Probability Distributions

Abel Isaias Gutierrez-Cruz

17/12/2021

Exercise 8.1

A company producing cereals offers a toy in every sixth cereal package in celebration of their 50th anniversary. A father immediately buys 20 packages.

```
pbinom(4, 20, prob = 1/6)
```

a. What is the probability of finding 4 toys in the 20 packages?

```
## [1] 0.7687492
```

```
pbinom(0, 20, prob = 1/6)
```

b. What is the probability of finding no toy at all?

```
## [1] 0.02608405
```

c. The packages contain three toys. What is the probability that among the 5 packages that are given to the family's youngest daughter, she finds two toys? This event follows a hypergeometric distribution

M = 3 x = 2 n =5 N = 20

```
dhyper(x = 2, m=3, n = 20, k=5)
```

```
## [1] 0.1016375
```

Exercise 8.2

A study on breeding birds collects information such as the length of their eggs 2 (in mm). Assume that the length is normally distributed with $\mu = 42.1$ mm and $\sigma = 20.82$. What is the probability of

a. finding an egg with a length greater than 50 mm? Given:

$P(X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$

$P(X \geq b) = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right)$

$P(X \geq b) = 1 - \Phi(\frac{50-42.1}{20.82}) = 1 - \Phi(0.37) \approx 0.35$

```
1 - pnorm(50, mean = 42.1, sd = 20.82)
```

```
## [1] 0.3521795
```

b. finding an egg between 30 and 40 mm in length? $P(30 \leq x \leq 40) = \Phi\left(\frac{40-42.1}{20.82}\right) - \Phi\left(\frac{30-42.1}{20.82}\right)$

```
pnorm(40, mean = 42.1, sd = 20.82) - pnorm(30, mean = 42.1, sd = 20.82)
```

```
## [1] 0.1792667
```

Exercise 8.3

A dodecahedron is a die with 12 sides. Suppose the numbers on the die are 1–12. Consider the random variable X which describes which number is shown after rolling the die once. What is the distribution of X ? Determine $E(X)$ and $Var(X)$

This event follows a Discrete Uniform Distribution $p = \frac{1}{12}$

$E(X) = \frac{k+1}{2} = \frac{13}{2} = 6.5$

$Var(X) = \frac{1}{12}(k^2-1) = \frac{1}{12}(12^2-1) = 11.9166667$

Exercise 8.4

Felix states that he is able to distinguish a freshly ground coffee blend from an ordinary supermarket coffee. One of his friends asks him to taste 10 cups of coffee and tell him which coffee he has tasted. Suppose that Felix has actually no clue about coffee and simply guesses the brand. What is the probability of at least 8 correct guesses?

The event follow a Binomial Distribution. The possible outputs are yes or not. $p = \frac{1}{2}$

The probability of being at least 8 times right is the same as not being wrong more than 2 times, so we could calculate only $P(x \geq 2)$

$P(x=0) = \left(\begin{array}{c}10\\0\end{array}\right)0.5^0(0.5)^{10} \approx 9.765625 \times 10^{-4}$

$P(x=1) = \left(\begin{array}{c}10\\1\end{array}\right)0.5^1(0.5)^{9} \approx 0.0097656$

$P(x=2) = \left(\begin{array}{c}10\\2\end{array}\right)0.5^2(0.5)^{8} \approx 0.0439453$

$P(x \geq 2) = P(x=0) + P(x=1) + P(x=2) = 0.0546875$

Code in R:

```
1-pbinom(7, 10, 0.5)
```

```
## [1] 0.0546875
```

Exercise 8.5

An advertising board is illuminated by several hundred bulbs. Some of the bulbs are fused or smashed regularly. If there are more than 5 fused bulbs on a day, the owner of the board replaces them, otherwise not. Consider the following data collected over a month which captures the number of days (n_i) on which i bulbs were broken:

Fused bulbs	0	1	2	3	4	5
n_i	6	8	8	5	2	1

a. Suggest an appropriate distribution for X : "number of broken bulbs per day". If we consider the high amount of bulbs and the low amount of bulbs fused, we can propose the Poisson distribution.

b. What is the average number of broken bulbs per day? What is the variance? $\bar{x} = \frac{1}{30}(0 \cdot 6 + 1 \cdot 8 + 2 \cdot 8 + 3 \cdot 5 + 4 \cdot 2 + 5 \cdot 1) = 1.7333333$

$s^2 = \frac{1}{30}(0 \cdot 6 + 1^2 \cdot 8 + 2^2 \cdot 8 + 3^2 \cdot 5 + 4^2 \cdot 2 + 5^2 \cdot 1) - 1.7^2 = 1.7290044$

c. Determine the probabilities P(X = x) using the distribution you chose in (a) and using the average number of broken bulbs you calculated in (b). Compare the probabilities with the proportions obtained from the data.

.	f_i	$Po(1.73)$
$P(X = 0)$	0.2	0.1772844
$P(X = 1)$	0.267	0.306702
$P(X = 2)$	0.267	0.2652973
$P(X = 3)$	0.167	0.1529881
$P(X = 4)$	0.067	0.0661673
$P(X = 5)$	0.033	0.0228939

```
1- ppois(5, 1.73)
```

d. Calculate the probability that at least 6 bulbs are fused, which means they need to be replaced.

```
## [1] 0.008666973
```

Almost 1% of the days the owner replace bulbs

e. Consider the random variable Y : "time until next bulb breaks". What is the distribution of Y ? Exponential distribution with $\lambda = 1.73$

f. Calculate and interpret $E(Y)$ $E(Y) = \frac{1}{\lambda} = \frac{1}{1.73} = 0.5780347$

Exercise 8.6

A country has a ratio between male and female births of 1.05 which means that 51.22 % of babies born are male.

a. What is the probability for a mother that the first girl is born during the first three births? Probabilities of getting a girl: $p = 1 - 0.5122 = 0.4878$

So:

```
pgeom(2, 0.4878)
```

```
## [1] 0.8656249
```

b. What is the probability of getting 2 girls among 4 babies? To answer this question we can use Binomial Distribution

```
dbinom(2, 4, 0.4878)
```

```
## [1] 0.3745536
```

Exercise 8.7

A fishermen catches, on average, three fish in an hour. Let Y be a random variable denoting the number of fish caught in one hour and let X be the time interval between catching two fishes. We assume that X follows an exponential distribution.

a. What is the distribution of Y ? It is a Poisson distribution

b. Determine $E(Y)$ and $E(X)$. $E(Y) = \lambda = 3$ $E(X) = \frac{1}{\lambda}$ 0.3333333

c. Calculate $P(Y = 5)$ and $P(Y < 1)$ $P(Y = 5)$

```
dpois(5, 3)
```

```
## [1] 0.1008188
```

$P(Y < 1)$

```
ppois(0, 3)
```

```
## [1] 0.04978707
```

Exercise 8.8

A restaurant sells three different types of dessert: chocolate, brownies, yogurt with seasonal fruits, and lemon tart. Years of experience have shown that the probabilities with which the desserts are chosen are 0.2, 0.3, and 0.5, respectively.

a. What is the probability that out of 5 guests, 2 guests choose brownies, 1 guest chooses yogurt, and the remaining 2 guests choose lemon tart? The event follows a Multinomial Distribution

```
dmultinom(c(2, 1, 2), prob = c(0.2, 0.3, 0.5))
```

```
## [1] 0.09
```

```

dmultinom(c(0, 0, 3), prob = c(0.2, 0.3, 0.5))
```

b. Suppose two out of the five guests are known to always choose lemon tart. What is the probability of the others choosing lemon tart as well?

```
## [1] 0.125
```

c. Determine the expectation assuming a group of 20 guests. $n = 20$ $E(X) = (20 \cdot 0.2, 20 \cdot 0.3, 20 \cdot 0.5) = (4, 6, 10)$

Exercise 8.9

A reinsurance company works on a premium policy for natural disasters. Based on experience, it is known that W = "number of natural disasters from October to March" (winter) is Poisson distributed with λ_W = 4. Similarly, the random variable S = "number of natural disasters from April to September" (summer) is Poisson distributed with $\lambda_S = 3$. Determine the probability that there is at least 1 disaster during both summer and winter based on the assumption that the two random variables are independent.

Consider independent both variables: $P(S \geq 1, W \geq 1) = P(S \geq 1) \cdot P(W \geq 1)$

```
(1- ppois(0, 4)) * 1-(ppois(0, 3))
```

```
## [1] 0.9318973
```

Exercise 8.10

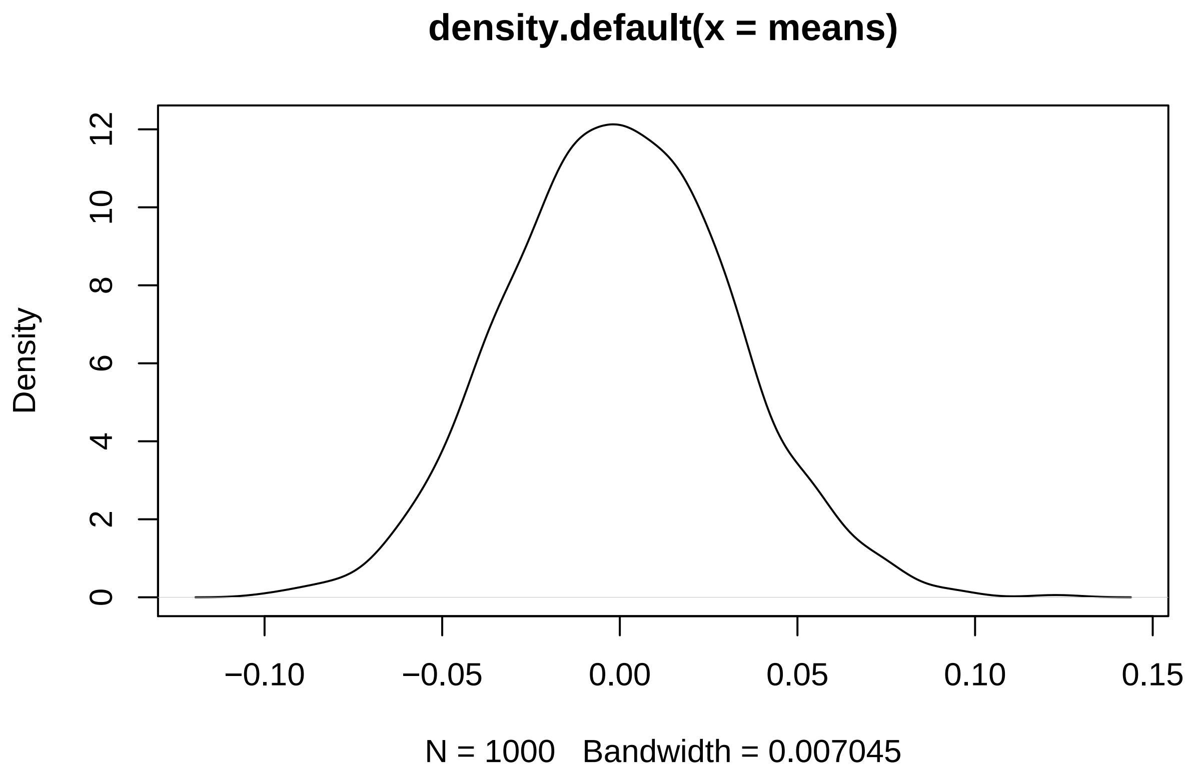
Read Appendix C.3 to learn about the Theorem of Large Numbers and the Central Limit Theorem.

```
set.seed(23)
repetitions <- 1000

means <- c(rep(NA, repetitions))
for (i in 1:repetitions){
    means[i] <- mean(rnorm(1000))
}

plot(density(means))
```

a. Draw 1000 realizations from a standard normal distribution using R and calculate the arithmetic mean. Repeat this process 1000 times. Evaluate the distribution of the arithmetic mean by drawing a kernel density plot and by calculating the mean and variance of it.




```
mean(means)
```

```
## [1] -4.76251e-05
```

```
var(means)
```

```
## [1] 0.0009899252
```

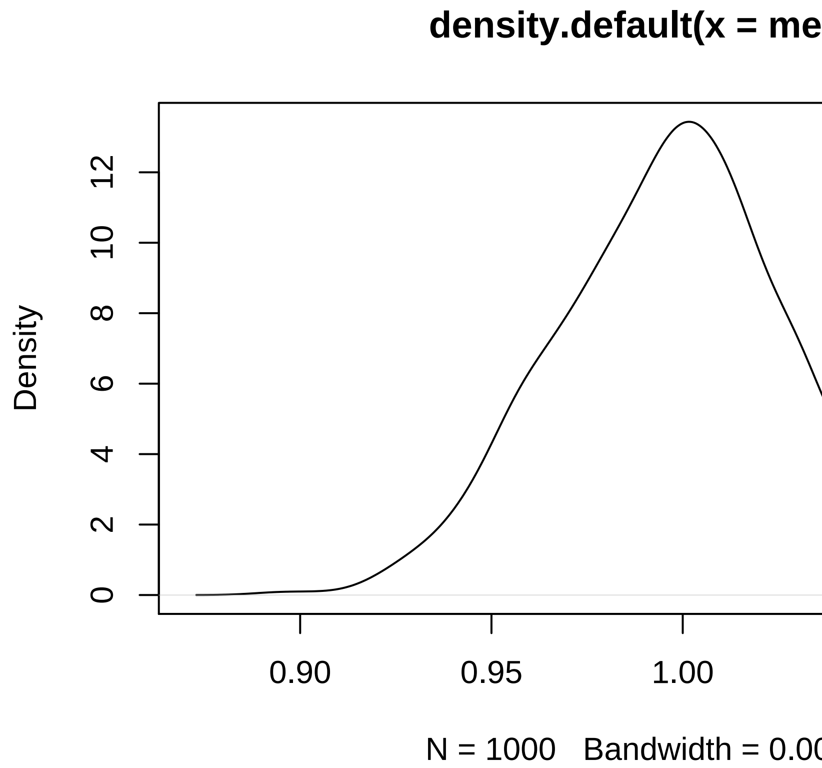
The arithmetic is close to zero. And the variance is close to $\sigma^2/n = 1/1000 = 0.001$

```
set.seed(23)
repetitions <- 1000

means2 <- c(rep(NA, repetitions))
for (i in 1:repetitions){
    means[i] <- mean(rexp(1000))
}

plot(density(means))
```

b. Repeat the procedure in (a) with an exponential distribution with $\lambda = 1$. Interpret your findings in the light of the Central Limit Theorem.




```
mean(means2)
```

```
## [1] NA
```

```
var(means2)
```

```
## [1] NA
```

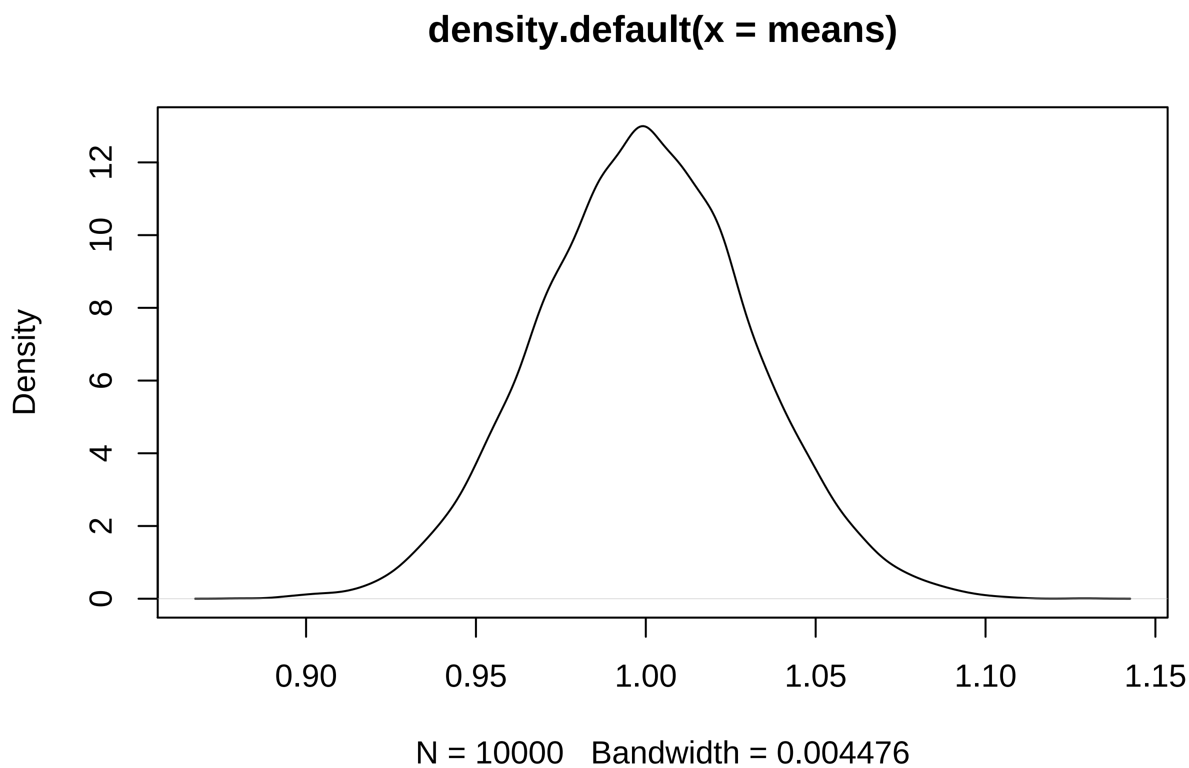
X_i do not necessarily need to follow a normal distribution for $\bar{X}$ to follow a normal distribution

```
set.seed(23)
repetitions <- 10000

means2 <- c(rep(NA, repetitions))
for (i in 1:repetitions){
    means[i] <- mean(rexp(1000))
}

plot(density(means))
```

c. Repeat the procedure in (b) using 10,000 rather than 1000 realizations. How do the results



change and why?

```
mean(means2)
```

```
## [1] NA
```

```
var(means2)
```

```
## [1] NA
```

Increasing the number of repetitions makes the distribution look closer to a normal distribution