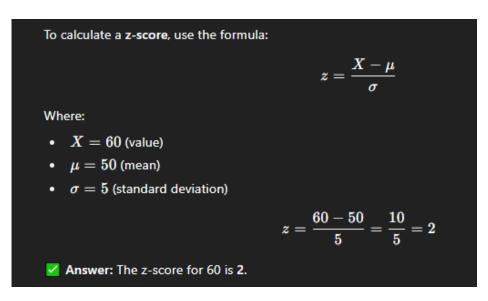
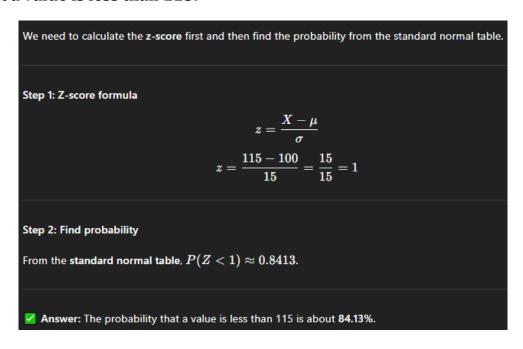
Day 25 | Discrete and Continuous Distribution

Normal Distribution

1. A normally distributed dataset has a mean of 50 and a standard deviation of, what is the z-score for a value of 60?



2. For a normal distribution with μ = 100 and σ = 15, find the probability that a value is less than 115.



3. Explain why the normal distribution is called a "bell curve" and name at least two real-world examples where it occurs.

The normal distribution is called a "bell curve" because its graph is symmetric and shaped like a bell—it rises to a peak at the mean (most frequent value) and tapers off evenly on both sides.

Real-world examples:

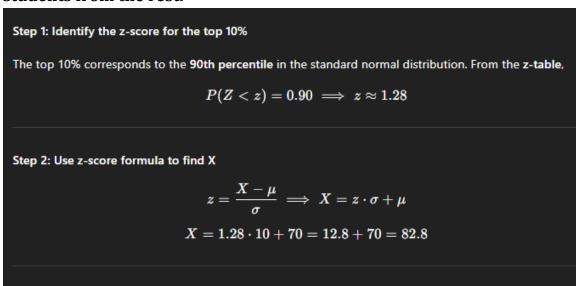
- 1. Human heights most people are around the average height, with fewer very short or very tall individuals.
- 2. Exam scores in large classes, most students score near the average, with few very high or very low scores.

Other examples include weights of newborns and IQ scores.

4. If the height of adult males is normally distributed with mean = 175 cm and σ = 8 cm, find the percentage of males taller than 180 cm.

We can solve this using **z**-scores and the standard normal distribution. Step 1: Compute **z**-score $z = \frac{X - \mu}{\sigma} = \frac{180 - 175}{8} = \frac{5}{8} = 0.625$ Step 2: Find probability for Z < 0.625 From the standard normal table: $P(Z < 0.625) \approx 0.734$ Step 3: Find probability for Z > 0.625 P(Z > 0.625) = 1 - 0.734 = 0.266 ✓ Answer: Approximately 26.6% of adult males are taller than 180 cm.

5. The marks of students in a class follow a normal distribution with mean = 70 and σ = 10. Find the mark that separates the top 10% of students from the rest.



Binomial Distribution

6. Define the binomial distribution and list its main assumptions.

Answer: The mark that separates the top 10% is approximately 82.8.

Binomial Distribution:

A binomial distribution models the number of successes in a fixed number of independent trials, where each trial has only two possible outcomes (success or failure) and the probability of success remains constant.

Main Assumptions:

- 1. Fixed number of trials (nn is known).
- 2. Two possible outcomes for each trial (success/failure).
- 3. Constant probability of success (pp) in each trial.
- 4. Independent trials the outcome of one trial does not affect another.

Example: Flipping a coin 10 times and counting the number of heads.

7. A coin is flipped 10 times. What is the probability of getting exactly 6 heads? (Use p = 0.5)

This is a binomial probability problem.

Step 1: Binomial formula

$$P(X=k)=inom{n}{k}p^k(1-p)^{n-k}$$

Where:

•
$$n = 10, k = 6, p = 0.5$$

$$\binom{10}{6} = \frac{10!}{6!(10-6)!} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$$

Step 2: Apply formula

$$P(X=6) = 210 \cdot (0.5)^6 \cdot (0.5)^4 = 210 \cdot (0.5)^{10}$$
 $(0.5)^{10} = \frac{1}{1024} \approx 0.0009765625$ $P(X=6) = 210 \cdot 0.0009765625 \approx 0.205$

- ✓ Answer: The probability of getting exactly 6 heads is approximately 0.205 (20.5%).
- 8. In a batch of 100 items, each has a 0.02 probability of being defective. What is the probability of finding exactly 3 defective items?

This is a binomial probability problem.

Given:

•
$$n = 100, k = 3, p = 0.02$$

Step 1: Binomial formula

$$P(X=k)=inom{n}{k}p^k(1-p)^{n-k}$$

$$\binom{100}{3} = \frac{100 \cdot 99 \cdot 98}{3 \cdot 2 \cdot 1} = 161700$$

Step 2: Apply formula

$$P(X=3) = 161700 \cdot (0.02)^3 \cdot (0.98)^{97}$$

$$(0.02)^3 = 0.000008, \quad (0.98)^{97} \approx 0.142$$

$$P(X=3) \approx 161700 \cdot 0.000008 \cdot 0.142 \approx 0.183$$

Answer: The probability of finding exactly 3 defective items is approximately 0.183 (18.3%).

9. Differentiate between binomial and normal distribution in terms of shape and use cases.

Feature	Binomial Distribution	Normal Distribution
~110pc	Discrete, takes integer values (0, 1, 2,, n); shape can be skewed or symmetric depending on pp	Continuous, smooth, symmetric bell-shaped curve
Use Cases	number of trials (e.g., heads in coin tosses,	Modeling continuous data that clusters around a mean (e.g., heights, weights, test scores)
Outcomes	(success/failure)	Can take any value within a range
Parameters	n = number of trials, p = probability of success	$\mu = \text{mean}, \ \sigma = \text{standard deviation}$

10. A basketball player makes 70% of his free throws. If he takes 5 shots, find the probability that he makes exactly 4 of them.

Given: $\bullet \quad n=5, \, k=4, \, p=0.7$ Step 1: Binomial formula $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ $\binom{5}{4} = \frac{5!}{4!1!} = 5$ Step 2: Apply formula $P(X=4) = 5 \cdot (0.7)^4 \cdot (0.3)^1$ $(0.7)^4 = 0.2401, \quad (0.3)^1 = 0.3$ $P(X=4) = 5 \cdot 0.2401 \cdot 0.3 = 5 \cdot 0.07203 \approx 0.36015$

Central Limit Theorem

11. State the Central Limit Theorem in your own words.

Answer: The probability that he makes exactly 4 shots is approximately 0.36 (36%).

Central Limit Theorem (CLT):

The CLT states that the distribution of sample means becomes approximately normal as the sample size grows, even if the population isn't normal.

Example: Averages of students' heights from different classes form a normal distribution.

12. Why is the Central Limit Theorem important in statistics and hypothesis testing?

Importance of Central Limit Theorem (CLT):

1. **Allows normal approximation:** Enables using normal distribution methods even if the population isn't normal.

- 2. **Basis for hypothesis testing:** Justifies calculating probabilities and confidence intervals for sample means.
- 3. **Simplifies analysis:** Makes it easier to make inferences about population parameters from sample data.
- 13. A population has μ = 50 and σ = 10. If we take samples of size n = 25, what will be the mean and standard deviation of the sampling distribution?

For a sampling distribution of the sample mean:

- Mean of the sampling distribution ($\mu_{\bar{X}}$) = population mean = 50
- Standard deviation of the sampling distribution ($\sigma_{\bar{X}}$) = $\frac{\sigma}{\sqrt{n}}$

$$\sigma_{ar{X}} = rac{10}{\sqrt{25}} = rac{10}{5} = 2$$

Answer:

- Mean = 50
- Standard deviation = 2
- 14. A population's mean is unknown, but $\sigma = 12$. If we take samples of size 36 and find the sample mean is 80, explain how CLT helps estimate the population mean.

Using the Central Limit Theorem (CLT):

- Even though the population mean is unknown, the sampling distribution of the sample mean will be approximately normal for n=36 (since it's large enough).
- The sample mean (80) can be used as a point estimate of the population mean (μ) .
- The standard error of the mean is:

$$\sigma_{ar{X}}=rac{\sigma}{\sqrt{n}}=rac{12}{\sqrt{36}}=2$$

- · This allows constructing confidence intervals around 80 to estimate the population mean.
 - 15. Give one example each where: a) The population is normal but n is small. b) The population is not normal but n is large, and CLT applies.

Examples:

a) Population is normal, n is small:

- Heights of students in a small classroom (e.g., n = 5).
- b) Population is not normal, n is large, CLT applies:
 - Daily sales in a store (skewed distribution), taking the average of 50 days' sales the sample mean will be approximately normal.