How to Prepare Statistics for SSC CGL Tier II - Study Notes in PDF

SSC CGL Tier II Exam will be conducted soon! The SSC CGL Prelims Exam was conducted from 5th August to 24th August 2017. This year the exam pattern for Prelims Exam also went through a lot of chances. There are a total of 4733 vacancies that will be filled this year. If you are confident that you will make it through to the SSC CGL Tier II, then you can read the article given below. This article will help on **How to Prepare Statistics for SSC CGL Tier II.** In the following article, you will know in detail about Descriptive Statistics, viz. Central Tendency & Dispersion, Mean, Median & Mode, Skewness & Kurtosis, etc. You can also take our SSC CGL Online Mock Tests to boost up your preparation strategy.

Descriptive Statistics is the best way to describe and summarize the characteristics of a data set in terms of two of its properties i.e. **Central Tendency** and **Dispersion.**

Central Tendency - Prepare Statistics for SSC CGL Tier II

In order to describe and represent a set of data as a single number, we need measures of central tendency that intended to describe the performance of the group and centre of the data. It also tells us about the shape and nature of the distribution. Measures of central tendency include:

Mean:

The sum of all the observations divided by the number of observations

- Find the mean of 1, 2, 3, 4, 5, 6, 7, 8, 9
- $\frac{1+2+3+4+5+6+7+8+9}{9} = 5$

Median:

The score in the middle when the observations are ordered from the smallest to the largest. If the total number of observations n is an odd number, then the number on the position is the median. If n is an even number, then the average of the two numbers on the $\frac{n}{2}$ and $\frac{n}{2}+1$ positions is the median.

- o Find the median of 5, 6, 11, 10, 4, 9, 7
  - 4, 5, 6, **7**, 9, 10, 11 = 7
- o Find the median of 5, 17, 15, 3, 9, 18, 6, 10
  - 3, 5, 6, **9**, **10**, 15, 17, 18 = $\frac{9+10}{2} = 9.5$

Mode:

The number that occurs most frequently. If two numbers tie then the observation will have two modes and is called Bimodal

- Find the mode of 2, 6, 3, 9, 5, 6, 2, 6
- 2, 2, 3, 5, **6**, **6**, **6**, 9 = 6

Relation between Mean, Median and Mode

Mean – Mode = 3 (Mean – Median)

Scales of Measurement:-

1. **Nominal Scale** – That can simply be broken down into categories

2. **Ordinal Scale** – That can be categorized and can be placed in order or ranking
3. **Interval Scale** – That can be ranked but has no absolute zero point
4. **Ratio Scale** – That allows to compare and has meaningful zero values

For Nominal scale, the **mode** is the only measure that can be used. For Ordinal Scale, the **mode** and the **median** may be used. For Interval – Ratio Scale, the **mean**, **median** and **mode** all can be used.

Partition Values

If the samples are arranged in ascending or descending order, then the measures of central tendency divides the observations in two equal parts. In the same way, the given series can be divided into four, ten and hundred equal parts.

Quartiles

Quartiles divides a series into 4 equal parts i.e. Q1, Q2 and Q3. Q1 is known as first or lower Quartile covering 25% observations. Q2 is known as second Quartile is the same as Median of the series. Q3 is known as third or upper Quartile covering 75% observations.

$$Q_1 = l + \frac{\left(\frac{n}{4} - cf\right)}{f} \times i;\ Q_3 = l + \frac{\left(\frac{3n}{4} - cf\right)}{f} \times i$$

Where,

l = lower limit of median class; i = class interval

cf = total of all frequencies before median class

f = frequency of median class; n = total number of observations

Deciles: - Deciles divides a series into 10 equal parts i.e $D_1, D_2, D_3,$ etc.

$$D_1 = l + \frac{\left(\frac{n}{10} - cf\right)}{f} \times i;\ D_2 = l + \frac{\left(\frac{2n}{10} - cf\right)}{f} \times i;$$

$$D_3 = l + \frac{\left(\frac{3n}{10} - cf\right)}{f} \times i;\ D_4 = l + \frac{\left(\frac{4n}{10} - cf\right)}{f} \times i$$

Where,

l = lower limit of median class; i = class interval

cf = total of all frequencies before median class

f = frequency of median class; n = total number of observations

Percentiles: - Percentiles divides a series into 100 equal parts i.e., $P_1, P_2, P_3, \ldots, P_{99}, P_{100}$ etc.

$$P_1 = l + \frac{\left(\frac{n}{100} - cf\right)}{f} \times i;\ P_2 = l + \frac{\left(\frac{2n}{100} - cf\right)}{f} \times i;$$

$$P_{99} = l + \frac{\left(\frac{99n}{10} - cf\right)}{f} \times i;\ P_{100} = l + \frac{\left(\frac{100n}{100} - cf\right)}{f} \times i$$

Where,

l = lower limit of median class; i = class interval

cf = total of all frequencies before median class

f = frequency of median class; n = total number of observations

Measures of Dispersion - Prepare Statistics for SSC CGL Tier II

Measures of central tendency gives an idea of an average but it is important to know how the data are clustered or scattered away from the average and the degree to which it spread is called dispersion.

Various measures of dispersion are:-

1. **Range: -** Range is the difference between the largest and the smallest value of the series.
   - Range = Largest value – Smallest Value
   - Co – efficient of range = $\frac{L-S}{L+S}$
     - L = Largest value in the observation
     - S = Smallest value in the observation

2. **Inter–Quartile Range**: - It is the difference between the third quartile and the first quartile.
   - Inter–Quartile Range = $Q_3 - Q_1$
   - It is also known as range of middle 50% values
3. **Percentile Range**: - It is the difference between the 90th and 10th percentile.
   - Percentile Range = $P_{90} - P_{10}$
   - It is also known as range of middle 80% values
4. **Quartile Deviation**: - It is the average of the difference between the third quartile and the first quartile. It is an absolute measure of dispersion.
   - Quartile Deviation = $\frac{Q_3 - Q_1}{2}$

- Co–efficient of Quartile Deviation = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

5. **Mean Deviation**: - It is an average or the mean of the deviations of the values from a fixed point. It is a calculative measures of dispersion
   - Mean Absolute Deviation = $\frac{\frac{\sum_{i=1}^{i=N}|X_i - \bar{X}|}{N}}{N}$; where, N = Number of observations, $\bar{X}$ = Mean
6. **Standard Deviation**: - It is defined as the square root of the mean of the squared deviations of individual values around their mean. If the values of the observations are same, then standard deviation is zero and it is least affected by fluctuations.

- $\sigma = \sqrt{\frac{\sum_{i=1}^{i=N}|X_i - \bar{X}|^2}{N}}$; $S^2 = \frac{\sum_{i=1}^{i=N}|X_i - \bar{X}|^2}{N}$;

Where,

- $\sigma =$ Standard Deviation
- $S^2 =$ = Variance
- $\sum_{i=1}^{i=N}|X_i - \bar{X}|^2 =$ sum of the square of deviations from the mean
- N = total number of observations

Measures of dispersion help to describe the width of the distribution, but they don't give any information about the shape of the distribution. There are further statistics that give information about the shape of the distribution. They are:-

- **1st moment** – Mean (describes central value)

- 1st moment = $\frac{\sum_{i=1}^{i=N}|X_i-\bar{X}|^1)}{n}$, is equal to zero

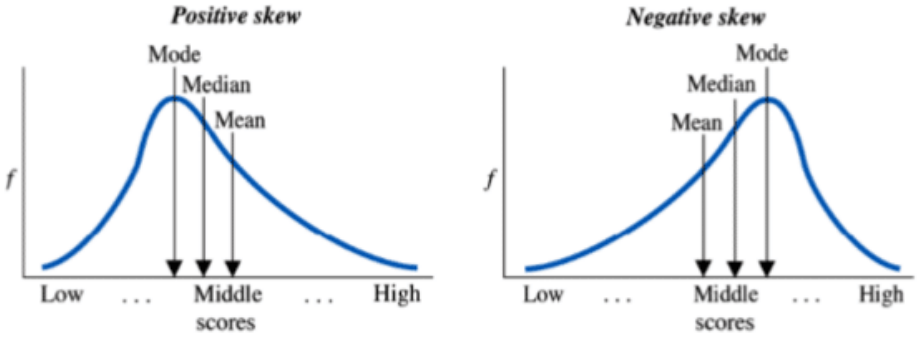
- **2nd moment** – Variance (describes dispersion)
  - $\frac{\sum_{i=1}^{i=N}|X_i-\bar{X}|^2}{n}$, gives information on the spread or scale of the distribution of numbers
- **3rd moment** – Skewness (describes asymmetry)
  - $\frac{\sum_{i=1}^{i=N}|X_i-\bar{X}|^3}{n}$, gives information on the Skewness of the distribution
- **4th moment** – Kurtosis (describes peakedness)
  - $\frac{\sum_{i=1}^{i=N}|X_i-\bar{X}|^4}{n}$, gives information on the Kurtosis of the distribution

Skewness - Prepare Statistics for SSC CGL Tier II

Measures of dispersion tells us about the variation of the data set whereas, Skewness tells us about the direction of variation of data set. It is a measure of symmetry i.e. same to the right and the left of the center point.

Skewness can be positive or negative or zero

- When the values of mean, median and mode are equal, there is no Skewness.
- When mean > median > mode, Skewness will be positive.
- When mean < median < mode, Skewness will be negative.



Pearson's coefficient of Skewness = $\frac{Mean-Mode}{SD} = \frac{3(Median-Mean)}{SD}$

It ranges between –3 to +3

Bowley's measure of Skewness = $\frac{Q_3-2M_d+Q_1}{Q_3-Q_1}$

Where, Q1 = first Quartile, Q2 = second Quartile, Q3 = third Quartile, = Median

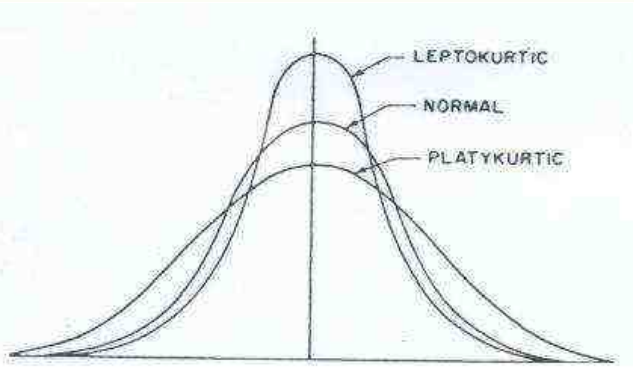
Kelly's measure of Skewness = $\frac{P_{90}-2P_{50}+P_{10}}{P_{90}-P_{10}}$

Where,

P_{90} = 90th Percentile; P_{50} = 50th Percentile; P_{10} = 10th Percentile

Kurtosis - Prepare Statistics for SSC CGL Tier II

It measures the relative peakedness or flatness of a distribution compared to the normal distribution



- When Kurtosis > 0, the peak of a curve becomes relatively high and that curve is called Leptokurtic. The positive Kurtosis indicates a flat distribution with long tails
- When Kurtosis < 0, the curve is flat-topped, then it is called Platykurtic. The negative Kurtosis indicates a peaked distribution with short tails
- A normal curve is neither very peaked nor very flat-topped, so it is taken as a basis for comparison. The normal curve is called Mesokurtic. For a normal distribution, kurtosis is equal to 3.

The measure of Kurtosis, known as Percentile coefficient of kurtosis is:

Kurtosis = $\frac{QD}{P_{90}-P_{10}}$

Where,

Q.D is semi-interquartile range=Q.D= $\frac{Q3-Q1}{2}$

P_{90} = 90th Percentile; P_{10}= 10th Percentile

Now that you know the details about SSC CGL Tier II Statistics, you can click on the link given below and start your preparation for the upcoming exams.

SSC CGL Recruitment 2017

Detailed SSC CGL Tier II Syllabus

If you want to begin your preparation for SSC CGL Tier II & want to practice for the same, then you can click on the link given below and quickly start!

Solve Free Practice Questions on Testbook

Do you want to discuss with us or fellow aspirants about the coming SSC CGL Tier II Exam? If yes, then click on the link given below, which will take you to Testbook Discuss!

Take me to Testbook Discuss!