Stat Definitions

Sreejith Sreekumar

December 22, 2021

Contents

1 Confidence Interval

- A confidence interval gives the PROBABILITY that our true value lies within the range of values. Bigger interval = higher probability

2 Probability

- Area under an interval of a distribution curve.

3 Likelihood vs Probability

- Answer from Cross Validated Likelihood: What is the best values of the parameters so that the data that we observed follows a <some> distribution? Probability: Assuming that the data comes from a certain distribution what is the chance. Probability is the area under the PDF curve

4 Percentage of normal distribution lies within 1 std of mean? 2, 3 st

- 68%, 95%, 99.7%

5 SGD Update Rule

$$\theta = \theta - \alpha \Delta J(\theta)$$

$$(Current\ \theta\ vector) - learning\ rate * (Gradient\ of\ Slope)$$

6 Probability and Statistics

The problems considered by probability and statistics are inverse to each other. In probability theory we consider some underlying process which has some randomness or uncertainty modeled by random variables, and we figure out what happens. In statistics we observe something that has happened, and try to figure out what underlying process would explain those observations.

7 Law of Large Numbers

If you repeat an experiment independently a large number of times and average the result, what you obtain should be close to the expected value

8 Central Limit Theorem

9 Type I and Type II Error

Type 1: False Positive, Type 2: False Negative

10 Inverse Document Frequency

$$idf = log \frac{|D|}{d : ti \in d}$$

where | D | is the number of documents in our corpus, and | {d : ti ∈ d} | is the number of documents in which the term appears.

11 Kolmogorov - Smirnov Test

Tests whether sample fits a distribution well.

12 Confidence Interval

There are 100 products and 25 of them are bad. What is the confidence interval?

p = 25/100 = 0.25

CI = 0.25 +/- 1.96 sqrt((0.25(1-0.25)) * 100) CI = p +/- Z * sqrt(variance of binom dist)

CI = (16.5,33.5)

95% confidence = plus or minus 1.96 STDEV

13 Isolation Forest

Creates splits like a random forest, but find out how difficult is it to isolate the path to split an instance. Answer: Quora