Lecture 18
Confidence Intervals

Review

There are two types of estimation

1. **Point estimation** - is estimation of the value of a parameter with the value of a statistic (i.e estimating μ with $\bar{x}$ or p with $\hat{p}$)

2. **Interval estimation** - is the estimation of the value of a parameter with an interval of values. The device we will be using for interval estimation is a *confidence* interval.

Confidence Interval for $\bar{x}$

$$\bar{x} \pm z\,\frac{\sigma}{\sqrt{n}}$$

Confidence Interval for $\hat{p}$

$$\hat{p} \pm z\sqrt{\frac{p(1-p)}{n}}$$

Example

Let x be the yield of a chemical reaction under certain circumstances. Let $\bar{x}$ be the mean yield for a sample of 25 observations of x. Suppose we obtain a sample and find that $s = 0.4$. What are the (estimated) *standard error* and the *margin of error* of $\bar{x}$? Compute the 95% confidence interval for the population mean yield.

Confidence Level

- We can control the confidence level of a confidence interval by adjusting the **standard score**

- Since both $\bar{x}$ and $\hat{p}$ are normal random variables, the standard score comes from the standard normal distribution and is represented by z

Level	z
68%	0.994
90%	1.645
95%	1.960
99%	2.576

For one standard score: $P\left(\bar{x} - 0.99\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 0.99\frac{\sigma}{\sqrt{n}}\right) \approx 0.68$

For two standard scores: $P\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right) \approx 0.95$

For three standard scores: $P\left(\bar{x} - 2.57\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 2.57\frac{\sigma}{\sqrt{n}}\right) \approx 0.99$

- The same applies to confidence intervals for $\hat{p}$

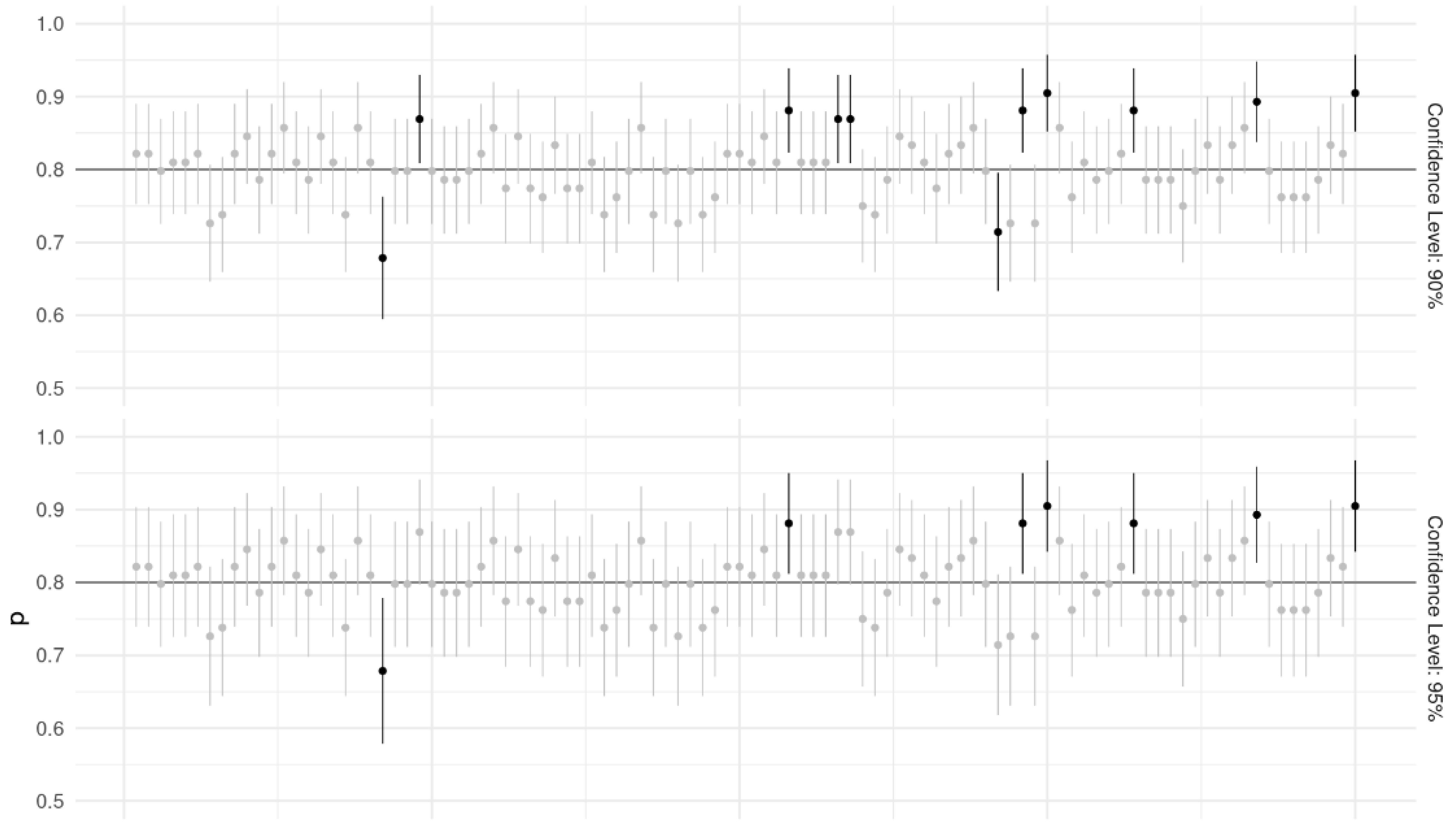
Example

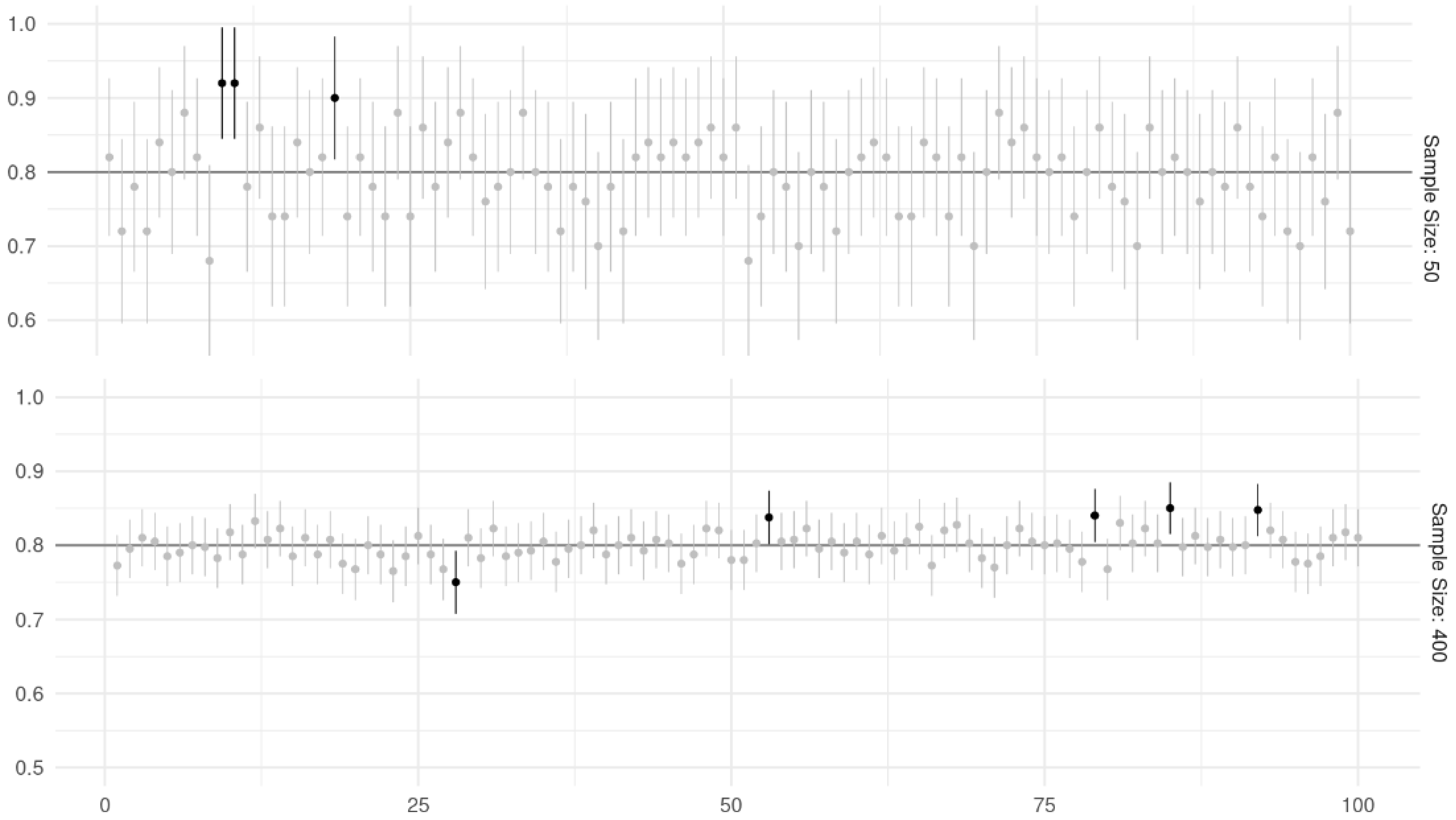
- The platyfish is a common type of aquarium fish. Female platyfish have demonstrated sexual selection based on male tail color. Suppose that females show a preference toward males with yellow tails. Out of a sample of 84 observations, the yellow-tailed male was preferred on 67 observations. Let be the probability that a female will prefer the yellow-tailed male. What is our estimate of p using the confidence interval with confidence level 95%? What about 90%?

Effects of sample size and confidence level

- How does changing the confidence level effect a confidence interval?

- How does changing the sample size change effect a confidence interval?


1.0
0.9
0.8
0.7
0.6
0.5
p
Confidence Level: 90%
Confidence Level: 95%



1.0
0.9
0.8
0.7
0.6
0.5
0
25
50
75
100
Sample Size: 50
Sample Size: 400

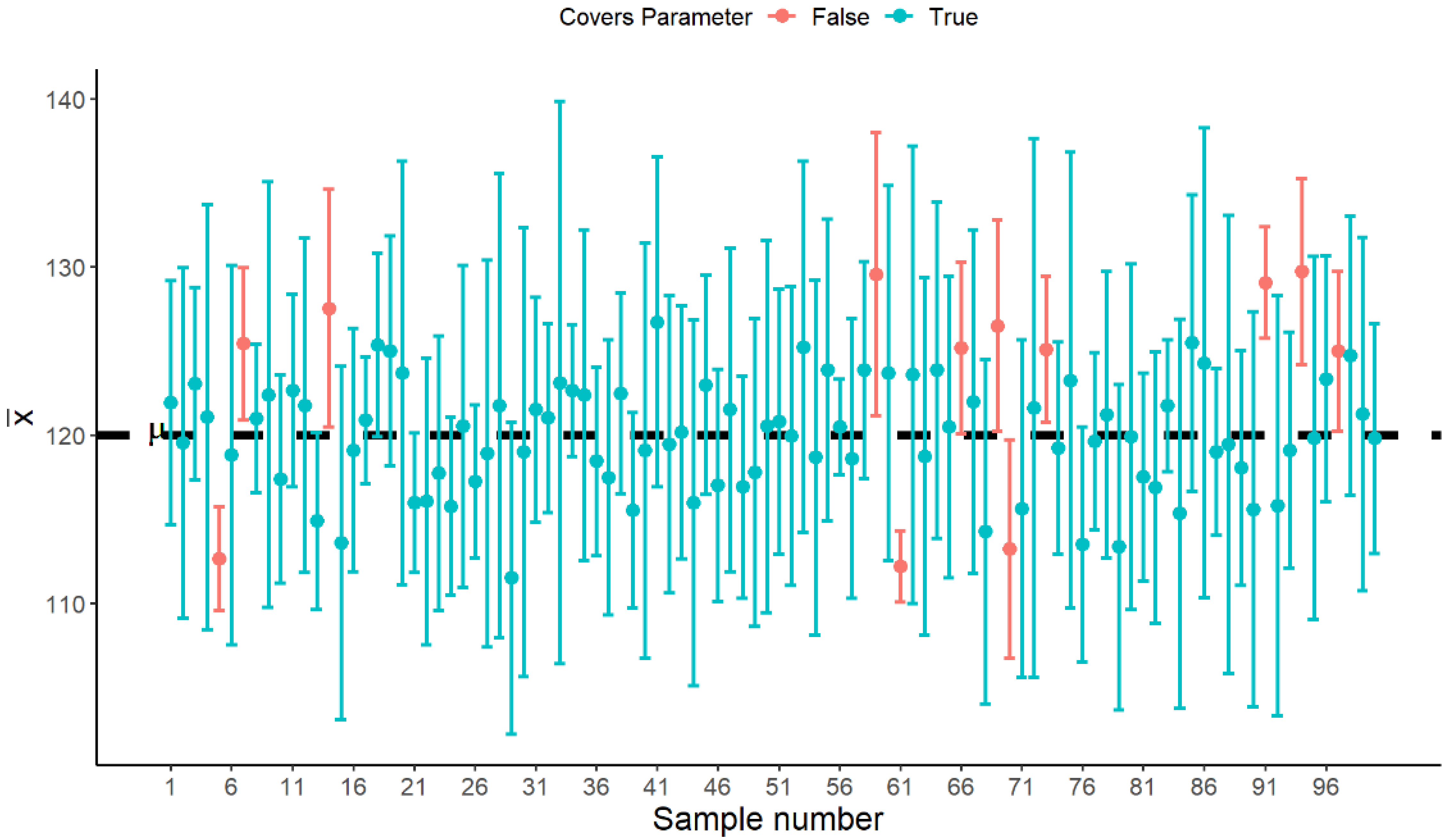
Confidence Interval for μ

- Unfortunately, the actual confidence level of the confidence interval

$$\bar{x} \pm z \frac{s}{\sqrt{n}}$$

Is less than specified confidence level, especially is n is small

This is because when n is small the Central Limit Theorem no longer applies



Covers Parameter
False
True
140
130
120
110
x̄
μ
1
6
11
16
21
26
31
36
41
46
51
56
61
66
71
76
81
86
91
96
Sample number

Confidence Interval for μ

- A solution is modify the confidence interval of μ to

$$\bar{x} \pm t \frac{s}{\sqrt{n}}$$

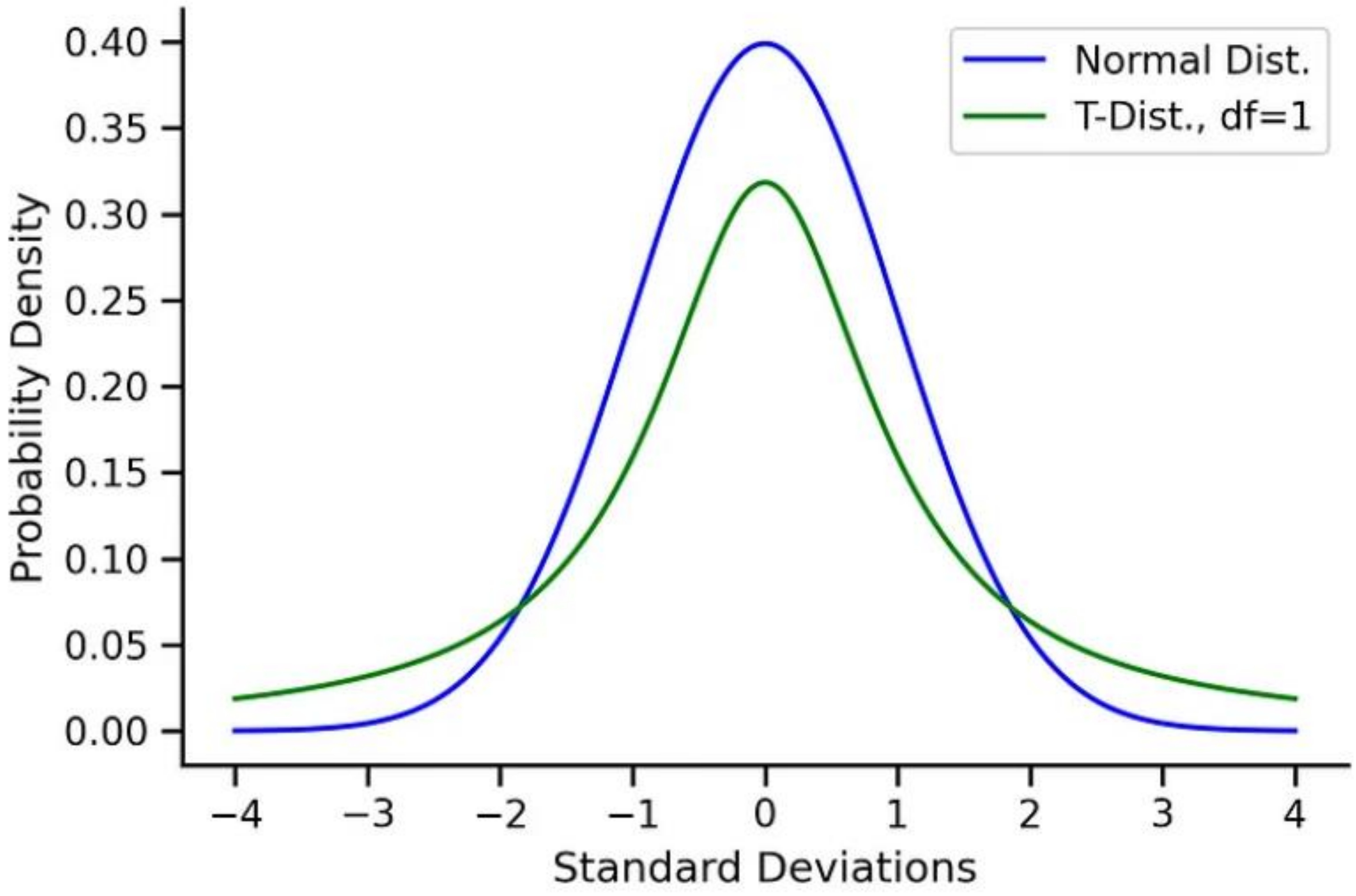
Where t denotes a t-score from the t-distribution with $n - 1$ degrees of freedom

The t-distribution resembles a standard normal distribution but with heavier tails

its standard deviation is a bit larger than 1 and depends on the *degrees of freedom*

It requires the assumption that the population distribution of x is normal

Despite this assumption it is still valid under most violations of normality


Normal Dist.
T-Dist., df=1
Probability Density
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−4
−3
−2
−1
0
1
2
3
4
Standard Deviations

Important!

- From now on we will not necessarily being using 2 as our standard score in confidence intervals.

- For confidence intervals for p look up the value of z corresponding to the desired confidence level.

- For confidence intervals for μ look up the value of corresponding to the desired confidence level and degrees of freedom $n - 1$.

Example

- A Social Survey asked “What do you think is the ideal number of children for a family to have?” 30 females were asked to give a likert response between 0 and 6 and the mean score of the recorded responses was $\bar{x} = 2.56$ with standard deviation $s = 0.84$