DATA 605 - Assignment 5

Joshua Sturm

03/04/2018

```r
library(ggplot2)
```

Problem 1

Choose independently two numbers B and C at random from the interval $[0, 1]$ with uniform density. Prove that B and C are proper probability distributions.

Note that the point (B, C) is then chosen at random in the unit square. Find the probability that:

Part 1

(a) $B + C < \frac{1}{2}$

Part 2

(b) $BC < \frac{1}{2}$

Part 3

(c) $|B - C| < \frac{1}{2}$

Part 4

(d) $max\{B, C\} < \frac{1}{2}$

Part 5

(e) $min\{B, C\} < \frac{1}{2}$

Solutions

There are two conditions for the functions to be proper probability distributions:

- $\sum P(X = x) = 1$
- $P(X = x) \geq 0 \quad \forall x$

B and C each have a 100% probability of $0 \leq B/C \leq 1$, so $B + C$ has a 100% probability of $0 \leq B + C \leq 2$.

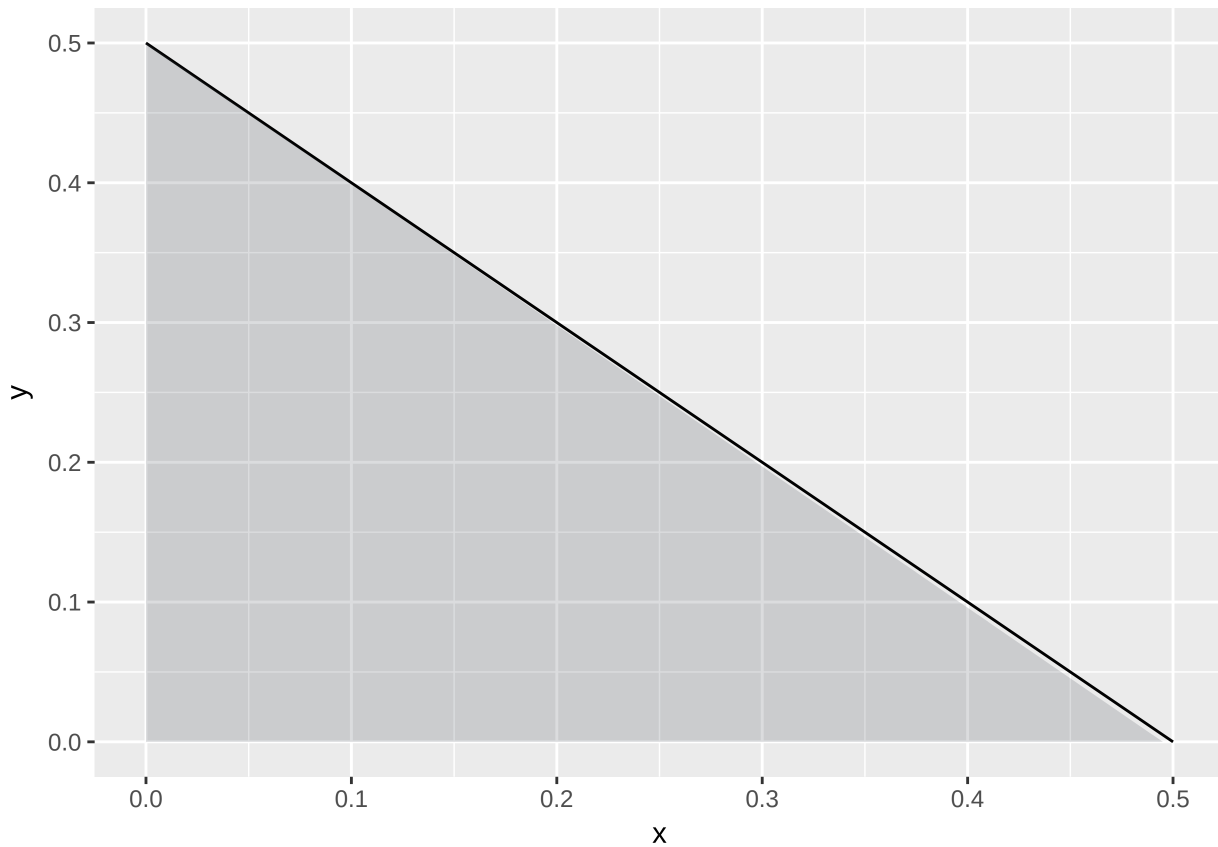
```r
B <- runif(n = 5000, min = 0, max = 1)
C <- runif(n = 5000, min = 0, max = 1)
```

Part 1

1. This is easier to understand visually, so I'll graph the equation. If we let $B = x$ and $C = y$, we can plot the function $f(x) = 0.5 - x$.

```
funcShaded <- function(x) {
    y <- seq(0.5,0, length.out = 100)
    y[x < 0 | x > 0.5] <- NA
    return(y)
}

function1 <- function(x) 0.5 - x
ggplot(data = data.frame(x = 0), mapping = aes(x = x)) +
  stat_function(fun = function1) +
  xlim(0, 0.5) +
  ylim(0, 0.5) +
  stat_function(fun=funcShaded, geom="area", fill="#4d535e", alpha=0.2)
```



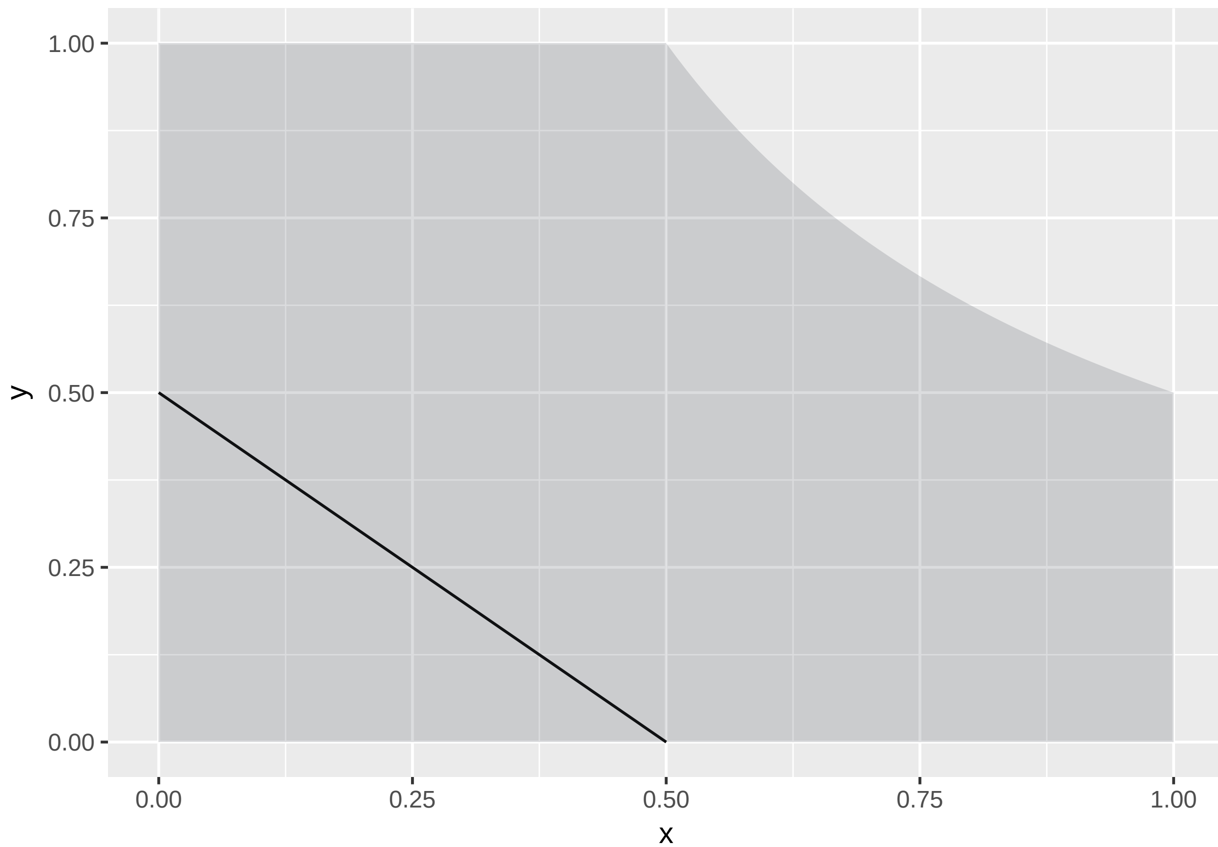
The probability that $x + y < \frac{1}{2}$ is just the area under the function. In this case, since it's essentially a triangle, we can simply apply the area formula: $\frac{1}{2} * \text{base} * \text{height}$.

```
base <- 0.5
height <- 0.5
0.5*base*height
## [1] 0.125
```

Part 2

$xy = \frac{1}{2} \to y = \frac{0.5}{x} \to y = \frac{1}{2x}$.

```
function2 <- function(x) 1/(2*x)
ggplot(data = data.frame(x = 0), mapping = aes(x = x)) +
  stat_function(fun = function1) +
  xlim(0, 1) +
  ylim(0, 1) +
  geom_rect(aes(xmin=0, xmax=0.5, ymin=0,ymax=1), fill="#4d535e", alpha=0.2) +
  stat_function(fun=function2, geom="area", fill="#4d535e", alpha=0.2)
```



Similarly, we need to calculate the area under this curve. We can break it up into two parts: a rectangle (x=0:0.5, y=0:1) and out function $\frac{1}{2x}(x = 0.5 : 1, y = 0 : \frac{1}{2x}.\backslash\ 0.5 * 1 + \int_{0.5}^{1} \frac{1}{2x} dx = 0.846574$.

```
integrate(function2, 0.5, 1)[[1]] + 0.5
## [1] 0.8465736
```

Part 3

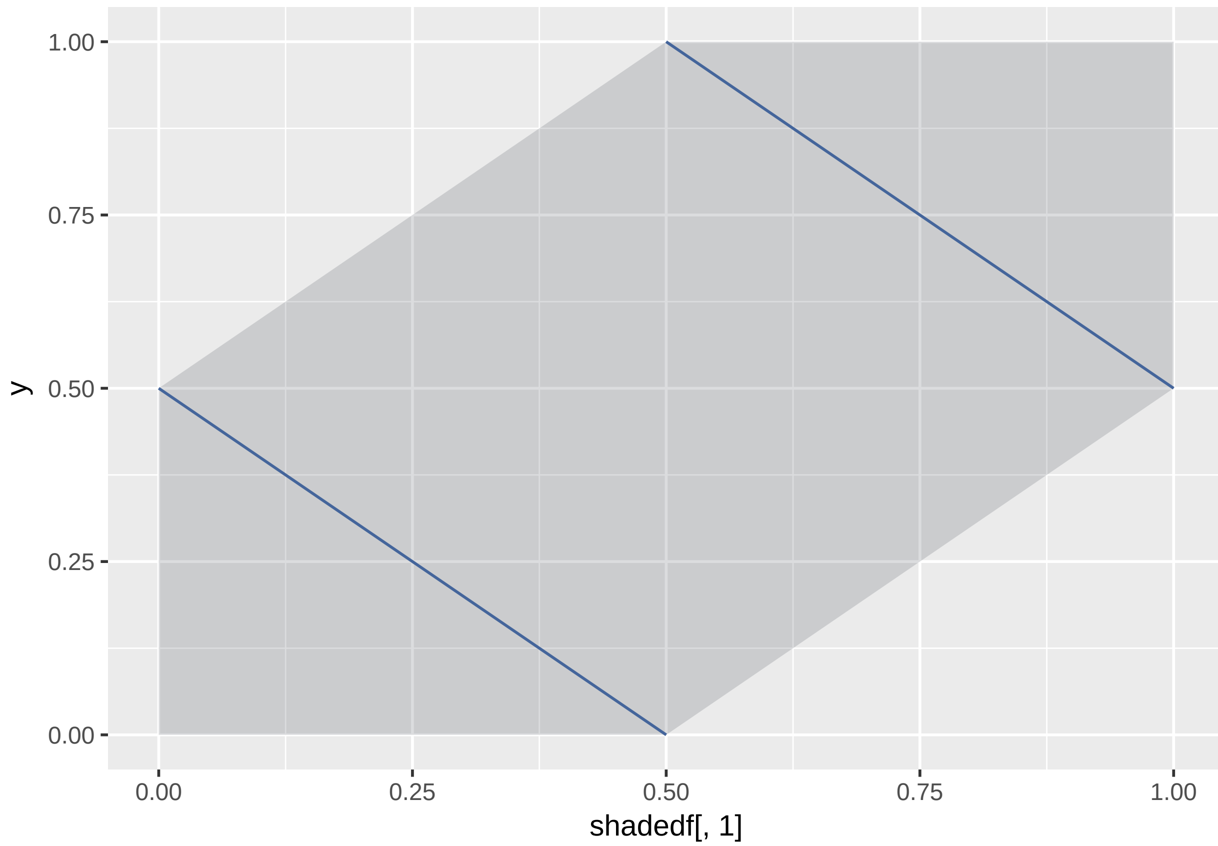
$|x - y| = \frac{1}{2}$ gives us two lines: $-x - y = \frac{1}{2}$ and $x - y = \frac{1}{2}$.

```
function3 <- function(x) x + 1/2
function31 <- function(x) x - 1/2

shadedx <- c(0,0,0.5,1,1,0.5)
shadedy <- c(0,0.5,1,1,0.5,0)
```

```r
shadedf <- cbind(shadedx, shadedy)

ggplot() +
  stat_function(fun = function3) +
  stat_function(fun = function31) +
  xlim(0, 1) +
  ylim(0, 1) +
  geom_polygon(aes(shadedf[,1],shadedf[,2]), fill="#4d535e", alpha=0.2) +
  geom_line(aes(x=seq(0,0.5,length.out = 3), y=seq(0.5,0, length.out = 3)), colour = "#44659b") +
  geom_line(aes(x=seq(0.5,1,length.out = 3), y=seq(1,0.5,length.out = 3)), colour="#44659b")
```



The blue lines are not part of the graph, but I included them to better understand how I approached this problem. Each blue line forms a triangle, similar to the one we encountered in part a, and also outlines our unit square. Thus, the area is

$$\text{Area of square - 2 * area of triangle} = 1 - 2 * 0.25 = 0.75.$$

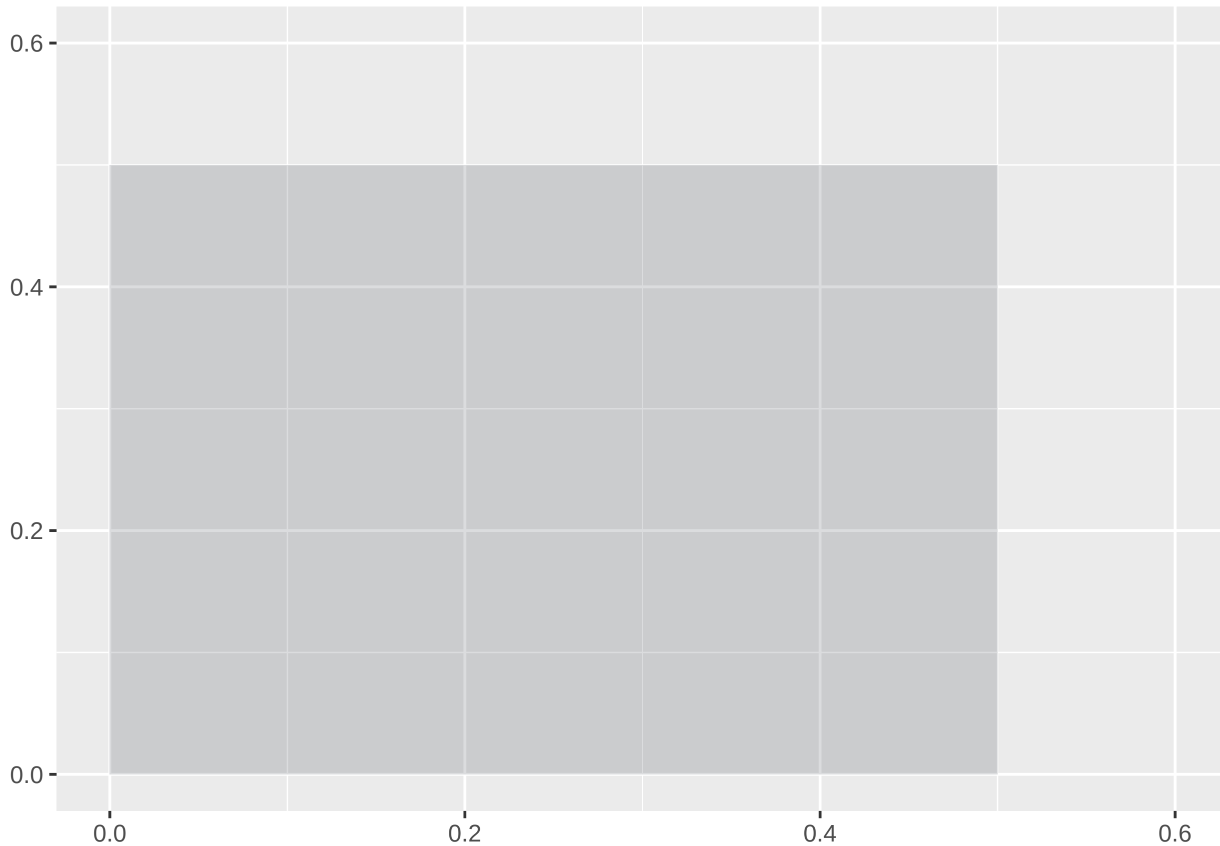
Part 4

$$P(B \leq \frac{1}{2}) * P(C \leq \frac{1}{2}) = (\frac{1}{2})^2 = \frac{1}{4}.$$

Alternatively, continuing with the area theme, we could plot the 0.5 by 0.5 square, and its area would be given by $0.5 * 0.5 = 0.25$.

```r
ggplot() +
  xlim(0, 0.6) +
```

```r
ylim(0, 0.6) +
geom_rect(aes(xmin=0, xmax=0.5, ymin=0,ymax=0.5), fill="#4d535e", alpha=0.2)
```



Part 5

Either $B < \frac{1}{2}$ or $C < \frac{1}{2}$ is needed to satisfy $min\{B, C\} < \frac{1}{2}$.

$P(B < \frac{1}{2} \cup C < \frac{1}{2}) = P(B < \frac{1}{2}) + P(C < \frac{1}{2}) - P(B < \frac{1}{2} \cap C < \frac{1}{2})$.

$P(B < \frac{1}{2}) = \frac{1}{2}$

$P(C < \frac{1}{2}) = \frac{1}{2}$

$P(B < \frac{1}{2} \cap C < \frac{1}{2}) = \frac{1}{4}$

Thus, $\frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} = 0.75$.

References

- https://stackoverflow.com/questions/45301798/ggplot2-how-to-shade-an-area-above-a-function-curve-and-below-a-line