Biostatistics – Internal-1
For MSc11, IBAB

R. Srivatsan

April 27, 2023

Answer all questions $(10 \times 6 = 60$ marks)

1. A list of 20 students consists of 10 boys and 10 girls. In a random draw, three names are drawn successively without replacement. What is the probability that the first draw is a girl, second draw is a boy and the third draw is a girl?

2. Suppose it is known that 5% of adults who take a medication feel the side effects. If 8 people take the medication, compute the probability that more than 5 of them will experience a side effect.

3. A field biologist observes a particular bird in a park everyday between 8 and 10 AM for many yesrs. Based on data acquired over several years, if 9 is the mean number of birds visiting the park between 8 and 10 AM, compute the probability of observing 14 birds in the same time period. What is the probability of observing less than 4 birds in the same interval?

4. In a Gaussian distribution of mean=40 and standard deviation 5, compute the probability of a random value
   (1) more than 52
   (2) less than 26
   (3) between 45 to 55
   (4) between 20 to 60

5. In a t distribution compute the probability of a random deviate with a value
   (1) more than 2.3
   (2) less than -2.65
   (3) between 2.1 to 3.1
   (4) between -2.0 to 2.0

6. Given a normally distributed population with a mean of 100 and standard deviation of 20, compute the probability $96 \leq \overline{x} \leq 108$ for an observed mean $\overline{x}$ based on a sample size of 25.

7. A data set of n samples has a sample mean of 29.0 and sample standard deviation of 3.8. For this data set, construct a
   (1) 90% confidence interval
   (2) 95% confidence interval
   (3) 99% confidence interval

8. The data set given below consists of measurements of $CaCO_3$ concentrations (mg/L) in 20 water samples.
   X = {130.8, 129.9, 131.5, 131.2, 129.5, 132.7, 131.5, 127.8, 133.7, 132.2,
   134.8, 131.7, 133.9, 129.8, 131.4, 128.8, 132.7, 132.8, 131.4, 131.3 }

(i) Draw a box-and-whisker diagram for this data by computing the required statistical parameters.

(ii) Compute a 95% Confidence interval on population mean for this data. Mark it on the bar plot.

9. A clinical study measured the weight of diabetic patients insured by two different insurance companies. The data is given below where the weight (in pounds of) 12 patients from each group are tabulated:

Company-1 : $252, 240, 205, 200, 170, 170, 320, 148, 214, 270, 265, 203$
Company-2 : $185, 310, 212, 238, 184, 136, 200, 270, 200, 212, 182, 225$

Compute a 95% confidence intervals on population means for these two data sets.

Plot a Box-Whisker diagram to compare these two data sets side by side, by marking the 95% confidence interval on the bars. Commnet on the result.

10. A quantity S is a function of two independent variables x and y through the following relation:

$$S = ax^2 + by^2.$$

(a) The variable x is measured to be 10.0 with an uncertainity of 2.6 and the variable y is measured to be 15.1 with an undertainity of 2.9. The value of the constants are $a = 5.0$ and $b = 12.0$. Compute the uncertainity in S.

(b) In another measurement, x and y are measured to be
x = 10.5 with a standard deviation 2.2
y = 14.9 with a standard deviation 2.6
Compute the standard deviation on S.