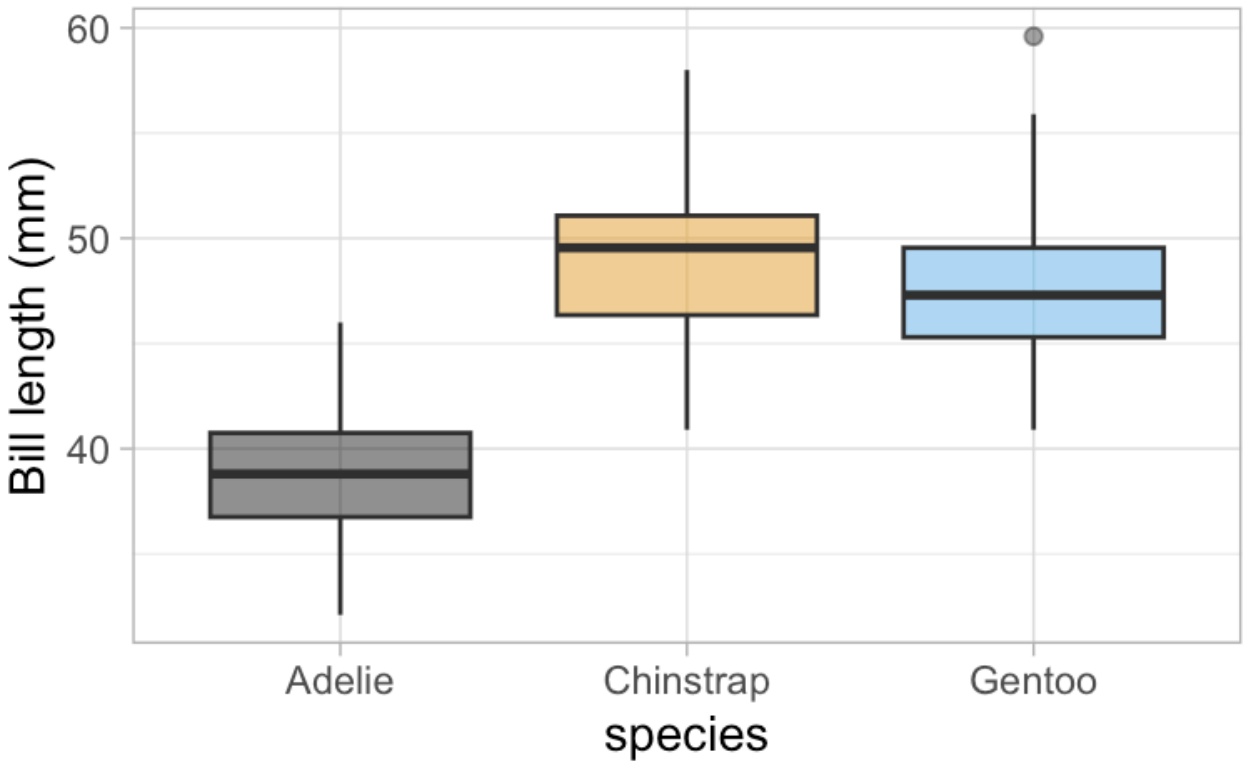
Two-Sample Confidence Intervals

Stat 250

Click here for PDF version

Bill lengths by species



species	min	Q1	median	Q3	max	mean	sd	n	missing
Adelie	32.10	36.75	38.80	40.75	46.00	38.79	2.66	151	1
Chinstrap	40.90	46.35	49.55	51.08	58.00	48.83	3.34	68	0
Gentoo	40.90	45.30	47.30	49.55	59.60	47.50	3.08	123	1

Review: Pivotal quantity

A statistic that depends on the random sample, $X_1, \ldots, X_n$, and the parameter, θ, but whose distribution does not depend on θ.

Your turn

Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$ and

$Y_1, \ldots, Y_m \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$. X_i and Y_j are independent.

1. Find $E(\bar{X}_n - \bar{Y}_m)$
2. Find $Var(\bar{X}_n - \bar{Y}_m)$
3. What distribution will does Z have?

$$Z = \frac{(\bar{X}_n - \bar{Y}_m) - E(\bar{X}_n - \bar{Y}_m)}{SD(\bar{X}_n - \bar{Y}_m)}$$

4. Is Z a pivotal quantity?

Your turn

Calculate and interpret a 90% confidence interval for the difference in mean bill length between Gentoo and Chinstrap penguins.

species	min	Q1	median	Q3	max	mean	sd	n	missing
Adelie	32.10	36.75	38.80	40.75	46.00	38.79	2.66	151	1
Chinstrap	40.90	46.35	49.55	51.08	58.00	48.83	3.34	68	0
Gentoo	40.90	45.30	47.30	49.55	59.60	47.50	3.08	123	1

Underlying validity conditions

- Each random sample is taken from a normal population distribution
- The random samples are independent

Ask Yourself...

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

Checking conditions

- Are the penguins independent within the species?
- Are the species independent?
- Are the bill lengths for each species approximately normal?

Robustness two-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)
- If the two population distributions have the same shape and the sample sizes are too different, skewness isn't a problem
- t-procedure is in trouble if the two population distributions have very different shapes (but would we even want to compare means then?)
- t-procedures are not resistant to outliers
- If observations are not independent, the results can be misleading

Revisiting matched pairs

Is it safe to look at social media while driving?

- Previous research on smart phone use while driving has primarily focused on phone calls and texting.
- Study looked at the effects of different smart phone tasks on car-following performance in a driving simulator.
- Drivers performed driving only baseline simulation
- Drivers performed other phone tasks: texting, reading Facebook posts, exchanging photos on Snapchat, viewing updates on Instagram
- Brake reaction times (in seconds) recorded

CIs for matched pairs data

- For a matched pairs experiment, look at the differences between responses for each unit (pair)
- Compute a new variable for differences, then use a one-sample t procedure

Subject	Baseline	Facebook	Texting	Instagram	SnapChat
1	0.863	1.254	1.011	0.963	0.865
2	0.847	1.100	0.900	0.600	0.783
3	0.836	1.021	1.064	0.947	0.808
4	0.655	0.864	0.974	0.726	1.010
5	0.900	0.793	0.856	0.817	0.837
6	0.957	1.252	1.178	1.134	1.175
7	0.780	0.856	1.010	0.861	0.817
And so on for 10 more rows...					

Subject	Baseline	Facebook	Texting	Instagram	SnapChat
8	0.954	0.814	1.250	1.022	0.861
And so on for 10 more rows...					

CIs for matched pairs data

Summary statistics for `Facebook - Baseline` brake reaction times

min	Q1	median	Q3	max	mean	sd	n	missing
−0.14	0.12	0.21	0.37	1.51	0.30	0.39	18	0

Calculate a 92% confidence interval for the mean difference in reaction times.