Frequently Asked Questions

1. Give examples of continuous random variable.
   Answer:
   We have many examples, where the random variable follows continuous distribution.
   - The speed of a car
   - The concentration of a chemical in a water sample
   - Tensile strengths
   - Heights of people in a population
   - Lengths or areas of manufactured components
   - Life time of a battery
   - The annual rain fall in a region

2. Mention different properties of a random variable.
   Answer:
   Two important properties by which we may characterize random variables, regardless of the probability distribution are location and the dispersion.

3. What do you mean by property of location?
   Answer:
   The property of location is the central tendency of the variable, which describes a value around which the variable tends to cluster successes in n trials, and so on.

4. What do you mean by property of Dispersion?
   Answer:
   Dispersion is the typical range of values that might be expected to be observed in experiments. This gives some idea of the spread in values that might result from our experiment(s).

5. Mention the different measures of location.
   Answer:
   The central tendency, or location, of a variable can be indicated by any (or all) of the following: the mode, the median, or the mean.

6. Explain mode.
   Answer:
   The mode is the most probable value of a discrete variable. More generally, it is the maximum of the probability distribution function. In continuous distribution, it is the value of x for which the pdf attains its maximum. The value of x_{mode} such that

   $f(x_{mode}) = F_{max}$

7. What is multimodal distribution?
   Answer:
   Multi-modal probability distributions have more than one mode — distributions with two modes are bimodal, and so on. Although multi-modal distributions may have several local maxima, there is usually a single global maximum that is the single most probable value of the random variable.

8. Explain Median of the distribution.
   Answer:

Median is the value Q_2 such that
$P(x < Q_2) = P(x > Q_2)$
In other words, there is an equal probability of observing a value greater than or less than the median.
The median is also the second quartile— hence the origin of the symbol Q_2.

9. Explain mean of the distribution.
Answer:
The mean or expected value E(X) of the continuous variable is defined as follows.

$E(X) = \mu_x = \int_{-\infty}^{\infty} xf(x)dx$, where f(x) is a mathematical function that defines the probability distribution of the random variable x.

10. Compare different measures of location.
Answer:
We have defined three different indicators of the location of a random variable: the mean, median and mode. Each of these has a slightly different meaning.

Imagine that you are betting on the outcome of a particular experiment:

- If you choose the mode, you are essentially betting on the single most likely outcome of the experiment
- If you choose the median, you are equally likely to be larger or smaller than the outcome
- If you choose the mean, you have the best chance of being closest to the outcome

A random variable cannot be predicted exactly, but each of the three indicators gives a sense of what value the random variable is likely to be near. In most applications, the mean gives the best single description of the location of the variable.

11. Explain the nature of the curve using the measures of location.
Answer:
It turns out that the three values are different only for asymmetric distributions. If a distribution is skewed to the right (or positively skewed) then
$\mu_x > Q_2 > x_{mode}$
While for distributions skewed to the left (negatively skewed)
$\mu_x < Q_2 < x_{mode}$
For symmetrical ("bell-shaped") distributions, the mean, median and mode all have exactly the same value.

12. Explain dispersion of a random variable.
Answer:
Some variables are more "variable," more uncertain, than others. Of course, theoretically speaking, a variable may assume any one of the range of values in the domain. However, when speaking of the variability of a random variable, we generally mean the range of values that would commonly (i.e., most probably) be observed in an experiment. This property is called the dispersion of the random variable. Dispersion refers to the range of values that are commonly assumed by the variable.

13. Write different measures of dispersion.

Answer:

It is convenient to describe variability with a single value. Three common ways to do so are:

i. The inter-quartile range and the semi-inter-quartile range

ii. The mean absolute deviation

iii. The variance and the standard deviation

14. Explain variance and standard deviation.

Answer:

Like the mean absolute deviation, the variance and standard deviation measure the dispersion of a random variable about its mean μ_x. The variance of a random variable x, σ^2_x, is the expected value of $(x - \mu_x)^2$, which is the squared deviation of x from its mean value:

$\sigma^2_x = E(x - \mu_x)^2$

The variance for continuous variables is given by

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx$$

Higher variance signifies greater variability of a random variable.

One problem with using the variance to describe the dispersion of a random variable is that the units of variance are the squared units of the original variable. For example, if x is a length measured in m, then σ^2_x has units of m^2. The standard deviation, σ_x, has the same units as x, and so is a little more convenient at times. The standard deviation is simply the positive square root of the variance.

Sometimes the variability of a random variable is specified by the relative standard deviation, RSD:

$$RSD = \frac{\sigma_x}{\mu_x}$$

15. Compare the different measures of dispersion.

Answer:

We have described three common ways to measure a random variable's dispersion: semi-inter-quartile range, QR, mean absolute deviation, MD, and the standard deviation, σ. These measures are related to each other, so, in a sense, it makes no difference which we use.

In fact, for distributions that are only moderately skewed, MD ≈ 0.8σ and QR ≈ 0.67σ. For a variety of reasons, the variance and standard deviation are the best measures of dispersion of a random variable.