**2. Consider a dataset containing the heights (in centimeters) of 1000 individuals. The**

**mean height is 170 cm with a standard deviation of 10 cm. The dataset is approximately**

**normally distributed, and its skewness is approximately zero. Based on this information,**

**answer the following questions:**

**a. What percentage of individuals in the dataset have heights between 160 cm**

**and 180 cm?**

**b. If we randomly select 100 individuals from the dataset, what is the probability**

**that their average height is greater than 175 cm?**

**c. Assuming the dataset follows a normal distribution, what is the z-score**

**corresponding to a height of 185 cm?**

**d. We know that 5% of the dataset has heights below a certain value. What is**

**the approximate height corresponding to this threshold?**

**e. Calculate the coefficient of variation (CV) for the dataset.**

**f. Calculate the skewness of the dataset and interpret the result.**

**Solution:**

a. To calculate the percentage of individuals with heights between 160 cm and 180 cm, we need to find the z-scores for these heights and then use the standard normal distribution table.

The z-score formula is given by: z = (x - μ) / σ

For 160 cm:

z1 = (160 - 170) / 10 = -1

For 180 cm:

z2 = (180 - 170) / 10 = 1

Using a standard normal distribution table, we can find the percentage associated with these z-scores. From the table, the percentage corresponding to a z-score of -1 is approximately 0.1587, and the percentage corresponding to a z-score of 1 is also approximately 0.8413.

Therefore, the percentage of individuals with heights between 160 cm and 180 cm is:

Percentage = (0.8413 - 0.1587) \* 100 = 68.26%

b. To calculate the probability that the average height of 100 randomly selected individuals is greater than 175 cm, we need to use the Central Limit Theorem. According to the theorem, the distribution of sample means will be approximately normal, regardless of the shape of the original distribution, as long as the sample size is sufficiently large.

Since the population mean is 170 cm and the population standard deviation is 10 cm, the standard deviation of the sample means (standard error) can be calculated as σ / √n, where n is the sample size. In this case, n = 100.

Standard error = 10 / √100 = 1

To find the probability, we need to calculate the z-score for a height of 175 cm:

z = (x - μ) / σ = (175 - 170) / 1 = 5

Using the standard normal distribution table, we find that the percentage associated with a z-score of 5 is very close to 1 (approximately 0.99999).

Therefore, the probability that the average height of 100 randomly selected individuals is greater than 175 cm is almost 1 (or 100%).

c. To find the z-score corresponding to a height of 185 cm, we can use the formula:

z = (x - μ) / σ = (185 - 170) / 10 = 1.5

Therefore, the z-score corresponding to a height of 185 cm is 1.5.

d. To find the approximate height corresponding to the threshold where 5% of the dataset has heights below it, we need to find the z-score associated with the 5th percentile.

From the standard normal distribution table, the z-score corresponding to the 5th percentile is approximately -1.645.

Using the z-score formula: z = (x - μ) / σ, we can rearrange it to solve for x:

x = z \* σ + μ = -1.645 \* 10 + 170 = 153.55

Therefore, the approximate height corresponding to the threshold is 153.55 cm.

e. The coefficient of variation (CV) is a measure of relative variability and is calculated as the ratio of the standard deviation (σ) to the mean (μ). The formula is:

CV = (σ / μ) \* 100

In this case, the mean height is 170 cm and the standard deviation is 10 cm.

CV = (10 / 170) \* 100 = 5.88%

Therefore, the coefficient of variation for the dataset is approximately 5.88%.

f. The skewness of a dataset measures the asymmetry of its distribution. A skewness value of zero indicates a symmetric distribution. Since the skewness is approximately zero in this case, it suggests that the heights in the dataset are relatively symmetrically distributed around the mean of 170 cm.