From Estimators to Confidence Intervals

Stat 250

Click here for PDF version

Strategy

- We have an estimator, $\hat{\theta}$, in hand
- Use $\hat{\theta}$ to find a range of plausible values for θ with known long-run properties

$$P(\hat{\theta}_L \leq \theta \leq \hat{\theta}_U) = 1 - \alpha$$

Your turn

Finding normal quantiles in R

`qnorm(p)` will calculate the p quantile of $N(0, 1)$

Find the value of `q` that is needed for the following $(1 - \alpha)100\%$ normal-based CIs:

1. 90%
2. 95%
3. 97%

03:00

Your turn

Find a 90% confidence interval for the mean bill length of Gentoo penguins.

Assume that $\sigma = 3.08$.

Sample statistics:

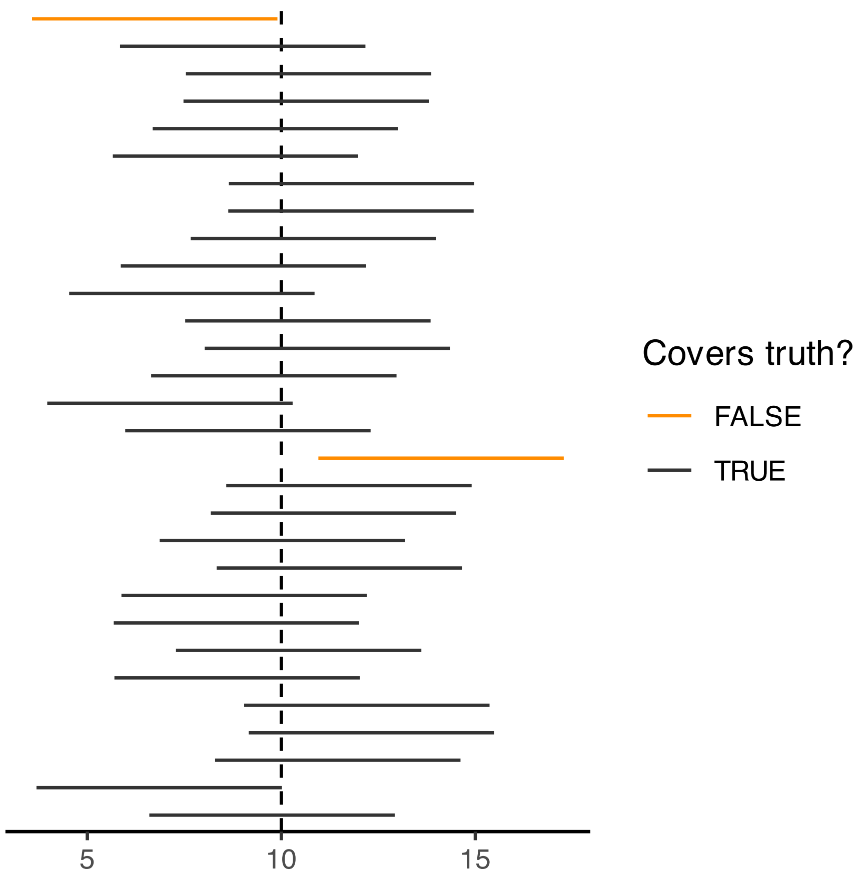
- $n = 123$
- $\bar{x} = 47.5$

Interpreting CIs redux

"Stat 101" answer

> We are $(1-\alpha)100\%$ confident that the true parameter of interest is between L and U

Interpreting CIs redux



- (L, U) is a random interval **before** data are observed
- The **process** by which the interval constructed is a random process
- $(1-\alpha)100\%$ is the long-run proportion of intervals that will capture the parameter
- In practice, we don't know which "type" of interval we have (good/bad)

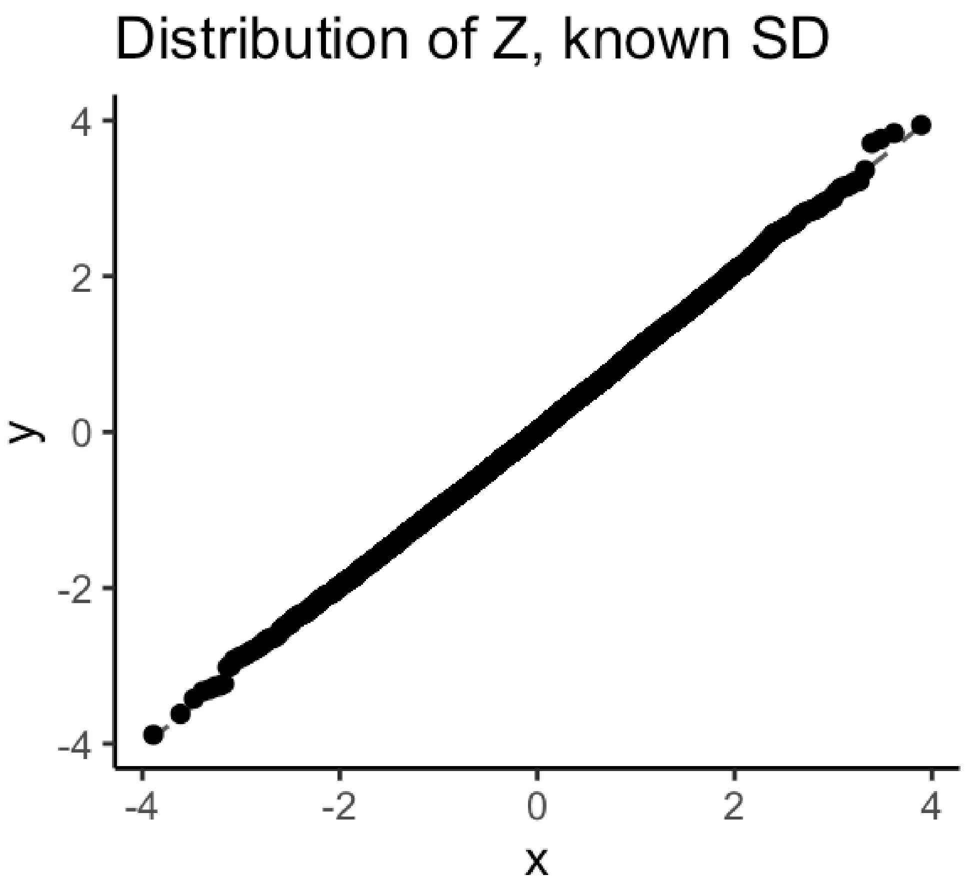
Plug-in principle

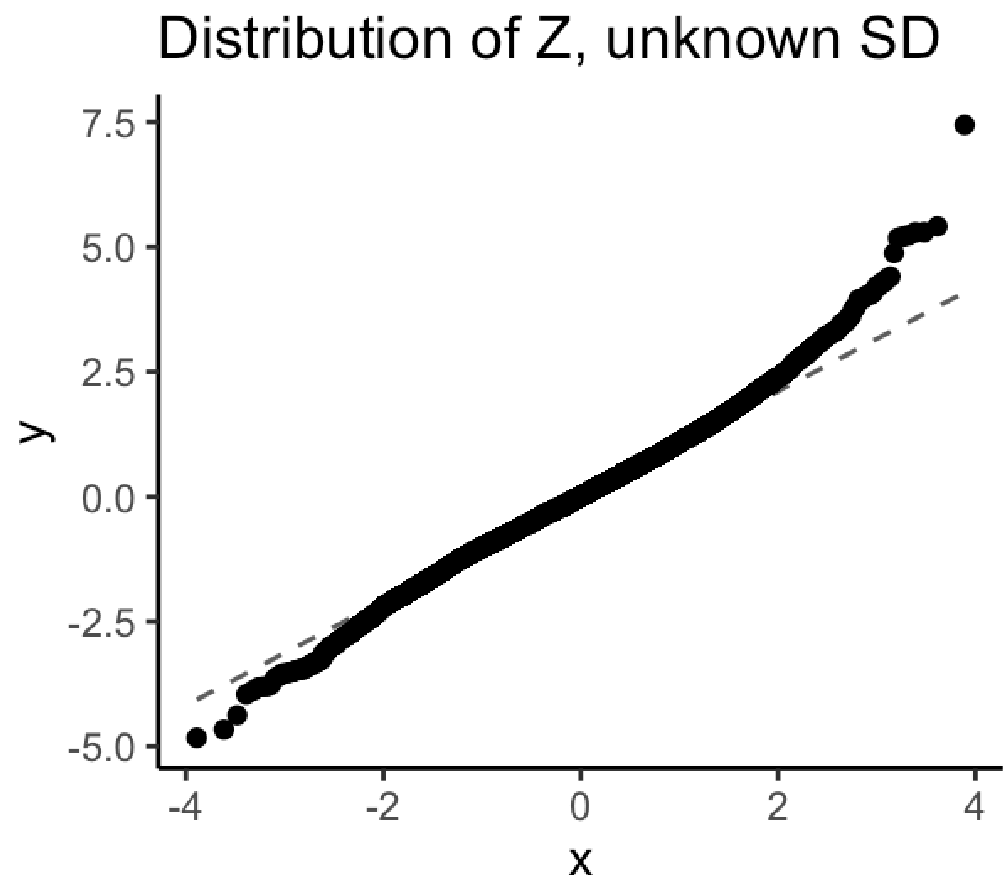
Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

- In practice both μ and σ^2 are unknown
- $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$
- PROBLEM: $\bar{x} \pm z_{1-\alpha/2} \left(\dfrac{\sigma}{\sqrt{n}} \right)$

Proposed solution: plug in the sample standard deviation

Estimating σ impacts the distribution

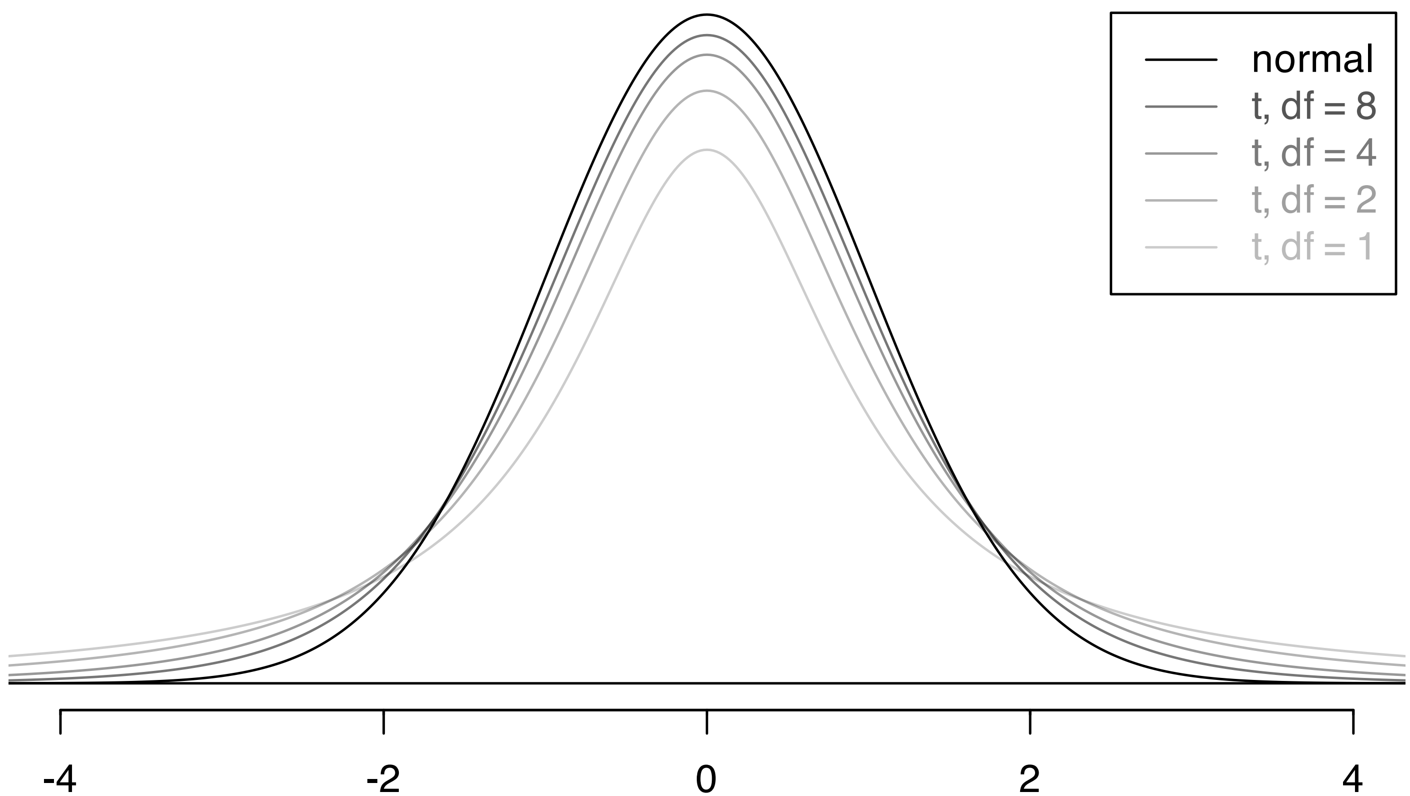




(Student's) t distribution

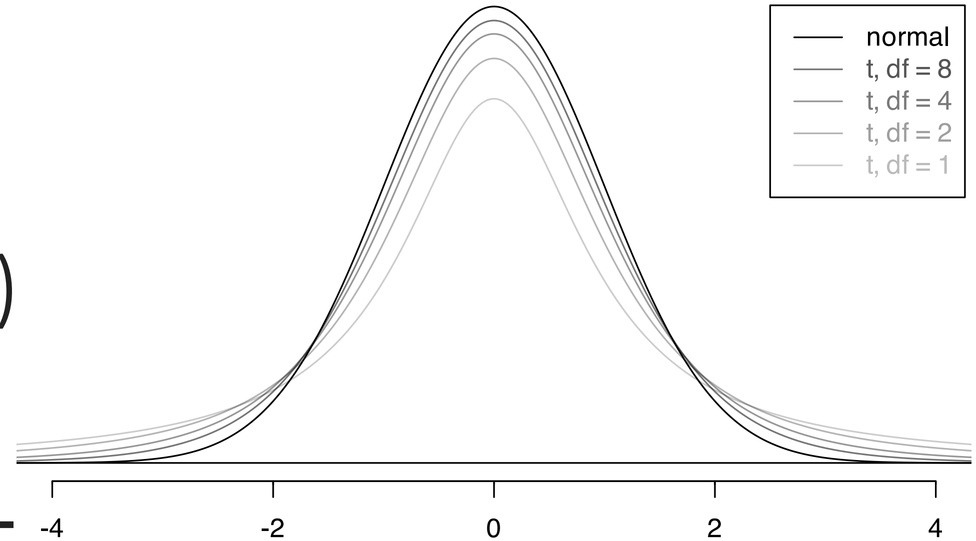
Let $T = \frac{Z}{\sqrt{V/df}}$ where $Z \sim N(0,1)$, $V \sim \chi^2_{df}$, and

$Z \perp V \implies T \sim t_{df}$



t distribution properties

- Symmetric around 0
- For $\mathrm{df} = 1$, mean doesn't exist (Cauchy distribution)
- For $\mathrm{df} \geq 2$, $E(T) = E(Z)E\left(1/\sqrt{V/n}\right) = 0$
- Heavier tails than normal distribution
- $t_{\mathrm{df}} \longrightarrow N(0,1)$ as $\mathrm{df} \longrightarrow \infty$



Your turn

Finding t quantiles in R

`qt(p, df)` will calculate the p quantile of t_{df}

Find the value of `q` that is needed for the following $(1-\alpha)100\%$ normal-based CIs:

1. 90%, n = 123
2. 95%, n = 25
3. 99%, n = 34

03:00

Your turn

Find a 90% confidence interval for the mean bill length of Gentoo penguins.

Assume that $\sigma = 3.08$.

Sample statistics:

- $n = 123$
- $\bar{x} = 47.5$
- $s = 3.08$

Underlying validity conditions

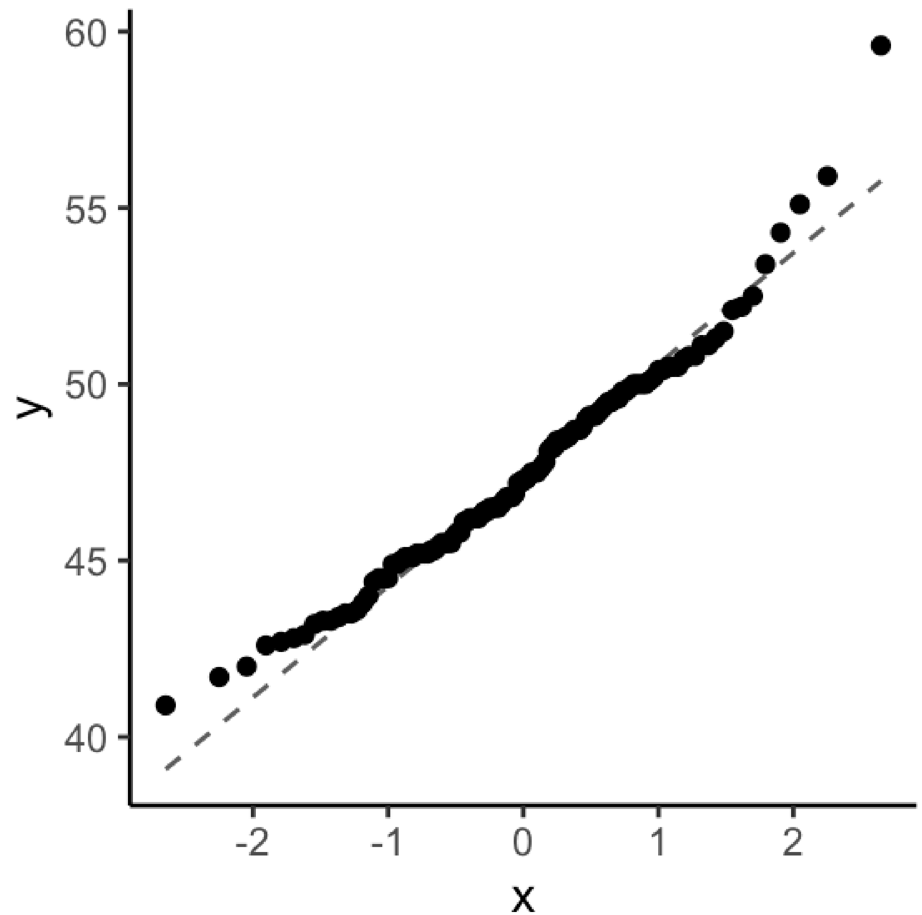
We have a random sample from a normal population distribution

Ask Yourself...

- Are the observations independent?
- Are the observations approximately normal?

Checking conditions

- Are the penguins independent?
- Are the bill lengths approximately normal?



Robustness

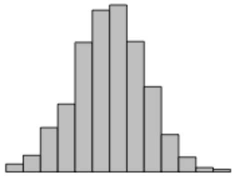
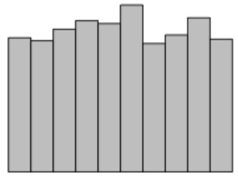
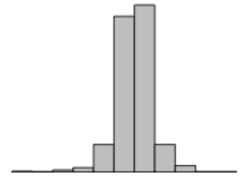
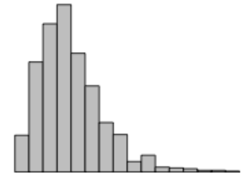
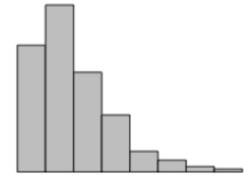
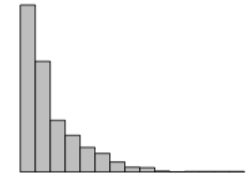
If the a procedure “perform well” even if some of the assumptions under which they were developed do not hold, then they will be called **robust.**

Simulation study

To check whether a procedure is robust, we can use simulation:

1. Simulate data from a variety of different probability distributions
2. Run the procedure (e.g., build a one-sample t-interval)
3. Compare the results of the procedure to what should have happened.

   for a large number of CIs, approximately 95% of 95% CIs should capture the parameter value

n	Bell-shaped 	Short-tailed 	Long-tailed 	Mild Skew 	Moderate Skew 	Strong Skew 
5	95.3	**94**	96.3	**91.6**	**91.8**	**89.8**
10	95.9	**94**	96.3	**93.3**	**93.2**	**90.8**
25	95.3	95.4	95.9	**93.8**	**93.5**	**90.3**
50	94.8	94.3	96.3	**94.1**	**94**	**93.8**
100	95.3	95.7	94.9	95.1	95.9	**94.6**

Robustness one-sample t

- If the population distribution is roughly symmetric and unimodal, then the procedure works well for sample sizes of at least 10–15 (just a rough guide)
- For skewed population distributions, the t-procedure can be substantially affected, depending on the severity of the skew and the sample size.
- t-procedures are not resistant to outliers.
- If observations are not independent, the results can be misleading.