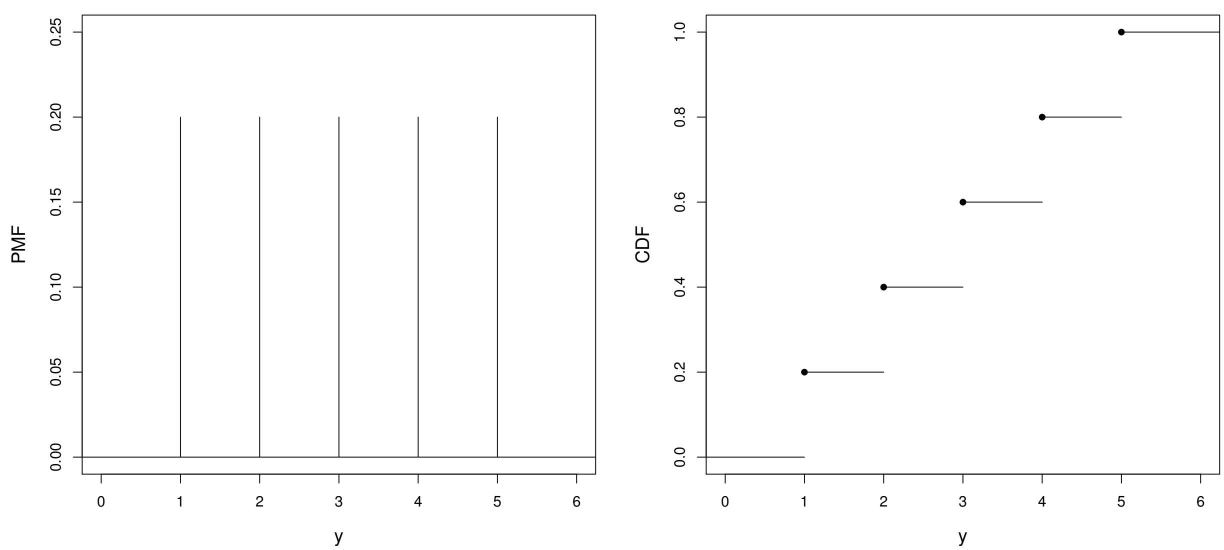
4.2. Note that Y is a discrete random variable with support $R = \{1, 2, 3, 4, 5\}$.
(a) Only 1 of the 5 keys work; it can be selected on the first try, the second, the third, and so on. Because useless keys are discarded (and because keys are selected at random), the probability mass function of Y is simply

y	1	2	3	4	5
$p_Y(y)$	0.2	0.2	0.2	0.2	0.2

(b) The cumulative distribution function (cdf) of Y is

$$F_Y(y) = \begin{cases} 0, & y < 1 \\ 0.2, & 1 \leq y < 2 \\ 0.4, & 2 \leq y < 3 \\ 0.6, & 3 \leq y < 4 \\ 0.8, & 4 \leq y < 5 \\ 1, & y \geq 5. \end{cases}$$

Here are the pmf and cdf of Y graphed side by side (like in the notes):



Here is the R code I used to make these graphs:

```
y = c(1,2,3,4,5)
prob = c(0.2,0.2,0.2,0.2,0.2)
# Plot PMF
plot(y,prob,type="h",xlab="y",ylab="PMF",xlim=c(0,6),ylim=c(0,0.25),xaxt='n',
    cex.lab=1.25)
abline(h=0)
axis(side=1,at=c(0,1,2,3,4,5,6),labels=c("0","1","2","3","4","5","6"))
# Plot CDF
cdf = c(0,cumsum(prob))
cdf.plot = stepfun(y,cdf,f=0)
plot.stepfun(cdf.plot,xlab="y",ylab="CDF",verticals=FALSE,do.points=TRUE,
    main="",pch=16,cex.lab=1.25)
```

(c) We have

$$\begin{aligned} P(Y<3) &= P(Y=1)+P(Y=2)=0.2+0.2=0.4 \\ P(Y\leq 3) &= P(Y=1)+P(Y=2)+P(Y=3)=0.2+0.2+0.2=0.6 \\ P(Y=3) &= 0.2 \end{aligned}$$

(d) No, Y is a discrete random variable−not continuous.

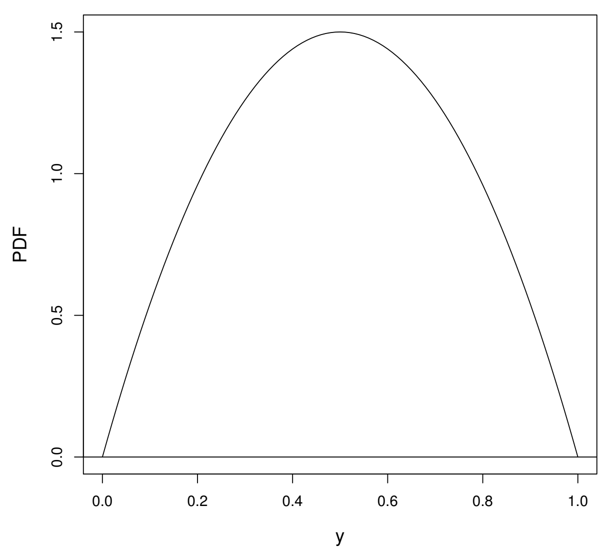
4.8. Note that Y is a continuous random variable with support $R=\{y: 0\leq y\leq 1\}$.
(a) We know that

$$1=\int_{\mathbb{R}} f_Y(y)dy=\int_0^1 ky(1-y)dy=k\left(\frac{y^2}{2}-\frac{y^3}{3}\right)\bigg|_{y=0}^{1}=k\left(\frac{1}{2}-\frac{1}{3}\right)=\frac{1}{6}k.$$

Therefore, $k=6$ and the pdf of Y is

$$f_Y(y)=\begin{cases} 6y(1-y), & 0\leq y\leq 1 \\ 0, & \text{otherwise.} \end{cases}$$

Here is a graph of the pdf that I made in R:




```
# Plot PDF
y = seq(0,1,0.01)
pdf = 6*y*(1-y)
plot(y,pdf,type="l",xlab="y",ylab="PDF",xlim=c(0,1),cex.lab=1.25)
abline(h=0)
```

(b) We have

$$P(0.4\leq Y\leq 1)=\int_{0.4}^1 6y(1-y)dy=6\left(\frac{y^2}{2}-\frac{y^3}{3}\right)\bigg|_{y=0.4}^{1}=6\left[\left(\frac{1}{2}-\frac{1}{3}\right)-\left(\frac{0.4^2}{2}-\frac{0.4^3}{3}\right)\right]=0.648.$$

In R,

```
> integrand <- function(y){6*y*(1-y)}
> integrate(integrand,lower=0.4,upper=1)
0.648 with absolute error < 7.2e-15
```

(c) Because Y is continuous, $P(0.4 \leq Y \leq 1) = P(0.4 \leq Y < 1)$.
(d) Recall the definition of conditional probability

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

Here, we note that $A = \{Y \leq 0.4\}$ and $B = \{Y \leq 0.8\}$ and write

$$P(Y \leq 0.4|Y \leq 0.8) = \frac{P(Y \leq 0.4 \text{ and } Y \leq 0.8)}{P(Y \leq 0.8)}.$$

Note that $A = \{Y \leq 0.4\} \subset B = \{Y \leq 0.8\}$. Therefore, $A \cap B = A = \{Y \leq 0.4\}$ and

$$P(Y \leq 0.4|Y \leq 0.8) = \frac{P(Y \leq 0.4)}{P(Y \leq 0.8)}.$$

Let's calculate both of these probabilities:

$$\begin{aligned} P(Y \leq 0.4) &= \int_0^{0.4} 6y(1-y)dy = 6\left(\frac{y^2}{2} - \frac{y^3}{3}\right)\Bigg|_{y=0}^{0.4} = 6\left(\frac{0.4^2}{2} - \frac{0.4^3}{3}\right) = 0.352 \\ P(Y \leq 0.8) &= \int_0^{0.8} 6y(1-y)dy = 6\left(\frac{y^2}{2} - \frac{y^3}{3}\right)\Bigg|_{y=0}^{0.8} = 6\left(\frac{0.8^2}{2} - \frac{0.8^3}{3}\right) = 0.896. \end{aligned}$$

Therefore,

$$P(Y \leq 0.4|Y \leq 0.8) = \frac{P(Y \leq 0.4)}{P(Y \leq 0.8)} = \frac{0.352}{0.896} \approx 0.393.$$

(e) Same as part (d). Because Y is continuous, the endpoints do not matter.

4.15. (a) If b denotes the minimum possible time, we know $b \geq 0$ (time cannot be negative). Also, $y^2 \geq 0$ always. Therefore, $f_Y(y) \geq 0$ for all $y \in \mathbb{R}$. Now we need to show $f_Y(y)$ integrates to 1. Note that

$$\int_{\mathbb{R}} f_Y(y)dy = \int_b^{\infty} \frac{b}{y^2}dy = b\left(-\frac{1}{y}\Bigg|_{y=b}^{\infty}\right) = -b\left(\lim_{y\to\infty} \frac{1}{y} - \frac{1}{b}\right) = -b\left(0 - \frac{1}{b}\right) = 1.$$

Therefore, $f_Y(y)$ is a valid pdf.
(b) We have 2 cases to consider.

Case 1: When $y < b$,

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

Case 2: When $y \geq b$,

$$\begin{aligned} F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \int_{-\infty}^{b} 0dt + \int_{b}^{y} \frac{b}{t^2}dt &= 0 + b\left(-\frac{1}{t}\Big|_{t=b}^{y}\right) \\ &= b\left(\frac{1}{t}\Big|_{t=y}^{b}\right) = b\left(\frac{1}{b} - \frac{1}{y}\right) = 1 - \frac{b}{y}. \end{aligned}$$

Therefore, the cdf of Y is

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

(c) Recall the cdf in part (b) is $F_Y(y) = P(Y \leq y)$ for all $y \in \mathbb{R}$. Therefore, by the complement rule,

$$P(Y > b + c) = 1 - \underbrace{P(Y \leq b + c)}_{=\, F_Y(b+c)} = 1 - F_Y(b + c) = 1 - \left(1 - \frac{b}{b + c}\right) = \frac{b}{b + c}.$$

(d) This is a conditional probability. If $d > c$, then the event $\{Y > b + d\} \subset \{Y > b + c\}$. Therefore,

$$P(Y > b + d|Y > b + c) = \frac{P(Y > b + d \text{ and } Y > b + c)}{P(Y > b + c)} = \frac{P(Y > b + d)}{P(Y > b + c)} = \frac{\frac{b}{b+d}}{\frac{b}{b+c}} = \frac{b + c}{b + d}.$$

4.19. (a) Recall that the pdf

$$f_Y(y) = \frac{d}{dy}F_Y(y),$$

so we need to calculate the derivative of $F_Y(y)$ on each of the four regions where $F_Y(y)$ is defined.

Region 1: $y \leq 0$.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}0 = 0$$

Region 2: $0 < y < 2$.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(\frac{y}{8}\right) = \frac{1}{8}$$

Region 3: $2 \leq y < 4$.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(\frac{y^2}{16}\right) = \frac{y}{8}$$

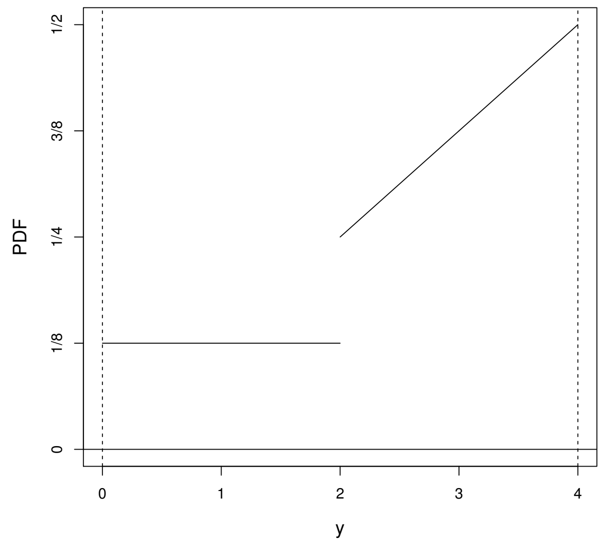
Region 4: $y \geq 4$.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}0 = 0.$$

Therefore, the pdf of Y is

$$f_Y(y) = \begin{cases} \dfrac{1}{8}, & 0 < y < 2 \\ \dfrac{y}{8}, & 2 \leq y < 4 \\ 0, & \text{otherwise.} \end{cases}$$

Here is a graph of the pdf that I made in R. It is easy to show this pdf is valid; i.e., $f_Y(y)$ integrates to 1.




```
y = seq(2,4,0.001)
pdf = y/8
plot(y,pdf,type="l",xlab="y",ylab="PDF",xlim=c(0,4),ylim=c(0,0.5),yaxt="n",
    cex.lab=1.25)
abline(h=0)
abline(v=0,lty=2)
abline(v=4,lty=2)
lines(c(0,2),c(1/8,1/8),lty=1)
axis(2,at=c(0,1/8,1/4,3/8,1/2),labels=c("0","1/8","1/4","3/8","1/2"))
```

(b) Using the pdf of Y, we have

$$\begin{aligned} P(1 \leq Y \leq 3) = \int_1^3 f_Y(y)dy &= \int_1^2 \frac{1}{8}dy + \int_1^3 \frac{y}{8}dy \\ &= \left(\frac{y}{8}\right)\Big|_1^2 + \left(\frac{y^2}{16}\right)\Big|_2^3 = \left(\frac{2}{8} - \frac{1}{8}\right) + \left(\frac{9}{16} - \frac{4}{16}\right) = \frac{7}{16}. \end{aligned}$$

We could also calculate this probability from the cdf of Y:

$$P(1 \leq Y \leq 3) = F_Y(3) - F_Y(1) = \frac{9}{16} - \frac{1}{8} = \frac{7}{16}.$$

(c) Using the pdf of Y, we have

$$P(Y \geq 1.5) = 1 - P(Y < 1.5) = 1 - \int_0^{1.5} f_Y(y)dy = 1 - \int_0^{1.5} \frac{1}{8}dy = 1 - \left(\frac{1.5}{8}\right) = \frac{13}{16}.$$

We could also calculate this probability from the cdf of Y:

$$P(Y \geq 1.5) = 1 - P(Y < 1.5) = 1 - P(Y \leq 1.5) = 1 - F_Y(1.5) = 1 - \left(\frac{1.5}{8}\right) = \frac{13}{16}.$$

(d) This is a conditional probability:

$$P(Y \geq 1|Y \leq 3) = \frac{P(Y \geq 1 \text{ and } Y \leq 3)}{P(Y \leq 3)} = \frac{P(1 \leq Y \leq 3)}{P(Y \leq 3)}.$$

In part (b), we calculated $P(1 \leq Y \leq 3) = \frac{7}{16}$. We can get $P(Y \leq 3)$ using the cdf:

$$P(Y \leq 3) = F_Y(3) = \frac{3^2}{16} = \frac{9}{16}.$$

Therefore,

$$P(Y \geq 1|Y \leq 3) = \frac{P(1 \leq Y \leq 3)}{P(Y \leq 3)} = \frac{\frac{7}{16}}{\frac{9}{16}} = \frac{7}{9}.$$

4.25. In Exercise 4.19, we derived the pdf to be

$$f_Y(y) = \begin{cases} \frac{1}{8}, & 0 < y < 2 \\ \frac{y}{8}, & 2 \leq y < 4 \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, the mean of Y is

$$\begin{aligned} E(Y) = \int_{\mathbb{R}} y f_Y(y) dy &= \int_0^2 \frac{y}{8} dy + \int_2^4 \frac{y^2}{8} dy \\ &= \left(\frac{y^2}{16}\right)\Big|_0^2 + \left(\frac{y^3}{24}\right)\Big|_2^4 = \frac{4}{16} + \left(\frac{64}{24} - \frac{8}{24}\right) \approx 2.583. \end{aligned}$$

Let's get the second moment $E(Y^2)$. We have

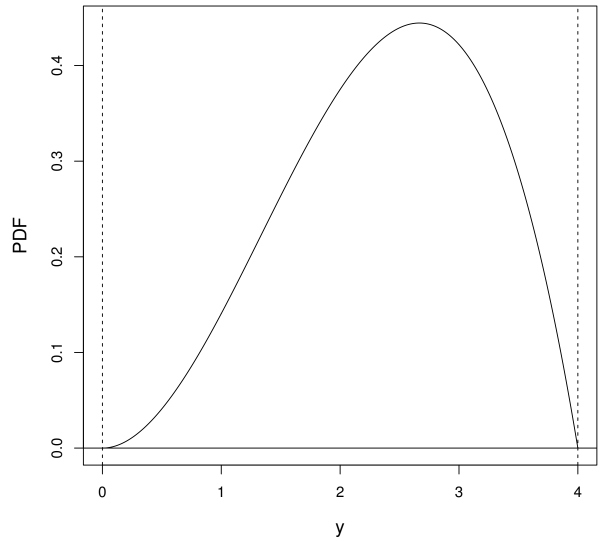
$$\begin{aligned} E(Y^2) = \int_{\mathbb{R}} y^2 f_Y(y) dy &= \int_0^2 \frac{y^2}{8} dy + \int_2^4 \frac{y^3}{8} dy \\ &= \left(\frac{y^3}{24}\right)\Big|_0^2 + \left(\frac{y^4}{32}\right)\Big|_2^4 = \frac{8}{24} + \left(\frac{256}{32} - \frac{16}{32}\right) \approx 7.833. \end{aligned}$$

From the variance computing formula, we have

$$V(Y) = E(Y^2) - [E(Y)]^2 \approx 7.833 - (2.583)^2 \approx 1.161.$$

4.32. Let Y denote the weekly CPU time (in hours). I made a graph of the pdf of Y in R (see next page) using the following code:

```
y = seq(0,4,0.001)
pdf = (3/64)*y^2*(4-y)
plot(y,pdf,type="l",xlab="y",ylab="PDF",xlim=c(0,4),ylim=c(0,max(pdf)),
    cex.lab=1.25)
abline(h=0)
abline(v=0,lty=2)
abline(v=4,lty=2)
```



(a) The mean of Y is

$$\begin{aligned} E(Y) = \int_{\mathbb{R}} y f_Y(y) dy &= \int_0^4 \frac{3}{64} y^3 (4-y) dy = \frac{3}{64} \int_0^4 (4y^3 - y^4) dy \\ &= \frac{3}{64} \left(y^4 - \frac{y^5}{5} \right) \Big|_0^4 = \frac{3}{64} \left(256 - \frac{4^5}{5} \right) = 12 - \frac{48}{5} = 2.4 \text{ hours.} \end{aligned}$$

This is the expected amount of CPU time (in hours) used per week.

Check:

```
> integrand <- function(y) {(3/64)*y^3*(4-y)}
> integrate(integrand,lower=0,upper=4)
2.4 with absolute error < 2.7e-14
```

Let's get the second moment $E(Y^2)$. We have

$$\begin{aligned} E(Y^2) = \int_{\mathbb{R}} y^2 f_Y(y) dy &= \int_0^4 \frac{3}{64} y^4 (4-y) dy = \frac{3}{64} \int_0^4 (4y^4 - y^5) dy \\ &= \frac{3}{64} \left(\frac{4y^5}{5} - \frac{y^6}{6} \right) \Big|_0^4 = \frac{3}{64} \left(\frac{4^6}{5} - \frac{4^6}{6} \right) = 6.4. \end{aligned}$$

Check:

```
> integrand.2 <- function(y) {(3/64)*y^4*(4-y)}
> integrate(integrand.2,lower=0,upper=4)
6.4 with absolute error < 7.1e-14
```

From the variance computing formula, we have

$$V(Y) = E(Y^2) - [E(Y)]^2 = 6.4 - (2.4)^2 = 0.64 \text{ (hours}^2\text{)}.$$

(b) Let $C = 200Y$ denote the cost per week. The expected weekly cost is

$$E(C) = E(200Y) = 200E(Y) = 200(2.4) = 480 \text{ dollars.}$$

The variance is

$$V(C) = V(200Y) = (200)^2 V(Y) = (200)^2(0.64) = 25600 \text{ (dollars)}^2.$$

(c) Let's calculate $P(C > 600)$. We have

$$P(C > 600) = P(200Y > 600) = P(Y > 3) = \int_3^4 \frac{3}{64} y^2(4-y)dy \approx 0.262.$$

I did this integral in R:

```
> integrand.c <- function(y) {(3/64)*y^2*(4-y)}
> integrate(integrand.c,lower=3,upper=4)
0.2617188 with absolute error < 2.9e-15
```

I would expect the weekly cost to exceed 600 dollars roughly 1 out of every 4 weeks.

4.34. This problem is the "continuous version" of Exercise 3.29 (see HW4). I don't understand the hint (I think they want you to use Fubini's Theorem). It is easier to use integration by parts. Suppose Y is a continuous random variable that is also nonnegative (i.e., with support $0 \leq y < \infty$). Suppose Y has pdf $f_Y(y)$ and cdf $F_Y(y)$. Start with

$$\int_0^\infty [1 - F_Y(y)]dy$$

and let

$$\begin{array}{ll} u = 1 - F_Y(y) & du = -f_Y(y)dy \\ dv = dy & v = y. \end{array}$$

With these selections, using integration by parts, we have

$$\begin{aligned} \int_0^\infty [1 - F_Y(y)]dy &= y[1 - F_Y(y)]\Big|_0^\infty - \int_0^\infty y[-f_Y(y)]dy \\ &= y[1 - F_Y(y)]\Big|_0^\infty + \underbrace{\int_0^\infty y f_Y(y)dy}_{=\,E(Y)}. \end{aligned}$$

Therefore, it suffices to show

$$y[1 - F_Y(y)]\Big|_0^\infty = \lim_{y\to\infty} y[1 - F_Y(y)] - 0[1 - F_Y(0)] = \lim_{y\to\infty} y[1 - F_Y(y)] = 0.$$

Recall that one of the validity properties for a cdf $F_Y(y)$ was that $\lim_{y\to\infty} F_Y(y) = 1$. Therefore,

$$\lim_{y\to\infty} [1 - F_Y(y)] = 1 - 1 = 0.$$

Therefore, provided that $1 - F_Y(y)$ converges to 0 "faster" than y diverges to $+\infty$, the result holds.

4.35. I don't know why the authors want you to assume Y is continuous. This statement is true for any random variable Y, so long as $E(Y^2)$ exists. Also, although I understand why the authors gave the hint they did, I think my proof is easier. Define the function

$$h(a) = E[(Y-a)^2] = E(Y^2 - 2aY + a^2) = E(Y^2) - 2aE(Y) + a^2.$$

Note that

$$\frac{d}{da}h(a) = -2E(Y) + 2a \overset{\text{set}}{=} 0 \implies a = E(Y).$$

This argument shows that $a = E(Y)$ is a first-order critical point of $h(a)$. Because $(d^2/da^2)h(a) = 2 > 0$, the critical point $a = E(Y)$ is a minimizer of $h(a)$ by the second-derivative test.

4.37. Suppose Y is a continuous random variable with pdf $f_Y(y)$ that satisfies

$$f_Y(y) = f_Y(-y), \text{ for all } y \in \mathbb{R};$$

i.e., the function $f_Y(y)$ is symmetric about 0, and suppose $E(Y)$ exists. We want to show $E(Y) = 0$. Using the hint, we have

$$E(Y) = \int_{\mathbb{R}} y f_Y(y)dy = \underbrace{\int_{-\infty}^{0} y f_Y(y)dy}_{(*)} + \int_0^\infty y f_Y(y)dy.$$

In the first integral $(*)$, let $w = -y \implies dw = -dy$. Therefore, $(*)$ equals

$$\int_\infty^0 (-w) f_Y(-w)(-dw) = \int_\infty^0 w f_Y(w)dw = -\int_0^\infty w f_Y(w)dw.$$

Therefore, the first integral $(*)$ is the additive inverse of the second integral. Thus, the result.

4.145. The pdf of Y is

$$f_Y(y) = \begin{cases} e^y, & y < 0 \\ 0, & \text{otherwise.} \end{cases}$$

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

$$\begin{aligned} E(e^{3Y/2}) = \int_{-\infty}^0 e^{3y/2} e^y dy &= \int_{-\infty}^0 e^{5y/2} dy \\ &= \frac{2}{5} e^{5y/2}\Big|_{-\infty}^0 = \frac{2}{5}\left(1 - \lim_{y \to -\infty} e^{5y/2}\right) = \frac{2}{5}(1-0) = \frac{2}{5}. \end{aligned}$$

(b) The mgf of Y is

$$\begin{aligned} E(e^{tY}) = \int_{-\infty}^0 e^{ty} e^y dy &= \int_{-\infty}^0 e^{y(t+1)} dy \\ &= \frac{1}{t+1} e^{y(t+1)}\Big|_{-\infty}^0 = \frac{1}{t+1}\left(1 - \lim_{y \to -\infty} e^{y(t+1)}\right). \end{aligned}$$

Now,

$$\lim_{y\to-\infty} e^{y(t+1)} = 0$$

only if $t+1 \geq 0$; i.e, if $t \geq -1$. Otherwise, if $t < -1$, then $t+1 < 0$ and the limit above does not exist. However, note that we cannot have $t = -1$ or else $t+1 = 0$ and $m_Y(t)$ doesn't exist (i.e., cannot divide by zero). Therefore, for $t > -1$, the mgf of Y exists and is equal to

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

Before we do part (c), recall part (a). Note that

$$E(e^{3Y/2}) = E(e^{tY}) = m_Y(t),$$

when $t = 3/2$. Thus,

$$m_Y(3/2) = \frac{1}{(3/2)+1} = \frac{2}{5} \quad \longleftarrow \quad \text{Cool!!}$$

(c) It's probably easiest to use the mgf. The first derivative of $m_Y(t)$ is

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

Therefore,

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

The second derivative of $m_Y(t)$ is

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

Therefore,

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

From the variance computing formula, we have

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