1. Use the basic counting rule. There are

$$\begin{aligned} n_1 &= \text{number of ways to choose 5 white balls from 70} = \binom{70}{5} \\ n_2 &= \text{number of ways to choose 1 yellow ball from 25} = \binom{25}{1}. \end{aligned}$$

Therefore, there are

$$n_1 \times n_2 = \binom{70}{5}\binom{25}{1} = 12103014 \times 25 = 302575350$$

possible ways to select 6 numbers in the lottery. There is only 1 way to select all numbers correctly. Assuming all possible ways are equally likely, the probability of winning is

$$\frac{1}{302575350}.$$

2. This problem is very easy if you remember the exponential cdf; i.e.,

$$F_Y(y) = \begin{cases} 0, & y \leq 0 \\ 1 - e^{-y/\beta}, & y > 0. \end{cases}$$

To find the median $\phi_{0.5}$, we set $F_Y(\phi_{0.5}) = P(Y \leq \phi_{0.5}) = 0.5$ and solve for $\phi_{0.5}$. We have

$$\begin{aligned} 0.5 \stackrel{\text{set}}{=} F_Y(\phi_{0.5}) &= 1 - e^{-\phi_{0.5}/\beta} \\ &\Longrightarrow e^{-\phi_{0.5}/\beta} = 0.5 \\ &\Longrightarrow -\frac{\phi_{0.5}}{\beta} = \ln(0.5) = \ln\left(\frac{1}{2}\right) = -\ln 2 \Longrightarrow \phi_{0.5} = \beta \ln 2. \end{aligned}$$

Note: If you did not remember the exponential cdf, then you can integrate the pdf. We have

$$\begin{aligned} 0.5 \stackrel{\text{set}}{=} F_Y(\phi_{0.5}) = P(Y \leq \phi_{0.5}) &= \int_0^{\phi_{0.5}} \frac{1}{\beta} e^{-y/\beta} dy \\ &= \frac{1}{\beta}\left(-\beta e^{-y/\beta}\right)\Big|_0^{\phi_{0.5}} = \left(e^{-y/\beta}\right)\Big|_{\phi_{0.5}}^0 = 1 - e^{-\phi_{0.5}/\beta}. \end{aligned}$$

Then solve for $\phi_{0.5}$ in the same way as above.

3. Write

$$m_Y(t) = \frac{1}{1-t^2} = (1-t^2)^{-1}.$$

This makes it easier to take derivatives. The first derivative of $m_Y(t)$ is

$$\frac{d}{dt} m_Y(t) = (-1)(1-t^2)^{-2}(-2t) = 2t(1-t^2)^{-2}.$$

Therefore,

$$E(Y) = \frac{d}{dt} m_Y(t)\Big|_{t=0} = 2(0)(1-0)^{-2} = 0.$$

The second derivative of $m_Y(t)$ is

$$\frac{d^2}{dt^2}m_Y(t) = \frac{d}{dt}\left[2t(1-t^2)^{-2}\right] = 2(1-t^2)^{-2} + (2t)(-2)(1-t^2)^{-3}(-2t).$$

There is no need simplify this, because all we do is evaluate it at $t = 0$. We have

$$E(Y^2) = \frac{d^2}{dt^2}m_Y(t)\bigg|_{t=0} = 2(1-0)^{-2} + 2(0)(-2)(1-0)^{-3}(-2)(0) = 2.$$

The variance of Y is

$$V(Y) = E(Y^2) - [E(Y)]^2 = 2 - 0 = 2.$$

4. The support

$$R = \{(y_1, y_2) : y_1 > 0, y_2 > 0,\ y_1 + y_2 < 1\}$$

is shown at the top of the next page (left). The joint pdf $f_{Y_1,Y_2}(y_1, y_2)$ is a three-dimensional function which takes the value $24y_1y_2$ over this region and is otherwise equal to zero. Note that the boundary of the support is

$$y_1 + y_2 = 1 \implies y_2 = 1 - y_1,$$

a linear function of y_1 with intercept 1 and slope -1. We can find $P(Y_1 < 0.5)$ by finding the volume under $f_{Y_1,Y_2}(y_1, y_2)$ over the set

$$B = \{(y_1, y_2) : 0 < y_1 < 0.5,\ y_2 > 0,\ y_1 + y_2 < 1\},$$

which is shown at the top of the next page (right). The double integral limits come from this picture. We have

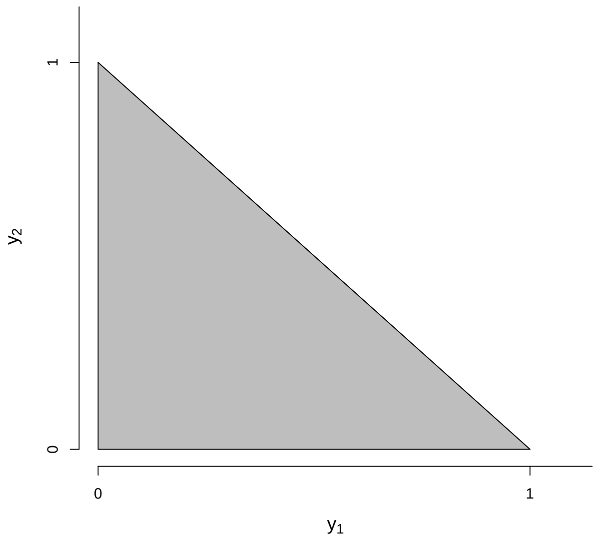
$$\begin{aligned}
P(Y_1 < 0.5) &= \int_{y_1=0}^{0.5}\int_{y_2=0}^{1-y_1} 24y_1y_2\ dy_2dy_1 \\
&= \int_{y_1=0}^{0.5} 24y_1\left[\left(\frac{y_2^2}{2}\right)\bigg|_{y_2=0}^{1-y_1}\right]dy_1 \\
&= \int_{y_1=0}^{0.5} 12y_1(1-y_1)^2\ dy_1 \\
&= 12\int_{y_1=0}^{0.5}(y_1 - 2y_1^2 + y_1^3)\ dy_1 \\
&= 12\left(\frac{y_1^2}{2} - \frac{2y_1^3}{3} + \frac{y_1^4}{4}\right)\bigg|_{y_1=0}^{0.5} = 12\left(\frac{1}{8} - \frac{1}{12} + \frac{1}{64}\right) = 0.6875.
\end{aligned}$$

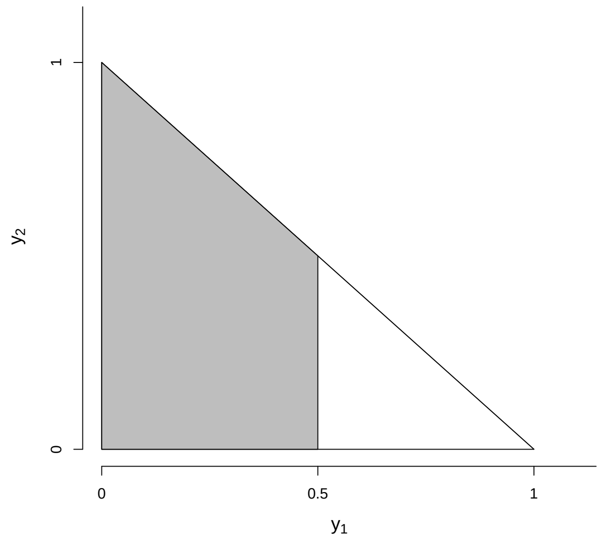
Note: You could also do this problem by finding the marginal pdf of Y_1 and then calculating $P(Y_1 < 0.5)$ as an area under this pdf. For $0 < y_1 < 1$, the marginal pdf of Y_1 is

$$f_{Y_1}(y_1) = \int_{y_2=0}^{1-y_1} 24y_1y_2\ dy_2 = 24y_1\left[\left(\frac{y_2^2}{2}\right)\bigg|_{y_2=0}^{1-y_1}\right] = 12y_1(1-y_1)^2.$$

Summarizing,

$$f_{Y_1}(y_1) = \begin{cases} 12y_1(1-y_1)^2, & 0 < y_1 < 1 \\ 0, & \text{otherwise.} \end{cases}$$





You would then calculate $P(Y_1 < 0.5)$ as

$$\int_{y_1=0}^{0.5} 12y_1(1-y_1)^2 \; dy_1 = 0.6875$$

just like in the previous solution.

5. Let Y denote the number of people arriving for treatment per hour. In the problem, we assume $Y \sim \text{Poisson}(\lambda = 5)$.

(a) The probability at most two people arrive for treatment is

$$\begin{aligned} P(Y \leq 2) &= P(Y=0) + P(Y=1) + P(Y=2) \\ &= \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} \;\; = \; e^{-5}\left(1 + 5 + \frac{25}{2}\right) \approx 0.125. \end{aligned}$$

(b) The time until the second person arrives, T, follows a gamma distribution with $\alpha = 2$ and

$$\beta = \frac{1}{\lambda} = \frac{1}{5}.$$

The variance of T is

$$\sigma^2 = 2\left(\frac{1}{5}\right)^2 \implies \sigma = \sqrt{2\left(\frac{1}{5}\right)^2} = \frac{\sqrt{2}}{5} \approx 0.283.$$

6. Let Y denote the number of events that occur (out of n) so that $Y \sim b(n,p)$. We want to calculate $P(Y = y|Y \geq 1)$. Using the definition of conditional probability, we have

$$P(Y = y|Y \geq 1) = \frac{P(Y = y \text{ and } Y \geq 1)}{P(Y \geq 1)}.$$

The denominator is

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \binom{n}{0} p^0 (1-p)^n = 1 - (1-p)^n.$$

- When $y = 0$, the numerator is 0 because $\{Y = 0 \text{ and } Y \geq 1\} = \emptyset$, the null set.
- When $y = 1, 2, ..., n$, then $\{Y = y\} \subseteq \{Y \geq 1\}$, so $\{Y = y \text{ and } Y \geq 1\} = \{Y = y\}$. Therefore, the numerator is

$$P(Y = y \text{ and } Y \geq 1) = P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}.$$

Summarizing,

$$P(Y = y|Y \geq 1) = \begin{cases} 0, & y = 0 \\ \dfrac{\binom{n}{y} p^y (1-p)^{n-y}}{1-(1-p)^n}, & y = 1, 2, ..., n. \end{cases}$$

7. The cdf $F_Y(y) = P(Y \leq y)$ is defined for all $y \in \mathbb{R}$. Therefore, we consider three cases:

Case 1: When $y \leq 0$,

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \int_{-\infty}^{y} 0dt = 0.$$

Case 2: When $0 < y < 2$,

$$\begin{aligned} F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt &= \int_{-\infty}^{0} 0dt + \int_0^y \frac{t}{6}(1+t^2)dt \\ &= \frac{1}{6}\int_0^y (t+t^3)dt \\ &= \frac{1}{6}\left(\frac{t^2}{2} + \frac{t^4}{4}\right)\Big|_0^y = \frac{1}{6}\left(\frac{y^2}{2} + \frac{y^4}{4}\right) = \frac{y^2}{24}(2+y^2). \end{aligned}$$

Case 3: When $y \geq 2$,

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \int_{-\infty}^{0} 0dt + \underbrace{\int_0^2 \frac{t}{6}(1+t^2)dt}_{=1} + \int_2^y 0dt = 1.$$

Summarizing,

$$F_Y(y) = \begin{cases} 0, & y \leq 0 \\ \dfrac{y^2}{24}(2+y^2), & 0 < y < 2 \\ 1, & y \geq 2. \end{cases}$$

8. The marginal pmfs are

y_1	1	2	3
$p_{Y_1}(y_1)$	0.25	0.45	0.30

and

y_2	1	2	3
$p_{Y_2}(y_2)$	0.30	0.40	0.30

The marginal means are

$$E(Y_1) = 1(0.25) + 2(0.45) + 3(0.30) = 2.05$$

and

$$E(Y_2) = 1(0.30) + 2(0.40) + 3(0.30) = 2.$$

We now calculate

$$\begin{aligned} E(Y_1Y_2) &= 1(1)(0.10) + 2(1)(0.15) + 3(1)(0.05) + 1(2)(0.10) + 2(2)(0.20) + 3(2)(0.10) \\ &\quad + 1(3)(0.05) + 2(3)(0.10) + 3(3)(0.15) = 4.25. \end{aligned}$$

Using the covariance computing formula, we have

$$\text{Cov}(Y_1, Y_2) = E(Y_1Y_2) - E(Y_1)E(Y_2) = 4.25 - 2.05(2) = 0.15.$$

9. Note that

$$\left(aY - \frac{1}{a}\right)^2 = a^2Y^2 - 2Y + \frac{1}{a^2}.$$

Now take the expectation of this:

$$Q(a) = E\left[\left(aY - \frac{1}{a}\right)^2\right] = a^2E(Y^2) - 2E(Y) + \frac{1}{a^2}.$$

To minimize $Q(a)$, we take its derivative, set it equal to zero, and solve for a. We have

$$\frac{d}{da}Q(a) = \frac{d}{da}\left[a^2E(Y^2) - 2E(Y) + \frac{1}{a^2}\right] = 2aE(Y^2) - \frac{2}{a^3}.$$

Setting this equal to zero, we have

$$2aE(Y^2) - \frac{2}{a^3} \stackrel{\text{set}}{=} 0 \implies 2aE(Y^2) = \frac{2}{a^3} \implies a^4E(Y^2) = 1 \implies a = \left[\frac{1}{E(Y^2)}\right]^{1/4}.$$

From the variance computing formula, we have

$$E(Y^2) = V(Y) + [E(Y)]^2 = \sigma^2 + \mu^2.$$

Therefore,

$$a = \left(\frac{1}{\sigma^2 + \mu^2}\right)^{1/4}.$$

To verify this solution minimizes $Q(a)$, note that

$$\frac{d^2}{da^2}Q(a) = 2E(Y^2) + \frac{6}{a^4} > 0.$$

Therefore, the solution $a = (\sigma^2 + \mu^2)^{-1/4}$ is a minimizer by the second derivative test.

10. Treating each candidate as a distinct "object," there are

$$N = 9! = 362880$$

possible permutations. We can think of one outcome (i.e., ordering of the candidates) in the underlying sample space S as having the following structure:

$$(\_\ \_\ \_\ \_\ \_\ \_\ \_\ \_\ \_).$$

Define the event

$$A = \{\text{candidates within each party grouped together}\}.$$

The number of outcomes in A can be found using the basic counting rule with

$$\begin{aligned}
n_1 &= \text{number of ways to permute R, D, and I ordering} = 3! \\
n_2 &= \text{number of ways to permute R candidates} = 4! \\
n_3 &= \text{number of ways to permute D candidates} = 3! \\
n_4 &= \text{number of ways to permute I candidates} = 2!
\end{aligned}$$

The number of outcomes in A is

$$n_a = 3! \times 4! \times 3! \times 2! = 1728.$$

Therefore,

$$P(A) = \frac{n_a}{N} = \frac{1728}{362880} \approx 0.0048.$$

Note: You could also treat candidates within each party as indistinguishable; e.g., you cannot tell one R candidate from another, etc. Then there would be

$$N = \binom{9}{4\ 3\ 2} = 1260$$

outcomes in S. The number of outcomes in A is

$$n_a = n_1 = \text{number of ways to permute R, D, and I ordering} = 3! = 6.$$

The steps to calculate n_2, n_3, and n_4 in the other solution are now vacuous because you cannot tell candidates in the same party from each other. Therefore,

$$P(A) = \frac{n_a}{N} = \frac{6}{1260} \approx 0.0048.$$

11. For $y \geq 30$, the pdf of Y is

$$f_Y(y) = \frac{d}{dy}\, F_Y(y) = \frac{d}{dy}\left[1 - e^{-(y-30)/5}\right] = 0 - \left(-\frac{1}{5}\right) e^{-(y-30)/5} = \frac{1}{5} e^{-(y-30)/5}.$$

Summarizing,

$$f_Y(y) = \begin{cases} \frac{1}{5} e^{-(y-30)/5}, & y \geq 30 \\ 0, & \text{otherwise.} \end{cases}$$

The mean of Y is

$$E(Y) = \int_{\mathbb{R}} y f_Y(y)dy = \int_{30}^{\infty} \frac{y}{5} e^{-(y-30)/5} dy.$$

In the last integral, let

$$u = y - 30 \implies du = dy.$$

Note that with this substitution, the limits change as well; i.e.,

$$y : 30 \to \infty \implies u : 0 \to \infty.$$

Therefore,

$$E(Y) = \int_{30}^{\infty} \frac{y}{5} e^{-(y-30)/5} dy = \int_{0}^{\infty} (u + 30)\ \frac{1}{5} e^{-u/5} du.$$

Note that $\frac{1}{5}e^{-u/5}$ is the exponential pdf with mean $\beta = 5$ and we are integrating over $(0, \infty)$. Therefore, the last integral equals

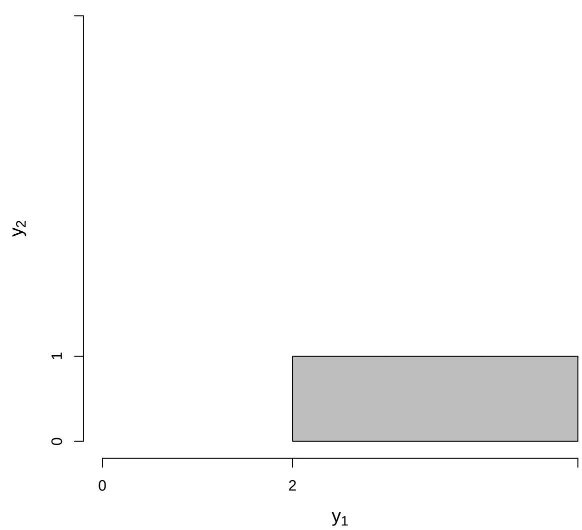
$$E(U + 30), \quad \text{where } U \sim \text{exponential}(5).$$

We have $E(U + 30) = E(U) + 30 = 5 + 30 = 35$ minutes.

12. The support

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

is shown below. The joint pdf $f_{Y_1,Y_2}(y_1, y_2)$ is a three-dimensional function which takes the value $16y_2/y_1^3$ over this region and is otherwise equal to zero.



The support is rectangular and does not involve a constraint between y_1 and y_2. Note that we can write

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{16y_2}{y_1^3} = \frac{16}{y_1^3} \times y_2 = g(y_1)h(y_2),$$

where $g(y_1) = 16y_1^{-3}$ and $h(y_2) = y_2$ are both are nonnegative functions. Because $f_{Y_1,Y_2}(y_1, y_2)$ can be written in this way, we know Y_1 and Y_2 must be independent.

Note: You could also derive the marginal pdfs $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$ and show that

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

For $y_1 > 2$, the marginal pdf of Y_1 is

$$f_{Y_1}(y_1) = \int_{y_2=0}^{1} \frac{16y_2}{y_1^3}\, dy_2 = \frac{16}{y_1^3}\left[\left(\frac{y_2^2}{2}\right)\Big|_{y_2=0}^{1}\right] = \frac{8}{y_1^3}.$$

For $0 < y_2 < 1$, the marginal pdf of Y_2 is

$$f_{Y_2}(y_2) = \int_{y_1=2}^{\infty} \frac{16y_2}{y_1^3}\, dy_1 = 16y_2\left[\left(-\frac{1}{2y_2^2}\right)\Big|_{y_1=2}^{\infty}\right] = 16y_2\left(0+\frac{1}{8}\right) = 2y_2.$$

Summarizing,

$$f_{Y_1}(y_1) = \begin{cases} \dfrac{8}{y_1^3}, & y_1 > 2 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad \begin{cases} 2y_2, & 0 < y_2 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Note that

$$\frac{16y_2}{y_1^3} = f_{Y_1,Y_2}(y_1,y_2) = f_{Y_1}(y_1)f_{Y_2}(y_2) = \frac{8}{y_1^3} \times 2y_2.$$

13. Define the two events

$$\begin{aligned} A &= \{\text{resident smoked}\} \\ D &= \{\text{lung cancer death}\}. \end{aligned}$$

The first bullet says $P(A) = 0.2$. The second bullet says $P(D) = 0.006$. The third bullet says $P(D|A) = 10P(D|\overline{A})$.

We want $P(A|D)$. Use Bayes' Rule:

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)} = \frac{P(D|A)(0.2)}{0.006}.$$

We get $P(D|A)$ by using LOTP; i.e.,

$$\begin{aligned} 0.006 = P(D) &= P(D|A)P(A) + P(D|\overline{A})P(\overline{A}) \\ &= P(D|A)(0.2) + 0.1P(D|A)(1-0.2) \\ &= P(D|A)[0.2 + 0.1(0.8)] \\ \Longrightarrow & \ P(D|A) = \frac{0.006}{0.28}. \end{aligned}$$

Therefore,

$$P(A|D) = \frac{0.2}{0.006} \times \frac{0.006}{0.28} = \frac{5}{7}.$$

14. We have the hierarchical model:

$$\begin{aligned} Y_2|Y_1 = y_1 &\sim \text{exponential}(2y_1) \\ Y_1 &\sim \mathcal{U}(0,5). \end{aligned}$$

To find $V(Y_2)$, we will use Adam's Rule; i.e.,

$$V(Y_2) = E[V(Y_2|Y_1)] + V[E(Y_2|Y_1)].$$

The conditional distribution of $Y_2|Y_1 = y_1 \sim \text{exponential}(2y_1)$. Therefore, remembering what we know about the mean and variance of the exponential distribution,

$$E(Y_2|Y_1 = y_1) = 2y_1 \implies E(Y_2|Y_1) = 2Y_1$$

and

$$V(Y_2|Y_1 = y_1) = (2y_1)^2 = 4y_1^2 \implies V(Y_2|Y_1) = 4Y_1^2.$$

Therefore,

$$V(Y_2) = E(4Y_1^2) + V(2Y_1) = 4E(Y_1^2) + 4V(Y_1),$$

where $Y_1 \sim \mathcal{U}(0,5)$. The mean of $Y_1 \sim \mathcal{U}(0,5)$ is $E(Y_1) = \frac{5}{2}$, the midpoint of 0 and 5. The second moment of Y_1 is

$$E(Y_1^2) = \int_0^5 \frac{y_1^2}{5} dy_1 = \frac{1}{5}\left(\frac{y_1^3}{3}\right)\Big|_0^5 = \frac{25}{3}.$$

Therefore,

$$V(Y_1) = E(Y_1^2) - [E(Y_1)]^2 = \frac{25}{3} - \left(\frac{5}{2}\right)^2 = \frac{25}{12}.$$

Finally,

$$V(Y_2) = 4E(Y_1^2) + 4V(Y_1) = 4\left(\frac{25}{3}\right) + 4\left(\frac{25}{12}\right) = \frac{125}{3} \approx 41.67.$$

15. The first moment is the mean $E(Y)$. Using the definition of mathematical expectation, we have

$$E(Y) = \sum_{y=1}^{\infty} y p_Y(y) = \sum_{y=1}^{\infty} \frac{y}{y \ln 2}\left(\frac{1}{2}\right)^y = \frac{1}{\ln 2} \underbrace{\sum_{y=1}^{\infty}\left(\frac{1}{2}\right)^y}_{=\ 1,\text{ see below}}.$$

We recognize

$$\sum_{y=1}^{\infty}\left(\frac{1}{2}\right)^y$$

as an infinite geometric sum with common ratio $r = 1/2$. Therefore,

$$\sum_{y=1}^{\infty}\left(\frac{1}{2}\right)^y = \left[\sum_{y=0}^{\infty}\left(\frac{1}{2}\right)^y\right] - 1 = \left(\frac{1}{1-\frac{1}{2}}\right) - 1 = 2 - 1 = 1.$$

Therefore,

$$E(Y) = \frac{1}{\ln 2}.$$