Normalizing constant:

The concept of normalizing constant is used in probability theory. This is used to reduce any probability function to a probability density function with total probability of one. Here the non-negative functions must be multiplied to make the area as 1.

Variance:

Variance measures the variation of a single random variable (like the height of a person in a population). Variance is the spread or the difference in a set of data. Mathematically, it is the average squared deviation from the mean score.

$$\mathrm{Var}(X) = \Sigma\ (X_i - X)^2 / N = \Sigma\ x_i^2 / N$$

Here,

N is the number of scores in a set of scores
X is the mean of the N scores.
X_i is the ith raw score in the set of scores
x_i is the ith deviation score in the set of scores
$\mathrm{Var}(X)$ is the variance of all the scores in the set

Covariance:

Covariance is used to measure the relationship of two random variables or the variance of two variables.   We use this formula to get covariance:

$$\mathrm{Cov}(X, Y) = \Sigma\ (X_i - X)\ (Y_i - Y) / N = \Sigma\ x_i y_i / N$$

Where,

N is the number of scores in each set of data
X is the mean of the N scores in the first data set
X_i is the $i$the raw score in the first set of scores
x_i is the ith deviation score in the first set of scores
Y is the mean of the N scores in the second data set
Y_i is the $i$the raw score in the second set of scores
y_i is the ith deviation score in the second set of scores
$\mathrm{Cov}(X, Y)$ is the covariance of corresponding scores in the two sets of data

If we consider just a random sample without knowing the population size, it will not be possible to find an unbiased estimator; here the sample variance will work as a biased estimator. As an example, let’s say we have only one sample and we want to calculate or estimate the mean and variance. Here, the mean needs to be the exact population mean. One reserved observation is "-1" and so you have N-1 in computing variance estimate.

Covariance matrix calculation: