605-HW05-Probability

Michael Y.

September 29, 2019

Contents

Setup **1**

Part 1 - Proper probability distributions **2**

Choose independently two numbers B and C at random from the interval $[0, 1]$ with uniform density. 2

Part 2 - Probability of operations on B,C **3**

(a) B + C $< \frac{1}{2}$. 3

(b) BC $< \frac{1}{2}$. 4

(c) |B - C| $< \frac{1}{2}$. 5

(d) max{B,C} $< \frac{1}{2}$. 6

(e) min{B,C} $< \frac{1}{2}$. 7

Setup

```
knitr::opts_chunk$set(echo = TRUE,
                      fig.pos='H')
directory = "C:/Users/Michael/Dropbox/priv/CUNY/MSDS/201909-Fall/DATA605_Larry/20190929_Week05/"
knitr::opts_knit$set(root.dir = directory)

### Make the output wide enough
options(scipen = 999, digits=12, width=150)

### no other libraries required this week
```

Part 1 - Proper probability distributions

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.

p.59: **Definition 2.1** Let X be a continuous real-valued random variable.
A **density function for** X is a real-valued function f which satisfies

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

for all $a, b \in R$.

The uniform PDF on the interval [0,1] is defined as

$$f_X(x) = \begin{cases} 0, & if \quad x \leq 0 \\ 1, & if \quad 0 \leq x \leq 1 \\ 0, & if \quad x \geq 1 \end{cases}$$

Thus, the density is always non-negative, as required.

For the uniform density function on the interval [0,1], $f(x) = 1$, so

$$P(a \leq X \leq b) = \int_a^b 1 \cdot dx = b - a$$

for $0 \leq a \leq b \leq 1$.

P.60: The **probability** that the outcome of the experiment falls in an interval $[a, b]$ is given by

$$P([a, b]) = \int_a^b f(x)dx$$

,
that is, by the area under the graph of the density function in the interval $[a, b]$.

For the interval [0,1], the probability is

$$P(0 \leq X \leq 1) = \int_0^1 1 \cdot dx = 1 - 0 = 1$$

thus the total probability equals 1, as required.

P.61: **Definition 2.2** Let X be a continuous real-valued random variable.
Then the **cumulative distribution function** (CDF) of X is defined by the equation $F_X(x) = P(X \leq x)$.

It is clear that X always takes on a value between 0 and 1, so the uniform **cumulative distribution function** of X is given by

$$F_X(x) = \begin{cases} 0, & if \quad x \leq 0 \\ x, & if \quad 0 \leq x \leq 1 \\ 1, & if \quad x \geq 1 \end{cases}$$

The distribution is always non-negative, as required.

Since B and C are independent and identially distributed to X, the above applies to each of B, C.

Part 2 - Probability of operations on B,C

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

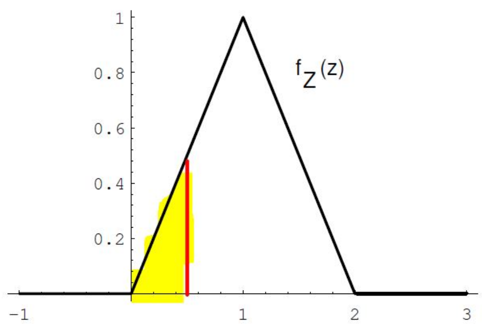
(a) $\mathbf{B + C < \frac{1}{2}}$.

The sum of two uniform variables corresponds to example 2.14 on pp. 63-64.

Let $Z = B + C$. Then the PDF for Z has positive density on the interval [0,2] and is defined as

$$f_Z(z) = \begin{cases} 0, & if \quad z \le 0 \\ z, & if \quad 0 \le z \le 1 \\ 2 - z, & if \quad 1 \le z \le 2 \\ 0, & if \quad z \ge 2 \end{cases}$$

Density of the sum z is the large triangle from (0,0) to (1,1) to (2,0):



The total area under this (large) triangle is $\frac{width*height}{2} = 2 \cdot 1 \cdot \frac{1}{2} = 1$.

The CDF for Z has positive density on the interval [0,2] and is defined as

$$f_Z(z) = \begin{cases} 0, & if \quad z \le 0 \\ \frac{z^2}{2}, & if \quad 0 \le z \le 1 \\ 1 - \frac{(2-z)^2}{2}, & if \quad 1 \le z \le 2 \\ 1, & if \quad z \ge 2 \end{cases}$$

The region in which density $= z < \frac{1}{2}$ is the smaller triangle from (0,0) to ($\frac{1}{2}$,0) to ($\frac{1}{2}$,$\frac{1}{2}$), shaded in yellow.

The area of this yellow triangle is $\frac{width*height}{2} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = 1(\frac{1}{2})^3 = 0.125 = \frac{1}{8}$, which can be computed from the CDF for $z = \frac{1}{2}$. Therefore, the **theoretical probability** is $\frac{1}{8} = 0.125$.

Simulate sum of $Z = B + C < \frac{1}{2}$:

```
n=1000000
B=runif(n,0,1)
C=runif(n,0,1)
Z=B+C
prob1 = sum(Z<=1/2)/n
# Probability that Z=B+C < 1/2 :
prob1
```

```
## [1] 0.125248
```

The **simulated probability** that $Z = B + C < \frac{1}{2}$ is 0.125248 .

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

This has to be separated into two cases, the results of which will be summed for the final answer:

(1) if $B \leq \frac{1}{2}$ then C can take on any value $\in [0,1]$ and $B \cdot C \leq \frac{1}{2}$.
On a grid plotting B vs. C, this will give a rectangle of width $\frac{1}{2}$ and height 1 for an area of $\frac{1}{2}$.

(2) if $B > \frac{1}{2}$ then we must have $C < \frac{0.5}{x}$ in order for $B \cdot C \leq \frac{1}{2}$.
On a grid plotting B vs. C, for $B \in [\frac{1}{2}, 1]$ this will give the following curve for C:

$$Z = \int_{b=\frac{1}{2}}^{b=1} \left[\frac{0.5}{c}\right] dc = \frac{1}{2}\log[c]\Bigg|_{b=\frac{1}{2}}^{b=1} = \frac{1}{2}\left[log(1) - log\left(\frac{1}{2}\right)\right] = \frac{1}{2}\left[0 - log\left(\frac{1}{2}\right)\right] = \frac{1}{2}\left[log\left(2\right)\right)] = log(sqrt(2))$$

So, the **theoretical** answer is $Pr\left(Z = B \cdot C \leq \frac{1}{2}\right) = \frac{1}{2} + \ln(\sqrt{2}) \approx 0.846574$

Simulate product $Z = B \cdot C < \frac{1}{2}$ **:**

```
n=1000000
B=runif(n,0,1)
C=runif(n,0,1)
Z=B*C
prob2 = sum(Z<=1/2)/n
# Probability that Z=B*C < 1/2 :
prob2
```

```
## [1] 0.846806
```

The **simulated probability** that $Z = B \cdot C < \frac{1}{2}$ is 0.846806 .

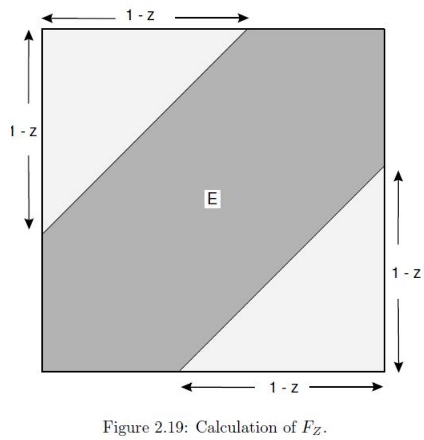


Figure 1: Grid of B vs C where |B-C| is less than z

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

This corresponds to Example 2.16 on pp.65-66, and figure 2.19 on p.67:

Let $Z = |B - C|$. Then the PDF for Z has positive density on the interval [0,1] and is defined as

$$f_Z(z) = \begin{cases} 0, & if \quad z \leq 0 \\ 2(1-z), & if \quad 0 \leq z \leq 1 \\ 0, & if \quad z > 1 \end{cases}$$

Then the CDF for Z is defined as

$$F_Z(z) = \begin{cases} 0, & if \quad z \leq 0 \\ 1-(1-z)^2, & if \quad 0 \leq z \leq 1 \\ 1, & if \quad z > 1 \end{cases}$$

For $z = \frac{1}{2}$, the **theoretical probability** that $Pr\left(|B - C| < z = \frac{1}{2}\right)$ is

$$Pr\left(z = |B - C| \leq \frac{1}{2}\right) = 1 - \left(1 - \frac{1}{2}\right)^2 = 1 - \left(\frac{1}{2}\right)^2 = 1 - \frac{1}{4} = \frac{3}{4} = 0.75$$

.

Removing the triangles of size $\frac{1}{8}$ from the square above yields the same result: $1 - \frac{1}{8} - \frac{1}{8} = \frac{6}{8} = \frac{3}{4} = 0.75$.

Simulate difference $Z = |B - C| < \frac{1}{2}$ **:**

```
n=1000000
B=runif(n,0,1)
C=runif(n,0,1)
Z=abs(B-C)
prob3 = sum(Z<=1/2)/n
# Probability that Z=|B-C| < 1/2 :
prob3
```

```
## [1] 0.74991
```

The **simulated probability** that $Z = |B - C| < \frac{1}{2}$ is 0.74991 .

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

$$
\begin{aligned}
Pr\left(Z = max(B,C) < \frac{1}{2}\right) &= Pr\left(B < \frac{1}{2} \wedge C < \frac{1}{2}\right) \\
&= Pr\left(B < \frac{1}{2}\right) \cdot Pr\left(C < \frac{1}{2}\right) \\
&= \frac{1}{2} \cdot \frac{1}{2} \\
&= \frac{1}{4} \\
&= 0.25
\end{aligned}
$$

where the second equality comes from independence of B and C .

The **theoretical probability** that $Z = max(B,C) < \frac{1}{2}$ is $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = 0.25$.

Simulate max: $Z = max(B,C) < \frac{1}{2}$ **:**

```r
n=1000000
B=runif(n,0,1)
C=runif(n,0,1)
Z1=cbind(B,C)
Z=apply(Z1,1,max)
prob4 = sum(Z<=1/2)/n
# Probability that Z=|B-C| < 1/2 :
prob4
```

```
## [1] 0.250632
```

The **simulated probability** that $Z = max(B,C) < \frac{1}{2}$ is 0.250632 .

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

$$
\begin{aligned}
Pr\left(Z = min(B,C) < \frac{1}{2}\right) &= Pr\left(\left(B < \frac{1}{2}\right) \vee \left(C < \frac{1}{2}\right)\right) \\
&= 1 - Pr\left(\left(B > \frac{1}{2}\right) \wedge \left(C > \frac{1}{2}\right)\right) \\
&= 1 - Pr\left(B > \frac{1}{2}\right) \cdot Pr\left(C > \frac{1}{2}\right) \\
&= 1 - \frac{1}{2} \cdot \frac{1}{2} \\
&= 1 - \frac{1}{4} \\
&= 0.75
\end{aligned}
$$

where the second equality stems from De Morgan's Law,

https://en.wikipedia.org/wiki/De_Morgan%27s_laws

and the third equality steps from independence of B and C .

The **theoretical probability** that $Z = max(B,C) < \frac{1}{2}$ is $1 - \frac{1}{2} \cdot \frac{1}{2} = 1 - \frac{1}{4} = \frac{3}{4} = 0.75$.

Simulate min: $Z = min(B,C) < \frac{1}{2}$ **:**

```r
n=1000000
B=runif(n,0,1)
C=runif(n,0,1)
Z1=cbind(B,C)
Z=apply(Z1,1,min)
prob5 = sum(Z<=1/2)/n
# Probability that Z=|B-C| < 1/2 :
prob5
```

```
## [1] 0.749477
```

The **simulated probability** that $Z = min(B,C) < \frac{1}{2}$ is 0.749477 .