Method for constructing a confidence interval based on estimating the profile likelihood

Julia Lynn, Timothy NeCamp, Edward Ionides, Susan Murphy
University of Michigan, Ann Arbor

Abstract

Our goal is to construct a confidence interval in order to understand reasonable values of that parameter given the data presented. In order to obtain these confidence intervals and the bounds for the parameter of interest, we use profile likelihood techniques. Profile likelihoods cannot be computed exactly due to two sources of error in profile likelihood evaluation: horizontal error and vertical shortfall. We present an estimation method for computing the profile likelihood where we control for both sources of error in the sampled points and use our estimated profile likelihood to construct confidence intervals. We predict that as our horizontal noise decreases, and the number of data points increase, that we will achieve a better estimated confidence interval. Through our simulation, we hope to attain a robust method, with good estimates of the profile likelihood and confidence intervals for various data generating scenarios. These methods have applications for predicting mobile health intervention outcomes.

Method of Estimating Parameters

Description of the use of Profile Likelihood

Our method of estimating complex parameters uses profile likelihood to construct a confidence interval for the parameter of interest. The profile likelihood is found by maximizing over the nuisance parameters. Statistical theory tells us that intersecting our profile likelihood with a horizontal line 1.92 vertical units down from the maximum constructs an asymptotic 95% confidence interval for our parameter of interest.

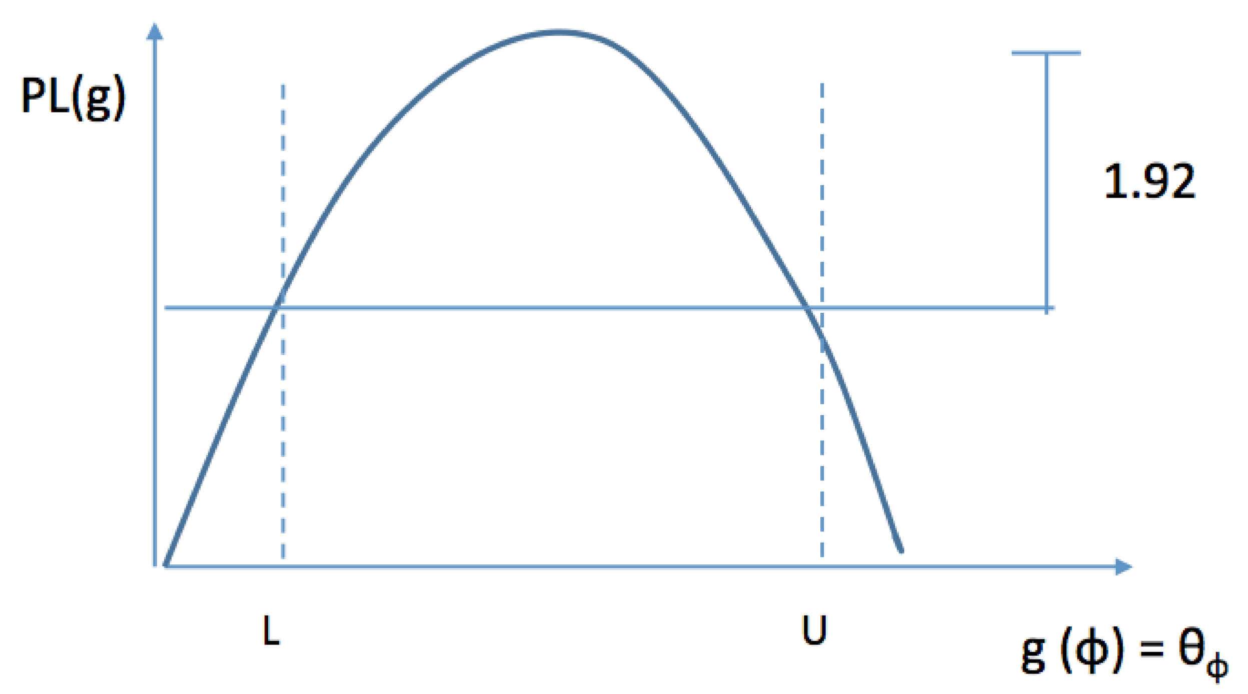

Figure 1: Constructing a 95% confidence interval S=$\{g : L \leqslant g \leqslant U\}$.

Estimating Profile Likelihood Curve

In our problem, we do not have the true profile likelihood curve due to two sources of error:
1. Horizontal error- We cannot accurately evaluate our implicit function, g(ϕ)
2. Vertical error- We cannot maximize to calculate our profile likelihood curve
Our method estimates the profile likelihood.

References: Ionides, E., et. al. (2017) Monte Carlo profile confidence intervals. arXiv:1612.02710

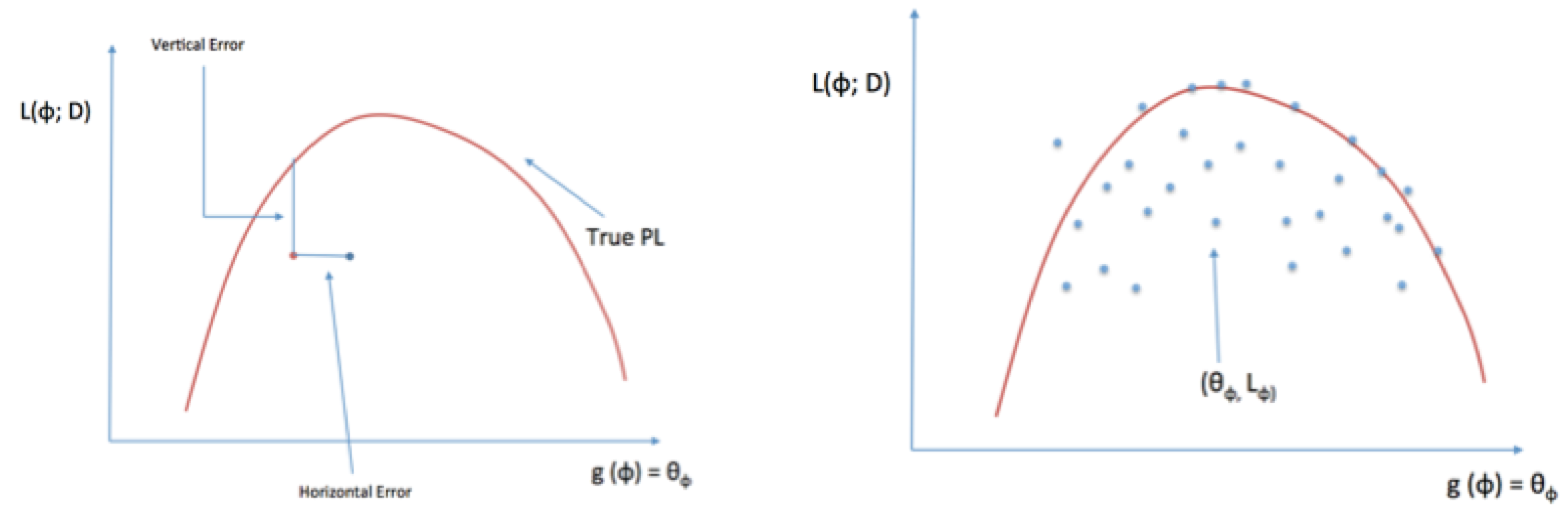

Figure 2: Estimating profile likelihood curve by accounting for horizontal and vertical error in our problem.

Simulation Description

Our simulation will be run on simple data for the estimation of simple parameters such as the mean. Eventually, we hope to create a robust method that has good prediction scores on complex data and parameters and can be used in a variety of different situations and problems.

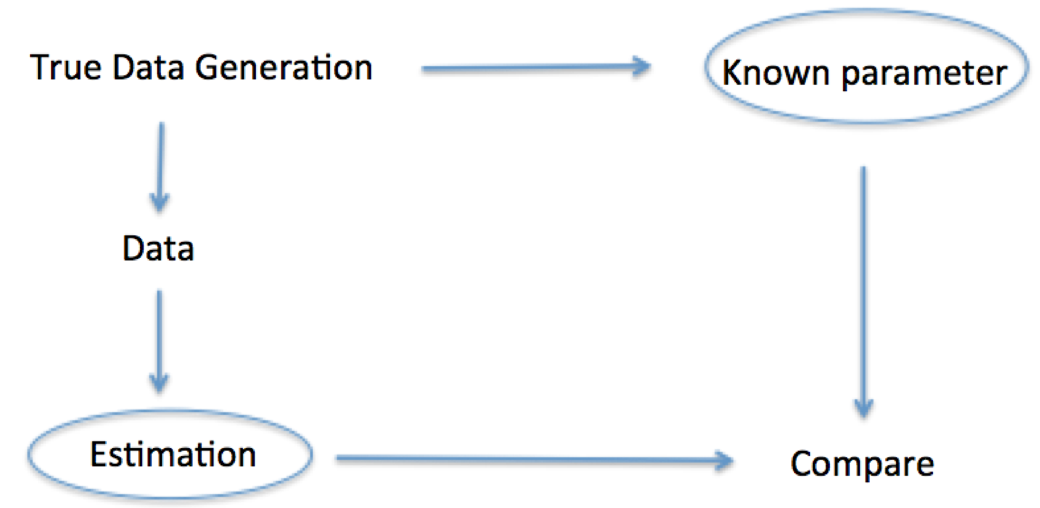

Figure 3: Progression of the simulation designed to evaluate and train our method to increase its accuracy and robustness.

Once we have a robust method, we will implement our method on new data with unknown parameter values so at to obtain an accurate estimate of these parameters.

Timing Comparison for our Simulation

We run our simulation to compare computational efficiency of different resources. The implementation with the highest number of iterations and the use of parallelization and Flux computing gives us the most accurate and computationally efficient estimation of complex parameters.

	Nonparallelized		Parallelized	
K iterations	Local	Flux	Local	Flux
5	26.15268	29.99499	7.368404	38.702 seconds
10	50.08322	1.003597 min	13.94509	39.428 seconds
50	4.299456 mins	5.565897 mins	1.245231 mins	43.279 seconds
100	8.606687 mins	10.04016 mins	2.734002 mins	55.747 seconds
500	41.982031 mins	50.67982 mins	10.003241 min	154.748 seconds

Figure 4: Progression of the simulation designed to evaluate and train our method to increase its accuracy and robustness.

Simulation Results

We run a simulation with data generated from a bivariate normal. Our parameter of interest is the maximum of the two means. We vary the mean values, the number of observations, n, and the number of points used to estimate the profile likelihood, s.

	n=20				n=100			
	s=20		s=100		s=20		s=100	
	Estimated	True	Estimated	True	Estimated	True	Estimated	True
μ = max(5.1, -5)	95.50%	96.40%	96.00%	96.40%	91.30%	94.70%	91.80%	94.70%
μ = max(5.1, 5)	97.10%	95.40%	97.10%	95.40%	96.70%	96.00%	97.00%	96.00%
μ = max(5.1, 5.09)	96.50%	94.70%	96.80%	94.60%	95.70%	95.00%	97.10%	95.00%

Figure 5: Percentage of true parameter found in the estimated and true confidence intervals.

Our results give the percentage of the constructed confidence intervals that contain the true parameter in comparison to the percentage of true confidence intervals that contain the true parameter. Our method did a good job of constructing confidence intervals, as the percentage of estimated confidence intervals that contain the true parameters is close to 95% for all the results. Visualization of possible results are below.

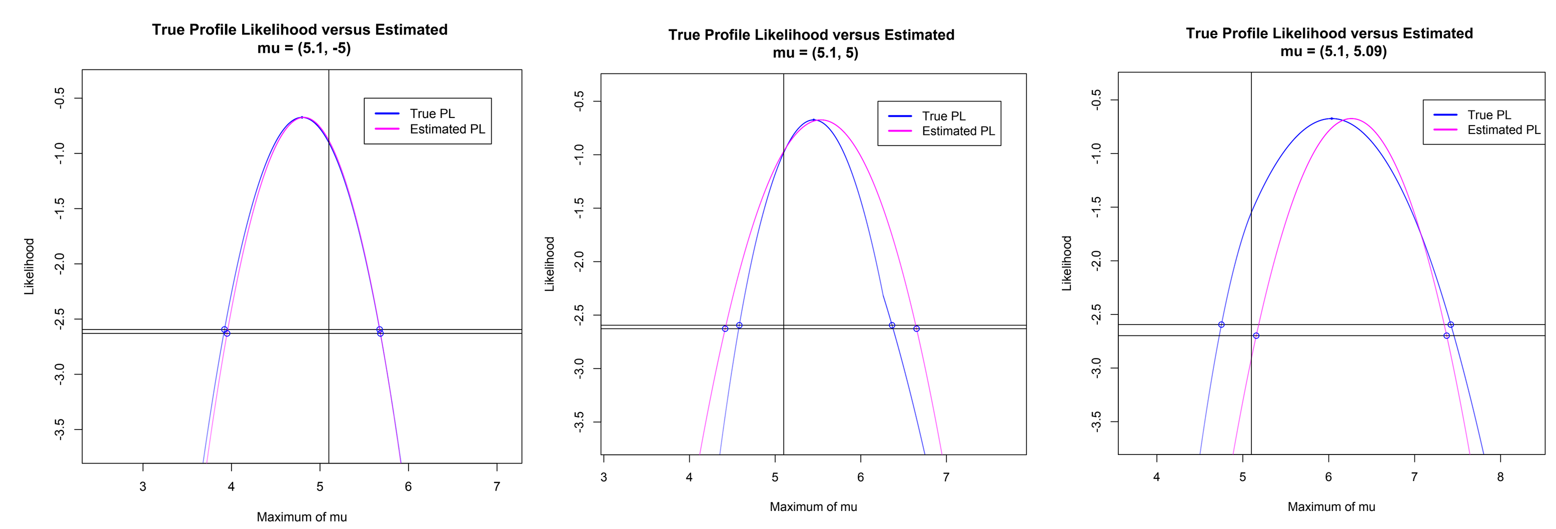

Figure 6: Plots of the true profile likelihood and the estimated profile likelihood for the maximum mean parameter.

Future Work

Our method of estimating the profile likelihood to construct confidence intervals for a complex function of the parameters performs well. There are two future directions we hope to work towards. First, we hope to improve our estimation method in situations when it does not perform well. Secondly, we want to use our method to construct confidence intervals for the health outcomes of patients in mobile health trials.