# Basic Probability Theory and Statistics

**Statistics:** You can use descriptive statistical methods to transform raw observations into information that you can understand and share. deriving correlations between couples of features/columns tells you whether they have a relationship (linear at the very least). That's a statistic. And an important one in many cases. To have a well-functioning linear model you need a good amount of linear relationships between your predictors and the outcome.

**Probability:** Probability Measures the event occurrences in a Random Experiment. Probability is a number between 0 and 1, where, 0 indicates impossibility and 1 indicates certainty. The higher the probability of an event, the more likely it is that the event will occur.

**Conditional Probability:** Conditional Probability is a measure of the probability of an event given that evidence another event has already occurred. If the event of interest is A and the event B is known or assumed to have occurred, "the conditional probability of A given B", is usually written as P(A|B).

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

**Independence:** Two independent events, if the probability of each event not effect on the probability of the other event occurring.

$$P(A, B) = P(A) * P(B) \quad where \quad P(A) \neq 0 \quad and \quad P(B) \neq 0$$

$$P(A \mid B) = P(A) \quad \text{and} \quad P(B \mid A) = P(A)$$

**Conditional Independence: Two event conditionally independent to occur the third event, A** and B are conditionally independent given C if and only if, given knowledge that C already occurred, knowledge of whether A occurs provides no additional information on the likelihood of B occurring, and knowledge of whether B occurs provides no additional information on the likelihood of A occurring.

$$P(A \mid C, B \mid C) = P(A \mid C) * P(B \mid C) \quad where \quad P(A \mid C) \neq 0 \quad and \quad P(B \mid C) \neq 0$$

**Variance:** The variance of a random variable X is a measure of how concentrated the distribution of a random variable X is around its mean. It's defined as

$$Var[X] = E[X - E[X]^2] = E[X^2] - [E[X]]^2$$

**Probability Distribution:** Is a mathematical function that maps the all possible outcomes of a random experiment with its associated probability. It depends on the Random Variable X , whether it's discrete or continues.

**Joint probability distribution:** the probability distribution that defines their simultaneous behavior during outcomes of a random experiment.

**Conditional probability distribution:** P(Z|X,Y) It means for every possible combination of random variables X, Y we represent a probability distribution over Z.

**Factor:** a function or a table which takes a number of random variables {X_1, X_2,...,X_n} as an argument and produces a real number as a output.

$$\phi : Val(X_1, X_2, \ldots X_n) \rightarrow \mathfrak{R} \qquad Scope = \{X_1, X_2, \ldots X_n\}$$