Assignment 3

May 5, 2022

1 Assignment 3 (Python)

1.1 Problem 1

An order for an automobile can specify either an automatic or a standard transmission, either with or without air conditioning, and with any one of the four colors red, blue, black, or white. **Use Python** to find the set of possible orders for this experiment.

1.2 Problem 2

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. We randomly choose 18 samples. **Using Python**, answer the following questions:

(a) What's the probability that exactly 4 of the chosen samples have the pollutant?

(b) What's the probability that at least 3 of the chosen samples have the pollutant?

(c) What's the expected number of samples having the pollutant? (use the basic definition $E[x] = \sum xf(x)$to calculate the answer & don't use any direct formulae for E(X))

(d) Plot the PMF of the random variable representing the number of samples having the pollutant. (Name the x-axis x & the y-axis $f(x)$)

(e) Plot the CDF of the random variable representing the number of samples having the pollutant. (Name the x-axis x & the y-axis $F(x)$)

1.3 Problem 3

A random variable has a continuous uniform distribution along the interval (3 to 12).

(a) **Use Python** to randomly get 5 observations of this RV, then plot the histogram of these values (try this for several times & observe that the output is changing each time).

(b) Is the output of the histogram plotting function as expected (i.e. representing a uniform distribution)? Explain how this plotting function works (specifying what's the x-axis & y-axis)?

(c) Generate 20, 100, 100000, 200000, 500000 observations of the RV & plot the histogram for each case including the case in part a. What do you observe as you increase the sample size?

1.4 Problem 4

Let a random variable X has a normal distribution of $\mu = 5$ and $\sigma = 2$. **Use Python** to find the probability that

(a) x lies between (2,8)

(b) x lies between (10,15).

(c) **Use Python** to draw the entire PDF of x and shade the area between the two values in (a) and (b). This can give you a graphical clue of the probability that X lies between two values without the need to calculate the probability exactly. Use this to visualize & compare the two required probabilities.

1.5 Problem 5

The breakdown time of an insulating fluid between electrodes at 34 kV. The times, in minutes, are as follows:

$0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, 72.89$

Using Python calculate the sample mean, sample median, sample range, and sample standard deviation.

1.6 Problem 6

A healthcare provider monitors the number of CAT scans performed each month in each of its clinics. The most recent year of data for a particular clinic follows (the reported variable is the number of CAT scans each month expressed as the number of CAT scans per thousand members of the health plan):

2.31, 2.09, 2.36, 1.95, 1.98, 2.25, 2.16, 2.07, 1.88, 1.94, 1.97, 2.02.

(a) **Use Python** to find a 95% CI on the mean number of CAT scans performed each month at this clinic.

(b) Repeat (a) using the standard normal distribution. Compare the results.