Estimating a Normal Mean

Suppose we have a sample of $N = 5$ values x_i,

$$x_i \sim N(\mu, 1)$$

We want to estimate μ, including some quantification of uncertainty in the estimate: an interval with a probability attached.

Frequentist approaches: method of moments, BLUE, least-squares/χ^2, maximum likelihood

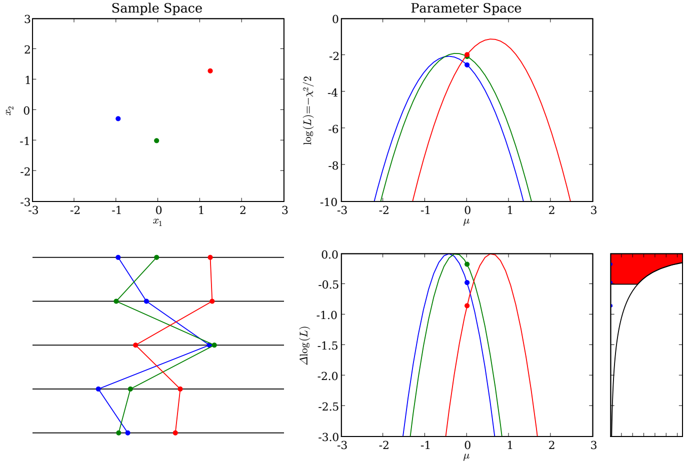
Focus on likelihood (equivalent to χ^2 here); this is closest to Bayes.

$$\begin{aligned} \mathcal{L}(\mu) &= p(\{x_i\}|\mu) \\ &= \prod_i \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_i-\mu)^2/2\sigma^2}; \qquad \sigma = 1 \\ &\propto e^{-\chi^2(\mu)/2} \end{aligned}$$

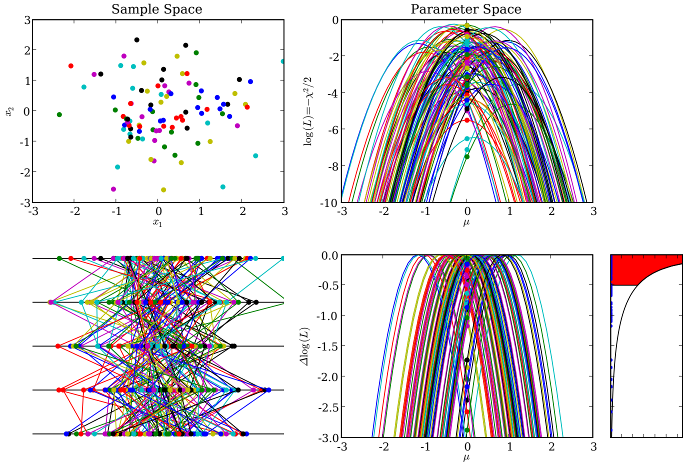
Estimate μ from maximum likelihood (minimum χ^2).
Define an interval and its coverage frequency from the $\mathcal{L}(\mu)$ curve.

Construct an Interval Procedure for Known μ

Likelihoods for 3 simulated data sets, $\mu = 0$

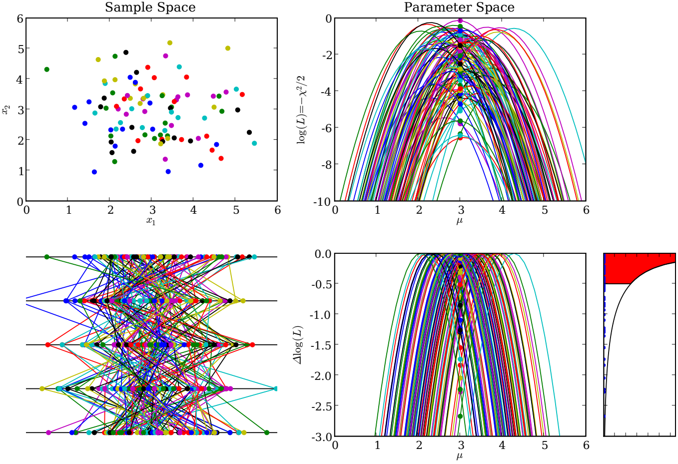


Likelihoods for 100 simulated data sets, $\mu = 0$



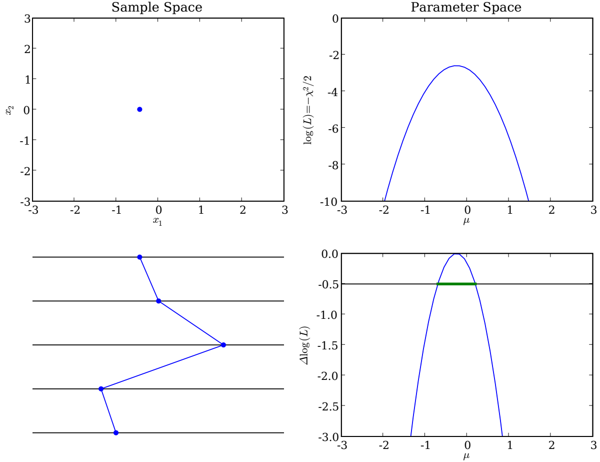
Explore Dependence on μ

Likelihoods for 100 simulated data sets, $\mu = 3$



Luckily the $\Delta \log \mathcal{L}$ distribution is the same!

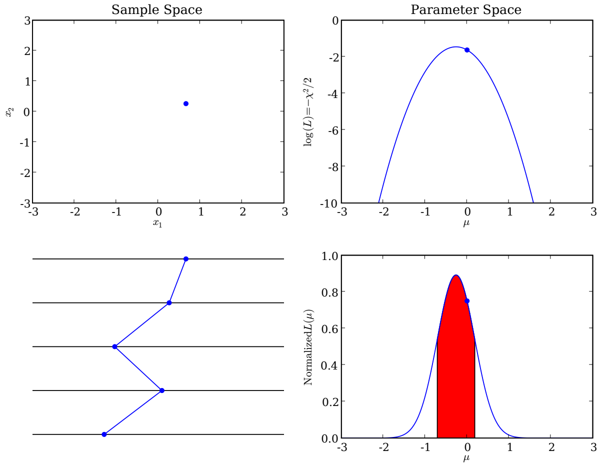
Apply to Observed Sample



Report the green region, with coverage as calculated for ensemble of hypothetical data (red region, previous slide).

Bayesian Approach

Normalize the likelihood for the observed sample; report the region that includes 68.3% of the normalized likelihood.



When They'll Differ

Both approaches report $\mu \in [\bar{x} - \sigma/\sqrt{N}, \bar{x} + \sigma/\sqrt{N}]$, and assign 68.3% to this interval (with different meanings).

This matching is a *coincidence*!

When might results differ? ($\mathcal{F}$ = frequentist, $\mathcal{B}$ = Bayes)

- If $\mathcal{F}$ procedure doesn't use likelihood directly
- If $\mathcal{F}$ procedure properties depend on params (nonlinear models, pivotal quantities)
- If $\mathcal{F}$ properties depend on likelihood shape (conditional inference, ancillary statistics, recognizable subsets)
- If there are extra uninteresting parameters (nuisance parameters, corrected profile likelihood, conditional inference)
- If $\mathcal{B}$ uses important prior information

Also, for a different task—comparison of parametric models—the approaches are qualitatively different (significance tests & info criteria vs. Bayes factors)