University of Glasgow

Statistical Inference, Lecture 10

Dr Benn Macdonald

Room 225 Maths and Stats Building,
Benn.Macdonald@glasgow.ac.uk

February 2021

- **Likelihood Intervals**
- **One and two-sample t-intervals and tests (exact coverage)**
- Wilks and Wald Intervals (approximate coverage)
- Multiparameter models
  - Wilks and Wald confidence regions

Interval Estimation

Rather than simply quoting point estimates of population parameters,

we want to identify an **interval estimate**, a range of plausible values, for an unknown parameter of interest,

to take into account the sampling variability in the estimators.

3.5 Interval Estimation using Likelihood

$$\frac{L(\theta)}{L(\hat{\theta}_{MLE})} \geq p$$

i.e.

$$R(\theta) \geq p$$

For p between 0 and 1, this is known as:

a 100p% likelihood interval for θ.

3.5 Interval Estimation using Likelihood

The log relative likelihood function:

$$r(\theta) = \ell(\theta) - \ell(\hat{\theta}_{MLE})$$

In terms of $r(\theta)$, a $100p\%$ likelihood interval for θ is defined by

$$R(\theta) \geq p$$

i.e.

$$r(\theta) \geq \log_e(p)$$

Example 7 continued: Air Conditioning Failures:
As mentioned previously, **a 50% likelihood interval is defined by**:

$$r(\theta) \geq \log_e(0.5) = -0.693$$

where,

$$r(\theta) = 24 \log_e(\theta) - 1539\theta + 123.9.$$

3.5 Interval Estimation using Likelihood

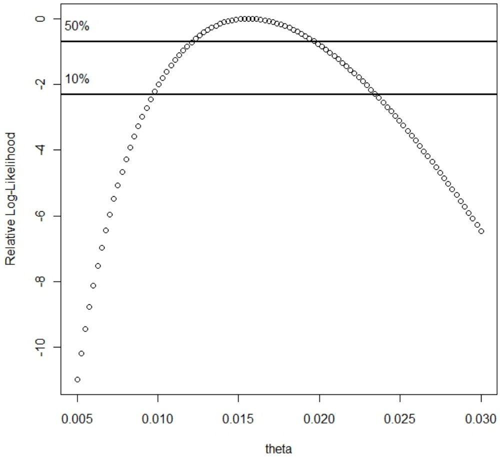


Figure: Relative log-likelihood function for Air Conditioning Data

3.5 Interval Estimation using Likelihood

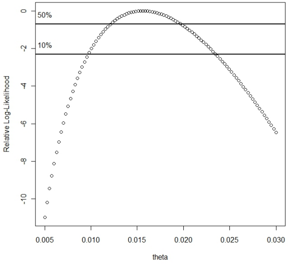


Figure: Relative log-likelihood function for Air Conditioning Data

For example, it appears as though a **50% interval for θ is approximately**:

$$(0.011, 0.019)$$

3.5 Interval Estimation using Likelihood

In this case the following iterative algorithm is used to find the bounds (B) for the interval, where θ_L is the lower bound and θ_U is the upper bound:

$$\theta_B^{(j+1)} = \theta_B^{(j)} - \frac{g(\theta_B^{(j)})}{g'(\theta_B^{(j)})}$$

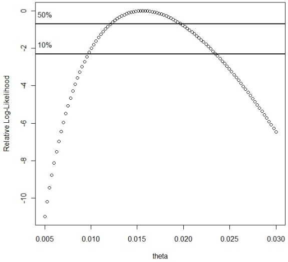
with $g(\theta_B) = r(\theta) - \log_e(0.5)$, where:

$$g(\theta_B) = 24\log_e(\theta_B) - 1539\theta_B + 123.9 + 0.693$$

$$g'(\theta_B) = \frac{24}{\theta_B} - 1539$$

3.5 Interval Estimation using Likelihood

To start the Newton-Raphson algorithm, **initial values** can be estimated from a plot of the relative log likelihood



i.e. for a lower bound start with 0.011 and for the upper bound start at 0.019.

3.5 Interval Estimation using Likelihood

Iteration	$\theta^{(j)}$	$g(\theta^{(j)})$	$g'(\theta^{(j)})$	$\frac{g(\theta^{(j)})}{g'(\theta^{(j)})}$
0	0.011	-0.573	642.82	-0.0009
1	0.012	-0.023	461	-0.0001
2	0.012	-0.023	461	-0.0001

Table: Newton-Raphson iterative algorithm to obtain the lower bound for an interval estimate around θ.

3.5 Interval Estimation using Likelihood

So, in Example 7, **a 50% likelihood interval for $\hat{\theta}_{MLE}$** is:

$$(0.012, 0.020)$$

This means it is likely that the unknown population parameter θ lies in this range.

It is likely that the average failure rate is in the range 0.012 to 0.020.

3.6 Confidence Intervals

What is a sensible choice for p?

In the **Normal model**, likelihood estimation is equivalent to another, older approach that produces interval estimates (known as **confidence intervals**).

Confidence intervals have a more specific interpretation.

This means that some values of p give more intuitively attractive interval estimates than others.

3.6.1 MLE for a Normal Model

Assume that $X_1, X_2, \ldots, X_n$, are independent $N(\mu, \sigma^2)$, where the standard deviation, σ, of the distribution is assumed to be known and always to be the same (no matter what the true population mean μ might be).

$$L(\mu; x_1, \ldots, x_n) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right\}$$

$$\hat{\mu}_{MLE} = \bar{x}$$

It can be shown (in pages 7-9 of your notes) that a **14.65% likelihood interval** for μ_T

$$r(\mu) \geq \log_{\mathrm{e}}(0.1465) = -1.92$$

is equivalent to a **95% confidence interval** for μ_T.

3.6.3.1 95% Confidence Intervals

Page 9

What is the interpretation of a 95% confidence interval (CI) for μ?

An interval estimate that has 95% coverage is called a 95% confidence interval.

On 95% of the occasions on which a 95% confidence interval is calculated from sample data, it will contain the true value of the parameter.

3.6.3.1 95% Confidence Intervals

A 95% confidence interval will not always contain the true value of the parameter.

In fact, on average only 95% of such intervals will do so.

3.6.3.1 Confidence Intervals

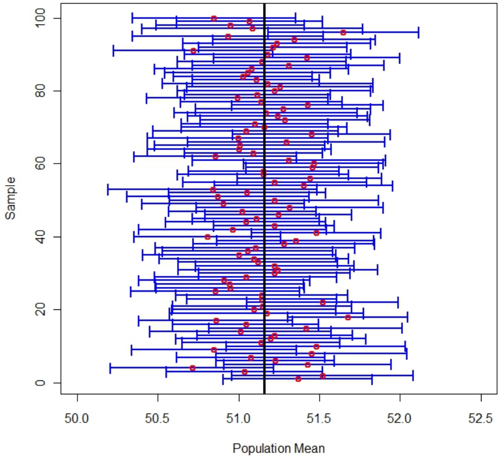


Figure: 100 95% CIs for the population mean based on $n = 25$.

3.6.3.1 Confidence Intervals

If we randomly choose one realisation, the probability is 95% that we choose an interval that contains the true population mean.

Averaging over many samples, 95% of the 95% confidence intervals constructed will capture the true population mean.

And 5% will not!

3.6 Confidence Intervals

Assume that $X_1, X_2, \ldots, X_n$, are independent $N(\mu, \sigma^2)$, where σ, is known.

$$\hat{\mu}_{MLE} = \bar{x}$$

We are interested in constructing a confidence interval for the population mean μ_T.

3.6.4 One-sample t-interval and t-test

How do we construct a confidence interval for μ?

Page 11

A **pivotal function** is a function of the data, $\mathbf{X}$, and the parameter of interest, θ, which, when regarded as a random variable calculated at θ_T (the true value of θ), has a probability distribution whose form does not depend on any unknown parameter.

We usually denote a pivotal function by $PIV(\theta_T, \mathbf{X})$.

3.6.4 One-sample t-interval and t-test

How do we construct a confidence interval for μ?

Page 11

A **pivotal function** for μ_T (when σ is **known**) is:

$$PIV(\mu_T, \mathbf{X}) = \frac{\bar{X} - \mu_T}{\sigma/\sqrt{n}} = Z$$

(see pages 7 and 8).

This pivotal function has an $N(0, 1)$ distribution whatever the value of μ_T.

3.6.4 One-sample t-interval and t-test

Now assume that $X_1, X_2, \ldots, X_n$ are independent random variables, each with a $N(\mu_T, \sigma_T^2)$ distribution, and that we wish to estimate μ_T, but that **both $\boldsymbol{\mu_T}$ and $\boldsymbol{\sigma_T}$ are unknown**.

3.6.4 One-sample t-interval and t-test

A **pivotal function for** $\boldsymbol{\mu_T}$ (when σ_T is unknown) is found by replacing σ_T with its estimator, s.

It can be shown that:

$$t = \frac{\bar{X} - \mu_T}{s/\sqrt{n}} \sim t(n-1)$$

where $t(n-1)$ is the Student's t distribution with $n-1$ degrees of freedom.

3.6.4 One-sample t-interval and t-test

A pivotal function for estimating μ_T, when σ is **known**:

$$Z = \frac{\bar{X} - \mu_T}{\sigma/\sqrt{n}} \sim N(0,1)$$

A pivotal function for estimating μ_T, when σ is **unknown**:

$$t = \frac{\bar{X} - \mu_T}{s/\sqrt{n}} \sim t(n-1)$$

3.6.4 One-sample t-interval and t-test

Page 12

Having identified a suitable pivotal function (if one exists), then **we can construct a confidence interval** for our parameter of interest.

It is possible to produce $100c\%$ confidence intervals for any value of c in the range $0 < c < 1$.

3.6.4 One-sample t-interval and t-test

Let $t_{1-\frac{(1-c)}{2}}(n-1)$ denote the value of a t random variable such that: $P(t \leq t_{1-\frac{(1-c)}{2}}(n-1)) = 1 - \frac{(1-c)}{2}$.

For example, let $t_{0.975}(n-1)$ denote the value of a t random variable such that: $P(t \leq t_{0.975}(n-1)) = 0.975, c = 0.95$.

3.6.4 One-sample t-interval and t-test

Since the $t(n-1)$ distribution is symmetric around 0, there is probability 0.95 (or 95%) that:

$$-t_{0.975}(n-1) \leq t = \frac{\bar{X} - \mu_T}{s/\sqrt{n}} \leq t_{0.975}(n-1)$$

i.e.

$$\mu_T \leq \bar{X} + t_{0.975}(n-1)s/\sqrt{n}$$

and

$$\mu_T \geq \bar{X} - t_{0.975}(n-1)s/\sqrt{n}$$

simultaneously.

3.6.4 One-sample t-interval and t-test

A 95% confidence interval for $\boldsymbol{\mu_T}$ (when σ_T is unknown) is:

$$\left(\bar{x} - t_{0.975}(n-1)\frac{s}{\sqrt{n}}, \bar{x} + t_{0.975}(n-1)\frac{s}{\sqrt{n}}\right)$$

i.e.

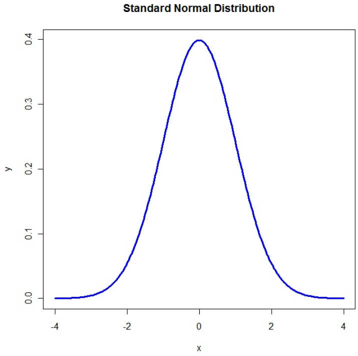
$$\left(\bar{x} \pm t_{0.975}(n-1)\frac{s}{\sqrt{n}}\right)$$

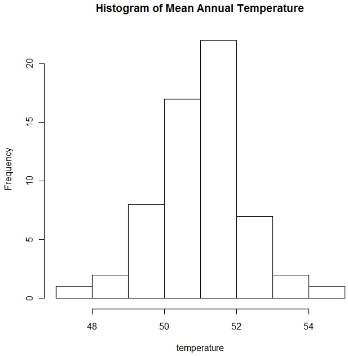
estimate $\pm\, t_{0.975}(n-1) \times$ estimated standard error

Example 9 - Mean Annual Temperature in New Haven

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{3069.6}{60} = 51.16, s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1.602$$

3.6.4 One-sample t-interval and t-test





3.6.4 One-sample t-interval and t-test

Example 9 - Mean Annual Temperature in New Haven

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{3069.6}{60} = 51.16, s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1.602$$

A 95% confidence interval for μ_T (when σ_T is unknown) is:

$$\left(\bar{x} \pm t_{0.975}(n-1)\frac{s}{\sqrt{n}}\right)$$

$$\left(51.16 - t_{0.975}(60-1)\frac{\sqrt{1.602}}{\sqrt{60}}, 51.16 + t_{0.975}(60-1)\frac{\sqrt{1.602}}{\sqrt{60}}\right)$$

3.6.4 One-sample t-interval and t-test

Example 9 - Mean Annual Temperature in New Haven

$$t_{0.975}(59) = 2.001$$

$$\left(51.16 - 2.001\frac{\sqrt{1.602}}{\sqrt{60}}, 51.16 + 2.001\frac{\sqrt{1.602}}{\sqrt{60}}\right)$$

$$(51.16 - 2.001 \times (0.163), 51.16 + 2.001 \times (0.163))$$

$$(50.83, 51.49)$$

Example 9 - Mean Annual Temperature in New Haven

Conclusion

It can therefore be concluded that the population mean annual temperature is highly likely to lie in the range 50.8^oF to 51.5^oF, with a point estimate for $\hat{\mu}$ of 51.2^oF