DISTRIBUTION ASSIGNMENT

QUESTION 1: Simulate 30 rolls with =RANDBETWEEN(1,6). What is the rolling a 3 exactly 5 times? (Hint: Use BINOM.DIST).

ANSWER: This is a binomial probability problem because:

- There are a fixed number of trials (30 rolls).
- Each roll has two outcomes: success (rolling a 3) or failure (not a 3).
- Probability of success remains constant: $p = \frac{1}{6}$.
- Rolla are independent.

Formula:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

Where:

- $n = 30$
- $x = 5$
- $p = \frac{1}{6}$

QUESTION 2: Generate 100 values in Excel using the continuous uniform distribution RAND() and plot a histogram. Describe the shape of the distribution.

ANSWER: The histogram shows an approximately uniform distribution. The values are spread evenly between 0 and 1, with no central peak or skewness. Minor variations occur due to randomness, but the overall distribution remains flat.

QUESTION 3: A dataset has a mean of 50 and a Standard deviation of 5. What percentage of values lie between 45 and 55 if the data follows a normal distribution?

ANSWER: GIVEN:

- Mean μ = 50
- Standard deviation σ = 5

Convert to z-scores:

$$z_1 = \frac{45 - 50}{5} = -1, \qquad z_2 = \frac{55 - 50}{5} = +1$$

In a normal distribution, approximately **68.27%** of values lie within ±1 **standard deviation** of the mean.

QUESTION 4: What is the concept of standardization (z-score), and why is it important in data analysis? Explain the formula and how standardization transforms a dataset.

ANSWER: Standardization converts raw data into **z-scores**, which show how far a value lies from the mean in units of standard deviation.

Formula:

$$z = \frac{x - \mu}{\sigma}$$

Why is it important?

- Makes variables with different units comparable.
- Helps identify outliers.
- Allows use of standard normal tables.
- Converts ay normal distribution into a standard normal distribution with mean 0 and SD 1.

QUESTION 5: What is Kurtosis and their type?

ANSWER: Kurtosis measures the **tailedness and peakedness** of a distribution.

Types:

- **Mesokurtic:** Normal-shaped curve (kurtosis ≈ 3).
- **Leptokurtic:** High peak and heavy tails, more extreme values.
- **Platykurtic:** Flat peak and light tails, fewer extremes.

Kurtosis helps in understanding the likelihood of outliers.

QUESTION 6: Explain why the uniform distribution is a good model for the outcome of rolling a fair die.

ANSWER: A fair die has:

- Six possible outcomes.
- Each outcome has equal probability (1/6).

- Outcomes are independent.

This matches the definition of a **discrete uniform distribution**, where every outcome is equally likely.

QUESTION 7: Use Excel to compute the probability of getting at least 8 successes in 15 trails with a success probability of 0.5.

ANSWER: Given:

- $n = 15$
- $p = 0.5$

"At least 8 successes" means $X \geq 8$.

Reason: When probability is 0.5 and trials are odd, the distribution is symmetric, so half the probability lies above the center.

QUESTION 8: How does log transformation help in stabilizing variance and making data more normally distributed?

ANSWER: Log transformation is used when data are positively skewed or when variance increases with the mean.

Benefits:

- Stabilizes variance.
- Reduces skewness.
- Makes data more normally distributed.
- Converts multiplicative relationships into additive ones.

It improves model accuracy and validity in statistical analysis.