Biostatistics – Assignment-2
For MSc12, IBAB

R. Srivatsan

March 16, 2023

(Last date for submission : April 6 , 2023

1.If the arithmetic mean of 4 numbers 8.1, 5.9, 6.4, X is given as 7.0, find X.

2. Find the ratio of the arithmetic mean to the geometric mean of the following data set:
$A = \{12, 14, 13, 15, 16, 24, 17, 11\}$

3. Find the expectation value (ie., population mean) of the function $f(x) = 3x^3 + 7x$ in the interval [0,2].

4. Explain the concept of *probability density* in a continuous probability distribution. In what way is it different from the discrete probability ditribution.

5. The fasting serum triglyceride levels (in mg/dl) of 20 patients is given in the data set below:
{44, 46, 49, 52, 55, 62, 67, 72, 77, 80, 83, 86, 88, 90, 92, 94, 99, 100, 101, 106}
      (a) Compute the mean and standard deviation of the data set.
      (b) Draw a frequency histogram of the data by choosing a suitable bin width.

6. In a hospital, 90% of the patients who undergo a particular type of operation survive. If the outcomes of operations on patients are independent of each other, what is the probability that exactly 5 operations are required to observe a failure?

7. The number of single nuclotide polymorphisms (SNP) in the genome of a strange species was found to follow Poisson distribution with a mean value of 2 SNPs per 1000Kb interval. Compute the probability of observing 12 SNPs in a genomic region of length 3000Kb.

8. The probability that a student studies for her math test is 0.6. If she studies, the probability she will pass the test is 0.8. If she does not study, the probability she will pass the test is 0.4. If she passes the test, what is the probability she studied?

9. For a particular population, the lifetime probability of contracting glaucoma is 0.007 and the lifetime probability of contracting diabetes is 0.02 . It is also found that the lifetime probability of contracting both the diseases is 0.0008 .
      (a) What is the lifetime probability of contracting either glaucoma or diabetes?
      (b) What is the probability of contracting glaucoma for a person who has diabetes?
      (c) What is the probability of contracting diabetes for a person who has glaucoma?

10. The weight of pumpkins harvested from a farm are found to follow a Gaussian distribution with mean 3.2 Kg and variance 1.21 Kg.

(1) What is the probability of getting a pumpkin whose weight is more than 6.5 Kg?

(2) What is the probability of getting a pumpkin with weight less than 2.2 Kg?

(3) Compute the probability that the weight is between 2.8 to 3.9 Kg.

11. Suppose the gene expression levels in an experiment follow Normal distribution with a mean value of 350 (in some units) and standard deviation of 110.

What is the probability that a randomly selected gene has an expression level between 170 and 230 units?

12. The weight of pumpkins harvested from a farm are found to follow a Gaussian distribution with mean 3.2 Kg and variance 1.21 Kg. What is the probability of getting a pumpkin whose weight is more than 6.5 Kg?

13. Cancer tissues of types T1,T2 and T3 are prepared for an experiment and are put into three boxes.

The first box contains 8 samples of tissue T1, 6 samples of tissue T2 and 5 samples of tissue T3.

The second box has 11 samples of tissue T1, 10 samples of tissue T2 and 7 samples of tissue T3.

The third box has 6 samples of tissue T1, 8 samples of tissue T2 and 4 samples of tissue T3.

During the experiment, a box is chosen randomly and from the selected group, two tissue samples are randomly picked up. It is found that both the tissues are of type T2.

If all the three boxes have equal probability of being chosen, then using Bayes theorem compute the probability that both the selected T2 type tissues are (i) from the first box; (ii) from the second box; (iii) from the third box.

14. The blood pressure of 12 football players measured before a match are given below:

$$B = \{82, 86, 89, 80, 78, 83, 77, 94, 90, 88, 82, 82\}$$

Find the mean, standard deviation and the median of this data set.

15. The following data set contains the measurement of glomerular filtration rate (GFR) of 12 pediatric renal transplant recipients:

$$G = \{18, 42, 36, 32, 58, 48, 43, 67, 60, 62, 37, 88\}$$

(a) Construct the mean, median, standard deviation, first quartile and third quartile for this data.

(b) Construct a box-and-whisker plot.