Normal distribution

Multivariate normal distribution

Population parameters – the parameters that define the statistical distribution for the entire population, e.g. population mean and population standard deviation for a population that is normally distributed.

Given a sample set from the population, we want to estimate the population parameters (like how we want to estimate the model parameters from our training data in machine learning).

We estimate the population parameters from the sample parameters. The more data we have, the more confidence we have in the estimates.

P-values and confidence intervals quantify the confidence we have in population parameter estimates.

Given two different samples, we can use statistics to quantify our confidence in how different they are. P-values and confidence intervals can tell us if the estimates from the two samples are statistically significant.

MLE

MAP

Confidence interval

Distributions

Descriptive stats like mean, stdev, variance, correlation, covariance

Expected value

Integrating over a distribution

Calculating mean of a distribution (integral of x\*P(x) I think)

Central limit theorem – this explains why we often see the normal distribution in nature.

is a random variable (a random process, where each outcome is associated with a number).

Add samples of this variable,

The distribution of this sum looks more like a normal distribution as .

<https://www.youtube.com/watch?v=zeJD6dqJ5lo&t=255s>