STATS 1000 / STATS 1004 / STATS 1504
Statistical Practice 1
Assignment 3
2020

DEADLINE:

- Sunday 6th September 2020 (Week 6) 5:00pm

CHECKLIST

☐: Have you shown all of your working, including probability notation where necessary?

☐: Have you given all numbers to **3 decimal** places.

☐: Have you included all R outputs and plots to support your answers where necessary.

☐: Have you made sure that all plots and tables each have a caption.

☐: Have you submitted your assignment via the online submission on MyUni before the deadline?

☐: Is your submission a single word document or pdf file - correctly orientated, easy to read? If not, penalties apply.

☐: Penalties for more than one document - 10% of final mark for each extra document. Note that you may resubmit and your final version is marked, but the final document should be a single file.

☐: If after the deadline, but within 24 hours, have you contacted us via email or Piazza and then submitted your assignment online via the online submission on MyUni?

☐: Penalties for late submission - within 24 hours 40% of final mark. After 24 hours, assignment is not marked and you get zero.

☐: Assignments emailed instead of submitted by the online submission on MyUni will not be marked and will recieve zero.

☐: Have you checked that the assignment submitted is the correct one, as we cannot accept other submissions after the due date.

☐: Do not write directly on the question sheet.

1. **Random variable**

   This question may be hand-written. Remember to attach it to the rest of the assignment.

   Notation recap: to write the probability that the random variable X has the value a, we write $P(X = a)$, not $P(X) = a$.

   A dishonest casino has a biased six-sided die. It still has the faces 1 to 6, but instead of each face being equally likely it has the following probabilities of each face:

Face	1	2	3	4	5	6
Probability of face	0.12	0.176	0.176	0.176	0.176	0.176

   Let X denote the random variable that is the face rolled. Using probability notation and showing your working, answer the following questions:

   (a) Verify that this is a valid discrete probability distribution. Remember we need to check two things.

   [2 marks]

   (b) What is the probability of rolling a six or a five?

   [2 marks]

   (c) What is the probability of rolling an even number?

   [2 marks]

   (d) Calculate the mean value of X.

   [2 marks]

   [Total: 8]

2. **Linear combinations**

   This question may be hand-written. Remember to attach it to the rest of the assignment.

   For each of the following games use the weighted die in Question 1. Calculate your true mean winnings. Show your working for full marks. *Hint: Let your winnings be denoted by Y and write an equation relating Y to the number on the die X.*

(a) The casino pays you $2 for each dot on the face of the dice corresponding to the number you rolled (e.g. rolling a 2 would get you $4, rolling a six would get you $12).

[2 marks]

(b) The casino pays you $2 for each spot of the number you roll, but you have to pay $3 to play.

[2 marks]

[Total: 4]

3. **Random allocation of observations**

This question must be typed up in Word. Remember to attach it to the rest of the assignment.

In Practical 5, we randomly allocated 50 observations of the "SPIP5.xlsx" data set into two groups of size 25 each. We did the random allocation in R and then analysed the two groups (Group A) and (Group B). In your Word document, include the following:

(a) A table of summary statistics for each group.

The summary statistics needed are the sample mean, sample standard deviation, sample median, sample IQR.

[2 marks]

(b) A side-by-side box-plot of the observations for each group.

[1 mark]

(c) A panel histogram of the observations for each group.

[1 mark]

(d) Compare the distribution of the observations for each group.

Remember to consider shape, location, spread and outliers.

[4 marks]

(e) Assume that the observations given in spreadsheet "SPIP5.xlsx" are the percentage loss of timber in 50 wood cubes, when exposed to a particular fungus after two types of wood preservatives, "A" and "B" are applied. According to your randomized data, which type of preserva-

tive is better for the wood? Explain your answer using the summary statistics obtained above? [3 marks]

[Total: 11]

4. **Binomial question**

   For full marks, please show all working and use probability notation.

   This question may be hand-written.

   A weighted coin is tossed 5 times. It is biased such that the probability of a head is 0.6.

   Let X be the random variable that represents the total number of heads obtained.

   (a) What is the distribution of X? Justify your answer.

   [5 marks]

   (b) Using R, calculate the following.

       i. The probability of obtaining no heads.

       [2 marks]

       ii. The probability of at least one head.

       [2 marks]

       iii. The probability of between 1 and 3 heads (inclusive).

       [2 marks]

       [Total: 11]

5. **Normal question**

   For full marks, please show all working and use probability notation.

   This question may be hand-written.

   The weight of male red kangaroos is normally distributed with a mean of 78 kg and a standard deviation of 6 kg.

   Let X be the random variable that represents the weight of a randomly selected male red kangaroo.

   Using R, calculate the following:

(a) The probability that a randomly selected male red kangaroo weighs less than 70 kg?

[2 marks]

(b) The probability that a randomly selected male red kangaroo weighs more than 90 kg or less than 70 kg?

[2 marks]

(c) The probability that a randomly selected male red kangaroo weighs between 70 and 90 kg?

[2 marks]

(d) The weight c such that 15% of all male red kangaroo are less than this weight?

[2 marks]

(e) Presentation marks (1 for presenting it in Word, 1 for appropriate figure captions, 1 for appropriate table captions).

[3 marks]

[Total: 11]

[[Assignment total: 45]]