

Descriptive Statistics Using R AMPBA 2023

Assignment II Total Points: 100

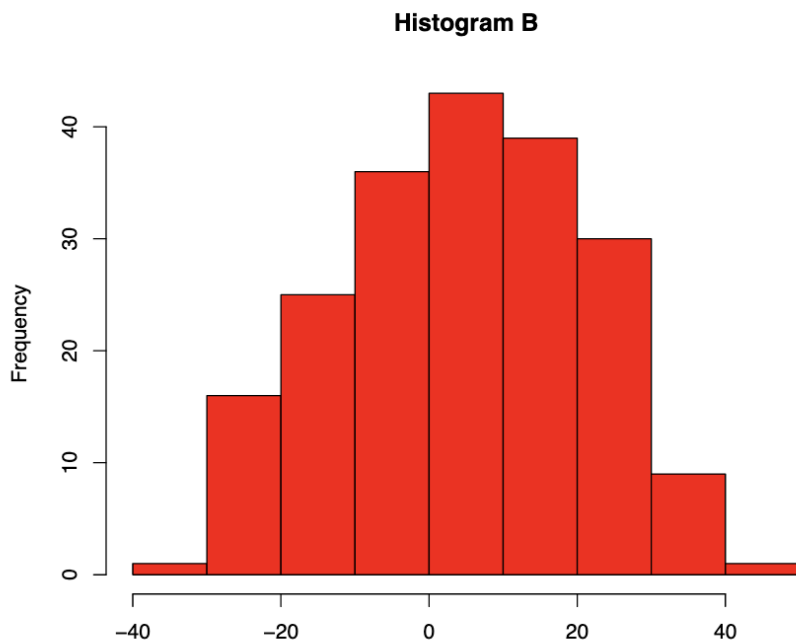
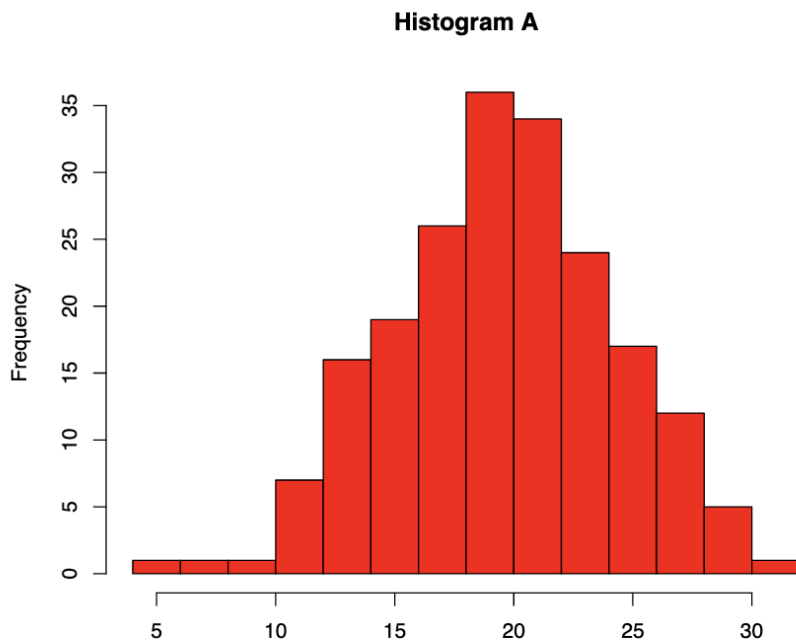
This deliverable has 30% weightage in the Consolidated Total Score.

Note: The Assignment submission form should be submitted as well. Your submission will not be considered without the Assignment Submission form being submitted.

1. What is the probability that, in a package of 12 tennis balls, 2 or more would be overweight if the probability of being overweight for a single ball is 10%? What is the RV here? What is its distribution? [5 points]

2. Consider the following two histograms:
 - a) Say which of the two histograms has a larger mean, or if they are about the same. Explain your answer. (2 points)
 - b) Say which of the two histograms has a larger standard deviation, or if they are about the same. Explain your answer. (2 points)
 - c) What simple change to this pair of graphs would have made the previous two questions easy to answer? (3 points)

[Total: 7 points]



3. For each of the following questions, say whether the random process is a binomial process or not, and explain your answer. As part of your explanation, comment on the

validity of each of the four conditions that must be true for a process to be a binomial process. [8 points, 4 each]

- a) One basketball player attempts 10 free throws, and the number of successful attempts is totaled.
- b) Ten different basketball players each attempt 1 free throw and the total number of successful attempts is totaled.

4. Let $B \sim \text{Bin}(20, 0.2)$. Compute the following probabilities. [10 points, a-c: 2 points each, d is worth 4 points]]

- a) $P(B = 4)$.
- b) $P(B \leq 1)$.
- c) $P(B > 1)$.
- d) Compute $E(B)$ and $\text{VAR}(B)$.

5. Hussain receives about six telephone calls from friends and family between 6 PM and 8:00 PM. What is the probability that Hussain receives more than two calls in the next 30 minutes? [10 points]

6. Consider a normal distribution with mean 5 and standard deviation 3. What is the probability that a single draw from this distribution has the value 5?

- a) 0
- b) 0.25
- c) 0.5
- d) 0.75

[3 points]

7. If $X_1 \sim N(\mu, \sigma^2)$ and $X_2 \sim N(\mu, \sigma^2)$ are iid (independent and identically distributed) normal random variables, then what's the difference between $2X_1$ and $X_1 + X_2$? Discuss both their distributions and parameters. [12 points, 6 points each]

Hint: Use properties of the Normal distribution in Chapter 4, Section 4–2 (Properties of the Normal Distribution) and ideas presented in Chapter 3, Section 3–3 (Sum and Linear Composites of Random Variables) of Complete Business Statistics by Aczel and Saundarpandian, 7th edition, to answer this question.

8. Let F be an RV that represents the operating temperature in Fahrenheit of one instance of a manufacturing process, and let $F \sim N(90, 25)$. Let C be an RV that represents the same process but measured in Celsius. Fahrenheit can be converted to Celsius using $C = \frac{5}{9}(F - 32)$. [20 points, 5 points each]

- a) Find the probability that one randomly selected instance of the process will have operating temperature greater than 93.8 Fahrenheit.
- b) C is also normally distributed. Use the properties of $E()$ and $Var()$ to find its mean and variance.
- c) Find the probability that one randomly selected instance of the process will have operating temperature below 29 Celsius.
- d) Find the Celsius temperature x such that the probability that the operating temperature in Celsius of one instance is less than x is .25

Hints:

1. Remember that the distribution of a Normal RV is represented as $N(\mu, \sigma^2)$.
2. In class, we saw that if we had an 'x' we could calculate $P(X < x)$ using the 'pnorm or pbinom or ppois' functions. What if you know the value of 'a' in $P[X < x] = a$ and want to find the value of 'x'? In R, use function 'qnorm'. For example, consider the below output to better understand this idea:

```

> # P(Z < 1): Z ~ N(0,1)
> pnorm(1)
[1] 0.8413447
> # P(Z < a) = 0.8413447. What is a?
> qnorm(0.8413447)
[1] 0.9999998
> #P(X < 2); X ~ N(3, 1)
> pnorm(2, mean = 3, sd = 1)
[1] 0.1586553
> #P(X < a) = 0.1586553?; What is a?
> qnorm(0.1586553, mean = 3, sd = 1)
[1] 2

```

9. Daily fluctuations of the French CAC-40 stock index from March to June 1997 seem to follow a normal distribution with mean of 2,600 and standard deviation of 50. Find the probability that the CAC-40 will be between 2,520 and 2,670 on a random day in the period of study. Solve this problem by converting the distribution of daily fluctuations of the French CAC-40 stock index from March to June 1997 to a standard normal distribution (Z). Show your work and verify that your calculations are correct using R. [10 points]

10. In the following studies, mention whether the study is observational or experimental. Give a very short (1-2 lines) justification. [15 points]

- a) A study took a random sample of teenagers and asked them about their bedtime habits. The data showed that teens stopped using electronic devices at least 2 hours before bedtime were more likely to go to sleep soundly than teens that used electronic devices right before going to bed.
- b) To study if completing all assignments in a course increases performance in the final exam, an instructor looked at data from her previous students. She found that in general, students that completed all the assignments, scored higher on the final exam than students that either didn't turn in their assignments or turned in incomplete assignments.
- c) To study if completing all assignments in a course increases performance in the final exam, an instructor randomly selected half her students and gave them assignments to turn in and readings to complete. To the other half, she only assigned the readings. She found that the students that completed the assignments scored higher on the final exam than the students who were randomized to receive the readings only.

General Instructions:

1. This is an Individual assignment.
2. Do NOT submit .zip files otherwise the submission will not be considered.
3. Please include your name and PGIDs in all files in the submission.
4. Any late submission will attract a penalty as mentioned in the course outline.
5. The honor code for this submission is **2N-b**.
6. **Please look through the honor code restrictions carefully before attempting the assignment as there will be strong consequences for breaking them.**

7. Please adhere to the given instructions, otherwise your submission will not be accepted, or a severe penalty will be applied

Deadline: 22nd January 2023, 11:55 pm