Sample Comparison & Hypothesis Testing

Logistics

- Conditional probability worksheet due now
- New homework will be out on Tuesday

Jury Panels

Eligible jurors in a county → List of eligible residents → Jury panel → Jury

Section 197 of California's Code of Civil Procedure: All persons selected for jury service shall be selected at random, from a source or sources inclusive of a representative cross section of the population of the area served by the court.

Sixth Amendment to the US Constitution: … the accused shall enjoy the right to a speedy and public trial, by an impartial jury of the state and district wherein the crime shall have been committed.

Robert Swain v. Alabama

1965 Supreme Court case about jury selection

- In Talladega, Alabama, 26% of residents were black
- In Swain's jury panel, 8 of 100 panelists were black
- All 8 were struck from the jury by the prosecution (using peremptory challenges)

Robert Swain v. Alabama

1965 Supreme Court case about jury selection

- In Talladega, Alabama, 26% of residents were black
- In Swain's jury panel, 8 of 100 panelists were black
- All 8 were struck from the jury by the prosecution (using peremptory challenges)

Ruling: "The overall percentage disparity has been small and reflects no studied attempt to include or exclude a specified number of [black men]."

Is the actual jury panel likely?

(Demo)

Empirical Distribution of a Statistic

- **Statistic**: a quantity computed for a particular sample
  - Ex sample: 1000 die rolls
  - Ex statistic: the number of 6's I see in that sample

Empirical Distribution of a Statistic

- **Statistic**: a quantity computed for a particular sample
  - Ex sample: 1000 die rolls
  - Ex statistic: the number of 6’s I see in that sample
- **Distribution**: the chance of each outcome of sampling
  - ***Not*** the chance of rolling a 6
  - The chance of getting some # of 6s in 1000 rolls
  - In between 0 and 6000, there is some # of 6s that might show up

Empirical Distribution of a Statistic

- **Statistic**: a quantity computed for a particular sample
    - Ex sample: 1000 die rolls
    - Ex statistic: the number of 6's I see in that sample
- **Distribution**: the chance of each outcome of sampling
    - ***Not*** the chance of rolling a 6
    - The chance of getting some # of 6s in 1000 rolls
    - In between 0 and 6000, there is some # of 6s that might show up
- **Probability Distribution:** chance of each value of a statistic computed from ***all*** possible samples

Empirical Distribution of a Statistic

- **Statistic**: a quantity computed for a particular sample
  - Ex sample: 1000 die rolls
  - Ex statistic: the number of 6’s I see in that sample
- **Distribution**: the chance of each outcome of sampling
  - ***Not*** the chance of rolling a 6
  - The chance of getting some # of 6s in 1000 rolls
  - In between 0 and 6000, there is some # of 6s that might show up
- **Probability Distribution:** chance of each value of a statistic computed from ***all*** possible samples
- **Empirical Distribution:** observations of a statistic computed from some samples drawn at random

Hypothesis Testing

Testing a Hypothesis

Step 1: The Hypotheses

- A test chooses between two views of how data were generated

Testing a Hypothesis

Step 1: The Hypotheses

- A test chooses between two views of how data were generated
- *Null hypothesis* proposes that data were generated at random

Testing a Hypothesis

Step 1: The Hypotheses

- A test chooses between two views of how data were generated
- *Null hypothesis* proposes that data were generated at random
- *Alternative hypothesis* proposes some effect other than chance

Testing a Hypothesis

Step 1: The Hypotheses

- A test chooses between two views of how data were generated
- *Null hypothesis* proposes that data were generated at random
- *Alternative hypothesis* proposes some effect other than chance

Step 2: The Test Statistic

- A value that can be computed for the data and for samples

Testing a Hypothesis

Step 1: The Hypotheses

- A test chooses between two views of how data were generated
- *Null hypothesis* proposes that data were generated at random
- *Alternative hypothesis* proposes some effect other than chance

Step 2: The Test Statistic

- A value that can be computed for the data and for samples

Step 3: The Probability Distribution of the Test Statistic

- What the test statistic might be if the null hypothesis were true

Testing a Hypothesis

Step 1: The Hypotheses

- A test chooses between two views of how data were generated
- *Null hypothesis* proposes that data were generated at random
- *Alternative hypothesis* proposes some effect other than chance

Step 2: The Test Statistic

- A value that can be computed for the data and for samples

Step 3: The Probability Distribution of the Test Statistic

- What the test statistic might be if the null hypothesis were true
- Approximate the probability distribution by an empirical distribution

Jury Hypothesis Test

Step 1: The Hypotheses

- *Null hypothesis:* The panel of jurors was selected at random from the eligible juror population. Any differences between the ethnicity distributions of panel and population are due to randomness.
- *Alternative hypothesis*: The panel was not selected at random.

Step 2: The Test Statistic

- TVD between a panel and the population distribution

Step 3: The Probability Distribution of the Test Statistic

- Simulate drawing panels from population. See if the actual panel TVD is likely.

Example: The Two Hypotheses

Gregor Mendel (1822 – 1884)

Austrian monk, founder of modern genetics

Tested the hypothesis that pea plants will bear purple or white flowers in the ratio 3:1

Example: The Two Hypotheses

Gregor Mendel (1822 – 1884)

Null hypothesis: Mendel's model describes the world. The distribution of the observed plants is different from the distribution in the model due to chance variation.

Alternative hypothesis: It doesn't.

(Demo)

Conclusion of a Test

- Choosing between the null and alternative hypothesis
  - Compare the observed test statistic to its empirical distribution under the null hypothesis
  - If the observed value is **inconsistent**, then we can reject the null hypothesis, leaving only the alternative
  - If the observed value is consistent with the distribution, then the test *does not* support the alternative hypothesis

Conclusion of a Test

- Choosing between the null and alternative hypothesis

Conclusion of a Test

- Choosing between the null and alternative hypothesis
  - Compare the observed test statistic to its empirical distribution under the null hypothesis

Conclusion of a Test

- Choosing between the null and alternative hypothesis
  - Compare the observed test statistic to its empirical distribution under the null hypothesis
  - If the observed value is **inconsistent**, then we can reject the null hypothesis, leaving only the alternative

Conclusion of a Test

- Choosing between the null and alternative hypothesis
  - Compare the observed test statistic to its empirical distribution under the null hypothesis
  - If the observed value is **inconsistent**, then we can reject the null hypothesis, leaving only the alternative
  - If the observed value is consistent with the distribution, then the test *does not* support the alternative hypothesis

Conclusion of a Test

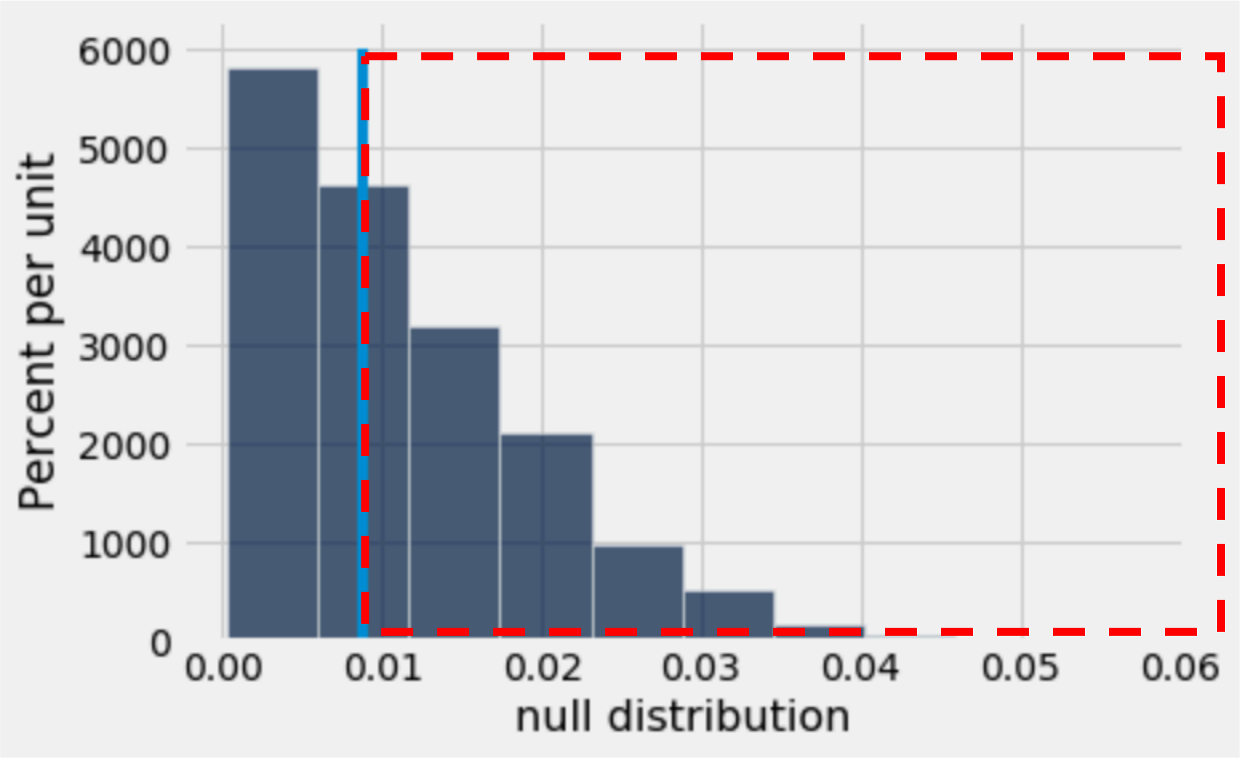
- Whether a value is consistent with a distribution:
  - A visualization may be sufficient

Conclusion of a Test

- Whether a value is consistent with a distribution:
  - A visualization may be sufficient
  - Convention: the observed significance value (P-value)
    - “The probability of getting my observed result if the null hypothesis were true”

P-Value

- The chance under the null hypothesis that the test statistic is equal to the value that was observed in the data, or is even further in the direction of the alternative.



Conventions of Consistency

- "**Inconsistent**": the test statistic is in the tail of the null distribution

- "**In the tail,"** first convention**:**
  - The area in the tail is less than 5%
  - The result is "statistically significant"

- "**In the tail**," second convention:
  - The area in the tail is less than 1%
  - The result is "highly statistically significant"

Error Probability

- Can the conclusion be wrong?

	Null is True	Alternative is True
Test rejects the null	X	
Test doesn't reject the null		X

Error Probability

- The cut off for the P-value is an error probability

- If your cut off is 5%
  - And the null hypothesis happens to be true
  - (but you don’t know that)

- Then there is about a 5% chance (1/20) that your test will reject the null hypothesis anyway.

Hypothesis Testing Logic

- Define 2 mutually exclusive descriptions: either this *or* that
- One of them can be evaluated using probability (the null)
- You can “reject the null” → accepting the alternative
- Otherwise, you’re still not sure, but the null looks reasonable
- Fisher (1935): “No isolated experiment, however significant in itself, can suffice for the experimental demonstration of any natural phenomenon.”

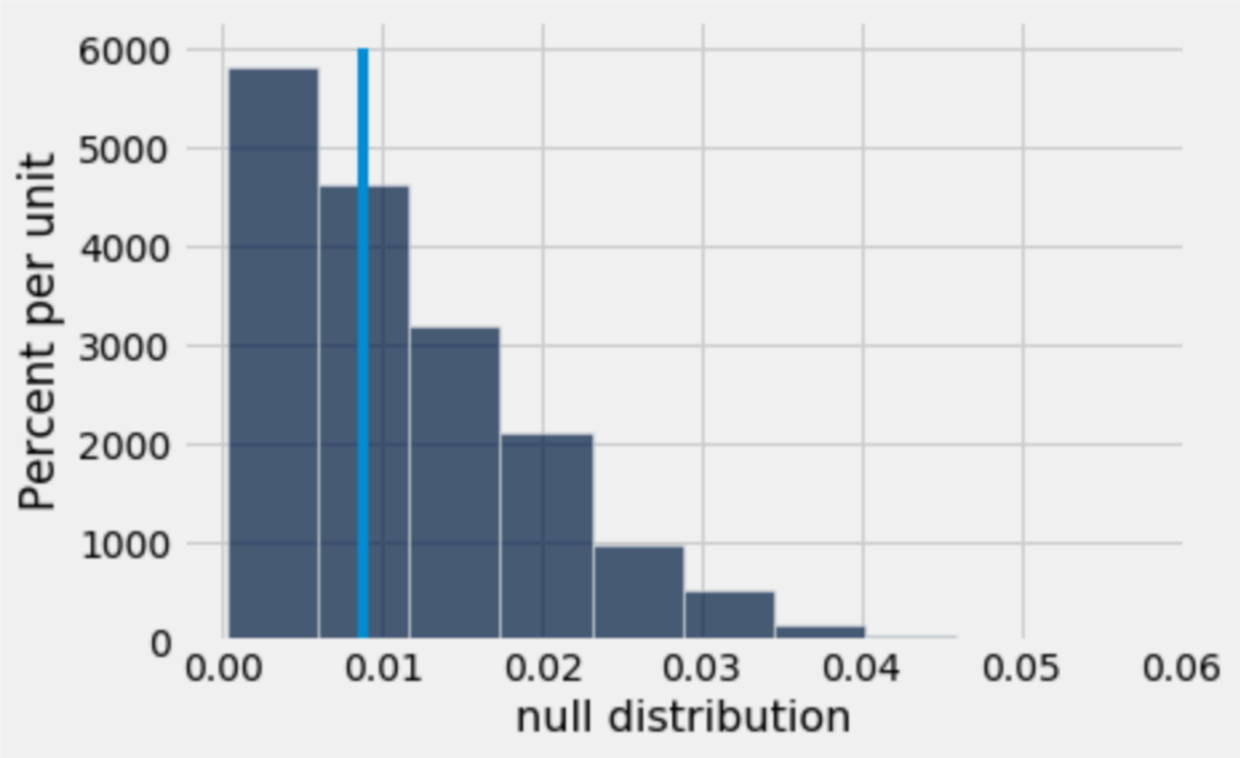
Quantifying Conclusions

1. Go find some data. These are *observations*.
2. Two descriptions of the world:
    1. Null: data come from a well-defined random process
    2. Alternative: Something else is going on
3. Evaluate how unusual the world would be under the null:
    1. Choose a test statistic to summarize the data
    2. Compute the P-value:

    P(the test statistic would be equal to or more extreme than the observed test statistic under the null hypothesis)

Quantifying Conclusions

P(the test statistic would be equal to or more extreme than the observed test statistic under the null hypothesis)



Discussion Question

- What if we changed our test statistic to not have the absolute value?

```
def old_test_statistic(sample):
    return abs(np.count_nonzero(sample == ‘Purple’)
                        / len(sample) - 0.75)


def new_test_statistic(sample):
    return np.count_nonzero(sample == ‘Purple’)
                        / len(sample) - 0.75
```