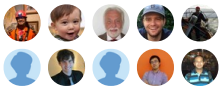
19 FOLLOWERS



Last asked: 31w ago

QUESTION TOPICS

Statistical Hypothesis Testing

Statistics (collected data)

Probability (statistics)

Statistics (academic discipline)

Machine Learning

QUESTION STATS

Views	3,259
Followers	19
Merged Questions	1
Edits	

3 Answers

Tarek El-Gaaly, PhD Student in Computer Vision
537 Views

Finding the optimum of a function in many cases involves taking the derivative. Differentiating is made easier when the function being maximized is the log-likelihood rather than the original likelihood. In addition it allows multiplications to be turned into addition for mathematical convenience (addition trick). In addition the logarithm is a monotonically increasing function and achieves its maximum at the same point as the original function. So in summary, its for mathematical convenience.

Written 38w ago • View Upvotes

More Answers Below. **Related Questions**

What is the intuition behind the likelihood-ratio test and how does it connect with the Neyman-Pearson lemma?

How are log likelihood ratio probabilities determined for LDPC codes?

Intuitively what is the difference among the Likelihood Ratio Test, the Wald Chi-square Test and the Lagrange Multiplier Test?

How can I remember the differences between odds ratio, hazard ratio, and likelihood ratio, and in what instances they should be applied?

What is an intuitive explanation for the log loss function?

Sumedha Sengupta, Research Statistician in Atmospheric Science , IAS, South Dakota School of Mi...
358 Views • Sumedha has 80+ answers in Statistics (academic discipline).

The log likelihood

It is little risky to answer intuitively the question asked. I like to use my rigorous Statistics training = 10yrs., practicing what I learned, in teaching, researching and applying for x yrs. and consulting for about y years with and without fees. Now in my retirement, I feel it is safe to keep on practicing it any way.
Having documented that, I will try to answer what you asked.
The Maximum Likelihood method is used in estimating a parameter θ of a distribution with a density function $f(x)$, from samples drawn at random from it. Naturally, one wants the estimate to be the best possible one of the unknown θ. So, a Likelihood function is defined such that it allows the sample to arrive at that desirable estimate. And we want that Likelihood to be Maximum for the best estimate. So, by definition the Likelihood Function is given by,

$L(\theta)=\prod f(xi), i=1,2....n$ for a sample of size n.
If one has a density function from the Type III family, such as, the Normal or Gamma, Weibull etc. the actual computation of the likelihood becomes rather long and complicated. So, we fall back on the natural log function and try to simplify the $L(\theta)=\prod f(xi), i=1,2....n$, in general. That means, the notation for the continuos product is changed to Summation.
or $\prod \rightarrow \Sigma$
And that is easier to handle. [note: if you don't agree, try with the normal density, you will definitely like it]

Then, we have the expression of log likelihood which needs to be maximized. So from the Freshman level Calculus, one maximizes by equating its first derivative (w.r.t θ) to zero and solving for θ . What you have is the MLE of

RELATED QUESTIONS

What is the intuition behind the likelihood-ratio test and how does it connect with the Neyman-Pearson lemma?

How are log likelihood ratio probabilities determined for LDPC codes?

Intuitively what is the difference among the Likelihood Ratio Test, the Wald Chi-square Test and the Lagrange Multiplier T...

How can I remember the differences between odds ratio, hazard ratio, and likelihood ratio, and in what instances they shou...

What is an intuitive explanation for the log loss function?

What is an intuitive explanation of F-score?

What is an intuitive explanation of Bootstrap resampling?

What is an intuitive explanation of the difference between parametric and nonparametric statistical tests?

What is an intuitive explanation of stochastic processes?

What is an intuitive explanation of covariance?

θ.

Written 31w ago

Benjamin Deonovic, graduate fellow
1.2k Views

the likelihood function measures how likelihood a set of parameters is given the observed data. In computing a likelihood ratio one is typically interested in how much more likely one parameter set is to another (e.g. is it more likely that the mean of this distribution is 0 [math]\mu=0[/math] or that it is greater than 0 [math]\mu>0[/math]). This ratio is typically constructed so that it is between 0 and 1. By taking the log of this ratio we end up with a concave function that is usually easy to maximize.

Written 111w ago • View Upvotes

Related Questions

What is an intuitive explanation of F-score?

What is an intuitive explanation of Bootstrap resampling?

What is an intuitive explanation of the difference between parametric and nonparametric statistical tests?

What is an intuitive explanation of stochastic processes?

What is an intuitive explanation of covariance?

What is an intuitive explanation of the Fisher exact test?

What is an intuitive explanation of a naive Bayes classifier?

What is an intuitive explanation of gradient descent?

What is an intuitive explanation of a moment (in probability)?

Machine Learning: What is an intuitive explanation of AUC?

What is an intuitive explanation of stochastic gradient descent?

How do you explain maximum likelihood estimation intuitively?

What is an intuitive explanation of Bayes' Rule?

What is the gradient of the log likelihood function in multinomial logistic regression?

What is an intuitive explanation of regularization?