Interval estimation

1. **The standard deviation as statistical error**
2. **Classical confidence intervals**
   (a) for a parameter with a Gaussian distributed estimator
   (b) for the mean of a Poisson distribution
3. **Approximate confidence intervals using the likelihood function or χ^2**

The standard deviation as statistical error

My experiment: data $x_1, \ldots, x_n \rightarrow$ estimate $\hat{\theta}_{\text{obs}}$

Also estimate variance of $\hat{\theta}$, $\widehat{\sigma^2}_{\hat{\theta}}$, often report something like

$$\hat{\theta}_{\text{obs}} \pm \hat{\sigma}_{\hat{\theta}} = 5.73 \pm 0.21$$

What does this really mean?

We know $\hat{\theta}$ will follow some pdf $g(\hat{\theta}; \theta)$,

estimate of θ is 5.73,

estimate of $\sigma_{\hat{\theta}}$ is $0.21 \rightarrow \sigma_{\hat{\theta}}$ measures width of $g(\hat{\theta}; \theta)$

Often $g(\hat{\vec{\theta}}; \vec{\theta})$ is multivariate Gaussian,

$\hat{\vec{\theta}}$, $\widehat{V} = \widehat{\text{cov}}[\hat{\theta}_i, \hat{\theta}_j]$ summarize our (estimated) knowledge

about $g(\hat{\vec{\theta}}; \vec{\theta})$, $\rightarrow$ input for error propagation, LS averaging, ...

We could stick with this as the convention for reporting errors,

regardless of the pdf of $g(\hat{\theta}; \theta)$.

Sometimes we do (e.g. for PDG averaging), but ...

if $g(\hat{\theta}; \theta)$ is Gaussian, then the interval

$$[\hat{\theta}_{\text{obs}} - \hat{\sigma}_{\hat{\theta}}, \hat{\theta}_{\text{obs}} + \hat{\sigma}_{\hat{\theta}}]$$

is a 68.3% central confidence interval (more later).

This is the more usual convention, and if $g(\hat{\theta}; \theta)$ not Gaussian,

central confidence interval $\rightarrow$ asymmetric errors

We have an estimator $\hat{\theta}$ for a parameter θ and an estimate $\hat{\theta}_{\text{obs}}$,

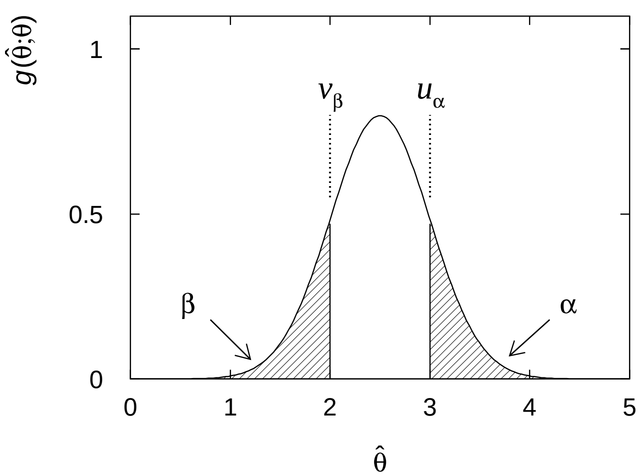
we also need $g(\hat{\theta};\theta)$ for all θ.

Specify 'upper and lower tail probabilities', e.g. $\alpha = \beta = 0.05$,

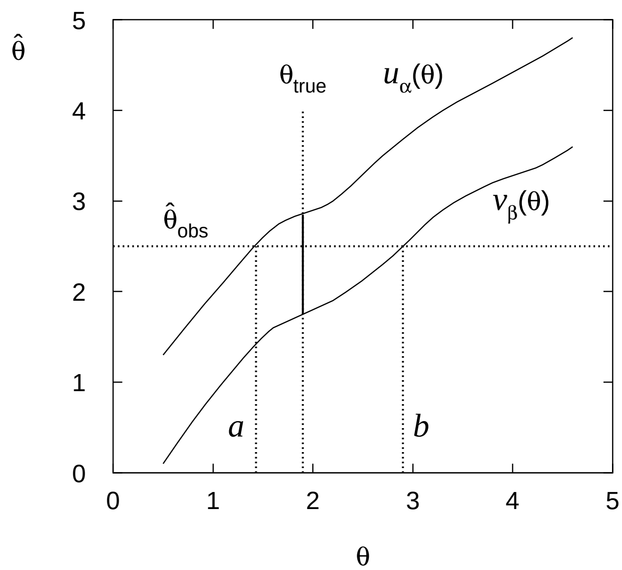
then, find functions $u_\alpha(\theta)$, $v_\beta(\theta)$ such that

$$\alpha = P(\hat{\theta} \geq u_\alpha(\theta)) = \int_{u_\alpha(\theta)}^{\infty} g(\hat{\theta};\theta)d\hat{\theta} = 1 - G(u_\alpha(\theta);\theta)\,,$$

$$\beta = P(\hat{\theta} \leq v_\beta(\theta)) = \int_{-\infty}^{v_\beta(\theta)} g(\hat{\theta};\theta)d\hat{\theta} = G(v_\beta(\theta);\theta)\,.$$



The region between $u_\alpha(\theta)$ and $v_\beta(\theta)$ is called the confidence belt.



Classical confidence intervals (2)

The probability to find $\hat{\theta}$ in the confidence belt, regardless of θ,

$$P(v_\beta(\theta) \leq \hat{\theta} \leq u_\alpha(\theta)) = 1 - \alpha - \beta.$$

Assume $u_\alpha(\theta)$, $v_\beta(\theta)$ monotonic, then

$$a(\hat{\theta}) \equiv u_\alpha^{-1}(\hat{\theta}) \ ,$$

$$b(\hat{\theta}) \equiv v_\beta^{-1}(\hat{\theta}) \ .$$

The inequalities

$$\hat{\theta} \geq u_\alpha(\theta),$$

$$\hat{\theta} \leq v_\beta(\theta),$$

imply

$$a(\hat{\theta}) \geq \theta \ ,$$

$$b(\hat{\theta}) \leq \theta \ .$$

$$\Rightarrow \quad P(a(\hat{\theta}) \geq \theta) = \alpha,$$

$$P(b(\hat{\theta}) \leq \theta) = \beta.$$

or together,

$$P(a(\hat{\theta}) \leq \theta \leq b(\hat{\theta})) = 1 - \alpha - \beta \ .$$

Classical confidence intervals (3)

The interval $[a(\hat{\theta}), b(\hat{\theta})]$ is called a confidence interval with confidence level or coverage probability $1 - \alpha - \beta$.

Its quintessential property:

probability to contain true parameter is $1 - \alpha - \beta$.

N.B. the interval is random, the true θ is an unknown constant.

Often report interval $[a, b]$ as $\hat{\theta}^{+d}_{-c}$, i.e. $c = \hat{\theta} - a$, $d = b - \hat{\theta}$.

So what does $\hat{\theta} = 80.25^{+0.31}_{-0.25}$ mean? It does not mean:

$P(80.00 < \theta < 80.56) = 1 - \alpha - \beta$, but rather:

repeat the experiment many times with same sample size,

construct interval according to same prescription each time,

in $1 - \alpha - \beta$ of experiments, interval will cover θ.

Sometimes only specify α or β, $\rightarrow$ one-sided interval (limit)

Often take $\alpha = \beta = \dfrac{\gamma}{2}$ $\rightarrow$ coverage probability $= 1 - \gamma$

$\rightarrow$ central confidence interval

N.B. 'central' confidence interval does not mean the interval is symmetric about $\hat{\theta}$, but only that $\alpha = \beta$.

The HEP error 'convention': 68.3% central confidence interval.

Classical confidence intervals (4)

Usually, we don't construct the confidence belt, but rather solve

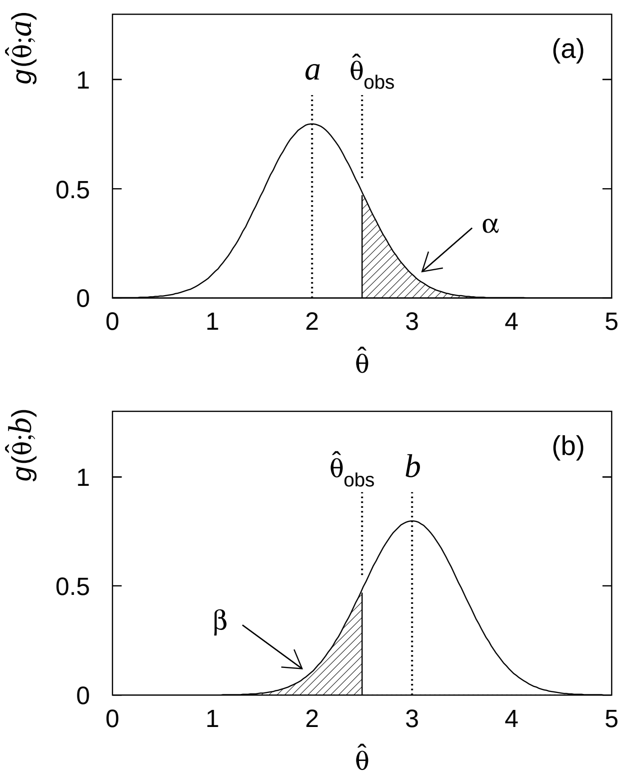
$$\alpha = \int_{\hat{\theta}_{\rm obs}}^{\infty} g(\hat{\theta}; a)\, d\hat{\theta} = 1 - G(\hat{\theta}_{\rm obs}; a)$$

$$\beta = \int_{-\infty}^{\hat{\theta}_{\rm obs}} g(\hat{\theta}; b)\, d\hat{\theta} = G(\hat{\theta}_{\rm obs}; b)$$

for interval limits a and b. (Gives same thing.)

→ a is hypothetical value of θ such that $P(\hat{\theta} > \hat{\theta}_{\rm obs}) = \alpha$;

b is hypothetical value of θ such that $P(\hat{\theta} < \hat{\theta}_{\rm obs}) = \beta$.



Confidence interval for Gaussian distributed estimator

Suppose we have $g(\hat{\theta};\theta) = \frac{1}{\sqrt{2\pi\sigma_{\hat{\theta}}^2}} \exp\left(\frac{-(\hat{\theta}-\theta)^2}{2\sigma_{\hat{\theta}}^2}\right)$.

To find confidence interval for θ, solve

$$\alpha = 1 - G(\hat{\theta}_{\rm obs}; a, \sigma_{\hat{\theta}}) = 1 - \Phi\left(\frac{\hat{\theta}_{\rm obs} - a}{\sigma_{\hat{\theta}}}\right),$$

$$\beta = G(\hat{\theta}_{\rm obs}; b, \sigma_{\hat{\theta}}) = \Phi\left(\frac{\hat{\theta}_{\rm obs} - b}{\sigma_{\hat{\theta}}}\right),$$

for a, b, where G is cumulative distribution for $\hat{\theta}$ and

$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-x'^2/2}\, dx'$ is cumulative of standard Gaussian.

$$\rightarrow \quad a = \hat{\theta}_{\rm obs} - \sigma_{\hat{\theta}}\, \Phi^{-1}(1-\alpha),$$

$$b = \hat{\theta}_{\rm obs} + \sigma_{\hat{\theta}}\, \Phi^{-1}(1-\beta).$$

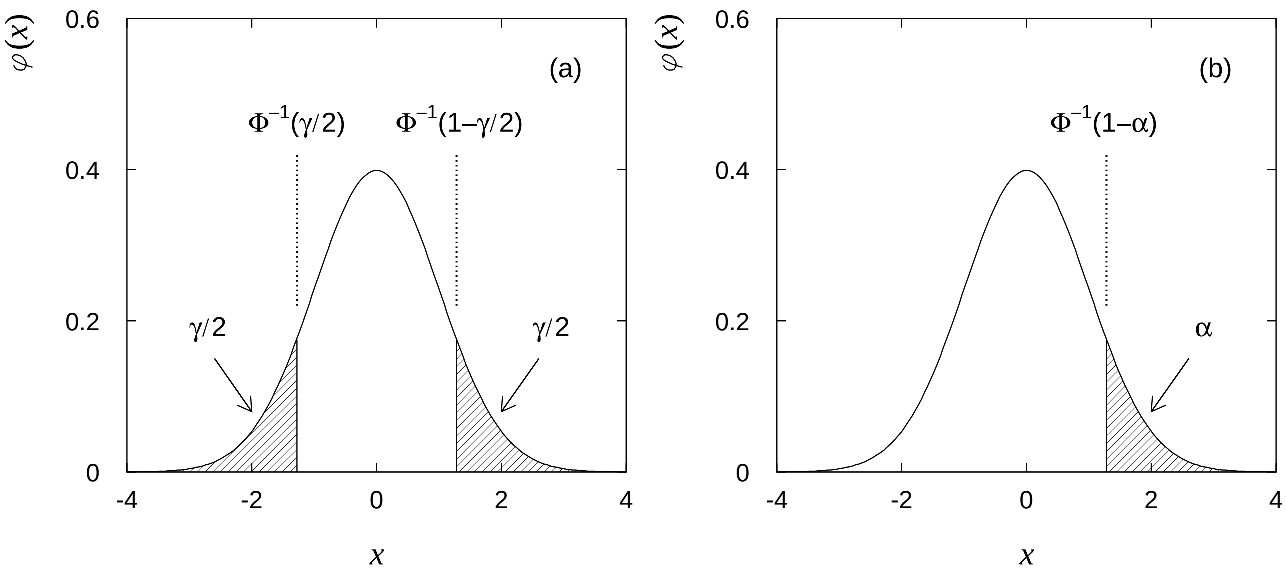
Φ^{-1} = quantile of standard Gaussian

(inverse of cumulative distribution, CERNLIB routine **GAUSIN**).

→ $\Phi^{-1}(1-\alpha)$, $\Phi^{-1}(1-\beta)$ give how many standard deviations a, b are from $\hat{\theta}$.

Quantiles of the standard Gaussian

To find the confidence interval for a parameter with a Gaussian, estimator we need the following quantiles:



Usually take a round number for the quantile ...

central		one-sided	
$\Phi^{-1}(1-\gamma/2)$	$1-\gamma$	$\Phi^{-1}(1-\alpha)$	$1-\alpha$
1	0.6827	1	0.8413
2	0.9544	2	0.9772
3	0.9973	3	0.9987

Sometimes take a round number for the coverage probability ...

central		one-sided	
$1-\gamma$	$\Phi^{-1}(1-\gamma/2)$	$1-\alpha$	$\Phi^{-1}(1-\alpha)$
0.90	1.645	0.90	1.282
0.95	1.960	0.95	1.645
0.99	2.576	0.99	2.326

Confidence interval for mean of Poisson distribution

Suppose n is Poisson, $\hat{\nu} = n$, estimate is $\hat{\nu}_{\rm obs} = n_{\rm obs}$,

$$P(n;\nu) = \frac{\nu^n}{n!} e^{-\nu}, \quad n = 0, 1, \ldots$$

Minor problem: for fixed α, β, confidence belt doesn't exist for all ν. No matter. Just solve

$$\alpha = P(\hat{\nu} \geq \hat{\nu}_{\rm obs}; a) = 1 - \sum_{n=0}^{n_{\rm obs}-1} \frac{a^n}{n!} e^{-a},$$

$$\beta = P(\hat{\nu} \leq \hat{\nu}_{\rm obs}; b) = \sum_{n=0}^{n_{\rm obs}} \frac{b^n}{n!} e^{-b},$$

for a, b. Use trick:

$$\sum_{n=0}^{m} \frac{\nu^n}{n!} e^{-\nu} = 1 - F_{\chi^2}(2\nu; n_{\rm d} = 2(m+1))$$

where F_{χ^2} is cumulative chi-square distribution for $n_{\rm d}$ dof,

$$a = \tfrac{1}{2} F_{\chi^2}^{-1}(\alpha; n_{\rm d} = 2n_{\rm obs}),$$

$$b = \tfrac{1}{2} F_{\chi^2}^{-1}(1-\beta; n_{\rm d} = 2(n_{\rm obs}+1)),$$

where $F_{\chi^2}^{-1}$ is the quantile of the chi-square distribution (CERNLIB routine `CHISIN`).

Interval for Poisson mean (continued)

Important special case: $n_{\rm obs} = 0$,

$$\rightarrow \beta = \sum_{n=0}^{0} \frac{b^n e^{-b}}{n!} = e^{-b}, \quad \rightarrow \quad b = -\log \beta .$$

For upper limit at confidence level $1 - \beta = 95\%$,

$$b = -\log(0.05) = 2.996 \approx 3.$$

Some more useful numbers...

$n_{\rm obs}$	lower limit a			upper limit b		
	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$	$\beta = 0.1$	$\beta = 0.05$	$\beta = 0.01$
0	–	–	–	2.30	3.00	4.61
1	0.105	0.051	0.010	3.89	4.74	6.64
2	0.532	0.355	0.149	5.32	6.30	8.41
3	1.10	0.818	0.436	6.68	7.75	10.04
4	1.74	1.37	0.823	7.99	9.15	11.60
5	2.43	1.97	1.28	9.27	10.51	13.11
6	3.15	2.61	1.79	10.53	11.84	14.57
7	3.89	3.29	2.33	11.77	13.15	16.00
8	4.66	3.98	2.91	12.99	14.43	17.40
9	5.43	4.70	3.51	14.21	15.71	18.78
10	6.22	5.43	4.13	15.41	16.96	20.14

Approximate confidence intervals from $\log L$ or χ^2

Recall trick for estimating $\sigma_{\hat{\theta}}$ if $\log L(\theta)$ parabolic:

$$\log L(\hat{\theta} \pm N\sigma_{\hat{\theta}}) = \log L_{\max} - \frac{N^2}{2}.$$

Claim: this still works even if $\log L$ not parabolic as an approximation for the confidence interval, i.e. use

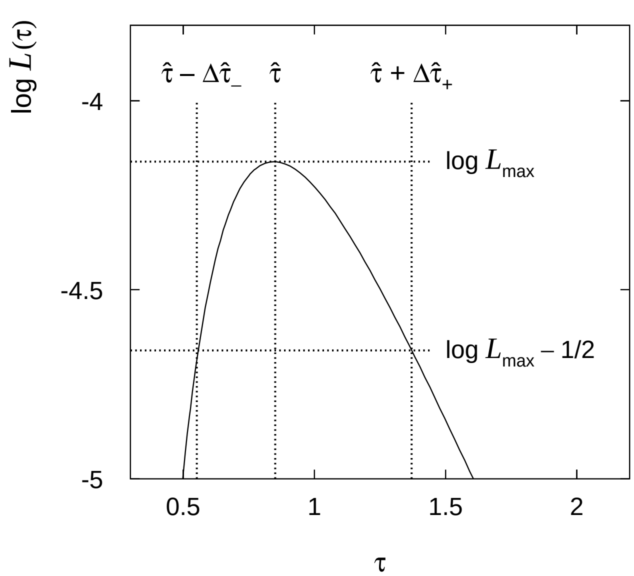
$$\log L(\hat{\theta}^{+d}_{-c}) = \log L_{\max} - \frac{N^2}{2},$$

$$\chi^2(\hat{\theta}^{+d}_{-c}) = \chi^2_{\min} + N^2,$$

where $N = \Phi^{-1}(1 - \gamma/2)$ is the quantile of the standard Gaussian corresponding to the confidence level $1 - \gamma$, e.g.

$$N = 1 \rightarrow 1 - \gamma = 0.683.$$

Our exponential example, now with $n = 5$ observations:



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

Interval estimation

1. **The standard deviation as statistical error:** tells how widely estimates $\hat{\theta}$ would be spread if experiment repeated. Needed for LS averaging, but sometimes want asymmetric error.

2. **Classical confidence intervals:** Complicated! Random interval which contains true parameter with fixed probability.

   (a) For a parameter with a Gaussian distributed estimator: $[\hat{\theta} - \sigma_{\hat{\theta}}, \hat{\theta} + \sigma_{\hat{\theta}}]$ is 68.3% central confidence interval.

   (b) For the mean of a Poisson distribution: observe n events, set limit on ν. If you observe none, your 95% upper limit is 3.

3. **Approximate confidence intervals using the likelihood function or χ^2:** take interval where $\log L$ within $1/2$ of maximum $\rightarrow$ approximate 68.3% confidence interval.