All of Statistics

My proposed solutions to the book:

All of Statistics - A Concise Course in Statistical Inference.

I really needed a proper refresher on statistics.

Let me know if you find any mistakes! I am sure there are plenty. :(

Contents

1 Chapter 2 - Random Variables

Exercises

2.1

Claim: $\mathbb{P}(X = x) = F(x^+) - F(x^-)$. (Discrete)
PROOF. By definition of the CDF:

$$F(x^+) = \lim_{z \downarrow x} F(z) = \lim_{z \downarrow x} \mathbb{P}(X \leqslant z), \qquad F(x^-) = \lim_{y \uparrow x} F(y) = \lim_{y \uparrow x} \mathbb{P}(X \leqslant y)$$

(so $y < x$ and $y \to x$, and $x < z$ and $x \leftarrow z$). By the right continuous property, we can deduce that $z > y$ and we can set $z = x$ and $y = x - 1$.

$$\mathbb{P}(X \leqslant x^+) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X = x) + \mathbb{P}(X \leqslant x - 1), \quad \mathbb{P}(X \leqslant x^-) = \mathbb{P}(X \leqslant x - 1)$$

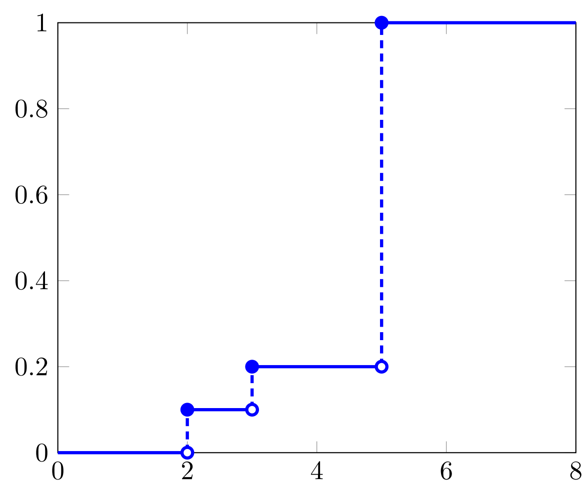
So:

$$\begin{aligned}
\mathbb{P}(X = x) &= \mathbb{P}(X = x) + \mathbb{P}(X \leqslant x - 1) - \mathbb{P}(X \leqslant x - 1) \\
&= \mathbb{P}(X \leqslant x) - \mathbb{P}(X \leqslant x - 1) \\
&= \mathbb{P}(X \leqslant x^+) - \mathbb{P}(X \leqslant x^-) \\
&= F(x^+) - F(x^-)
\end{aligned}$$

□

2.2

Let X be such that $\mathbb{P}(X = 2) = \mathbb{P}(X = 3) = 1/10$ and $\mathbb{P}(X = 5) = 8/10$. Here is a plot of the CDF.



By reading the plot, we can see that:

$$\mathbb{P}(2 < X \leqslant 4.8) = F(4.8) - F(2) = 2/10 - 1/10 = 1/10$$

$$\mathbb{P}(2 \leqslant X \leqslant 4.8) = F(4.8) = 2/10$$

2.3

Lemma 2.15 Let F be the CDF for a random variable X. Then:

1. $\mathbb{P}(X = x) = F(x) - F(x^-)$
2. $\mathbb{P}(x < X \leqslant y) = F(y) - F(x)$
3. $\mathbb{P}(X > x) = 1 - F(x)$
4. If X is continuous, then

$$F(b) - F(a) = \mathbb{P}(a < X < b) = \mathbb{P}(a \leqslant X < b) = \mathbb{P}(a < X \leqslant b) = \mathbb{P}(a \leqslant X \leqslant b)$$

PROOF. We will prove each statement in turn. (1.) was proved in exercise **2.1**. Doing (3) first, since we need it to prove (2).

(3) By definition of complements of sets $A = \{X > x\}$ means $A^c = \{X \leqslant x\}$, and it follows that:

$$\mathbb{P}(X > x) = \mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \mathbb{P}(X \leqslant x) = 1 - F(x).$$

(2) Assume $x < y$. We will need that $\{X > x\} \cup \{X \leqslant y\} = \Omega$, and we will also use Lemma 1.6 (in reverse).

$$\begin{aligned}
\mathbb{P}(x < X \leqslant y) &= \mathbb{P}(\{X > x\} \cap \{X \leqslant y\}) \\
&= \mathbb{P}(X > x) + \mathbb{P}(X \leqslant y) - \mathbb{P}(\{X > x\} \cup \{X \leqslant y\}) \\
&= 1 - F(x) + F(y) - 1 \\
&= F(y) - F(x)
\end{aligned}$$

(4) Similar argument for all cases, so will just do one. We just need to turn the inequalities into strict inequalities. For continuous random variables, pointwise probabilities are 0. Again, we will need to use $\{X > a\} \cup \{X < b\} = \Omega$.

Define $A := \{a \leqslant X\}$ and $B := \{X < b\}$. First, we make the following observation:

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(\{a \leqslant X\}) \\
&= \mathbb{P}(\{a = X\} \cup \{a < X\}) \\
&= \mathbb{P}(\{a = X\}) + \mathbb{P}(\{a < X\}) - \mathbb{P}(\{a = X\} \cap \{a < X\}) \\
&= 0 + \mathbb{P}(A') - 0 \\
&= \mathbb{P}(A')
\end{aligned}$$

where $A' = \{a < X\}$. We get 0 for the pointwise probability, since this is continuous, and we get 0 because the sets are disjoint. We have shown that $\mathbb{P}(A) = \mathbb{P}(A')$ and can use this to conclude the proof.

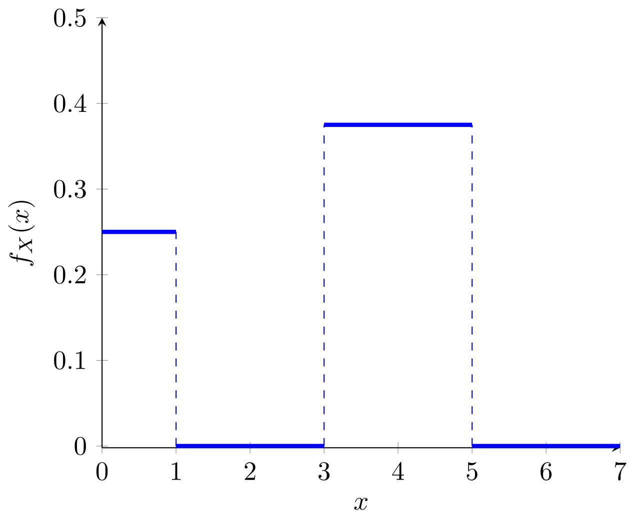
$$\begin{aligned}
\mathbb{P}(a \leqslant X < b) &= \mathbb{P}(\{a \leqslant X\} \cap \{X < b\}) \\
&= \mathbb{P}(A \cap B) \\
&= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(A \cup B) \\
&= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(\Omega) \\
&= \mathbb{P}(A') + \mathbb{P}(B) + \mathbb{P}(A' \cup B) \\
&= \mathbb{P}(A' \cap B) \\
&= \mathbb{P}(a < X < b)
\end{aligned}$$

□

2.4
X has the probability density (PDF):

$$f_X(x) = \begin{cases} 1/4 & 0 < x < 1 \\ 3/8 & 3 < x < 5 \\ 0 & \text{otherwise} \end{cases}$$

Plot of the PDF:



From the relatively simple structure, we can easily determine the area under the graph:

$$A = (1)\left(\frac{1}{4}\right) + (2)\left(\frac{3}{8}\right) = \frac{2}{8} + \frac{6}{8} = 1$$

(a) Finding the CDF by integrating the PDF. We will split up the integral in several parts. First for the case when $y \in (0, 1)$:

$$F_X(y) = \int_{-\infty}^{y} f_X(t)dt = \frac{1}{4}\int_0^y 1dt = \frac{1}{4}\Big[t\Big]_0^y = \frac{y}{4}$$

When $y = 1$ we have $F_X(1) = 1/4$. Next, we must consider the case $y \in (1, 3)$. Here the PDF is 0, so it doesn't increase. It remains constant at 1/4 (since the CDF doesn't decrease).

$$F_X(y) = \frac{1}{4}$$

Next is the case $y \in (3, 5)$. Consider the intermediary integral:

$$I_1 = \int_3^y \frac{3}{8}dt = \frac{3}{8}\Big[t\Big]_3^y = \frac{3y - 9}{8}$$

For values $y \in (3, 5)$ we start on 1/4, so the CDF in this region becomes:

$$F_X(y) = \frac{3y - 9}{8} + \frac{1}{4}$$

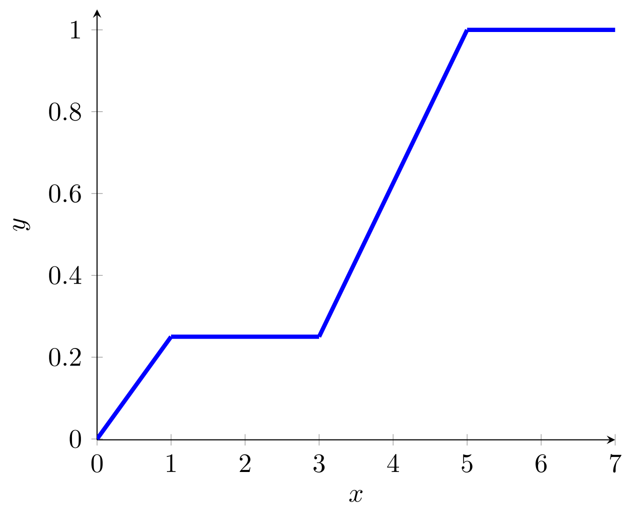
So, the full expression for the CDF becomes:

$$F_X(y) = \begin{cases} y/4 & y \in (0,1) \\ 1/4 & y \in (1,3) \\ \dfrac{3y-9}{8} + \dfrac{1}{4} & y \in (3,5) \\ 1 & y \geqslant 5 \end{cases}$$

Note that when $y = 5$ we get:

$$F_X(5) = \frac{3(5)-9}{8} + \frac{1}{4} = \frac{6}{8} + \frac{2}{8} = 1$$

Plot of the CDF:



(b) Defining $Y = 1/X$ and finding the PDF of Y. Following the hint we are given, we will consider the following three sets:

$$A_1 = \frac{1}{5} \leqslant y \leqslant \frac{1}{3}, \quad A_2 = \frac{1}{3} \leqslant y \leqslant 1, \quad A_3 = y \geqslant 1$$

Where A_1 corresponds to $(3,5)$, A_2 to $(1,3)$ and A_3 to $(0,1)$. We can express the CDF for $F_Y(y)$ in terms of $F_X(x)$:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(\frac{1}{X} \leqslant y) \\ &= \mathbb{P}(X \geqslant \frac{1}{y}) \\ &= 1 - \mathbb{P}(X \leqslant \frac{1}{y}) \\ &= 1 - F_X(\frac{1}{y}) \end{aligned}$$

First, we consider $A_1 : y \in [1/5, 1/3]$, and when we input $1/y$ to $F_X(\cdot)$, it will be in $(3, 5)$. So:

$$\begin{aligned}
F_Y(y) &= 1 - F_X(1/y) \\
&= 1 - \left(\frac{3(\frac{1}{y}) - 9}{8} + \frac{1}{4} \right) \\
&= 1 - \frac{3 - 9y}{8y} - \frac{1}{4} \\
&= \frac{3}{4} + \frac{9y - 3}{8y} \\
&= \frac{15y - 3}{8y}
\end{aligned}$$

Next, we consider $A_2 : y \in [1/3, 1]$. The input to $F_X(\cdot)$ will be in $(1, 3)$:

$$\begin{aligned}
F_Y(y) &= 1 - F_X(1/y) \\
&= 1 - \frac{1}{4} \\
&= \frac{3}{4}
\end{aligned}$$

Next, we consider $A_3 : y \geqslant 1$. The input to $F_X(\cdot)$ will be in $(0, 1)$:

$$\begin{aligned}
F_Y(y) &= 1 - F_X(1/y) \\
&= 1 - \frac{\frac{1}{y}}{4} \\
&= 1 - \frac{1}{4y}
\end{aligned}$$

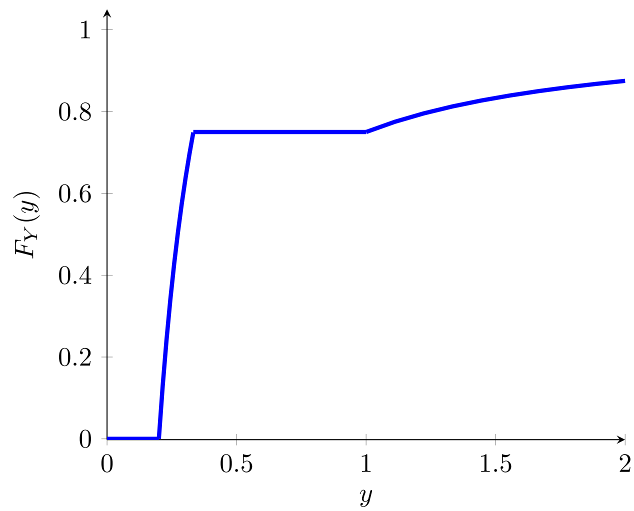
Also, whenever $y < 1/5$, then $1/y > 5$ which means $F_X(\cdot) = 1$, and so:

$$F_Y(y) = 1 - F_X(1/y) = 1 - 1 = 0.$$

This gives a full description of the CDF for $F_Y(y)$.

$$F_Y(y) = \begin{cases} 0 & y < 1/5 \\ \frac{15y - 3}{8y} & 1/5 \leqslant y \leqslant 1/3 \\ \frac{3}{4} & 1/3 \leqslant y \leqslant 1 \\ 1 - \frac{1}{4y} & y \geqslant 1 \end{cases}$$

Plot of CDF:



Finally, we can find the PDF of Y. We differentiate each of the parts in the CDF. When $y \in (1/5, 1/3)$:

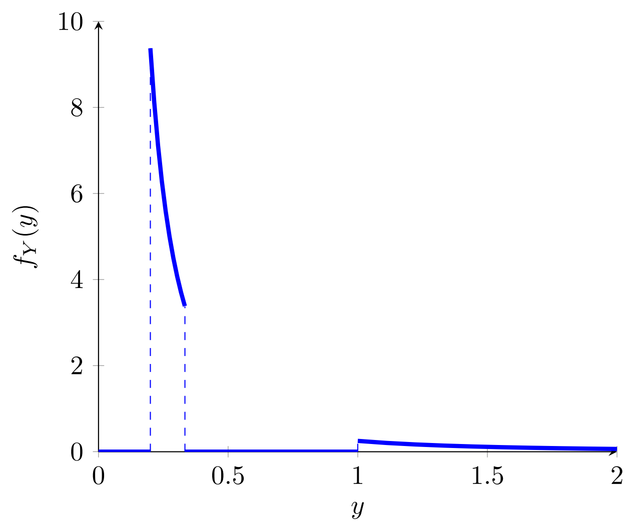
$$\frac{d}{dy}\left(\frac{15y-3}{8y}\right) = \frac{3}{8y^2}$$

When $y \geqslant 1$:

$$\frac{d}{dy}\left(1 - \frac{1}{4y}\right) = \frac{1}{4y^2}$$

(All other parts are constant, so they become 0). This gives us the PDF and its plot:

$$f_Y(y) = \begin{cases} 0 & y < 1/5 \\ \dfrac{3}{8y^2} & 1/5 \leqslant y \leqslant 1/3 \\ 0 & 1/3 < y < 1 \\ \dfrac{1}{4y^2} & y \geqslant 1 \end{cases}$$



2.5
Let X and Y be discrete RV. X and Y are independent if and only if $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all x and y.
PROOF.
$\Rightarrow$) Assume that X and Y are independent. That means that for any x, y, we have

$$\mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

Starting with the definition of the joint pdf:

$$\begin{aligned} f_{X,Y}(x,y) &= \mathbb{P}(X = x, Y = y) \\ &= \mathbb{P}(X = x \cap Y = y) \\ &= \mathbb{P}(X = x)\mathbb{P}(Y = y) \\ &= f_X(x)f_Y(y) \end{aligned}$$

Which shows that $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all x and y.

$\Leftarrow$) Assume that $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all x and y. By definition:

$$\begin{aligned} f_{X,Y}(x,y) &= \mathbb{P}(X = x, Y = y) \\ &= \mathbb{P}(X = x \cap Y = y) \end{aligned}$$

And,

$$f_X(x)f_Y(y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

From our assumption, these are equal, so $\mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ which shows that X and Y are independent.
By implication both ways, the statement is proved. $\square$

2.6
Let X have distribution F and density f, and let A be a subset of the real line, e.g. $A = (a, b)$ for some $a, b \in \mathbb{R}$ and $a < b$. We have the indicator function

$$I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

We will set $Y = I_A(X)$ and find the PDF and CDF of Y.

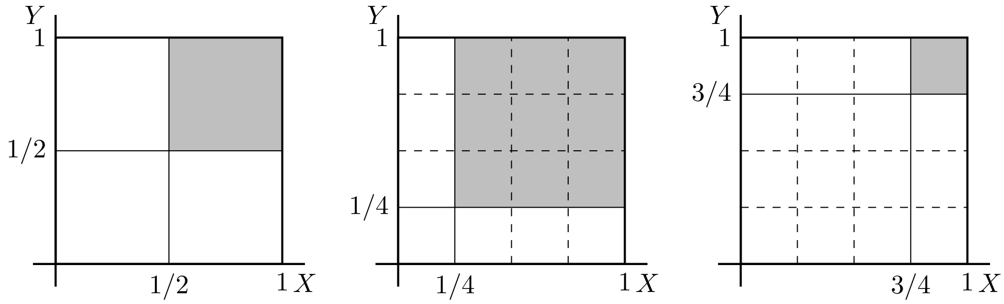
The exercise asks for a probability mass function, but that cannot be correct. Since X has a density f, it is a continuous RV. If $X \sim U(0,1)$ and $A = (0,1)$, then $Y = X$ and it will be a uniform variable with a continuous distribution, so not necessarily discrete.

And if it can be a continuous distribution, what happens if we define $A = \mathbb{Q} \subset \mathbb{R}$? There will be an infinite number of points in any interval with measure 0. Then we cannot define a PDF at all... Poorly formulated exercise in my opinion! Need to fill up with extra assumptions?

Skipping for now.

2.7

Let X and Y be independent and suppose that $X, Y \sim U(0,1)$. For $Z = \min(X, Y)$ we will find the density $f_Z(z)$ for Z. Following the hint, we will first find $\mathbb{P}(Z > z)$. Since any observations of $x, y \in (0,1)$, then we can immediately see that $\mathbb{P}(Z > 0) = 1$ and $\mathbb{P}(Z > 1) = 0$. But what happens for other values? Best way to find out is with some illustrations. Here are plots of the cases $\mathbb{P}(Z > 1/2)$, $\mathbb{P}(Z > 1/4)$ and $\mathbb{P}(Z > 3/4)$.



If we simulate lots of X and Y values, we see that about 1/4th of them will have both X and Y values larger than 1/2, so $\mathbb{P}(Z > 1/2) = 1/4$. Similarly, we get $\mathbb{P}(Z > 1/4) = 9/16$ and $\mathbb{P}(Z > 3/4) = 1/16$. Confirming this with a simulation.

```
# 2.7 - Simulating U(0,1)
N = 100000;X = runif(N);Y = runif(N)

Z = pmin(X, Y) # This is: Z = min{X, Y}

# Comparing simulated vs. theoretical results
sum(Z > 0.5)/N
1/4
sum(Z > 0.25)/N
9/16
sum(Z > 0.75)/N
1/16
```

```
> # Comparing simulated vs. theoretical results
> sum(Z > 0.5)/N
[1] 0.24888
> 1/4
[1] 0.25

> sum(Z > 0.25)/N
[1] 0.56156
> 9/16
[1] 0.5625

> sum(Z > 0.75)/N
[1] 0.06205
> 1/16
[1] 0.0625
```

By inspecting the images on the previous page, we can determine the 'shape' of the probabilities. For $Z > 1/4$ we remove the union of $X \leqslant 1/4$ and $Y \leqslant 1/4$. We define $A = \{X \leqslant z\}$ and $B = \{Y \leqslant z\}$, and can write the general case as:

$$\begin{aligned}
\mathbb{P}(Z > z) &= 1 - \mathbb{P}(\{X \leqslant z\} \cup \{Y \leqslant z\}) \\
&= 1 - \mathbb{P}(A \cup B) \\
&= 1 - \big[\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\big] \\
&= 1 - \big[\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B)\big]
\end{aligned}$$

Where we used Lemma 1.6, and the fact that X and Y are independent. By using the probability law of complements, we can find the expression for $\mathbb{P}(Z \leqslant z)$.

$$\mathbb{P}(Z \leqslant z) = \mathbb{P}(X \leqslant z) + \mathbb{P}(Y \leqslant z) - \mathbb{P}(X \leqslant z)\mathbb{P}(Y \leqslant z)$$

The CDF for a uniform distribution on $U(a, b)$ is:

$$F(z) = \frac{z-a}{b-a} \quad \Longrightarrow \quad F_X(z) = F_Y(z) = \frac{z-0}{1-0} = z$$

Which means:

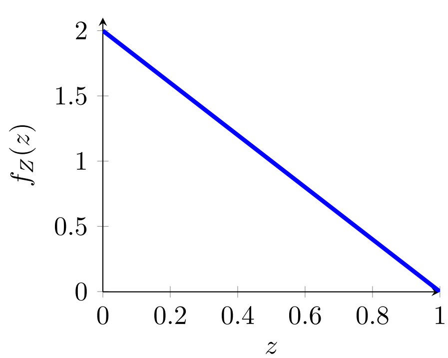
$$F_Z(z) = F_X(z) + F_Y(z) - F_X(z)F_Y(z) = 2z - z^2$$

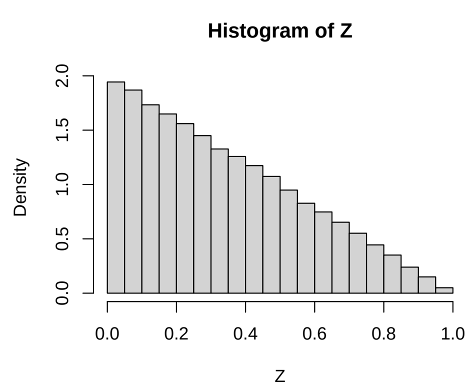
We can confirm our illustrations and simulated examples again by noting that:

$$\begin{aligned}
F_Z(1/2) &= \frac{1}{2} + \frac{1}{2} - \left(\frac{1}{2} \cdot \frac{1}{2}\right) = 1 - \frac{1}{4} = \frac{3}{4} \\
F_Z(1/4) &= \frac{1}{4} + \frac{1}{4} - \left(\frac{1}{4} \cdot \frac{1}{4}\right) = \frac{8}{16} - \frac{1}{16} = \frac{7}{16} \\
F_Z(3/4) &= \frac{3}{4} + \frac{3}{4} - \left(\frac{3}{4} \cdot \frac{3}{4}\right) = \frac{24}{16} - \frac{9}{16} = \frac{15}{16}
\end{aligned}$$

which gives us the opposite results as expected (since we simulated and illustrated $\mathbb{P}(Z > z)$). The PDF $f_Z(z)$ is the derivative of $F_Z(z)$. Including PDF-plot and histogram of the simulated Zs.

$$f_Z(z) = \frac{d}{dz}\Big(F_Z(z)\Big) = 2 - 2z.$$





2.8

The RV X has CDF F. Finding the CDF of $X^+ = \max\{0, X\}$.

From the definition of CDF:

$$\begin{aligned} F_{X^+}(u) &= \mathbb{P}\big(\max(0, X) \leqslant u\big) \\ &= \mathbb{P}\Big(\{\omega \in \Omega \ : \ X(\omega) \leqslant u \text{ and } u \geqslant 0\}\Big) \end{aligned}$$

We must consider two cases. When $u < 0$:

$$\begin{aligned} F_{X^+}(u) &= \mathbb{P}\Big(\{\omega \in \Omega \ : \ X(\omega) \leqslant u \text{ and } u \geqslant 0\}\Big) \\ &= \mathbb{P}(\emptyset) \\ &= 0 \end{aligned}$$

When $u \geqslant 0$:

$$\begin{aligned} F_{X^+}(u) &= \mathbb{P}\Big(\{\omega \in \Omega \ : \ X(\omega) \leqslant u \text{ and } u \geqslant 0\}\Big) \\ &= \mathbb{P}\Big(\{\omega \in \Omega \ : \ X(\omega) \leqslant u\}\Big) \\ &= \mathbb{P}(X \leqslant u) \\ &= F_X(u) \end{aligned}$$

So in summary:

$$F_{X^+}(u) = \begin{cases} 0 & u < 0 \\ F_X(u) & u \geqslant 0 \end{cases}$$

2.9

We have $X \sim \text{Exp}(\beta)$. The PDF is given by:

$$f(x) = \frac{1}{\beta} e^{-x/\beta}, \quad x > 0$$

Finding the CDF by integrating the PDF.

$$\begin{aligned} F(y) &= \int_{-\infty}^{y} f(x) dx \\ &= \frac{1}{\beta} \int_0^y e^{-x/\beta} dx \\ &= \frac{1}{\beta} \big[-\beta e^{-x/\beta} \big]_0^y \\ &= \big[-e^{-x/\beta} \big]_0^y \\ &= 1 - e^{-y/\beta} \end{aligned}$$

To find the inverse $F^{-1}(q)$ we set $q = F(y)$ and solve for y.

$$q = 1 - e^{-y/\beta} \quad \Longrightarrow \quad y = -\beta \log(1-q) \quad \Longrightarrow \quad F^{-1}(q) = -\beta \log(1-q)$$

2.10
If X and Y are independent, then $g(X)$ and $h(Y)$ are independent for some functions g and h.

PROOF. Let X and Y be some arbitrary random variables, and let x and y be values in the range of g and h such that $g(X) = x$ and $h(Y) = y$. Then:

$$\mathbb{P}(g(X) = x, h(Y) = y) = \mathbb{P}(X = g^{-1}(x), Y = h^{-1}(y))$$

By independence of X and Y.

$$= \mathbb{P}(X = g^{-1}(x))\mathbb{P}(Y = h^{-1}(y))$$
$$= \mathbb{P}(g(X) = x)\mathbb{P}(h(Y) = y)$$

which shows that g and h satisfies the condition for independence. □

2.11
Tossing a coin which has probability p of getting H. We let X denote the number of heads and Y the number of tails.

(a) Showing that X and Y are dependent. First it will be helpful to consider a simplified case where we have $N = 1$ and $N = 2$ coin tosses, and assuming $p = 1/2$.

$N = 1$ Tosses

	$Y = 0$	$Y = 1$	
$X = 0$	0	1/2	1/2
$X = 1$	1/2	0	1/2
	1/2	1/2	1

$N = 2$ Tosses

	$Y = 0$	$Y = 1$	$Y = 2$	
$X = 0$	0	0	1/4	1/4
$X = 1$	0	1/2	0	1/2
$X = 2$	1/4	0	0	1/4
	1/4	1/2	1/4	1

In the case of $N = 2$, we see that $f(0, 2) = 1/4$ while $f_X(0)f_Y(2) = 1/16$. This will be the inspiration for how we show it in the general case with N tosses and probability p of heads.

We only need to show one specific case where $\mathbb{P}(X = x, Y = y) \neq \mathbb{P}(X = x)\mathbb{P}(Y = y)$ to show that these values are dependent. The total number of tosses will be $N = X + Y$ and we compare the cases where we get N heads. In that case, using the Multinomial distribution:

$$\mathbb{P}(X = N, Y = 0) = \binom{N}{N, 0} = p^N(1-p)^0 = p^N,$$

but with the two Binomial distributions:

$$\mathbb{P}(X = N)\mathbb{P}(Y = 0) = \binom{N}{N}p^N(1-p)^0 \times \binom{N}{0}p^0(1-p)^N = p^N(1-p)^N.$$

These are not equal, showing that X and Y are dependent variables.

(b) Now we have $N \sim \text{Poisson}(\lambda)$, where $N = X + Y$ for X heads and Y tails. Show that these values are now independent.

Current solution, but not sure it's correct...
Assuming $X = x$ and $Y = y$. Then:

$$\begin{aligned}
\mathbb{P}(X = x \cap Y = y) &= \mathbb{P}(X = x \cap Y = y \cap N = x + y) \\
&= \mathbb{P}(X = x \cap Y = y \mid N = x + y) \cdot \mathbb{P}(N = x + y) \\
&= \binom{x+y}{x} p^x (1-p)^y \cdot e^{-\lambda} \frac{\lambda^{x+y}}{(x+y)!} \\
&= \frac{\cancel{(x+y)!}}{x!y!} p^x (1-p)^y \cdot e^{-\lambda} \frac{\lambda^x \lambda^y}{\cancel{(x+y)!}} \\
&= e^{-\lambda} \frac{p^x \lambda^x}{x!} \cdot \frac{(1-p)^y \lambda^y}{y!} \\
&= e^{-\lambda p} \frac{p^x \lambda^x}{x!} \cdot e^{-\lambda(1-p)} \frac{(1-p)^y \lambda^y}{y!} \\
&= \mathbb{P}(X = x | \lambda p) \cdot \mathbb{P}(Y = y | \lambda(1-p)) \\
&= \mathbb{P}(X = x)\mathbb{P}(Y = y)
\end{aligned}$$

Which shows we have independence. We used:

$$e^{-\lambda} = e^{-\lambda(p-p+1)} = e^{-\lambda p + \lambda p - \lambda} = e^{-\lambda p} e^{\lambda p - \lambda} = e^{-\lambda p} e^{-\lambda(1-p)}.$$

Also, we made the assumption that $\mathbb{P}(X = x \cap Y = y \mid N = x + y)$ is Binomial. But I think this is only true when we can assume that X and Y are independent... which is what we are trying to show. Will review this later... hopefully!

2.12 Theorem 2.33
Suppose that the range of X and Y is a (possibly infinite) rectangle. If $f(x, y) = g(x)h(y)$ for some functions g and h (not necessarily probability density functions) then X and Y are independent.

Proof. From the joint PDF, we can find the marginal distributions:

$$f_X(x) = \int f(x, y)dy, \quad f_Y(y) = \int f(x, y)dx$$

By applying these integrals to both sides of the equality:

$$\int f(x, y)dy = g(x) \int h(y)dy \quad \Longrightarrow \quad f_X(x) = g(x)$$

$$\int f(x, y)dx = h(y) \int g(x)dx \quad \Longrightarrow \quad f_Y(y) = h(y)$$

We get $g(x)$ and $h(y)$ since they are equal to $f(x, y)$ over the entire 'rectangle' and must therefore integrate to 1. This leaves us with: $f(x, y) = g(x)h(y) = f_X(x)f_Y(y)$, and by the results in exercise 2.5, this means that X and Y are independent. $\square$

2.13

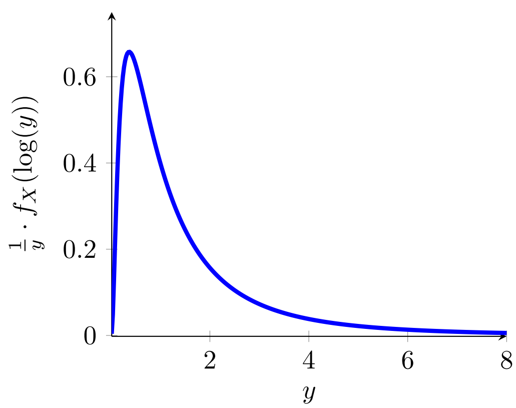
(a) Finding the PDF of $Y = e^X$ when $X \sim N(0,1)$.

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leqslant y) = \mathbb{P}(e^X \leqslant y) \\ &= \mathbb{P}(X \leqslant \log(y)) = F_X(\log(y)) \end{aligned}$$

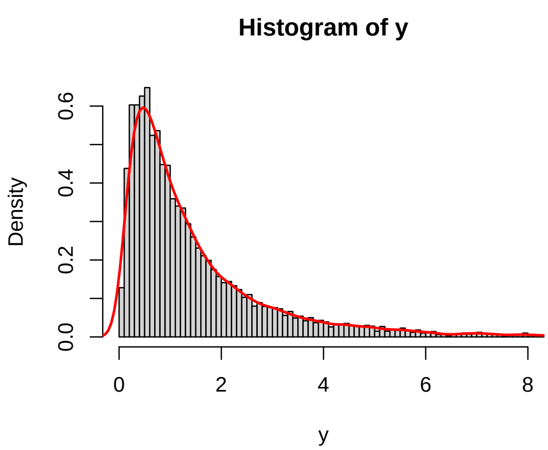
We simply get the standard normal distribution with $\log(y)$ as input. Differentiating to get the PDF:

$$f_Y(y) = \frac{d}{dy} F_X(\log(y)) = f_X(\log(y)) \cdot \frac{1}{y} = \frac{1}{y\sqrt{2\pi}} \exp\left(-\frac{1}{2}\big(\log(y)\big)^2\right)$$

Plotting the function:



(b) Plotting histogram of simulated results. Comparable to the plot above.




```r
x = rnorm(10000)
y = exp(x)
pdf("~/AllStatistics/files/ch2_2.13b.pdf",
     width = 4.7747, height = 4)
d = density(y)
hist(y, breaks = 200, xlim = c(0, 8),
     prob=TRUE)
lines(density(y), col="red", xlim=c(0,8))
dev.off()
```

2.14

We let (X, Y) be uniformly distributed on the unit disk: $\{(x, y) : x^2 + y^2 \leqslant 1\}$. Find the CDF and PDF of $R = \sqrt{X^2 + Y^2}$.

First, let's write some code in order to simulate the data, which are random points contained within the unit circle. And then use that to simulate R and view the histogram. Plots and code can be found on the next page. By inspection of the histogram it is clear that $f_R(r) = 2r$ which means $F_R(r) = r^2$.

Following the general recipe for transformation of multiple random variables. We will define the function $r(X, Y) = \sqrt{X^2 + Y^2}$.

$$F_R(r) = \mathbb{P}(R \leqslant r) = \mathbb{P}(r(X, Y) \leqslant r) = \mathbb{P}(X^2 + Y^2 \leqslant r^2) = \int\int_{A_r} f_{X,Y}(x, y)dxdy$$

Finding the set A_r which in this case is still the unit circle, since $\sqrt{x^2 + y^2} \leqslant 1 \implies x^2 + y^2 \leqslant 1$.

Since the area of the unit circle is π, and since a uniform distribution means the probability is equal everywhere, it follows that:

$$f_{X,Y}(x, y) = \frac{1}{\pi}.$$

As we saw, we must integrate this function over the circle: $\{(x, y) : x^2 + y^2 < r\}$ for some $0 \leqslant r \leqslant 1$. This is easiest when calculating the integral in polar coordinates, so we introduce $0 \leqslant \theta \leqslant 2\pi$ and $0 \leqslant \rho \leqslant r$ and change $dxdy = \rho d\rho d\theta$. So:

$$\begin{aligned}
F_R(r) &= \frac{1}{\pi}\int_0^{2\pi}\int_0^r \rho d\rho d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi}\left[\frac{\rho^2}{2}\right]_0^r d\theta \\
&= \frac{1}{\pi}\int_0^{2\pi}\frac{r^2}{2}d\theta \\
&= \frac{r^2}{2\pi}\int_0^{2\pi}(1)d\theta \\
&= \frac{r^2}{2\pi}\left[\theta\right]_0^{2\pi} \\
&= \frac{r^2}{2\pi}\cdot 2\pi \\
&= r^2
\end{aligned}$$

The CDF is $F_R(r) = r^2$ for $0 \leqslant r \leqslant 1$. By differentiating this, we get the PDF.

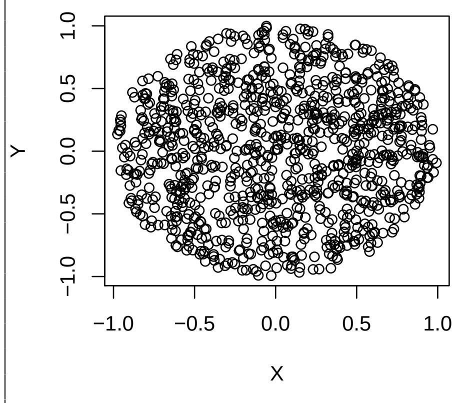
$$f_R(r) = \frac{d}{dr}F_R(r) = 2r,$$

also in the interval $0 \leqslant r \leqslant 1$. This confirms the results we found by simulating which are found on the next page.

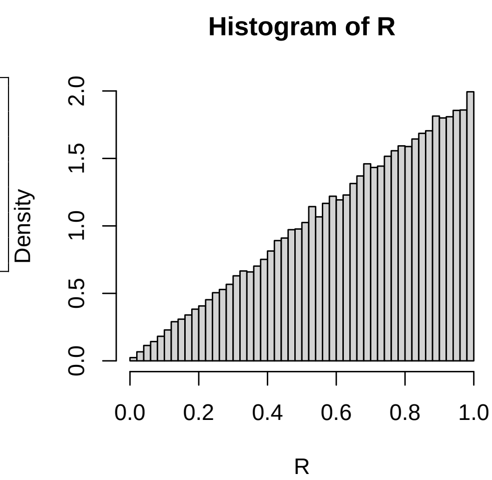
Plots and Code for 2.14

```
simulateUnitCircle = function(N) {
    TR = runif(N)
    UR = runif(N); VR = runif(N)
    t = 2*pi*TR
    u = UR+VR
    u[u > 1] = 2 - u[u > 1]
    r = u
    X = r*cos(t)
    Y = r*sin(t)
    retVal = list()
    retVal$X = X
    retVal$Y = Y
    return(retVal)
}
simVal = simulateUnitCircle(1000)
X = simVal$X
Y = simVal$Y
pdf("~/AllStatistics/files/ch2_2.14.pdf",
    width = 4, height = 4)
plot(X, Y)
dev.off()
```

Result from simulation:




```
simVal = simulateUnitCircle(50000)
R = sqrt(simVal$X^2 + simVal$Y^2)
pdf("~/AllStatistics/files/ch2_2.14b.pdf",
    width = 4, height = 4)
hist(R, breaks = 40, prob = TRUE)
plot(X, Y)
dev.off()
```



2.15

Let X have a continuous, strictly increasing CDF F, and set $Y = F(X)$. Find the density of Y.

By definition of the CDF, $F : \mathbb{R} \to [0,1]$ which will be the domain of Y. Since F is strictly increasing, then $F(X_1) < F(X_2)$ if $X_1 < X_2$, but the important property is that F must be *invertible*. For any $y \in [0,1]$ we get the following:

$$\begin{aligned} F_Y(y) = \mathbb{P}(Y \leqslant y) &= \mathbb{P}(F(X) \leqslant y) \\ &= \mathbb{P}(X \leqslant F^{-1}(y)) = F(F^{-1}(y)) = y \end{aligned}$$

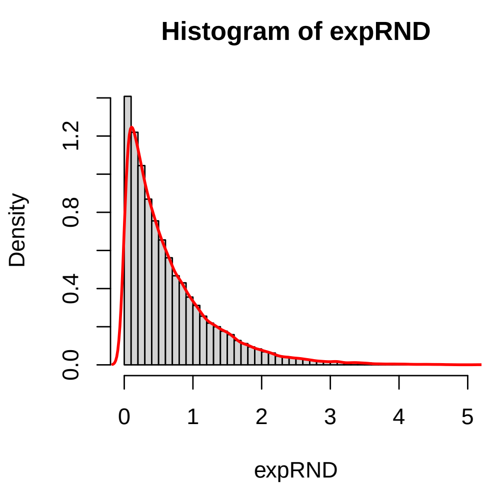
Since $F_Y(y) = y$, then the density is $f_Y(y) = 1$, the derivative. From these observations, we can conclude that $Y \sim U(0,1)$.

Next, if $U \sim U(0,1)$ and $X = F^{-1}(U)$, then X has F as its CDF.

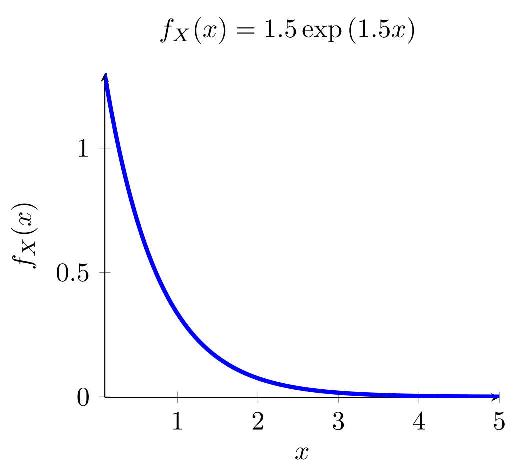
$$\mathbb{P}(F^{-1}(U) \leqslant x) = \mathbb{P}(X \leqslant x) = F(x) \quad \Longrightarrow \quad X \sim F$$

Writing a program that generates $\mathrm{Exp}(\beta)$ values. Using the result found in exercise 2.9.

```
simExp = function(beta, N){
    U = runif(N)
    return(-beta*log(1 - U))
}
# Simulating
expRND = simExp(beta = 2/3, N = 50000)
# Making plot
hist(expRND, breaks=80,
    xlim=c(0, 5), prob=TRUE)
lines(density(expRND), col="red", xlim=c
    (0,5), lwd=2)
```



Plot of the exponential distribution for $\beta = 2/3$.



2.16
If $X \sim \text{Poisson}(\lambda_1)$ and $X \sim \text{Poisson}(\lambda_2)$, and X and Y are independent, then the distribution of X given that $X+Y=n$ is $\text{Binomial}(n,\rho)$ where $\rho = \lambda_1/(\lambda_1+\lambda_2)$. (Changed some variable names).

PROOF. By definition of conditional probability, by applying hint 1, by independence and by using the fact that $X+Y$ is a Poisson distribution with $\lambda_1+\lambda_2$.

$$\begin{aligned}
\mathbb{P}(X=x|X+Y=n) &= \frac{\mathbb{P}(X=x \cap X+Y=n)}{\mathbb{P}(X+Y=n)} \\
&= \frac{\mathbb{P}(X=x \cap Y=n-x)}{\mathbb{P}(X+Y=n)} && \text{(Hint)} \\
&= \frac{\mathbb{P}(X=x)\mathbb{P}(Y=n-x)}{\mathbb{P}(X+Y=n)} && \text{(Independence)}
\end{aligned}$$

Each of these terms are known:

$$\begin{aligned}
\mathbb{P}(X=x) &= e^{-\lambda_1} \cdot \frac{\lambda_1^x}{x!} \\
\mathbb{P}(Y=n-x) &= e^{-\lambda_2} \cdot \frac{\lambda_2^{(n-x)})}{(n-x)!} \\
\mathbb{P}(X+Y=n) &= e^{-(\lambda_1+\lambda_2)} \cdot \frac{(\lambda_1+\lambda_2)^n}{n!}
\end{aligned}$$

We can invert the last equation:

$$\frac{1}{\mathbb{P}(X+Y=n)} = e^{\lambda_1+\lambda_2} \cdot \frac{n!}{(\lambda_1+\lambda_2)^n}$$

Putting it all together, we can complete the justification.

$$\begin{aligned}
\mathbb{P}(X=x|X+Y=n) &= \frac{\mathbb{P}(X=x)\mathbb{P}(Y=n-x)}{\mathbb{P}(X+Y=n)} \\
&= \left(e^{-\lambda_1} \cdot \frac{\lambda_1^x}{x!}\right)\left(e^{-\lambda_2} \cdot \frac{\lambda_2^{(n-x)})}{(n-x)!}\right)\left(e^{\lambda_1+\lambda_2} \cdot \frac{n!}{(\lambda_1+\lambda_2)^n}\right) \\
&= \left(\frac{\lambda_1^x}{x!}\right)\left(\frac{\lambda_2^{(n-x)})}{(n-x)!}\right)\left(\frac{n!}{(\lambda_1+\lambda_2)^x(\lambda_1+\lambda_2)^{(n-x)}}\right) \\
&= \left(\frac{n!}{x!(n-x)!}\right)\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{(n-x)} \\
&= \binom{n}{x}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{(n-x)} \\
&= \binom{n}{x}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^x\left(1-\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{(n-x)} \\
&= \binom{n}{x}\rho^x(1-\rho)^{(n-x)} \sim \text{Binomial}(n,\rho)
\end{aligned}$$

for $\rho = \lambda_1/(\lambda_1+\lambda_2)$. □

2.17
The joint probability density is:

$$f_{X,Y}(x,y) = \begin{cases} c(x+y^2) & 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1 \\ 0 & \text{otherwise} \end{cases}$$

Calculating $\mathbb{P}(X < 1/2 \,|\, Y = 1/2)$.

We begin by calculating the marginal distribution for $f_Y(y)$.

$$f_Y(y) = \int_0^1 cx + cy^2 dx = c\int_0^1 xdx + c\int_0^1 y^2 dx = c\Big[\frac{x^2}{2}\Big]_0^1 + cy^2\big[x\big]_0^1 = c\left(\frac{1}{2} + y^2\right)$$

Following the steps in example 2.38.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{c(x+y^2)}{c\left(\frac{1}{2}+y^2\right)} = \frac{x+y^2}{\frac{1}{2}+y^2}$$

Calculating the conditional probability:

$$\begin{aligned} \mathbb{P}(X < 1/2|Y = 1/2) &= \int_0^{1/2} f_{X|Y}(1/2|1/2)dx \\ &= \int_0^{1/2} \frac{x+\frac{1}{4}}{\frac{1}{2}+\frac{1}{4}}dx \\ &= \frac{4}{3}\int_0^{1/2} x + \frac{1}{4}dx \\ &= \frac{4}{3}\left[\frac{x^2}{2} + \frac{x}{4}\right]_0^{1/2} = \frac{4}{3}\cdot\frac{1}{4} = \frac{1}{3} \end{aligned}$$

2.18
Let $X \sim N(3, 16)$. Solve the following using the Normal table and using a computer package.

(a)

$$\mathbb{P}(X < 7) = \mathbb{P}(3 + 4Z < 7) = \mathbb{P}\left(Z < \frac{7-3}{4}\right) = \mathbb{P}(Z < 1) = 0.8413$$

From **R**:

```
> pnorm(1)
[1] 0.84134475
```

```
> 1 - pnorm(-0.25)
[1] 0.59870633
```

(b)

$$\begin{aligned} \mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \leqslant 2) = 1 - \mathbb{P}(3 + 4Z \leqslant 2) = 1 - \mathbb{P}\left(Z \leqslant \frac{2-3}{4}\right) \\ &= 1 - \mathbb{P}(Z < -1/4) = 1 - (1 - 0.5987) = 0.5987 \end{aligned}$$

(c) Find x such that $\mathbb{P}(X > x) = 0.05$. By finding 0.95 in the table, we see that it is generated by the value 1.645, so:

$$\mathbb{P}(Z \leqslant 1.645) = 0.95 \quad \Longrightarrow \quad \mathbb{P}(Z > 1.645) = 1 - 0.95 = 0.05$$

$$\mathbb{P}(4Z > 6.58) = \mathbb{P}(3 + 4Z > 9.58) = \mathbb{P}(X > 9.58) = 0.05$$

So, $x = 9.58$. From **R**:

```
1 - pnorm(9.5794, mean = 3, sd = 4)
[1] 0.05000037
```

(d)

$$\begin{aligned}\mathbb{P}(0 \leqslant X < 4) &= \mathbb{P}(0 \leqslant 3 + 4Z < 4) = \mathbb{P}(-3 \leqslant 4Z < 1) \\ &= \mathbb{P}(-3/4 \leqslant Z < 1/4) = \mathbb{P}(Z < 1/4) - \mathbb{P}(Z < -3/4) \\ &= 0.5987 - (1 - 0.7734) = 0.5987 - 0.2266 \\ &= 0.3721\end{aligned}$$

From **R**:

```
> pnorm(0.25) - pnorm(-0.75)
[1] 0.37207897
#### Alternatively
> CH = rnorm(10000000, mean = 3, sd = 4)
> sum(CH < 4 & CH > 0)/10000000
[1] 0.3719199
```

(e) Find x such that $\mathbb{P}(|X| > |x|) = 0.05$.

We can split this up in two sections if we assume $x > 0$. Using that the sets are disjoint:

$$\begin{aligned}\mathbb{P}(|X| > |x|) &= \mathbb{P}(X > x \cap X < -x) = \mathbb{P}(X > x) + \mathbb{P}(X < -x) \\ &= 1 - \mathbb{P}(X \leqslant x) + \mathbb{P}(X < -x) = 0.05 \\ \mathbb{P}(X \leqslant x) - \mathbb{P}(X < -x) &= 0.95\end{aligned}$$

Note we get minus. Translating to standard normal:

$$\mathbb{P}\left(Z \leqslant \frac{x-3}{4}\right) - \mathbb{P}\left(Z < \frac{-x-3}{4}\right) = 0.95$$

I am not able to see how to solve this by just looking up values because of the dependence. By using R, we can find that $x = 9.611$ which makes the lookup values 1.65275 and -3.15275. Verifying in **R**.

```
> X = rnorm(10000000, mean = 3, sd = 4)
> x = 9.611
> sum(X < -x | X > x)/10000000
[1] 0.0500769
> sum(abs(X) > abs(x))/10000000
[1] 0.0500769
```

2.19

Proving equation (2.12). Let X be a random variable with pdf $f_X(x)$, and define the random variable $Y = r(X)$, where r is either strictly monotone increasing or decreasing, so $r(\cdot)$ has the inverse function $s(\cdot) = r^{-1}(\cdot)$. Then, we can express the pdf of Y as:

$$f_Y(y) = f_X(s(y))\left|\frac{ds(y)}{dy}\right|.$$

Proof. The proof will be in two parts. Assume first that r is strictly increasing, which means that s is strictly increasing and its derivative is positive. From the CDF of Y, we get:

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) \\
&= \mathbb{P}(r(X) \leqslant y) \\
&= \mathbb{P}(X \leqslant r^{-1}(y)) \\
&= \mathbb{P}(X \leqslant s(y)) \\
&= F_X(s(y))
\end{aligned}$$

Differentiating the CDF to get the PDF.

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(s(y)) = f_X(s(y))\frac{ds(y)}{dy}$$

Now assume that r is strictly decreasing, which means s is strictly decreasing and thus has a negative derivative. From the CDF of Y, we now get:

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) \\
&= \mathbb{P}(r(X) \leqslant y) \\
&= \mathbb{P}(X \geqslant r^{-1}(y)) \\
&= 1 - \mathbb{P}(X \leqslant s(y)) \\
&= 1 - F_X(s(y))
\end{aligned}$$

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\big[1 - F_X(s(y))\big] = -f_X(s(y))\frac{ds(y)}{dy} = f_X(s(y))\left(-\frac{ds(y)}{dy}\right)$$

Since the derivative will be negative, the negative of the derivative will be positive. Generalizing to both cases, we can write:

$$f_Y(y) = f_X(s(y))\left|\frac{ds(y)}{dy}\right| \qquad \square$$

A short note on why we get $1 - \mathbb{P}(X \leqslant s(y))$ when r is strictly decreasing. As an example, say that $X \sim U(0, 5)$ and we define $Y = r(X)$ where $r(x) = 2 - 2x$, which is strictly decreasing. Then $r^{-1}(y) = s(y) = 1 - \frac{y}{2}$, which is also strictly decreasing. Then:

$$\mathbb{P}(r(X) \leqslant y) = \mathbb{P}(2 - 2X \leqslant y) = \mathbb{P}(2 - y \leqslant 2X) = \mathbb{P}(1 - \frac{y}{2} \leqslant X) = \mathbb{P}(X \geqslant s(y))$$

In order to get the expression in terms of $F_X(x)$, we rewrite it:

$$\mathbb{P}(X \geqslant s(y)) = 1 - \mathbb{P}(X \leqslant s(y)) = 1 - F_X(s(y)).$$

2.20
Let $X, Y \sim U(0,1)$ be independent. Find the PDF for $X-Y$ and X/Y. The joint density of X, Y is

$$f(x,y) = 1, \quad x, y \in (0,1)$$

and 0 elsewhere.

We define $Z = X - Y$. We note that Z can assume values between -1 and 1, which are the two extremes. This gives us $A_z = (-1,1)$. We split this into two cases: when $X > Y$, then $A_z = (0,1)$ and when $X < Y$, which means $A_z = (-1,0)$.

When $X > Y$ and $A_z = (0,1)$.

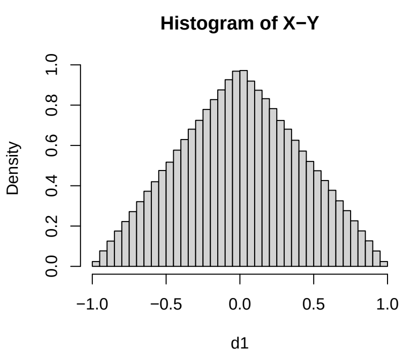
$$\begin{aligned} F_Z(z) &= \int_0^z \int_y^1 1 dx dy = \int_0^z [x]_y^1 dy \\ &= \int_0^z 1 - y dy = \left[y - \frac{y^2}{2} \right]_0^z \\ &= z - \frac{z^2}{2} \end{aligned}$$

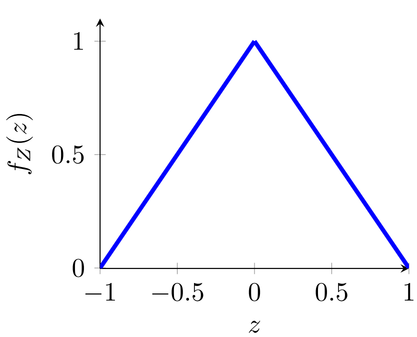
When $X < Y$ and $A_z = (-1,0)$.

$$\begin{aligned} F_Z(z) &= \int_{-1}^z \int_{-1}^y 1 dx dy = \int_{-1}^z [x]_{-1}^y dy \\ &= \int_{-1}^z y + 1 dy = \left[\frac{y^2}{2} + y \right]_{-1}^z \\ &= \frac{z^2}{2} + z + \frac{1}{2} \end{aligned}$$

Summarizing the CDF, and differentiating to find the PDF. (The 1/2 carries over).

$$F_Z(z) = \begin{cases} \frac{z^2}{2} + z + \frac{1}{2} & z \in (-1,0) \\ z - \frac{z^2}{2} + \frac{1}{2} & z \in (0,1) \end{cases} \implies f_Z(z) = \begin{cases} z+1 & z \in (-1,0) \\ 1-z & z \in (0,1) \end{cases}$$





Now we define $Z = X/Y$.

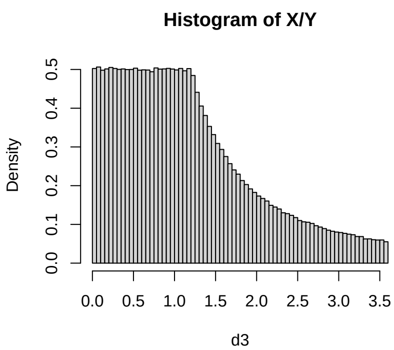
$$F_Z(z) = \mathbb{P}(Z \leqslant z) = \mathbb{P}(X/Y \leqslant z) = \mathbb{P}(X \leqslant zY)$$

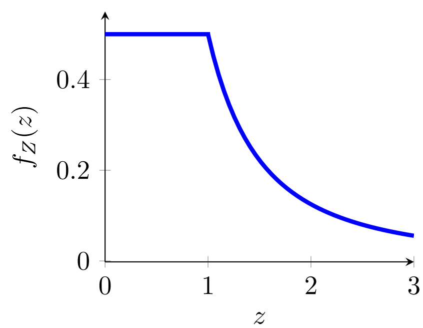
$$F_Z(z) = \int_0^1 \int_0^{\min(z,y)} 1 dx dy = \int_0^1 \min(z,y) dy = \begin{cases} \int_0^{\frac{1}{z}} zy dy + \int_{\frac{1}{z}}^1 1 dy & z \geqslant 1 \\ \int_0^1 zy dy & z < 1 \end{cases}$$

Evaluating the integrals to get the CDF, and differentiating to get the PDF:

$$F_Z(z) = \begin{cases} 1 - \frac{1}{2z} & z \geqslant 1 \\ \frac{z}{2} & z < 1 \end{cases} \implies f_Z(z) = \begin{cases} \frac{1}{2z^2} & z \geqslant 1 \\ \frac{1}{2} & z < 1 \end{cases}$$

Plots and simulation code.






```
# Simulating X-Y
N = 10000
X = runif(N);Y = runif(N)
d1 = X - Y
hist(d1, breaks = 50, prob = TRUE,
        main="Histogram of X-Y")

# Simulating X/Y
N = 1000000
X = runif(N);Y = runif(N)
d2 = X/Y
d3 = d2[abs(d2) < 3] # Removing very rare large vals
d3 = d3*length(d2)/length(d3) # Normalizing probability
hist(d3, breaks=60, prob=TRUE,
        main = "Histogram of X/Y")
```

R

2.21
Let $X_1, \dots, X_n \sim \text{Exp}(\beta)$ be IID. Define $Y = \max(X_1, \dots, X_n)$. Find the PDF of Y.