ARTIFICIAL INTELLIGENCE

Lab3. *Random variables and sampling distributions*

1. Generate randomly a single variable from binomial distribution (`rbinom` function) with $n = 30$ and $p = 0.3$. Verify, what will be the result if probability of success, p, increases.

2. Generate randomly 20 variables from binomial distribution with $n = 30$ and $p = 0.3$. Draw probability histogram (line graph) of the distribution. Compare the graphs if probability of success, p, increases.

3. The study on the drinking water pollution shows that 30% of all wells in the city are contaminated. Five wells were randomly chosen and the quality of the water was verified. Let the random variable S denote the number of contaminated wells among chosen. What is the probability distribution of S? Using functions available in R compute probability that:
   (a) exactly 3 wells are contaminated;
   (b) at least 3 wells are contaminated;
   (c) less than 3 wells are contaminated.

4. The probability that a fluorescent bulb burns for at least 500 hours is 0.9. Let B denote the number of bulbs among 8 chosen, which burns for at least 500 hours. Give the distribution of B. Compute:
   (a) $P(B=8)$; (b) $P(B=7)$; (c) $P(B>5)$; (d) $E(B)$; (e) $SD(B)$.
   Interpret tasks (d) and (e).

5. Power cell failures in a satellite have exponentially distributed times with $\lambda = 0.01$. Presently only 2 cells remain functioning. They are arranged in parallel and have independent lives, so that the satellite may function as long as at least 1 power cell works. Draw the density function using `curve` and `dexp` functions (choose the interval for the arguments properly to see the shape of exponential distribution). Apply `pexp` function to find the probability that a particular cell:
   (a) will survive for 200 or more days;
   (b) will fail on or before 100 days;
   (c) will fail before 500 days.

6. The magnitudes of earthquakes recorded in a region of North America can be modeled by an exponential distribution with mean 2.4, as measured on the Richter scale. Draw the density function (choose the interval for the arguments properly to see the shape of exponential distribution). Find the probability that the next earth quake to strike this region will
   (a) exceed 3.0 on the Richter scale;
   (b) fall between 2.0 and 3.0 on the Richter scale.
   Verify the results using integrals. Verify expectation using integrals.

7. Wires manufactured for use in a certain computer system are specified to have resistances between 0.12 and 0.14 ohm. The actual measured resistances of the wires produced by Company A have a normal probability distribution with a mean of 0.13 ohm and a standard deviation of 0.005 ohm. Draw the density function (`dnorm`) using relevant interval for arguments to see the shape of the distribution. What is the probability that a randomly selected wire from Company A production will meet the specifications?

8. Several sample observations will be obtained of paint-drying times. The drying time may be assumed to be normally distributed with a mean of 2 hours and the standard deviation 15 min. Draw the density function. Find the probability that the drying time is between 1h 50min and 2h 15min. Compare the results if you use minutes as parameters with the results obtained for decimal coding.

9. Mopeds (small motorcycles with an engine capacity below 50 cm^3) are very popular in Europe because of their mobility, ease of operation, and low cost. The article in a specialist journal described a rolling bench test for determining maximum vehicle speed. A normal distribution with mean value 46.8 km/h and standard deviation 1.75 km/h is postulated. Consider randomly selecting a single such moped. What is the probability that the maximum speed
   (a) is at most 50 km/h?
   (b) is at least 48 km/h?

10. Generate 500 numbers related to the random sample from binomial distribution with $n = 20$ and $p = 0.2$.
    (a) Divide the window with graphics into three parts. Draw probability histogram of generated data, binomial distribution and normal distribution (with relevant parameters). Compare the results. Use `discrete.histogram`, `plot` and `curve` functions.
    (b) On the same graph draw probability histogram of generated data and normal distribution (use `hist` and `curve` functions).
    (c) Verify approximation fit for various values of n, p and number of generated data.

11. Suppose that 25% of all students at a large public university receive financial aid. Let X be the number of students in a random sample of size 100 who applied for financial aid. Compute exact and approximate probability, that at most 15 students receive the aid.

12. Generate 200 samples of size 30 from standard normal distribution and for each sample compute mean. On the same graph draw the probability histogram of computed means and the density function of normal distribution with relevant parameters. Use `hist` and `curve` functions. Compare the results.

13. Generate 200 samples of size 10 from binomial distribution with $n = 30$ and $p = 0.4$ and for each sample compute mean. On the same graph draw the probability histogram of computed means and the density function of normal distribution with relevant parameters. Use `hist` and `curve` functions. Verify the results if the sample size increases, for example 10 (original), 50, 100.

14. Resistors of a certain type have resistances that average 200 ohms with a standard deviation of 10 ohms. 50 of these resistors are to be used in a circuit. Assuming that the resistance is normally distributed find the probability that
    (a) the <u>average</u> resistance of the 50 resistors is between 199 and 202 ohms;
    (b) the <u>total</u> resistance of the 50 resistors does not exceed 10,020 ohms.

15. The blood cholesterol levels of a population of workers have mean 202 and standard deviation 14. If a sample of 64 workers is selected, approximate the probability that the <u>sample mean</u> of their blood cholesterol levels is between 198 and 206.

16. The strength of a thread is a random variable with mean 0.5 lb and standard deviation of 0.2 lb. Assume the strength of a rope is the sum of the strengths of the threads in the rope. Find the probability that a rope consisting of 100 threads will hold 47 pounds.