Homework 1

Ashutosh Ekade

2023-10-09

Question 1: Expectations

1.a) Define both and explain the difference between (1) the expectation of a random variable and (2) the sample average.

Answer:

a) The expectation of a random variable is a measure of the central tendency of the random variable's distribution and is the average value with probability weighted averaging. It is defined as:

$\mathbb{E}[g(X)] = \sum_i g(X_i)p_x(x_i)$ for discrete random variable x if the series is convergent, and,

$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(X_i)f_x(x_i)dx$ for continuous random variable x.

b) The Sample Average

Sample average or the sample mean of a random draw from a larger population of X is defined as:

$\bar{X}_n = \frac{1}{n}\sum_i^n X_i$ where n is the sample size.

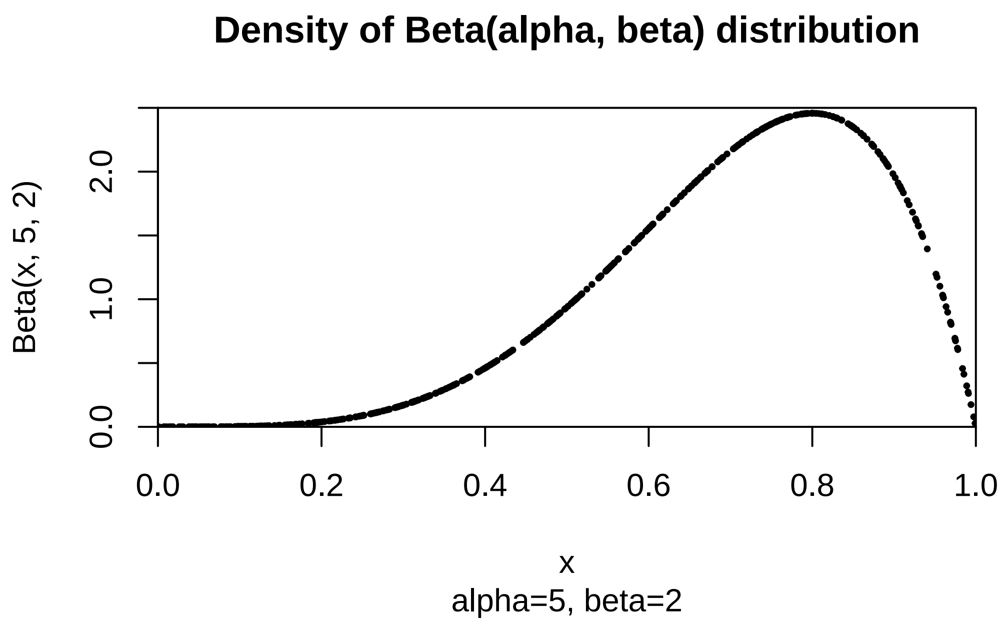
Difference:

Sample average is a function of the sample drawn from the population whereas the expected value is not. Expected value is a weighted average whereas sample average is not.

Question 2: LLN & CLT

2.a) Plot the density of a Beta(5,2) distribution over it's domain [0,1]. Make a publication-quality plot by changing any unwanted default plotting behavior and by adding relevant titles and labels. The first parameter of the Beta distribution is often labeled α and the second β. The `dbeta()` and related functions in R label these parameters as `shape1` and `shape2`.

```
# ...add answer here...
x <- runif(500) # generates 500 uniformly distributed random numbers between 0-1 (default
y <- dbeta(x, 5, 2)
plot(x=x, y=y, xlab="x", ylab="Beta(x, 5, 2)", main="Density of Beta(alpha, beta) distribu
     col = "black", pch = 20, cex = 0.5,
     xlim = c(0, 1),
     ylim = c(0, 2.5),
     yaxs="i", xaxs="i",
     sub = "alpha=5, beta=2")
```



2.b) State the Law of Large Numbers as simply as you can.

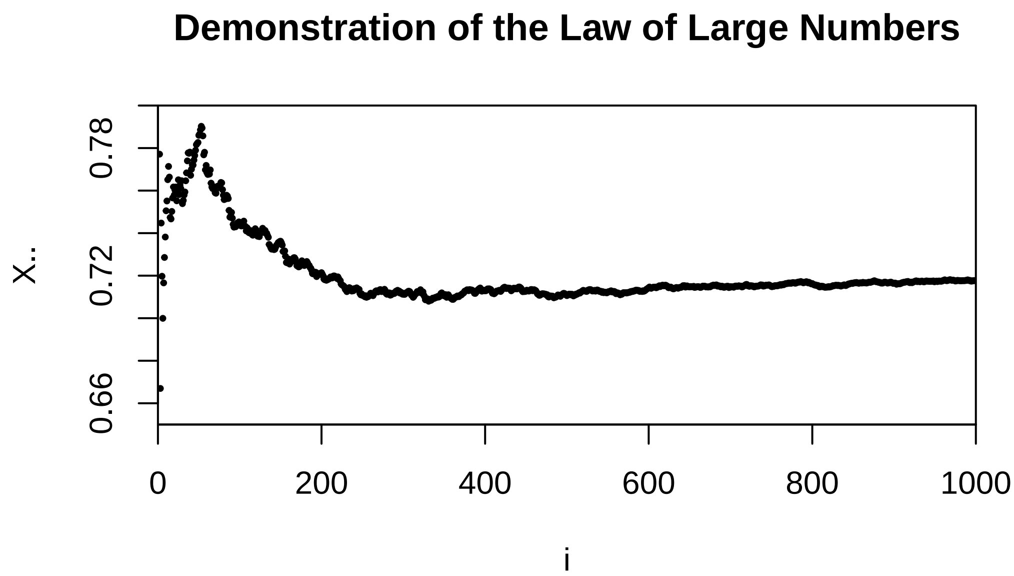
The law of large numbers states that if we repeat an experiment a large number of times and take the average of all the results then this average value approaches the expected value of the experiment.

2.c) Set the seed to the value 1234 (`set.seed(1234)`). Then take 1,000 random draws from the Beta(5,2) distribution using `rbeta()`. Calculate a running sample average. Specifically: calculate $\bar{X}_1 = x_1$, then calculate $\bar{X}_2 = (1/2)\sum_{i=1}^{2} X_i$, then calculate $\bar{X}_3 = (1/3)\sum_{i=1}^{3} X_i$. Continue until you have calculated $\bar{X}_{1000} = (1/1000)\sum_{i=1}^{1000} X_i$. Create a scatterplot with the values 1–1,000 on the horizontal axis and the 1,000 cumulative average values of $\bar{X}_i$ for $i = 1, \ldots, 1000$ you calculated on the vertical axis. Compare your value for $\bar{X}_{1000}$ to the $\mathbb{E}[X] = \alpha/(\alpha + \beta) = 5/7 = 0.7143$.

```r
set.seed(1234)
y <- vector(mode="list", length = 1000)
x <- rbeta(1000, 5, 2)
for (i in 1:1000) {
  y[i] <- mean(x[1:i])
}
plot(x=1:1000, y=y, xlab="i", ylab="X̄", main="Demonstration of the Law of Large Numbers",
     col = "black", pch = 20, cex = 0.5,
     xlim = c(0, 1000),
     ylim = c(0.65, 0.8),
     yaxs="i", xaxs="i")
```

```
Warning in title(...): conversion failure on 'X̄' in 'mbcsToSbcs': dot
substituted for <cc>

Warning in title(...): conversion failure on 'X̄' in 'mbcsToSbcs': dot
substituted for <84>
```




```r
print(5 / 7)
```

```
[1] 0.7142857
```

```r
print(y[1000])
```

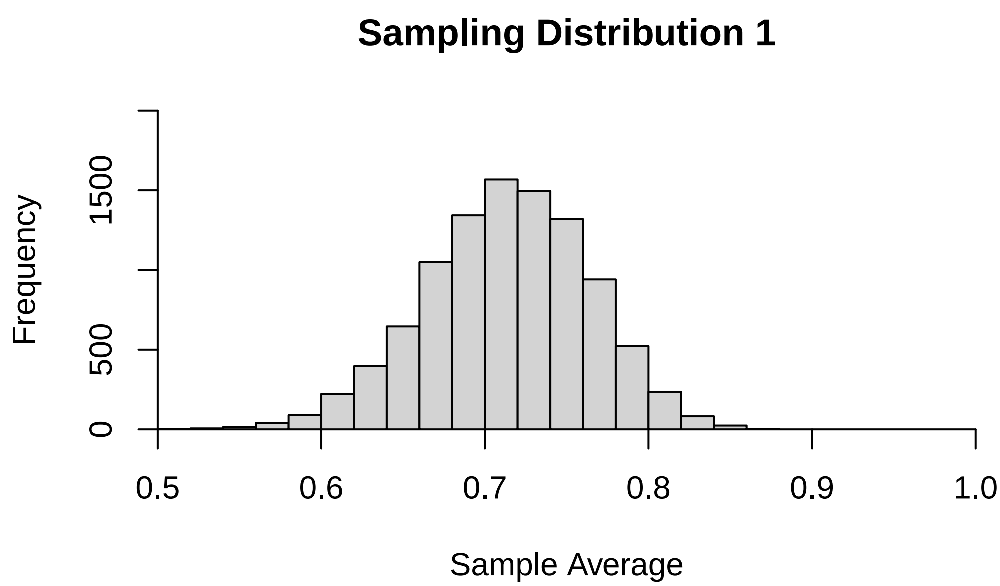
```
[[1]]
[1] 0.7176864
```

2.d) State the Central Limit Theorem as simply as you can.

The Central limit Theorem states that if we draw multiple large samples with replacement from a population then the sample mean will follow a normal distribution irrespective of the distribution of the population as long as the sample we draw is sufficiently large.

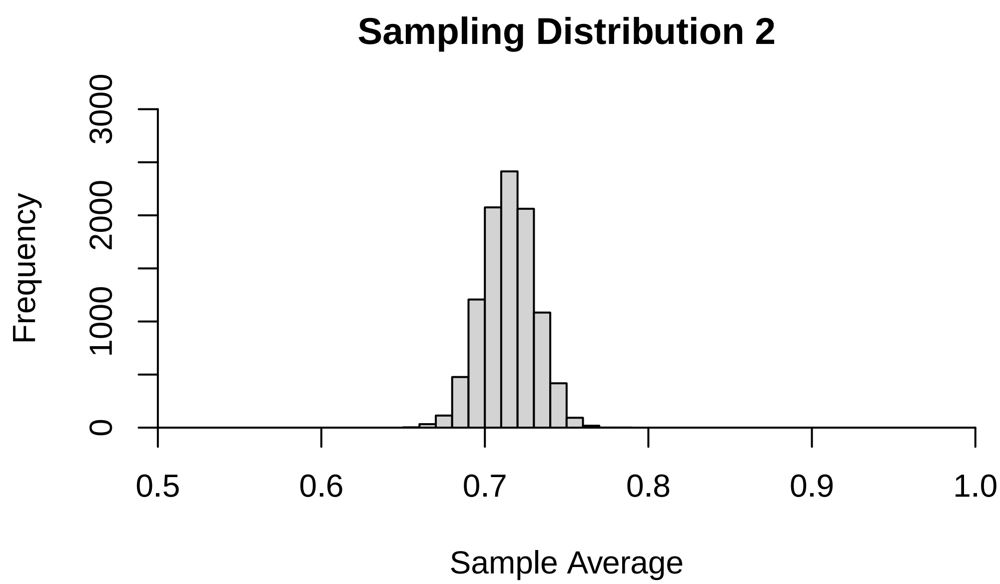
2.e) Make two plots. For the first plot, take D=10 draws from the Beta(5,2) distribution and calculate the sample average. Repeat the process of taking D=10 draws and finding the sample average R=10,000 times. Plot a histogram of the 10,000 sample averages. For the second plot, repeat the process with D=100 draws. These two histograms are called "sampling distributions."

```r
draws <- 10
iterations <- 10000
y <- numeric(length=iterations)
for (i in 1:iterations){
  x <- rbeta(draws, 5, 2)
  y[i] <- mean(x)
}
hist(y, main = "Sampling Distribution 1", xlab = "Sample Average", ylab = "Frequency",
     xlim = c(0.5, 1),
     ylim = c(0, 2000),
     yaxs="i", xaxs="i")
```




```r
draws <- 100
iterations <- 10000
y <- numeric(length=iterations)
for (i in 1:iterations){
  x <- rbeta(draws, 5, 2)
  y[i] <- mean(x)
}
hist(y, main = "Sampling Distribution 2", xlab = "Sample Average", ylab = "Frequency",
     xlim = c(0.5, 1),
```

```
     ylim = c(0, 3000),
     yaxs="i", xaxs="i")
```



Question 3: Linear Algebra

Suppose X and Y are defined as follows.

$$\mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 1 & 5 \\ 1 & 8 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 6 \\ 5 \\ 3 \\ 2 \end{bmatrix}$$

3.a) What is the rank of X? Provide a brief (approx 1 sentence) explanation. Check your work via R code using `Matrix::rankMatrix()`.

Rank of Matrix X is 2. Rank of a matrix is the highest number of independent rows or columns in a matrix. If we see in matrix X, row 3 can be created by adding row 1 & row 2, whereas row 4 is 2 times row 1. So we have 2 independent rows in X. We can verify that too in the code below.

```r
library(Matrix)
X <- matrix(c(1, 1, 1, 1, 1, 4, 5, 8), 4, 2); X
```

```
     [,1] [,2]
[1,]    1    1
[2,]    1    4
[3,]    1    5
[4,]    1    8
```

```r
t(X)
```

```
     [,1] [,2] [,3] [,4]
[1,]    1    1    1    1
[2,]    1    4    5    8
```

```r
y <- matrix(c(6, 5, 3, 2), 4, 1); y
```

```
     [,1]
[1,]    6
[2,]    5
[3,]    3
[4,]    2
```

```r
r <- rankMatrix(X)
print(r[1])
```

```
[1] 2
```

3.b) Calculate $\mathbf{X}'\mathbf{X}$. Use the `bmatrix` environment in Latex to typeset your answer. Check your work via R code.

$$\text{Matrix } \mathbf{X} = \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 1 & 5 \\ 1 & 8 \end{bmatrix} \text{ Matrix } \mathbf{X}' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \end{bmatrix} \mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 1 & 5 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 18 \\ 18 & 106 \end{bmatrix}$$

```r
t(X)%*%X
```

```
     [,1] [,2]
[1,]    4   18
[2,]   18  106
```

3.c) What is the rank of $\mathbf{X'X}$? Provide a brief explanation. Check your work via R code.

As we can see, in matrix $\mathbf{X'X}$, both the rows (or columns) are independent and one cannot be expressed as a multiple of the other. Hence the rank of $\mathbf{X'X}$ is 2.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 4 \\ 1 & 5 \\ 1 & 8 \end{bmatrix}$$

*Snippet for LaTex bmatrix is used with the help of ChatGPT.

```
r <- rankMatrix(t(X)%*%X)
print(r[1])
```

```
[1] 2
```

3.d) Find $(\mathbf{X'X})^{-1}$ "by hand" (as you would with paper and pencil) using the approach outlined on slide 47 of the Class 1 slides. Check your work via R code.

$$\mathbf{X'X} = \begin{bmatrix} 4 & 18 \\ 18 & 106 \end{bmatrix}$$

$$\mathbf{X'X}^{-1} = \frac{1}{|\mathbf{X'X}|} \begin{bmatrix} 106 & -18 \\ -18 & 4 \end{bmatrix} = \frac{1}{(4*106-18*18)} \begin{bmatrix} 106 & -18 \\ -18 & 4 \end{bmatrix} = \frac{1}{(100)} \begin{bmatrix} 106 & -18 \\ -18 & 4 \end{bmatrix} = \begin{bmatrix} 1.06 & -0.18 \\ -0.18 & 0.04 \end{bmatrix}$$

```
z <- t(X)%*%X
d <- solve(z); d
```

```
      [,1]  [,2]
[1,]  1.06 -0.18
[2,] -0.18  0.04
```

3.e) What is the rank of $(\mathbf{X}'\mathbf{X})^{-1}$? Provide a brief explanation. Check your work via R code.

All rows (or columns) of matrix $(\mathbf{X}'\mathbf{X})^{-1}$ are independent, so its rank is 2.

```
rankMatrix(d)[1]
```

```
[1] 2
```

3.f) Calculate $\mathbf{X}'\mathbf{y}$. Check your work via R code.

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 8 \end{bmatrix} \begin{bmatrix} 6 \\ 5 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} (1*6+1*5+1*3+1*2) \\ (1*6+4*5+5*3+8*2) \end{bmatrix} = \begin{bmatrix} 16 \\ 57 \end{bmatrix}$$

```
f <- t(X) %*% y; f
```

```
     [,1]
[1,]   16
[2,]   57
```

3.g) Use your results from 3d and 3f to calculate "by hand" $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Check your work via R code.

$$= \begin{bmatrix} 4 & 18 \\ 18 & 106 \end{bmatrix} \begin{bmatrix} 16 \\ 57 \end{bmatrix} = \begin{bmatrix} (4*16+18*57) \\ (18*16+106*57) \end{bmatrix} = \begin{bmatrix} 1090 \\ 6330 \end{bmatrix}$$

```
g <- z %*% f; g
```

```
     [,1]
[1,] 1090
[2,] 6330
```