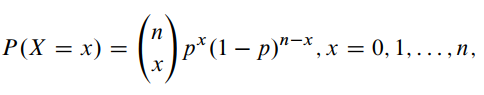
**Q1. Write a note on any two-standard discrete distribution functions.**

**Discrete Uniform Distribution:** The discrete uniform distribution represents a finite number of equally likely values. The simplest real-life example is the face obtained when a fair die is rolled once. It can also occur in some other physical phenomena, particularly when the number of possible values is small and the scientist feels that they are just equally likely. If we let the values of the random variable be 1, 2, …n, then the *probability mass function (pmf)* of the discrete uniform distribution is



**Binomial Distribution:** The binomial distribution represents a sequence of independent coin-tossing experiments. Suppose a coin with probability p; 0 < p < 1, for heads in a single trial is tossed independently a pre-specified number of times, say n times, n>=1. Let X be the number of times in these n tosses that a head is obtained. Then the *probability mass function (pmf)* of X is

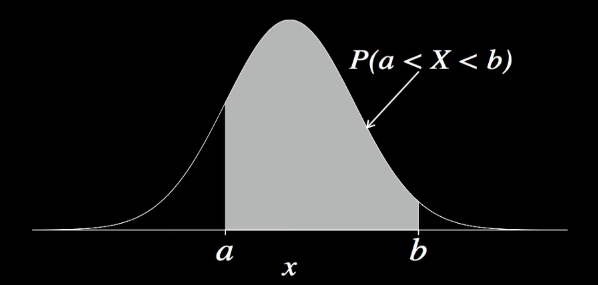


The term giving the choice of the x tosses out of the n tosses in which the heads occur.

Suppose a trial can result in only one of two outcomes, called a success (S) or a failure (F), the probability of obtaining a success being p in any trial. Such a trial is called a ***Bernoulli trial***. Suppose a Bernoulli trial is repeated independently a pre-specified number of times, say n times. Let X be the number of times in the n trials that a success is obtained. Then X has the pmf given above, and we say that X has a binomial distribution with parameters n and p and write 

**Q2. Write a note on any two-continuous distribution functions (other than Normal distribution function).**

We cannot model continuous random variables with the same methods we used for Discrete random variables. We model a continuous random variable with a curve f(x), called a *Probability Density Function (pdf).*

For continuous random variables, probabilities are areas under the curve. 

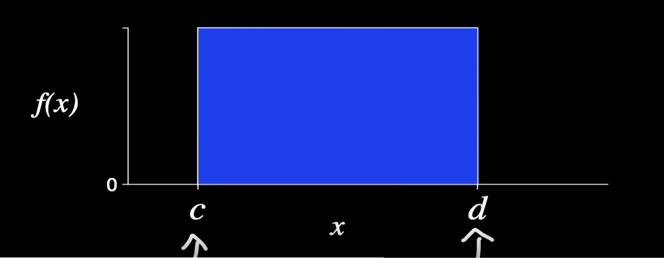
Probability of any random variable is 0 i.e., P(X=a) = 0

For any continuous probability distribution:

* 1. F(x) >=0 for all x.
  2. The area under the entire curve is equal to one.

**Uniform Distribution:**

For the uniform distribution, f(x) is constant over the possible values of x.



c 🡪 minimum value X can take on x

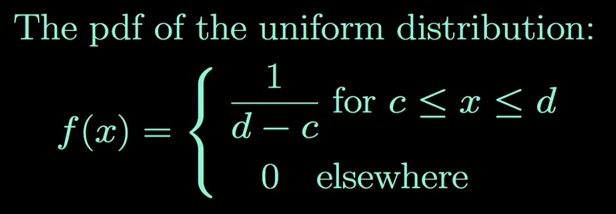
d 🡪 maximum value X can take on x

c and d are allowed to be any finite values.

Area under the curve should be 1. But, here the area under the curve is rectangle.

Therefore, Area = base x height = (d-c) \* f(x)

i.e. f(x) = 1/(d-c)



Median of the curve should be mid-point of base of the curve. That point/value should distribute the area under the curve into two halves.

Since it’s a symmetric distribution, mean = median = midpoint between c and d

Median = (c + d)/2

Mean (μ) = (c + d)/2

For continuous probability distribution, finding areas usually requires integration.

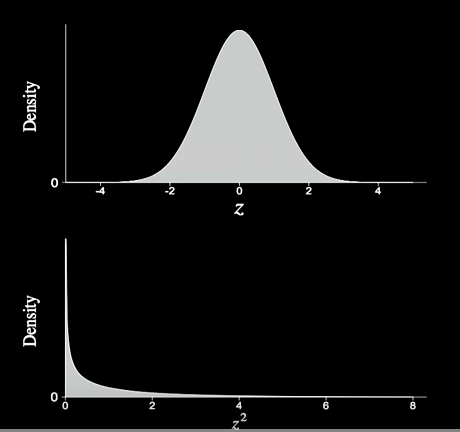
But for the uniform distribution, areas under the curve are simply rectangles.

**Chi-Square Distribution:**

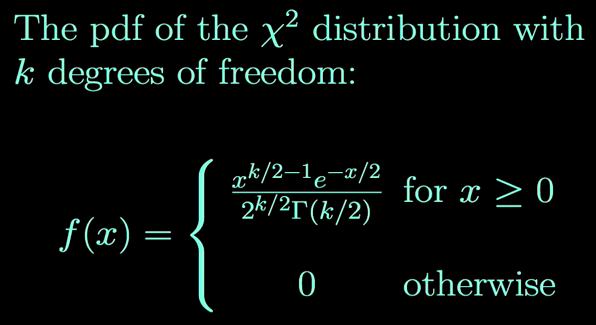
The χ2 distribution is a continuous probability distribution that is widely used in statistical inference.

The χ2 distribution is related to the standard normal distribution

If a random variable Z has the standard normal distribution, then Z2 has a χ2 distribution with one degree of freedom.

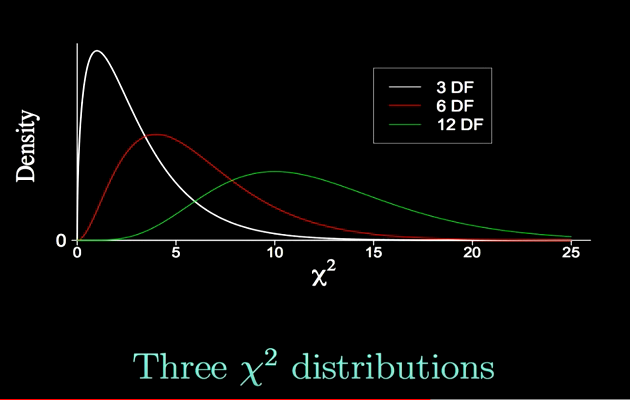


If Z1, Z2,…Zk are independent standard normal random variables, then Z12 + Z22 + … + Zk2 has a χ2 distribution with k degrees of freedom.



Mean (μ) = Degrees of freedom (k)

Variance (σ2) = Twice the degrees of freedom (2k)



As the degrees of freedom is increasing, the mean is increasing and variance is increasing as well.

Skewness will decrease as degrees of freedom is increasing (called as central limit theorem at work).

Areas and percentiles for the χ2 distribution can be found using computer software or a χ2 table.

Probabilities and percentiles are found by integrating the probability density function.

Deriving the mean and variance also requires integration.

**Q3. If X is a random variable and E(X) is Expectation of X. If another random variable, Y defined in terms of X:**

**Y = (X - E(X))**

**Then prove that E(Y) = 0**

**Expected value:** The expected value (or expectation) of a random variable is the arithmetic mean of that variable, i.e. E(X) = µ.

Expected value of random variable sometimes called as Population mean.

**Discrete case:** The expected value of a discrete random variable, X, is found by multiplying each X-value by its probability and then summing over all values of the random variable. That is, if X is discrete,



**Continuous case:** For a continuous variable X ranging over all the real numbers, the expectation is defined by



**Variance of X:** The variance of a random variable X is defined as the expected (average) squared deviation of the values of this random variable about their mean. That is,

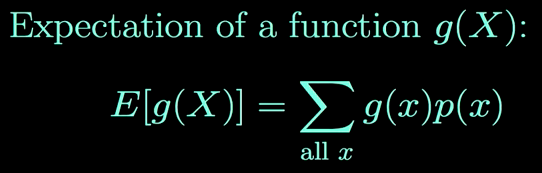


In the discrete case, this is equivalent to



**Standard deviation of X:** The standard deviation is the positive square root of the variance, i.e.



 g(x) represents the function of Random Variable

Given than:

Y = (X - E(X))

E(Y) = E[X – E(X)]

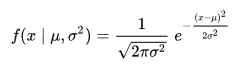
= E(X) – E(E(X))

= E(X) – E(X)

= 0

**Q4. Prove that the area under normal distribution is 1**

The probability density of the normal distribution is:



μ = mean

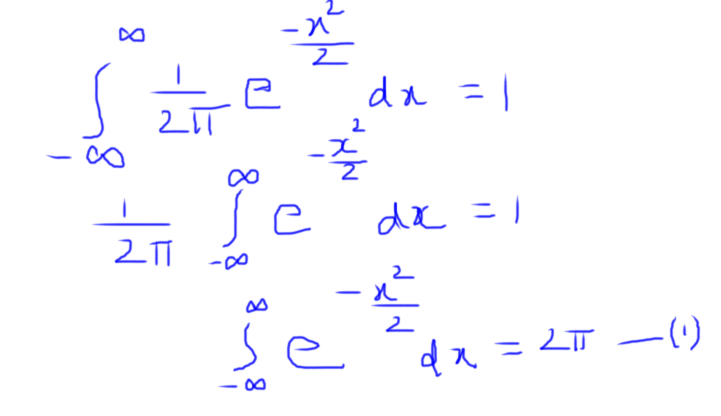
σ = standard deviation

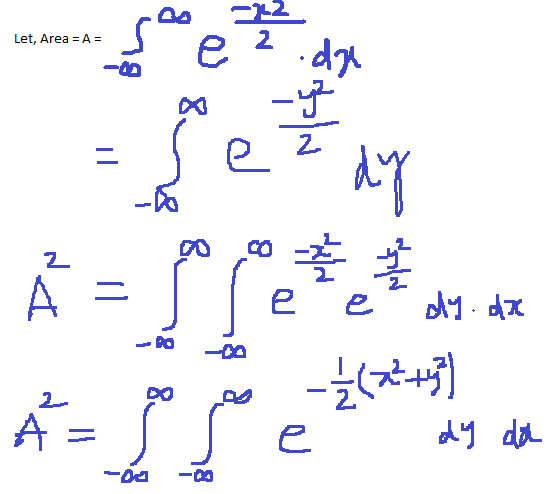
σ² = variance

The simplest case of a normal distribution is known as the standard normal distribution. This is a special case when μ = 0 and σ = 1 , and it is described by this [probability density function](https://en.wikipedia.org/wiki/Probability_density_function):

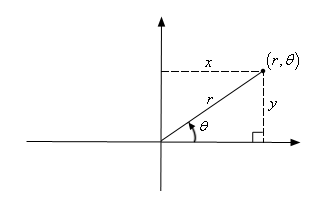


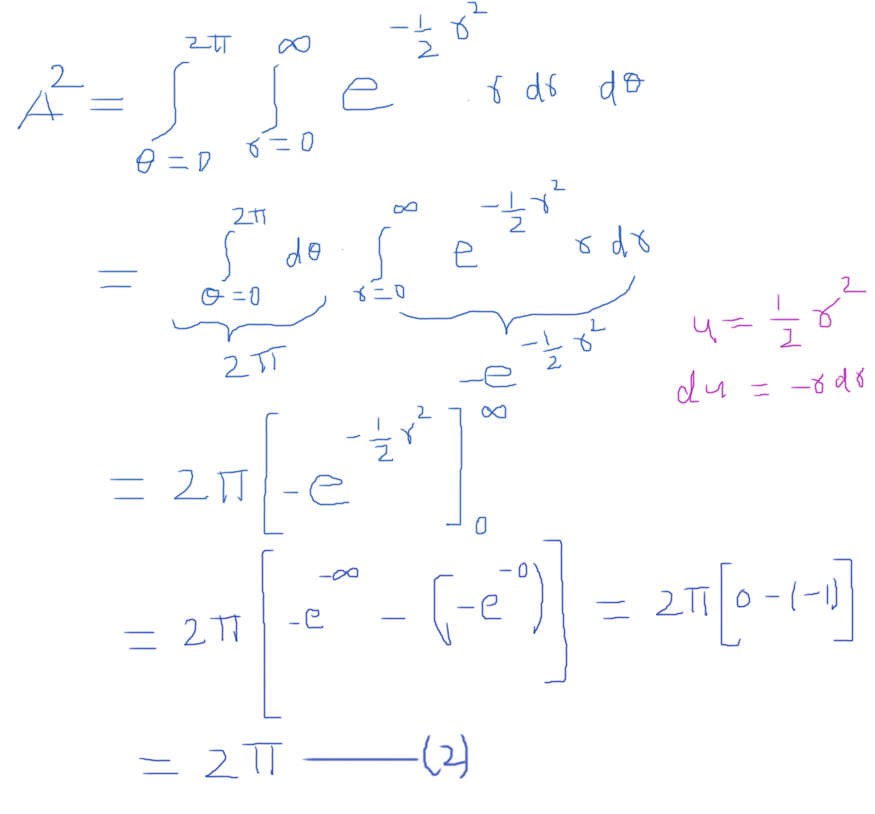
We have to prove that:





Let’s convert Cartesian (rectangular) co-ordinates x, y into polar (r, θ) form i.e. x² + y² = r²





From (1) and (2), we proved that, Area under curve is 1

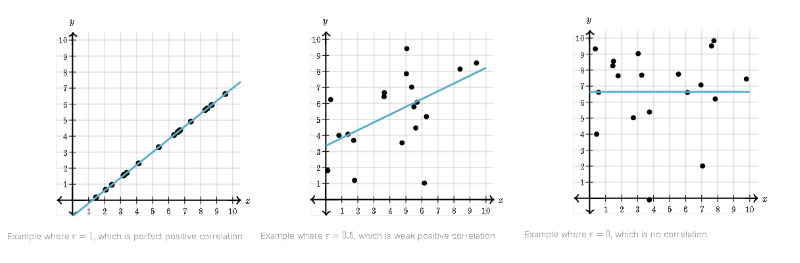
**Q5. Prove that the correlation coefficient, r lies between -1 and 1.**

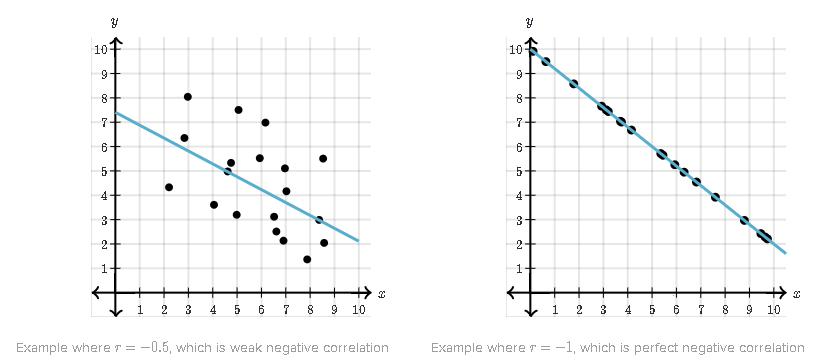
Correlation is the relationship between the variables.

The typical correlation statistic is the Pearson product moment correlation. For most people this is known as the correlation coefficient and is usually represented by the symbol 'r'. Correlation coefficients(r) lie between the values -1.00 and +1.00. A high absolute value indicates a high degree of relatedness, whereas the sign denotes whether the relation is positive (+) or negative (-).

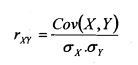
**Here are some facts about r:**

1. It always has a value between -1 and 1.
2. Strong positive linear relationships have values of r closer to 1.
3. Strong negative linear relationships have values of r closer to -1.
4. Weaker relationships have values of r closer to 0





Correlation can be defined as a quantitative measure of the degree or strength of relationship that may exist between two variables. You are already familiar with the concept of Karl Pearson's coefficient of correlation. If X and Yare two variables, we know that this correlation coefficient is given by the ratio of the covariance between X and Y to the product of the standard deviation of X and that of Y. In ' symbols:



The symbols have usual meaning. Here, the covariance in the numerator is important. This in fact, gives a measure of the simultaneous change in the two variables. It is divided by product of the standard deviation of X and Y to make the measure free of any unit in order to facilitate a comparison between more than one set of bi-variate data which may be expressed in different units. It may be noted here that this measure! of correlation coefficient is independent of a shift in the origin and a change of scale. The correlation coefficient lies between +1 and -1.

In symbol, -1 ≤ rxy ≤ 1

**Proof: -**

Let X1, X2, …, Xn be ‘n’ random variables defined on the same probability space, and

let Y=ɸ(X1, X2, …, Xn).

Then,

E[Y] = E[ɸ(X1, X2, …, Xn)]

Let X be uniformly distributed over the interval (-1,1) and let Y=X2, so that Y is completely dependent on X. Noting that for all odd values of k>0, the kth moment E[Xk] = 0, we have

E[XY] = E[X3] =0 and E[X]E[Y] = 0 \* E[Y] = 0

Therefore, X and Y are uncorrelated!

We have declared that Cov(X,Y) = 0 means X and Y are uncorrelated. On the other hand, if X and Y are linearly related – that is, X = aY for some constant a≠0 then, since E[X]=aE[Y], we have

Cov(X,Y) = a Var[Y] = Var[X]

* Cov2(X,Y) = Var[X] Var[Y]

In the general cas, it can be shown that

0 ≤ Cov2(X,Y) ≤ Var[X] Var[Y] ----------------------(1)

Using the following Cauchy-Schwarz inequality:

(E[XY])2 = E[X2]E[Y2]

Cov(X,Y) measures the degree of linear independence (or the degree of correlation) between the two random variables.

It is often useful to define a measure of this dependence in a scale-independent fashion. The correlation coefficient ρ(X,Y) is defined as

ρ(X,Y) =

=

Whenever σx and σy are defined

Using (1) we conclude that -1≤ρ(X,Y)≤1

**References:**

[**https://www.springer.com/cda/content/document/cda\_downloaddocument/9781441957795-c1.pdf?SGWID=0-0-45-907737-p173954517**](https://www.springer.com/cda/content/document/cda_downloaddocument/9781441957795-c1.pdf?SGWID=0-0-45-907737-p173954517)

[**https://www.khanacademy.org/math/statistics-probability/describing-relationships-quantitative-data/scatterplots-and-correlation/a/correlation-coefficient-review**](https://www.khanacademy.org/math/statistics-probability/describing-relationships-quantitative-data/scatterplots-and-correlation/a/correlation-coefficient-review)

[**http://egyankosh.ac.in/bitstream/123456789/23320/1/Unit-8.pdf**](http://egyankosh.ac.in/bitstream/123456789/23320/1/Unit-8.pdf)

Probability & Statistics with Reliability, Queuing and Computer Science Applications – by Kishor S. Trivedi