Confidence Interval Basics

- Interval estimation
- Confidence interval from inverting a test
- Example: limits on mean of Gaussian
- Confidence intervals from the likelihood function

Confidence intervals by inverting a test

In addition to a 'point estimate' of a parameter we should report an interval reflecting its statistical uncertainty.

Confidence intervals for a parameter θ can be found by defining a test of the hypothesized value θ (do this for all θ):

Specify values of the data that are 'disfavoured' by θ (critical region) such that $P(\text{data in critical region} \,|\, \theta) \leq \alpha$ for a prespecified α, e.g., 0.05 or 0.1.

If data observed in the critical region, reject the value θ.

Now invert the test to define a confidence interval as:

set of θ values that are not rejected in a test of size α (confidence level CL is $1-\alpha$).

Relation between confidence interval and p-value

Equivalently we can consider a significance test for each hypothesized value of θ, resulting in a p-value, p_θ.

If $p_\theta \leq \alpha$, then we reject θ.

The confidence interval at CL $= 1 - \alpha$ consists of those values of θ that are not rejected.

E.g. an upper limit on θ is the greatest value for which $p_\theta > \alpha$.

In practice find by setting $p_\theta = \alpha$ and solve for θ.

For a multidimensional parameter space $\boldsymbol{\theta} = (\theta_1, \dots \theta_M)$ *use* same idea – result is a confidence "region" with boundary determined by $p_{\boldsymbol{\theta}} = \alpha$.

Coverage probability of confidence interval

If the true value of θ is rejected, then it's not in the confidence interval. The probability for this is by construction (equality for continuous data):

$$P(\text{reject } \theta | \theta) \leq \alpha = \text{type-I error rate}$$

Therefore, the probability for the interval to contain or "cover" θ is

$$P(\text{conf. interval "covers" } \theta | \theta) \geq 1 - \alpha$$

This assumes that the set of θ values considered includes the true value, i.e., it assumes the composite hypothesis $P(\boldsymbol{x}|H,\theta)$.

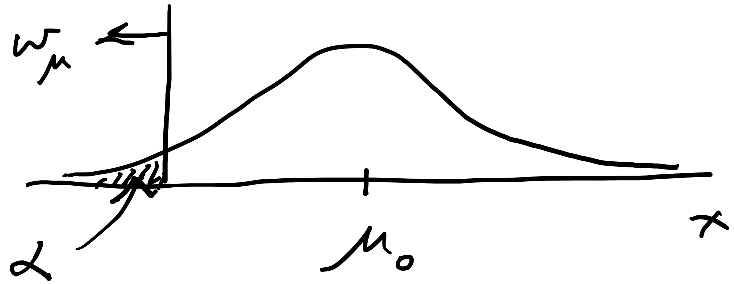
Example: upper limit on mean of Gaussian

When we test the parameter, we should take the critical region to maximize the power with respect to the relevant alternative(s).

Example: $x \sim \text{Gauss}(\mu, \sigma)$ (take σ known)

Test $H_0 : \mu = \mu_0$ versus the alternative $H_1 : \mu < \mu_0$

→ Put w_μ at region of x-space characteristic of low μ (i.e. at low x)

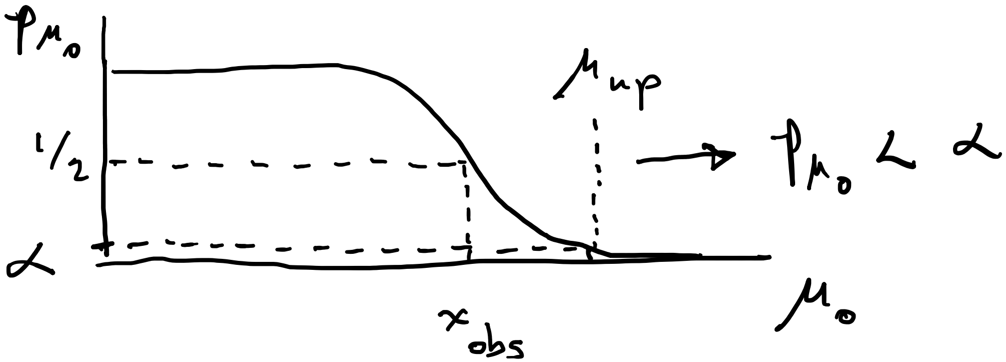


Equivalently, take the p-value to be

$$p_{\mu_0} = P(x \leq x_{\rm obs} | \mu_0) = \int_{-\infty}^{x_{\rm obs}} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu_0)^2/2\sigma^2}\, dx = \Phi\left(\frac{x_{\rm obs} - \mu_0}{\sigma}\right)$$

Upper limit on Gaussian mean (2)

To find confidence interval, repeat for all μ_0, i.e., set $p_{\mu 0} = \alpha$ and solve for μ_0 to find the interval's boundary



$$\mu_0 \to \mu_{\rm up} = x_{\rm obs} - \sigma\Phi^{-1}(\alpha) = x_{\rm obs} + \sigma\Phi^{-1}(1-\alpha)$$

This is an upper limit on μ, i.e., higher μ have even lower p-value and are in even worse agreement with the data.

Usually use $\Phi^{-1}(\alpha) = -\Phi^{-1}(1-\alpha)$ so as to express the upper limit as $x_{\rm obs}$ plus a positive quantity. E.g. for α = 0.05, $\Phi^{-1}(1-0.05)$ = 1.64.

Approximate confidence intervals/regions from the likelihood function

Suppose we test parameter value(s) $\boldsymbol{\theta} = (\theta_1, ..., \theta_N)$ using the ratio

$$\lambda(\boldsymbol{\theta}) = \frac{L(\boldsymbol{\theta})}{L(\hat{\boldsymbol{\theta}})} \qquad 0 \le \lambda(\boldsymbol{\theta}) \le 1$$

Lower $\lambda(\boldsymbol{\theta})$ means worse agreement between data and hypothesized $\boldsymbol{\theta}$. Equivalently, usually define

$$t_{\boldsymbol{\theta}} = -2 \ln \lambda(\boldsymbol{\theta})$$

so higher t_θ means worse agreement between $\boldsymbol{\theta}$ and the data.

p-value of $\boldsymbol{\theta}$ therefore

$$p_{\boldsymbol{\theta}} = \int_{t_{\boldsymbol{\theta},\mathrm{obs}}}^{\infty} f(t_{\boldsymbol{\theta}}|\boldsymbol{\theta})\, dt_{\boldsymbol{\theta}}$$

need pdf

Confidence region from Wilks' theorem

Wilks' theorem says (in large-sample limit and provided certain conditions hold...)

$$f(t_{\boldsymbol{\theta}}|\boldsymbol{\theta}) \sim \chi^2_N$$

chi-square dist. with # d.o.f. = # of components in $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$.

Assuming this holds, the p-value is

$$p_{\boldsymbol{\theta}} = 1 - F_{\chi^2_N}(t_{\boldsymbol{\theta}}|\boldsymbol{\theta})$$ ← set equal to α

To find boundary of confidence region set $p_\theta = \alpha$ and solve for t_θ:

$$t_{\boldsymbol{\theta}} = F^{-1}_{\chi^2_N}(1-\alpha)$$

Recall also

$$t_\theta = -2 \ln \frac{L(\theta)}{L(\hat{\theta})}$$

Confidence region from Wilks' theorem (cont.)

i.e., boundary of confidence region in $\boldsymbol{\theta}$ space is where

$$\ln L(\boldsymbol{\theta}) = \ln L(\hat{\boldsymbol{\theta}}) - \tfrac{1}{2} F^{-1}_{\chi^2_N}(1-\alpha)$$

For example, for $1-\alpha = 68.3\%$ and $n = 1$ parameter,

$$F^{-1}_{\chi^2_1}(0.683) = 1$$

and so the 68.3% confidence level interval is determined by

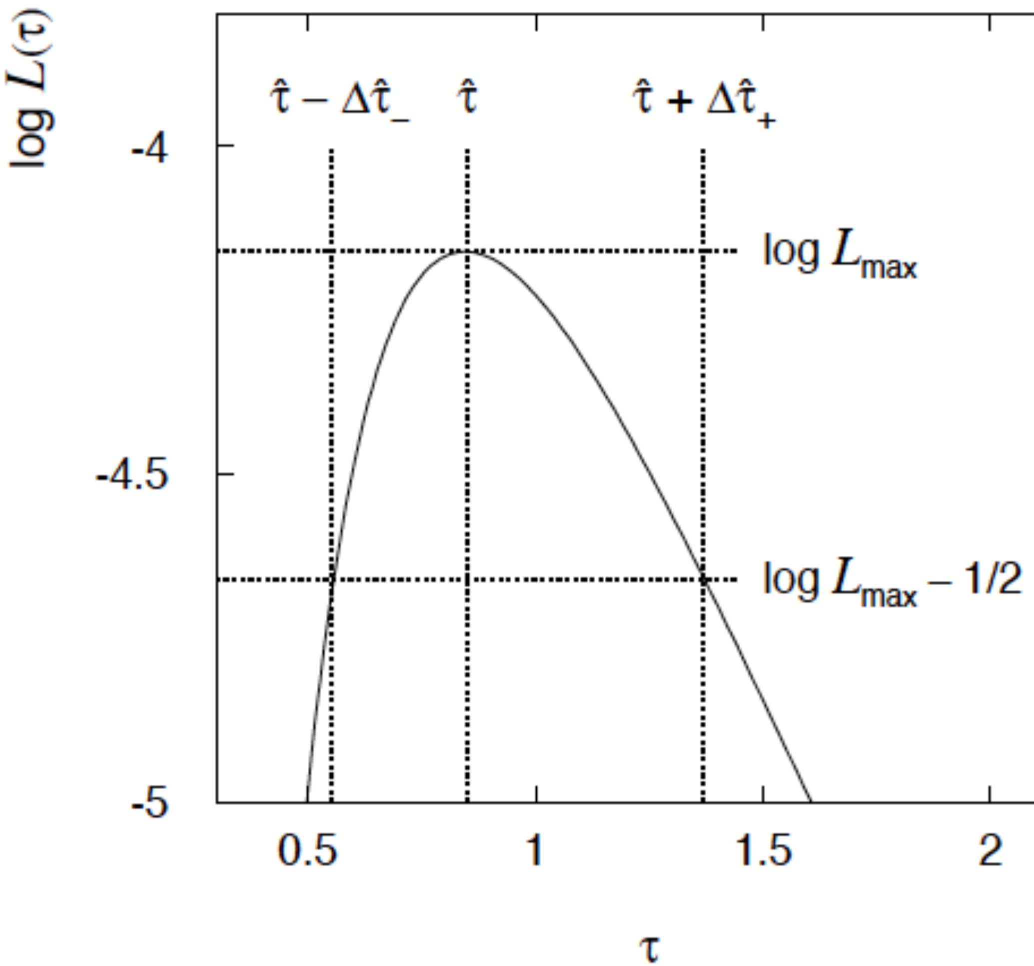
$$\ln L(\theta) = \ln L(\hat{\theta}) - \frac{1}{2}$$

Same as recipe for finding the estimator's standard deviation, i.e.,

$[\hat{\theta} - \sigma_{\hat{\theta}}, \hat{\theta} + \sigma_{\hat{\theta}}]$ is a 68.3% CL confidence interval.

Example of interval from $\ln L(\theta)$

For $N = 1$ parameter, CL = 0.683, $Q_\alpha = 1$.



Our exponential example, now with only $n = 5$ events.

Can report ML estimate with approx. confidence interval from $\ln L_{\max} - 1/2$ as “asymmetric error bar”:

$$\hat{\tau} = 0.85^{+0.52}_{-0.30}$$

Multiparameter case

For increasing number of parameters, CL $= 1 - \alpha$ decreases for confidence region determined by a given

$$Q_\alpha = F^{-1}_{\chi^2_n}(1-\alpha)$$

Q_α	$1-\alpha$				
	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
1.0	0.683	0.393	0.199	0.090	0.037
2.0	0.843	0.632	0.428	0.264	0.151
4.0	0.954	0.865	0.739	0.594	0.451
9.0	0.997	0.989	0.971	0.939	0.891

← # of par.

Multiparameter case (cont.)

Equivalently, Q_α increases with n for a given CL $= 1 - \alpha$.

$1-\alpha$	Q_α				
	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$ ← # of par.
0.683	1.00	2.30	3.53	4.72	5.89
0.90	2.71	4.61	6.25	7.78	9.24
0.95	3.84	5.99	7.82	9.49	11.1
0.99	6.63	9.21	11.3	13.3	15.1