Y_1}(y_1) = 0^2(0.05) + 1^2(0.25) + 2^2(0.50) + 3^2(0.20) = 4.05 \\
E(Y_2^2) &= \sum_{y_2=0}^{3} y_2^2 p_{Y_2}(y_2) = 0^2(0.05) + 1^2(0.20) + 2^2(0.48) + 3^2(0.27) = 4.55.
\end{aligned}$$

Therefore,

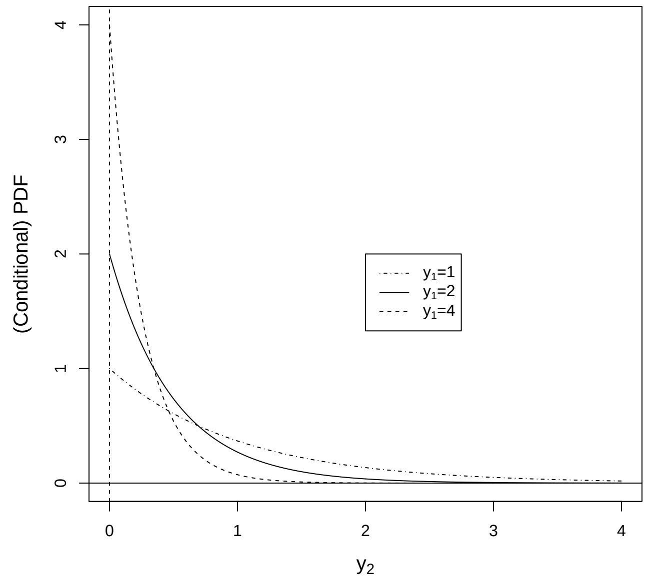
$$\begin{aligned}
V(Y_1) &= E(Y_1^2) - [E(Y_1)]^2 = 4.05 - (1.85)^2 = 0.6275 \\
V(Y_2) &= E(Y_2^2) - [E(Y_2)]^2 = 4.55 - (1.97)^2 = 0.6691.
\end{aligned}$$

To get $\text{Cov}(Y_1, Y_2)$, we use the covariance computing formula $\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2)$. We already calculated $E(Y_1) = 1.85$ and $E(Y_2) = 1.97$. We get $E(Y_1Y_2)$ from the joint pmf; i.e.,

$$\begin{aligned}
E(Y_1Y_2) &= \sum\sum_{(y_1,y_2)\in R} y_1 y_2 p_{Y_1,Y_2}(y_1, y_2) \\
&= (0)(0)(0.01) + (0)(1)(0.02) + (0)(2)(0.01) + (0)(3)(0.01) \\
&\quad + (1)(0)(0.01) + (1)(1)(0.06) + (1)(2)(0.12) + (1)(3)(0.06) \\
&\quad + (2)(0)(0.02) + (2)(1)(0.08) + (2)(2)(0.30) + (2)(3)(0.10) \\
&\quad + (3)(0)(0.01) + (3)(1)(0.04) + (3)(2)(0.05) + (3)(3)(0.10) = 3.76.
\end{aligned}$$

Therefore,

$$\text{Cov}(Y_1, Y_2) = 3.76 - (1.85)(1.97) \approx 0.1155$$



(this shows again Y_1 and Y_2 are dependent because the covariance is nonzero) and

$$\begin{aligned} V(Y_1 - 2Y_2) &= V(Y_1) + 4V(Y_2) - 4\text{Cov}(Y_1, Y_2) \\ &= 0.6275 + 4(0.6691) - 4(0.1155) \approx 2.84. \end{aligned}$$

3. (a) I'll pick a few values of $y_1 \in (0, 10)$. For example, when $y_1 = 1$, we have

$$f_{Y_2|Y_1}(y_2|y_1 = 1) = \begin{cases} e^{-y_2}, & y_2 > 0 \\ 0, & \text{otherwise,} \end{cases}$$

an exponential pdf with mean 1. When $y_1 = 2$, we have

$$f_{Y_2|Y_1}(y_2|y_1 = 2) = \begin{cases} 2e^{-2y_2}, & y_2 > 0 \\ 0, & \text{otherwise,} \end{cases}$$

an exponential pdf with mean 1/2. When $y_1 = 4$, we have

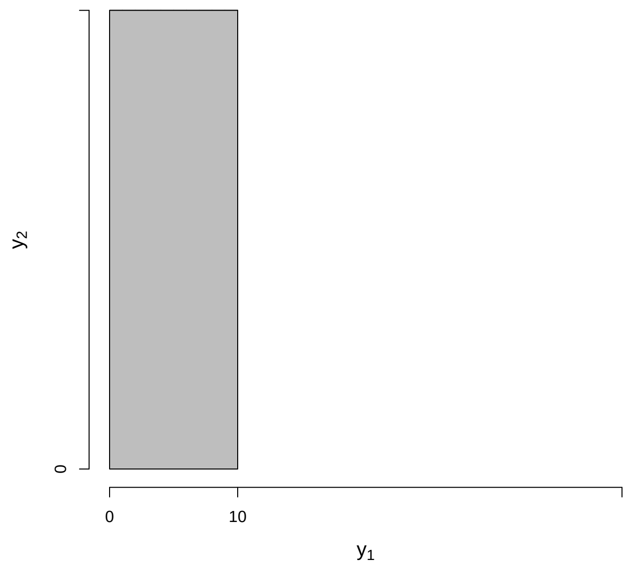
$$f_{Y_2|Y_1}(y_2|y_1 = 4) = \begin{cases} 4e^{-4y_2}, & y_2 > 0 \\ 0, & \text{otherwise,} \end{cases}$$

an exponential pdf with mean 1/4. In general, the conditional distribution of Y_2, given $Y_1 = y_1$, is exponential with mean $\beta = 1/y_1$; i.e.,

$$Y_2|Y_1 = y_1 \sim \text{exponential}\left(\frac{1}{y_1}\right).$$

Graphs of the three pdfs above are shown at the top of this page.
(b) We are given the conditional pdf $f_{Y_2|Y_1}(y_2|y_1)$ and the marginal pdf

$$f_{Y_1}(y_1) = \begin{cases} \frac{1}{10}, & 0 < y_1 < 10 \\ 0, & \text{otherwise.} \end{cases}$$



The joint pdf of Y_1 and Y_2 is

$$f_{Y_1,Y_2}(y_1,y_2) = f_{Y_2|Y_1}(y_2|y_1)f_{Y_1}(y_1).$$

This joint pdf is nonzero when $0 < y_1 < 10$ and $y_2 > 0$. Therefore, we have

$$f_{Y_1,Y_2}(y_1,y_2) = \begin{cases} \dfrac{y_1}{10}e^{-y_1 y_2}, & 0 < y_1 < 10,\ y_2 > 0 \\ 0, & \text{otherwise.} \end{cases}$$

A graph of the support $R = \{(y_1,y_2) : 0 < y_1 < 10,\ y_2 > 0\}$ is shown above. The marginal pdf of Y_2 is found by integrating the joint pdf $f_{Y_1,Y_2}(y_1,y_2)$ over y_1. For $y_2 > 0$, we have

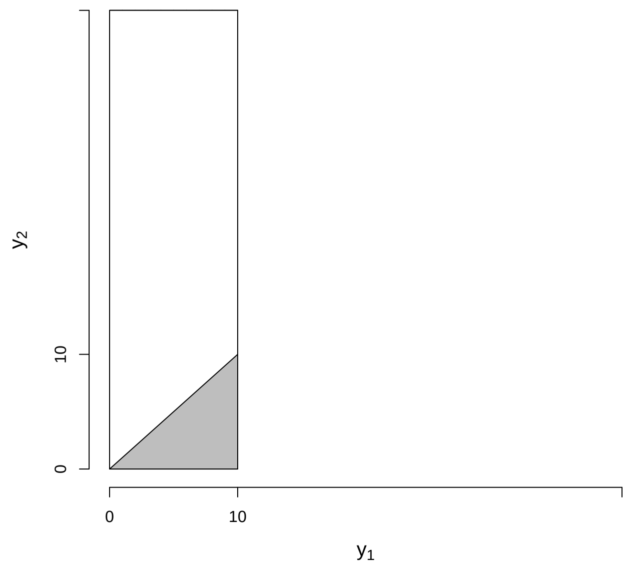
$$f_{Y_2}(y_2) = \int_{\mathbb{R}} f_{Y_1,Y_2}(y_1,y_2)dy_1 = \int_{y_1=0}^{10} \frac{y_1}{10}e^{-y_1 y_2}dy_1 = \frac{1}{10}\int_{y_1=0}^{10} y_1 e^{-y_1 y_2}dy_1.$$

The last integral can be computed using integration by parts. Let

$$\begin{aligned} u &= y_1 & du &= dy_1 \\ dv &= e^{-y_1 y_2}dy_1 & v &= -\frac{1}{y_2}e^{-y_1 y_2}. \end{aligned}$$

With these selections,

$$\begin{aligned} f_{Y_2}(y_2) = \frac{1}{10}\int_{y_1=0}^{10} y_1 e^{-y_1 y_2}dy_1 &= \frac{1}{10}\left(-\frac{y_1}{y_2}e^{-y_1 y_2}\Big|_{y_1=0}^{10} - \int_{y_1=0}^{10} -\frac{1}{y_2}e^{-y_1 y_2}dy_1\right) \\ &= \frac{1}{10}\left[\left(-\frac{10}{y_2}e^{-10y_2} - 0\right) - \left(-\frac{1}{y_2}\right)\left(-\frac{1}{y_2}\right)e^{-y_1 y_2}\Big|_{y_1=0}^{10}\right] \\ &= -\frac{1}{y_2}e^{-10y_2} - \frac{1}{10y_2^2}\left(e^{-10y_2} - 1\right) \\ &= \frac{1 - e^{-10y_2} - 10y_2 e^{-10y_2}}{10y_2^2}. \end{aligned}$$



Summarizing,

$$f_{Y_2}(y_2) = \begin{cases} \dfrac{1 - e^{-10y_2} - 10y_2 e^{-10y_2}}{10y_2^2}, & y_2 > 0 \\ 0, & \text{otherwise.} \end{cases}$$

I used R to ensure $f_{Y_2}(y_2)$ is a valid pdf; i.e., it integrates to one.

```
> integrand <- function(y){(1-exp(-10*y)-10*y*exp(-10*y))/(10*y^2)}
> integrate(integrand,lower=0,upper=Inf)
1 with absolute error < 2.9e-07
```

(c) No, Y_1 and Y_2 are not independent because

$$\begin{aligned} f_{Y_2|Y_1}(y_2|y_1) &= y_1 e^{-y_1 y_2} \\ f_{Y_2}(y_2) &= \frac{1 - e^{-10y_2} - 10y_2 e^{-10y_2}}{10y_2^2} \end{aligned}$$

when $y_2 > 0$. That is, conditioning on $Y_1 = y_1$ changes the distribution of Y_2. Equivalently, you could argue

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{y_1}{10} e^{-y_1 y_2} \neq \frac{1}{10} \times \frac{1 - e^{-10y_2} - 10y_2 e^{-10y_2}}{10y_2^2} = f_{Y_1}(y_1) f_{Y_2}(y_2).$$

That is, the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$ does not factor into the product of the marginal pdfs.
(d) To find $P(Y_1 - Y_2 > 0)$, we integrate the joint pdf $f_{Y_1,Y_2}(y_1, y_2)$ over the set

$$B = \{(y_1, y_2) : 0 < y_1 < 10,\ y_2 > 0,\ y_1 - y_2 > 0\}.$$

This set is shown in the graph at the top of the page, and the limits on the double integral come from this picture. The boundary of B is determined as follows:

$$y_1 - y_2 = 0 \implies y_2 = y_1.$$

Therefore,

$$P(Y_1 - Y_2 > 0) = \int_{y_1=0}^{10} \int_{y_2=0}^{y_1} \frac{y_1}{10} e^{-y_1 y_2} \, dy_2 dy_1 \approx 0.911.$$

I did this integral numerically in R:

```
> library(pracma)
> joint.pdf <- function(y1,y2) (y1/10)*exp(-y1*y2)
> y2max <- function(y1) y1
> integral2(joint.pdf,0,10,0,y2max)
$‘Q‘
[1] 0.91116514
$error
[1] 8.543103e-08
```

4. (a) If Z_1 and Z_2 are independent, then so are Z_1^2 and Z_2^2 (functions of independent random variables are independent). Therefore,

$$\begin{aligned} P(Z_1^2 \leq 1, Z_2^2 \leq 1) &\stackrel{Z_1^2 \perp\!\!\!\perp Z_2^2}{=} P(Z_1^2 \leq 1) P(Z_2^2 \leq 1) \\ &= P(-1 \leq Z_1 \leq 1) P(-1 \leq Z_2 \leq 1). \end{aligned}$$

Because $Z_1 \sim \mathcal{N}(0,1)$, the probability $P(-1 \leq Z_1 \leq 1)$ is the area under the standard normal pdf between ± 1; i.e.,

$$P(-1 \leq Z_1 \leq 1) = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = F_Z(1) - F_Z(-1),$$

where F_Z is the $\mathcal{N}(0,1)$ cdf. We can calculate this probability using R:

```
> pnorm(1,0,1)-pnorm(-1,0,1)
[1] 0.6826895
```

The calculation for $P(-1 \leq Z_2 \leq 1)$ is identical. Therefore,

$$P(Z_1^2 \leq 1, Z_2^2 \leq 1) = P(-1 \leq Z_1 \leq 1) P(-1 \leq Z_2 \leq 1) \approx (0.683)^2 \approx 0.466.$$

(b) We have to work with the joint pdf of Z_1 and Z_2. Because Z_1 and Z_2 are independent, we have

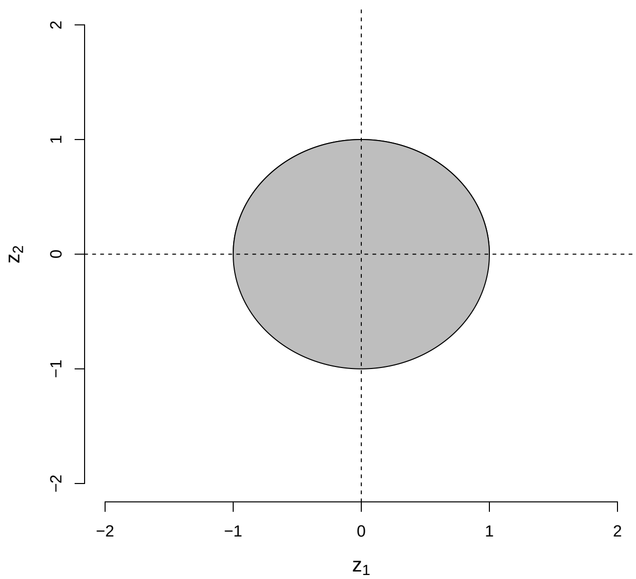
$$f_{Z_1,Z_2}(z_1, z_2) = f_{Z_1}(z_1) f_{Z_2}(z_2) = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-z_1^2/2}}_{\mathcal{N}(0,1) \text{ pdf}} \times \underbrace{\frac{1}{\sqrt{2\pi}} e^{-z_2^2/2}}_{\mathcal{N}(0,1) \text{ pdf}} = \frac{1}{2\pi} e^{-(z_1^2+z_2^2)/2}.$$

Summarizing, the joint pdf of Z_1 and Z_2 is given by

$$f_{Z_1,Z_2}(z_1, z_2) = \begin{cases} \frac{1}{2\pi} e^{-(z_1^2+z_2^2)/2}, & \begin{array}{l} -\infty < z_1 < \infty \\ -\infty < z_2 < \infty \end{array} \\ 0, & \text{otherwise.} \end{cases}$$

To calculate $P(Z_1^2 + Z_2^2 \leq 1)$, we need to integrate $f_{Z_1,Z_2}(z_1, z_2)$ over the set

$$B = \left\{ (z_1, z_2) : z_1^2 + z_2^2 \leq 1 \right\}.$$



This set is shown in the graph above. Note that the boundary of this set, that is, $z_1^2 + z_2^2 = 1$ is unit circle centered at the origin $(0, 0)$. We have

$$P(Z_1^2 + Z_2^2 \leq 1) = \int\int_B \frac{1}{2\pi} e^{-(z_1^2+z_2^2)/2} dz_1 dz_2.$$

Switch to polar coordinates. Let

$$\begin{aligned} z_1 &= r\cos\theta \\ z_2 &= r\sin\theta \end{aligned}$$

so that

$$z_1^2 + z_2^2 = r^2\cos^2\theta + r^2\sin^2\theta = r^2(\cos^2\theta + \sin^2\theta) = r^2$$

and $dz_1 dz_2 = r dr d\theta$. We have

$$\begin{aligned} P(Z_1^2 + Z_2^2 \leq 1) = \frac{1}{2\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{1} e^{-r^2/2}\ r dr d\theta &= \frac{1}{2\pi}\int_{\theta=0}^{2\pi}\left(\int_{r=0}^{1} re^{-r^2/2}dr\right)d\theta \\ &= \frac{1}{2\pi}\int_{\theta=0}^{2\pi}\left(-e^{-r^2/2}\Big|_{r=0}^{1}\right)d\theta \\ &= \frac{1}{2\pi}\int_{\theta=0}^{2\pi}(1 - e^{-1/2})d\theta \\ &= \frac{1 - e^{-1/2}}{2\pi}\int_{\theta=0}^{2\pi} d\theta \\ &= \frac{1 - e^{-1/2}}{2\pi} \times 2\pi = 1 - e^{-1/2} \approx 0.393. \end{aligned}$$

(c) If Z_1 and Z_2 are independent, then

$$V(aZ_1 + (1-a)Z_2) = a^2V(Z_1) + (1-a)^2V(Z_2) = a^2 + (1-a)^2,$$

because $V(Z_1) = V(Z_2) = 1$. Therefore, think of this as a function of a, say,

$$f(a) = a^2 + (1-a)^2$$

and now find the value of a that minimizes $f(a)$. Taking derivatives, we have

$$f'(a) = 2a + 2(1-a)(-1) = 4a - 2 \overset{\text{set}}{=} 0 \implies a = \frac{1}{2}.$$

Note that

$$f''(a) = 4 > 0$$

verifying that $a = 1/2$ is a minimizer by the second derivative test; i.e., the function $f(a)$ is concave up at $a = 1/2$. When Z_1 and Z_2 are not independent, then

$$V(aZ_1 + (1-a)Z_2) = a^2 V(Z_1) + (1-a)^2 V(Z_2) + 2a(1-a)\text{Cov}(Z_1, Z_2).$$

Note that $V(Z_1) = V(Z_2) = 1$ and

$$\rho = \frac{\text{Cov}(Z_1, Z_2)}{\sqrt{V(Z_1)}\sqrt{V(Z_2)}} = \text{Cov}(Z_1, Z_2).$$

Therefore,

$$V(aZ_1 + (1-a)Z_2) = a^2 + (1-a)^2 + 2a(1-a)\rho.$$

Therefore, think of this as a function of a, say,

$$g(a) = a^2 + (1-a)^2 + 2a(1-a)\rho$$

and now find the value of a that minimizes $g(a)$. Taking derivatives, we have

$$\begin{aligned} g'(a) &= 2a + 2(1-a)(-1) + (2-4a)\rho \\ &= 4a - 2 + 2\rho - 4a\rho \\ &= 4a(1-\rho) - 2(1-\rho) \\ &= (4a-2)(1-\rho). \end{aligned}$$

Provided $\rho \neq 1$, then setting $g'(a) = 0$ and solving for a gives $a = 1/2$ again. Note this solution minimizes $g(a)$ because

$$g''(a) = 4(1-\rho) > 0$$

as $1 - \rho$ is strictly positive. Note that if $\rho = 1$, then

$$V(aZ_1 + (1-a)Z_2) = a^2 + (1-a)^2 + 2a(1-a) = 1.$$

which is free of a; i.e., *all* linear combinations of the form $aZ_1 + (1-a)Z_2$ have the same variance. Interesting!

5. (a) First note that

$$\begin{aligned} [Y - (\beta_0 + \beta_1 X)]^2 &= (Y - \beta_0 - \beta_1 X)(Y - \beta_0 - \beta_1 X) \\ &= Y^2 - 2\beta_0 Y - 2\beta_1 XY + 2\beta_0\beta_1 X + \beta_1^2 X^2 + \beta_0^2. \end{aligned}$$

Therefore,

$$\begin{aligned} Q(\beta_0,\beta_1) = E\{[Y-(\beta_0+\beta_1 X)]^2\} &= E(Y^2 - 2\beta_0 Y - 2\beta_1 XY + 2\beta_0\beta_1 X + \beta_1^2 X^2 + \beta_0^2) \\ &= E(Y^2) - 2\beta_0 E(Y) - 2\beta_1 E(XY) + 2\beta_0\beta_1 E(X) + \beta_1^2 E(X^2) + \beta_0^2. \end{aligned}$$

To minimize $Q(\beta_0,\beta_1)$, we can take the partial derivatives and set them equal to zero; i.e.,

$$\begin{aligned} \frac{\partial}{\partial \beta_0} Q(\beta_0,\beta_1) &= -2E(Y) + 2\beta_1 E(X) + 2\beta_0 \overset{\text{set}}{=} 0 \\ \frac{\partial}{\partial \beta_1} Q(\beta_0,\beta_1) &= -2E(XY) + 2\beta_0 E(X) + 2\beta_1 E(X^2) \overset{\text{set}}{=} 0. \end{aligned}$$

Solving the first equation gives $\beta_0 = E(Y) - \beta_1 E(X)$ as desired. Now, for the second equation, we have

$$\begin{aligned} & -2E(XY) + 2[E(Y) - \beta_1 E(X)]E(X) + 2\beta_1 E(X^2) \overset{\text{set}}{=} 0 \\ & \quad \Longrightarrow -E(XY) + E(X)E(Y) - \beta_1 [E(X)]^2 + \beta_1 E(X^2) = 0 \\ & \quad \Longrightarrow \beta_1\{E(X^2) - [E(X)]^2\} = E(XY) - E(X)E(Y) \Longrightarrow \beta_1 = \frac{\text{Cov}(X,Y)}{\sigma_X^2}. \end{aligned}$$

Now note that

$$\beta_1 = \frac{\text{Cov}(X,Y)}{\sigma_X^2} = \frac{\rho\sigma_X\sigma_Y}{\sigma_X^2} = \rho\left(\frac{\sigma_Y}{\sigma_X}\right)$$

as claimed.

Note: To verify that the solution

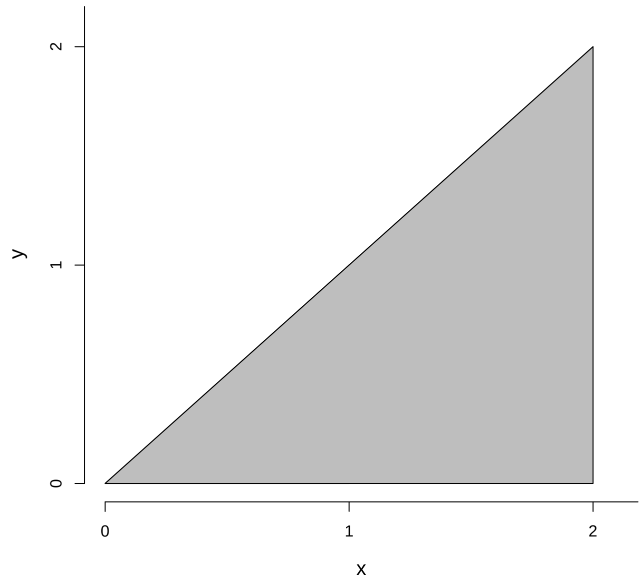
$$\begin{aligned} \beta_0 &= E(Y) - \beta_1 E(X) \\ \beta_1 &= \rho\left(\frac{\sigma_Y}{\sigma_X}\right) \end{aligned}$$

minimizes $Q(\beta_0,\beta_1)$, it suffices to show that $\mathbf{a}'\mathbf{H}\mathbf{a} > 0$ for all $\mathbf{a} \in \mathbb{R}^2$ ($\mathbf{a} \neq \mathbf{0}$), where $\mathbf{H}$ is the Hessian matrix of $Q(\beta_0,\beta_1)$; i.e.,

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2}{\partial \beta_0^2} Q(\beta_0,\beta_1) & \frac{\partial^2}{\partial \beta_0 \partial \beta_1} Q(\beta_0,\beta_1) \\ \frac{\partial^2}{\partial \beta_1 \partial \beta_0} Q(\beta_0,\beta_1) & \frac{\partial^2}{\partial \beta_1^2} Q(\beta_0,\beta_1) \end{pmatrix} = \begin{pmatrix} 2 & 2E(X) \\ 2E(X) & 2E(X^2) \end{pmatrix}.$$

Any (square) matrix $\mathbf{H}$ that satisfies $\mathbf{a}'\mathbf{H}\mathbf{a} > 0$ for all $\mathbf{a} \neq \mathbf{0}$ is said to be *positive definite*. In multivariable minimization problems, checking this condition for the Hessian (when evaluated at the first order critical point) is analogous to using the second derivative test in univariate minimization problems (such as in Problem 4). If we did not check this condition, then we would not know for sure if the solution for β_0 and β_1 above is a minimizer, a maximizer, or a so-called "saddle point." Let $\mathbf{a} = (a_1, a_2)'$. Suppose $\mathbf{a} \neq \mathbf{0} = (0,0)'$. We have

$$\begin{aligned} \mathbf{a}'\mathbf{H}\mathbf{a} &= \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} 2 & 2E(X) \\ 2E(X) & 2E(X^2) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\ &= 2\begin{pmatrix} a_1 + a_2 E(X) & a_1 E(X) + a_2 E(X^2) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 2\left[a_1^2 + 2a_1a_2E(X) + a_2^2E(X^2)\right]. \end{aligned}$$



It suffices to show the last quantity is > 0. Add and subtract $a_2^2[E(X)]^2$. We have

$$\mathbf{a}'\mathbf{H}\mathbf{a} = 2\left\{a_1^2 + 2a_1a_2E(X) + a_2^2[E(X)]^2 + a_2^2E(X^2) - a_2^2[E(X)]^2\right\}$$
$$= 2\left\{[a_1 + a_2E(X)]^2 + a_2^2V(X)\right\} > 0.$$

(b) A graph of the support

$$R = \{(x, y) : 0 < x < 2,\ 0 < y < x\}$$

is shown at the top of this page. The equation of the diagonal line (boundary of the support) is $y = x$. Because we need to get the (marginal) means and variances of X and Y, let's get the marginal distributions. For $0 < x < 2$,

$$f_X(x) = \int_{y=0}^{x} \frac{xy}{2}\, dy \;=\; \frac{x}{2}\left(\frac{y^2}{2}\Big|_{y=0}^{x}\right) \;=\; \frac{x}{2} \times \frac{x^2}{2} = \frac{x^3}{4}.$$

For $0 < y < 2$,

$$f_Y(y) = \int_{x=y}^{2} \frac{xy}{2}\, dx \;=\; \frac{y}{2}\left(\frac{x^2}{2}\Big|_{x=y}^{2}\right) \;=\; \frac{y}{2}\left(2 - \frac{y^2}{2}\right) = \frac{4y - y^3}{4}.$$

Summarizing, we have

$$f_X(x) = \begin{cases} \dfrac{x^3}{4}, & 0 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_Y(y) = \begin{cases} \dfrac{4y - y^3}{4}, & 0 < y < 2 \\ 0, & \text{otherwise.} \end{cases}$$

It is straightforward to verify each of these (marginal) pdfs integrate to 1.

Let's get the mean and variance of X. We have

$$\begin{aligned} E(X) = \int_{\mathbb{R}} x f_X(x)dx &= \int_0^2 \frac{x^4}{4}\ dx = \frac{1}{4}\left(\frac{x^5}{5}\bigg|_0^2\right) = \frac{1}{4}\left(\frac{32}{5}\right) = \frac{8}{5} \\ E(X^2) = \int_{\mathbb{R}} x^2 f_X(x)dx &= \int_0^2 \frac{x^5}{4}\ dx = \frac{1}{4}\left(\frac{x^6}{6}\bigg|_0^2\right) = \frac{1}{4}\left(\frac{64}{6}\right) = \frac{8}{3}. \end{aligned}$$

Therefore,

$$V(X) = E(X^2) - [E(X)]^2 = \frac{8}{3} - \left(\frac{8}{5}\right)^2 = \frac{8}{75}.$$

Now, let's get the mean and variance of Y. We have

$$\begin{aligned} E(Y) = \int_{\mathbb{R}} y f_Y(y)dy &= \int_0^2 \frac{4y^2 - y^4}{4}\ dx = \frac{1}{4}\left(\frac{4y^3}{3} - \frac{y^5}{5}\bigg|_0^2\right) = \frac{1}{4}\left(\frac{32}{3} - \frac{32}{5}\right) = \frac{16}{15} \\ E(Y^2) = \int_{\mathbb{R}} y^2 f_Y(y)dy &= \int_0^2 \frac{4y^3 - y^5}{4}\ dx = \frac{1}{4}\left(y^4 - \frac{y^6}{6}\bigg|_0^2\right) = \frac{1}{4}\left(16 - \frac{64}{6}\right) = \frac{4}{3}. \end{aligned}$$

Therefore,

$$V(Y) = E(Y^2) - [E(Y)]^2 = \frac{4}{3} - \left(\frac{16}{15}\right)^2 = \frac{44}{225}.$$

Now, let's get the covariance of X and Y. From the covariance computing formula, we have

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = E(XY) - \left(\frac{8}{5}\right)\left(\frac{16}{15}\right).$$

We calculate $E(XY)$ as follows:

$$\begin{aligned} E(XY) = \int_{\mathbb{R}^2} xy f_{X,Y}(x,y)dxdy &= \int_{x=0}^2 \int_{y=0}^x \frac{x^2y^2}{2}\ dydx \\ &= \int_{x=0}^2 \frac{x^2}{2}\left(\frac{y^3}{3}\bigg|_{y=0}^x\right)dx = \frac{1}{6}\int_{x=0}^2 x^5\ dx = \frac{1}{6}\left(\frac{x^6}{6}\bigg|_0^2\right) = \frac{16}{9}. \end{aligned}$$

Therefore,

$$\text{Cov}(X,Y) = \frac{16}{9} - \left(\frac{8}{5}\right)\left(\frac{16}{15}\right) = \frac{48}{675} \implies \rho = \frac{\text{Cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{\frac{48}{675}}{\sqrt{\frac{8}{75}}\sqrt{\frac{44}{225}}} = \frac{4}{\sqrt{66}}.$$

Finally, the intercept and slope of the best linear predictor are

$$\begin{aligned} \beta_0 &= \frac{16}{15} - \frac{2}{3}\left(\frac{8}{5}\right) = 0 \\ \beta_1 &= \rho\left(\frac{\sigma_Y}{\sigma_X}\right) = \frac{4}{\sqrt{66}}\left(\frac{\sqrt{\frac{44}{225}}}{\sqrt{\frac{8}{75}}}\right) = \frac{2}{3}. \end{aligned}$$

This gives us the equation

$$Y = \beta_0 + \beta_1 X = 0 + \left(\frac{2}{3}\right)X = \frac{2X}{3}.$$