Homework 4 Answers

EPI 525

Book exercises

3.38 Stenographer's typos

A very skilled court stenographer makes one typographical error (typo) per hour on average.

a

What are the mean and the standard deviation of the number of typos this stenographer makes in an hour?

Mean: 1, standard deviation: 1

b

Calculate the probability that this stenographer makes at most 3 typos in a given hour.

0.981

c

Calculate the probability that this stenographer makes at least 5 typos over 3 hours.

0.185

3.40 Osteosarcoma in NYC

Osteosarcoma is a relatively rare type of bone cancer. It occurs most often in young adults, age 10 - 19; it is diagnosed in approximately 8 per 1,000,000 individuals per year in that age group. In New York City (including all five boroughs), the number of young adults in this age range is approximately 1,400,000.

a

What is the expected number of cases of osteosarcoma in NYC in a given year?

11.2

b

What is the probability that 15 or more cases will be diagnosed in a given year?

0.161

c

The largest concentration of young adults in NYC is in the borough of Brooklyn, where the population in that age range is approximately 450,000. What is the probability of 10 or more cases in Brooklyn in a given year?

0.004

d

Suppose that in a given year, 10 cases of osteosarcoma were observed in NYC, with all 10 cases occurring among young adults living in Brooklyn. An official from the NYC Public Health Department claims that the probability of this event (that is, the probability of 10 or more cases being observed, and all of them occurring in Brooklyn) is what was calculated in part c). Is the official correct? Explain your answer. You may assume that your answer to part c) is correct. This question can be answered without doing any calculations.

Official is not correct

e

Suppose that over five years, there was one year in which 10 or more cases of osteosarcoma were observed in Brooklyn. Is the probability of this event equal to the probability calculated in part c)? Explain your answer.

0.01979

R exercises

R exercise 1

Below you will be using a dataset from Gapminder to complete a few R exercises.

Load all the packages you need below here

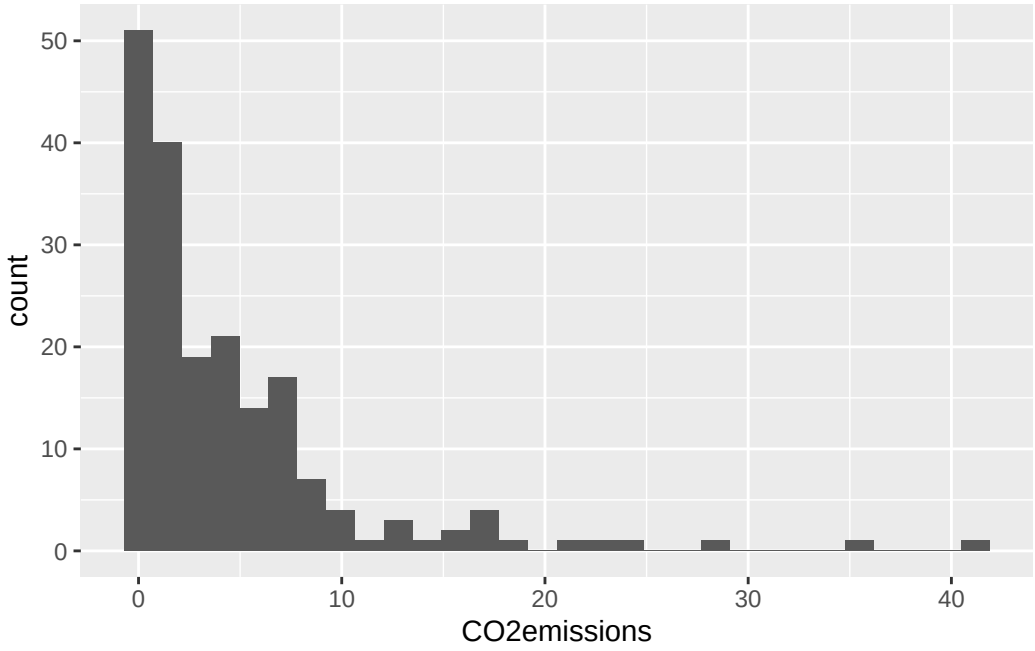
You don't need to do it all at once, you can add more libraries as you realize you need them.

Import dataset

Import the dataset called "Gapminder_2011_LifeExp_CO2.csv" You can find it in the student files under Data then Homework. You will need to download the file onto your computer, and use the correct file path to import the data.

Make a histogram

Using `ggplot2`, make a histogram of the variable `CO2emissions`.


50
40
30
20
10
0
count
0
10
20
30
40
CO2emissions