Likelihood based confidence intervals

Christine Borgen Linander

DTU Compute
Section for Statistics
Technical University of Denmark
chjo@dtu.dk

August 20th 2015

DTU Compute
Department of Applied Mathematics and Computer Science

DTU

A huge thank to a former colleague of mine Rune H B Christensen.

Outline

Motivation

We are interested in:

1. Which sensory difference (d') is most supported by the data?
2. Which interval of sensory differences is supported by the data?

We usually answer those with:

1. The maximum likelihood estimate (MLE), $\hat{d}'$
2. A confidence interval (CI) for d'

. . . but there are many ways to compute the CI, and which is best?

Problems with standard CIs

Standard (Wald) 95% confidence intervals:

- For binomial probability of a correct answer p_c: $\hat{p}_c \pm 1.96 \cdot \text{se}(\hat{p}_c)$
- For the Thurstonian δ: $\hat{\delta} \pm 1.96 \cdot \text{se}(\hat{\delta})$ (Bi et al, 1997)

Problems and solution:

- The standard CIs are incompatible and lead to contradictions
- The standard CIs do not cover the values of δ or p_c that are most supported by the data
- CIs based on the likelihood function have better properties

John and Dorothy's duo-trio experiment

- The guessing probability is $1/2$
- They obtain 13 correct answers to 20 samples
- John analyzes the probability of a correct answer, p_c: $\hat{p}_c = 0.65(0.11)$ and $CI_{95\%} = [0.44; 0.86]$ which covers $p_c = 1/2$
- Dorothy analyzes the Thurstonian δ: $\hat{\delta} = 1.42(0.63)$ and $CI_{95\%} = [0.18; 2.66]$ (Bi et al, 1997) which does NOT cover $\delta = 0$

Which method is (most) correct?

Should we trust John or Dorothy?

How much evidence is there really in the data about a difference between the products?

Peter and Sally's triangle experiment

- The guessing probability is 1/3
- They obtain 10 correct answers to 20 samples
- Peter analyzes the probability of a correct answer, p_c: $\hat{p}_c = 1/2(0.11)$ and $CI_{95\%} = [0.28; 0.72]$ which covers $p_c = 1/3$
- Sally analyzes the Thurstonian δ: $\hat{\delta} = 1.47(0.59)$ and $CI_{95\%} = [0.32; 2.62]$ (Bi et al, 1997) which does NOT cover $\delta = 0$

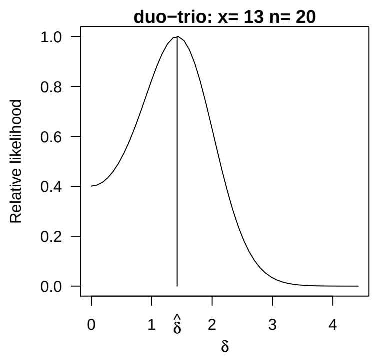
Which method is (most) correct?

Should we trust Peter or Sally?

How much evidence is there really in the data about a difference between the products?

Properties of the likelihood function

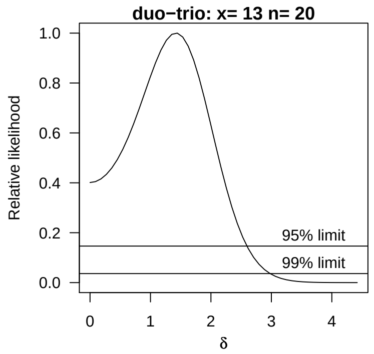
- Likelihood function = density: $L(\delta; x, n) = \binom{n}{x} p^x (1-p)^{n-x}, \; p = f_{\text{psy}}(\delta)$
- Measures support of values of δ relative to $\hat{\delta}$
- An objective way to measure information in the data about δ



- The maximum likelihood estimate (MLE)
- Likelihood CIs are given by horizontal lines
- Symmetric approximation produces standard (Wald) CIs proposed by Bi et al (1997)

Properties of the likelihood function

- Likelihood function = density:
  $L(\delta; x, n) = \binom{n}{x} p^x (1-p)^{n-x},\ p = f_{\mathsf{psy}}(\delta)$
- Measures support of values of δ relative to $\hat{\delta}$
- An objective way to measure information in the data about δ

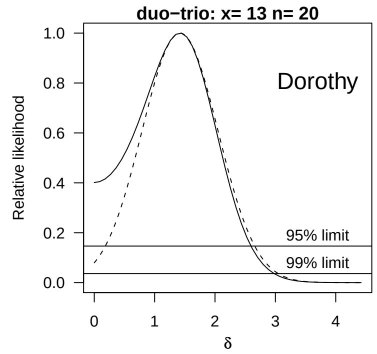


- The maximum likelihood estimate (MLE)
- Likelihood CIs are given by horizontal lines
- Symmetric approximation produces standard (Wald) CIs proposed by Bi et al (1997)

The likelihood function and confidence intervals

Properties of the likelihood function

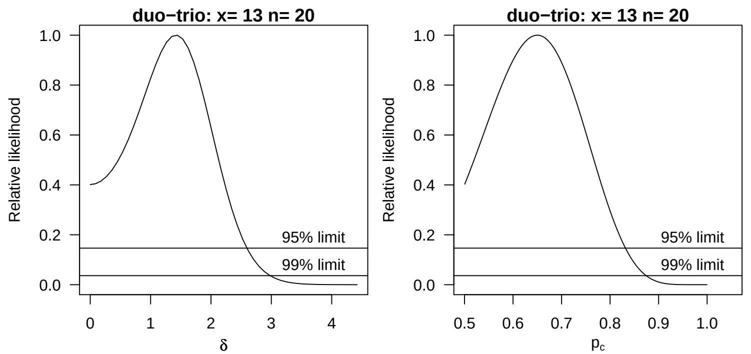
- Likelihood function = density:
  $L(\delta; x, n) = \binom{n}{x} p^x (1-p)^{n-x},\ p = f_{\mathsf{psy}}(\delta)$
- Measures support of values of δ relative to $\hat{\delta}$
- An objective way to measure information in the data about δ



- The maximum likelihood estimate (MLE)
- Likelihood CIs are given by horizontal lines
- Symmetric approximation produces standard (Wald) CIs proposed by Bi et al (1997)

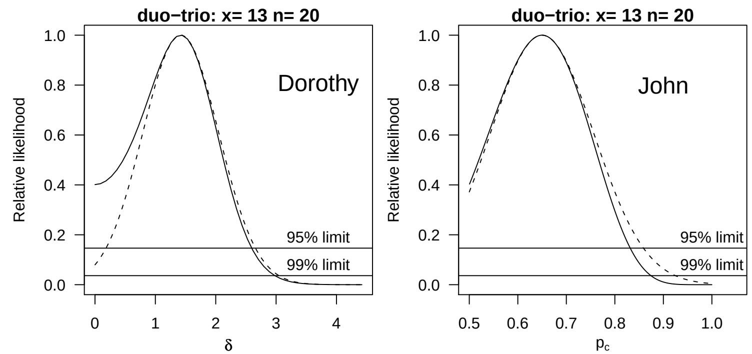
John and Dorothy's duo-trio example revisited

- "No difference" between the products has reasonably high likelihood
- An intermediate difference between products is most likely
- A large difference between products is unlikely



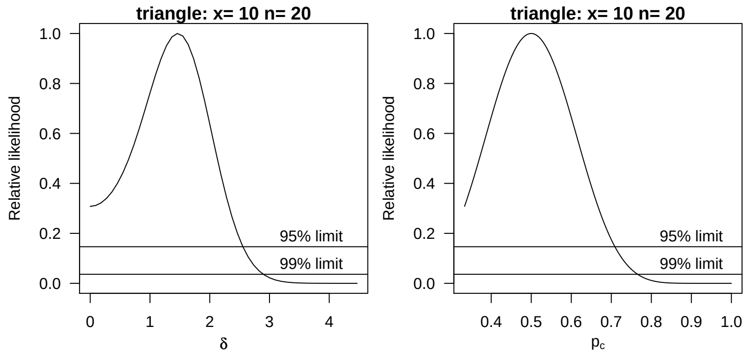
John and Dorothy's duo-trio example revisited (2)

- The symmetric approximations are inaccurate
- Neither John's nor Dorothy's CIs are appropriate
- Likelihood inference for δ and p_c is compatible



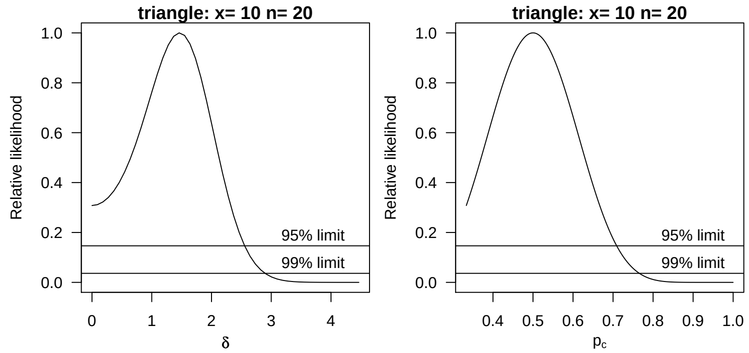
Peter and Sally's triangle example revisited (1)

- "No difference" between the products has reasonably high likelihood
- An intermediate difference between products is most likely
- A large difference between products is unlikely



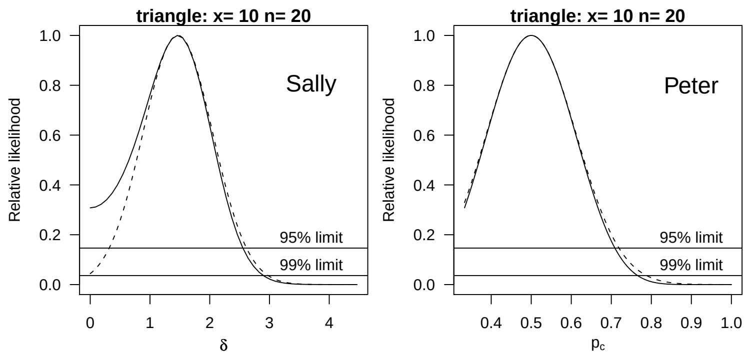
Peter and Sally's triangle example revisited (2)

- The likelihood curve tells the full story about the data
- The likelihood curve illustrates the effect of confidence level



Peter and Sally's triangle example revisited (3)

- The symmetric approximation for δ is quite inaccurate
- Sally's CI is very misleading



Coverage probability

- Boyles (2008) showed that likelihood CIs have the best coverage probability among common CIs for the binomial p.

Coverage probability in $\%$ for the binomial p with a nominal level of 95% (Boyles, 2008)

n	Standard	Exact	Likelihood
10	76.9	98.4	94.9
50	90.1	96.9	95.0
100	92.2	96.5	95.0
500	94.3	95.7	95.0

Likelihood methods in discrimination testing

Likelihood — a common framework for:

Estimation	Maximum likelihood
Testing	Likelihood ratio test
Confidence intervals	Profile likelihood

Gracefully handle boundary cases

Likelihood methods extend to complex situations