# A Comparison of Methods for Confidence Intervals 

Alexander D. Bukin<br>Budker Institute of Nuclear Physics<br>and<br>Stanford Linear Accelerator Center

Presented at PHYSTAT2003: Statistical Problems In Particle Physics, Astrophysics, And Cosmology , 9/8/2003-9/11/2003, Menlo Park, CA, USA

Work supported by Department of Energy contract DE-AC03-76SF00515.

# A Comparison of Methods for Confidence Intervals 

A.Bukin<br>Budker INP, 630090 Novosibirsk, Russia

Comparisons are carried out of the confidence intervals constructed with Neyman's frequentist method and with the $\Delta L=1 / 2$ likelihood method, using the example of low-statistics life time estimates.

## 1. P.D.F. FOR LIFE TIME ESTIMATORS

For a given value $\tau$ of the true life time, the P.D.F. of a measurement is

$$
\frac{\mathrm{d} W}{\mathrm{~d} t}=\frac{1}{\tau} \exp \left(-\frac{t}{\tau}\right)
$$

and so for an experiment with $n$ measurements

$$
\begin{equation*}
\mathrm{d} W=\frac{1}{\tau^{n}} \cdot \prod_{k=1}^{n} \mathrm{~d} t_{k} \cdot \exp \left(-\frac{t_{k}}{\tau}\right) \tag{1}
\end{equation*}
$$

The negative log likelihood function is

$$
\begin{equation*}
L=n \ln \tau+\frac{1}{\tau} \cdot \sum_{k=1}^{n} t_{k} \tag{2}
\end{equation*}
$$

The maximum likelihood estimator of the lifetime can easily be found minimizing $L$

$$
\begin{equation*}
\hat{\tau}=\frac{1}{n} \sum_{k=1}^{n} t_{k} ; \quad \min L=L_{0}=n+n \ln \hat{\tau} \tag{3}
\end{equation*}
$$

so the probability (1) can be transformed to

$$
\begin{equation*}
\frac{\mathrm{d} W}{\mathrm{~d} \hat{\tau}}=\frac{1}{(n-1)!} \cdot\left(\frac{n \hat{\tau}}{\tau}\right)^{n-1} \cdot \frac{n}{\tau} \cdot \exp \left(-\frac{n \hat{\tau}}{\tau}\right) \tag{4}
\end{equation*}
$$

Given some true value $\tau$ then for any algorithm that defines a confidence interval $\hat{\tau}_{-\Delta \tau^{(-)}}^{+\Delta \tau^{(+)}}$we can evaluate the coverage $P$ :

$$
\begin{equation*}
P=\frac{1}{(n-1)!} \cdot \frac{n}{\tau} \cdot \int_{\hat{\tau}_{1}}^{\hat{\tau}_{2}}\left(\frac{n \hat{\tau}}{\tau}\right)^{n-1} \cdot \exp \left(-\frac{n \hat{\tau}}{\tau}\right) \mathrm{d} \hat{\tau} \tag{5}
\end{equation*}
$$

where $\hat{\tau_{1}}+\Delta \tau^{(+)}=\tau ; \quad \hat{\tau_{2}}-\Delta \tau^{(-)}=\tau$.

## 2. LIKELIHOOD FUNCTION CONFIDENCE INTERVAL

The conventional Likelihood function method for finding a $68 \%$ confidence interval $[1,2]$ is to find the values of $\tau$ for which

$$
\Delta L=L-L_{0}=\frac{1}{2}
$$

In our case

$$
\begin{equation*}
\Delta L=n \cdot\left(\frac{\hat{\tau}}{\tau}-1+\ln \frac{\tau}{\hat{\tau}}\right)=\frac{1}{2} \tag{6}
\end{equation*}
$$

For example, for $n=5$ the limits are

$$
\begin{equation*}
\left[\tau_{1}, \tau_{2}\right]=[0.6595 \hat{\tau}, 1.6212 \hat{\tau}] \tag{7}
\end{equation*}
$$

The coverage of this interval, from Equation (5), is

$$
\frac{1}{4!} \cdot \frac{5}{\tau} \cdot \int_{\hat{\tau}_{1}}^{\hat{\tau}_{2}}\left(\frac{5 \hat{\tau}}{\tau}\right)^{4} \cdot \exp \left(-\frac{5 \hat{\tau}}{\tau}\right) \mathrm{d} \hat{\tau}=0.6747
$$

where the integration limits, corresponding to (7), are

$$
\left[\hat{\tau_{1}}, \hat{\tau_{2}}\right]=[0.6168 \tau, 1.5163 \tau]
$$

The coverage is close to, but significantly different from, the nominal value of 0.6827 .

Examples of confidence intervals obtained by this means are shown in Table I, as the values in parentheses. The $95 \%$ confidence interval was obtained using the rule $\Delta L=2$, and $90 \%$ upper limit using a one side interval for which

$$
\Delta L=\left[\operatorname{erf}^{-1}(2 \cdot 0.9-1)\right]^{2} \approx 0.821
$$

The coverage given by such intervals is shown in Fig. 1, evaluated using a Monte Carlo method.

## 3. BAYESIAN CONFIDENCE INTERVAL

For comparison we can estimate a Bayesian confidence interval for the same example of $n=5$. In the Bayesian approach [3-5], the likelihood function is considered to be a probability density for the true parameter $\tau$. Assuming a flat prior distribution for $\tau$ this is

$$
\frac{\mathrm{d} W}{\mathrm{~d} \tau}=P(\tau) \sim \mathcal{L}=\frac{1}{\tau^{n}} \cdot \exp \left(-\frac{n \hat{\tau}}{\tau}\right)
$$

After normalization (for $n \geq 2$ ) this becomes:

$$
P(\tau)=\frac{(n \hat{\tau})^{n-1}}{(n-2)!\cdot \tau^{n}} \cdot \exp \left(-\frac{n \hat{\tau}}{\tau}\right)
$$



Figure 1: Coverage for likelihood function confidence intervals, evaluated by Monte Carlo. Statistical errors are shown when they exceed the size of polymarker. $\mathbf{\Delta}$ $95 \%$ Conf.Interv., ■ - $90 \%$ Upper limit, • - $68 \%$ Conf.Interv.


Figure 2: Probability density function for the true value of the parameter $\tau$ in a Bayesian approach (Equation (8) with $n=5$ and $\hat{\tau}=1$ ). The shaded regions are the $16 \%$ "tails".
which for $n=5$ gives

$$
\begin{equation*}
\int_{0}^{\tau} \mathrm{d} W=\left[1+\frac{5 \hat{\tau}}{\tau}+\frac{1}{2}\left(\frac{5 \hat{\tau}}{\tau}\right)^{2}+\frac{1}{6}\left(\frac{5 \hat{\tau}}{\tau}\right)^{3}\right] \cdot e^{-\frac{5 \hat{\gamma}}{\tau}} \tag{8}
\end{equation*}
$$

The $68 \%$ central confidence region for this distribution is (see Fig. 2):

$$
\tau=\hat{\tau} \cdot\left(1_{-0.1552}^{+1.3974}\right)
$$

The coverage of this region is actually not $68.27 \%$ but $64.31 \%$.

## 4. NEYMAN'S CONFIDENCE INTERVAL

Neyman [5-7] proposed a frequentist construction of a confidence zone (or confidence belt) as follows


Figure 3: Illustration of the construction of a confidence zone (or confidence belt).
(see Figure 3):

1. One obtains functions $\hat{\tau}_{1}(\tau)$ and $\hat{\tau}_{2}(\tau)$ of the true parameter $\tau$ such that

$$
\begin{aligned}
& \int_{0}^{\hat{\tau}_{1}(\tau)} \frac{\mathrm{d} W(\hat{\tau} ; \tau)}{\mathrm{d} \hat{\tau}} \mathrm{~d} \hat{\tau}=\frac{1-\beta}{2} \\
& \int_{\hat{\tau}_{2}(\tau)}^{\infty} \frac{\mathrm{d} W(\hat{\tau} ; \tau)}{\mathrm{d} \hat{\tau}} \mathrm{~d} \hat{\tau}=\frac{1-\beta}{2}
\end{aligned}
$$

where $\beta$ is the confidence level required, here $\beta=0.6827$. For $n=5$ these are simply $\hat{\tau}_{1}(\tau)=$ $0.568 \tau, \hat{\tau}_{2}(\tau)=1.433 \tau$, as shown in Figure 3.
2. One defines the inverse functions

$$
\tau_{1}(\hat{\tau})=\hat{\tau}_{2}^{-1}(\hat{\tau}) ; \quad \tau_{2}(\hat{\tau})=\hat{\tau}_{1}^{-1}(\hat{\tau})
$$

which, for a given value of $\hat{\tau}$, define the borders of the confidence interval for $\tau$, with coverage $\beta$.
In our example, there are $\tau_{1}(\hat{\tau})=0.698 \hat{\tau}$, $\tau_{2}(\hat{\tau})=1.760 \hat{\tau}$.

Thus the result of a lifetime experiment of this type can be written

$$
\tau=\hat{\tau} \cdot\left(1_{-0.302}^{+0.760}\right)
$$

The coverage evaluated is 0.6826 - the difference of 0.0001 is purely due to rounding errors.

Table I shows these intervals for several values of $n$, with the likelihood approximation shown in parentheses for comparison.

Table II compares the coverage of all three methods for the $n=5$ case.

Table I Lifetime confidence intervals obtained by Neyman's method for various values of $n$, the number of measurements, and confidence levels.

| $n$ | $68 \%$ C.L. |  | $90 \%$ C.L. |  | 90\% C.L. |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\frac{\Delta \tau^{(-)}}{\hat{\tau}}$ | $\frac{\Delta \tau^{(+)}}{\hat{\tau}}$ | $\frac{\Delta \tau^{(-)}}{\hat{\tau}}$ | $\frac{\Delta \tau^{(+)}}{\hat{\tau}}$ | upper limit |
| 1 | $0.457(0.576)$ | $4.789(2.314)$ | $0.736(0.778)$ | $42.45(18.06)$ | $9.49 \hat{\tau}(8.49 \hat{\tau})$ |
| 2 | $0.394(0.469)$ | $1.824(1.228)$ | $0.648(0.682)$ | $7.690(5.305)$ | $3.76 \hat{\tau}(2.76 \hat{\tau})$ |
| 3 | $0.353(0.410)$ | $1.194(0.894)$ | $0.592(0.621)$ | $4.031(3.164)$ |  |
| 4 | $0.324(0.370)$ | $0.918(0.725)$ | $0.551(0.576)$ | $2.781(2.314)$ |  |
| 5 | $0.302(0.341)$ | $0.760(0.621)$ | $0.519(0.541)$ | $2.159(1.858)$ |  |
| 6 | $0.284(0.318)$ | $0.657(0.550)$ | $0.492(0.513)$ | $1.786(1.571)$ |  |
| 7 | $0.270(0.299)$ | $0.584(0.497)$ | $0.470(0.489)$ | $1.538(1.374)$ |  |
| 8 | $0.257(0.284)$ | $0.529(0.456)$ | $0.452(0.469)$ | $1.359(1.228)$ |  |
| 9 | $0.247(0.271)$ | $0.486(0.423)$ | $0.435(0.451)$ | $1.225(1.116)$ |  |
| 10 | $0.237(0.260)$ | $0.451(0.396)$ | $0.421(0.436)$ | $1.119(1.027)$ |  |
| 20 | $0.182(0.194)$ | $0.285(0.261)$ | $0.331(0.341)$ | $0.654(0.621)$ |  |
| 50 | $0.124(0.129)$ | $0.164(0.156)$ | $0.232(0.237)$ | $0.356(0.346)$ |  |

Table II Coverage of all three methods for $n=5$

| Method | Negative error <br> $\Delta \tau^{(-)} / \hat{\tau}$ | Positive error <br> $\Delta \tau^{(+)} / \hat{\tau}$ | Coverage, \% |
| :--- | :---: | :---: | :---: |
| Likelihood | 0.341 | 0.621 | 67.47 |
| Bayesian | 0.155 | 1.397 | 64.31 |
| Neyman's | 0.302 | 0.760 | 68.26 |

## 5. CONCLUSION

- Neyman's method for confidence intervals provides exact coverage, by construction.
- The intervals from $\Delta L=1 / 2$ agree well with the Neyman intervals for large $n$, but differ for small $n$, as can be seen in Table I. In such cases they undercover, i.e. the interval is smaller than the true one.
- Bayesian confidence intervals give very different results, and can undercover or overcover.


## References

[1] Derek J. Hudson. Lectures on elementary statistics and probability, CERN, Geneva, 1963
[2] A.G. Frodesen, O. Skjeggestad, H. Toffe. "Probability and statistics in particle physics". Universitetsforlaget, Oslo, 1979
[3] H. Jeffreys. "Theory of probability", 2nd ed., Oxford Univ. Press, 1948.
[4] B.P. Carlin and T.A. Louis. "Bayes and empirical Bayes methods for data analysis", Chapman \& Hall, London, 1996
[5] M. Kendall and A. Stuart. "The advanced theory of statistics", vol. 2, "Inference and relationship". Macmillan Publishing Co., New York, 1978.
[6] J. Neyman. "Outline of a theory of statistical estimation based on the classical theory of probability". Phil. Trans. A, 236 (1937) 333.
[7] R.J. Barlow. "Statistics. A guide to the use of statistical methods in the physical sciences." John Wiley \& Sons ltd., Chichester, England, 1989

