Central Limit Theorem Exercises

Marcar esta página

Homework fecha límite Jul 13, 2022 19:00 CEST

For these exercises, we will be using the following dataset:

```
library(downloader)
url <- "https://raw.githubusercontent.com/genomicsclass/dagdata/master/inst/extdata/mice_
filename <- basename(url)
download(url, destfile=filename)
dat <- na.omit( read.csv(filename) )
```

Central Limit Theorem Exercises #1

0/1 punto (calificado)

If a list of numbers has a distribution that is well approximated by the normal distribution, what proportion of these numbers are within one standard deviation away from the list's average?

Use the `pnorm()` function. You can look up more information with `?pnorm`.

1

1

Enviar Ha realizado 1 de 5 intentos

Central Limit Theorem Exercises #2

1 punto posible (calificable)

What proportion of these numbers are within two standard deviations away from the list's average?

Enviar Ha realizado 0 de 5 intentos

Central Limit Theorem Exercises #3

1 punto posible (calificable)

What proportion of these numbers are within three standard deviations away from the list's average?

Enviar Ha realizado 0 de 5 intentos

Central Limit Theorem Exercises #4

1 punto posible (calificable)

Define `y` to be the weights of males on the control diet.

What proportion of the mice are within one standard deviation away from the average weight?

Remember to use `popsd()` from **rafalib** for the population standard deviation.

Ocultar Notas

Enviar Ha realizado 0 de 5 intentos

Central Limit Theorem Exercises #5

1 punto posible (calificable)

What proportion of these numbers are within two standard deviations away from the list's average?

Enviar Ha realizado 0 de 5 intentos

Central Limit Theorem Exercises #6

1 punto posible (calificable)

What proportion of these numbers are within three standard deviations away from the list's average?

Enviar Ha realizado 0 de 5 intentos

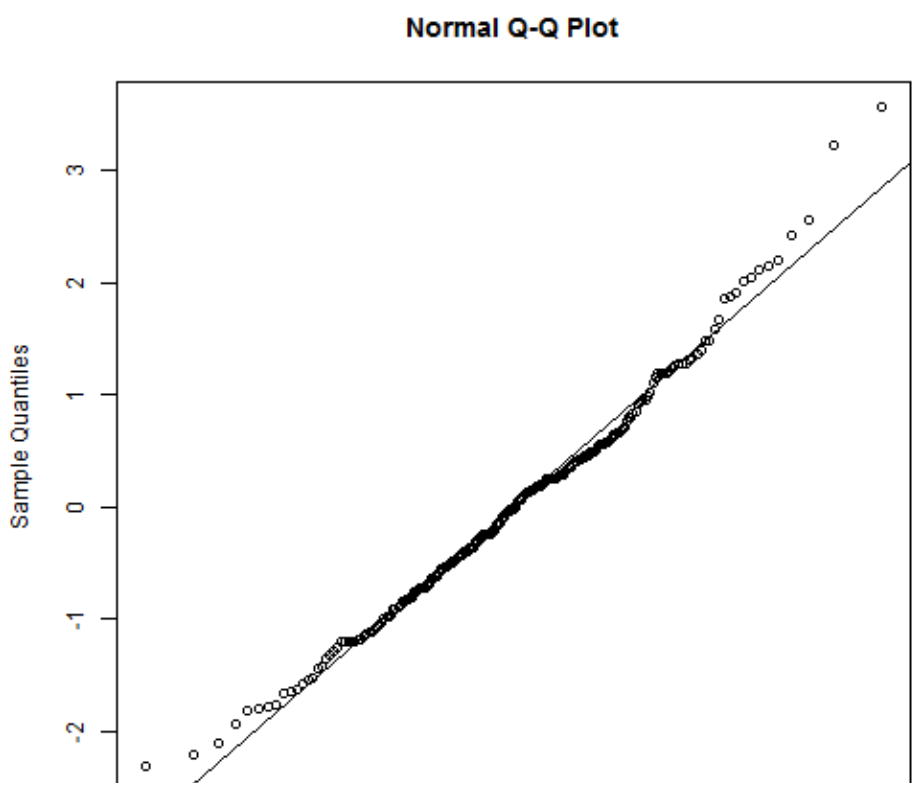
Central Limit Theorem Exercises #7

1 punto posible (calificable)

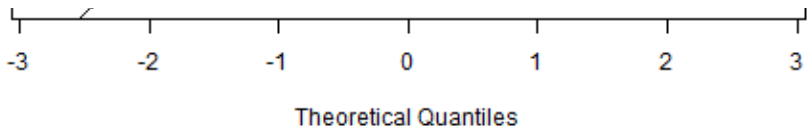
Note that the numbers for the normal distribution and our weights are relatively close. Also, notice that we are indirectly comparing quantiles of the normal distribution to quantiles of the mouse weight distribution. We can actually compare all quantiles using a qqplot.

```
qqnorm(z)
abline(0,1)
```

Which of the following best describes the qq-plot comparing mouse weights to the normal distribution?



Ocultar Notas



- The points on the qq-plot fall exactly on the identity line.
- The average of the mouse weights is not 0 and thus it can't follow a normal distribution.
- The mouse weights are well approximated by the normal distribution, although the larger values (right tail) are larger than predicted by the normal. This is consistent with the differences seen between question 3 and 6.
- These are not random variables and thus they can't follow a normal distribution.

Enviar Ha realizado 0 de 2 intentos

Central Limit Theorem Exercises #8

1 punto posible (calificable)

Here we are going to use the function `replicate()` to learn about the distribution of random variables. All the above exercises relate to the normal distribution as an approximation of the distribution of a fixed list of numbers or a population. We have not yet discussed probability in these exercises. If the distribution of a list of numbers is approximately normal, then if we pick a number at random from this distribution, it will follow a normal distribution. However, it is important to remember that stating that some quantity has a distribution does not necessarily imply this quantity is random. Also, keep in mind that this is not related to the central limit theorem. The central limit applies to averages of random variables. Let's explore this concept.

We will now take a sample of size 25 from the population of males on the chow diet. The average of this sample is our random variable. We will use the `replicate()` function to observe 10,000 realizations of this random variable. Set the seed at 1, then generate these 10,000 averages. Make a histogram and qq-plot of these 10,000 numbers against the normal distribution.

We can see that, as predicted by the CLT, the distribution of the random variable is very well approximated by the normal distribution.

```
y <- filter(dat, Sex=="M" & Diet=="chow") %>% select(Bodyweight) %>% unlist
set.seed(1)
avgs <- replicate(10000, mean( sample(y, 25)))
mypar(1,2)
hist(avgs)
qqnorm(avgs)
qqline(avgs)
```

What is the average of the distribution of the sample average?

Enviar Ha realizado 0 de 5 intentos

Central Limit Theorem Exercises #9

1 punto posible (calificable)

Ocultar Notas

1 punto posible (calificable)

What is the standard deviation of the distribution of sample averages (use `popsd()`)?

Enviar

Ha realizado 0 de 5 intentos

‹ Anterior

Siguiente ›

Ocultar Notas