Stat 13, Intro. to Statistical Methods for the Life and Health Sciences.

1. Heart transplant example.
2. Standardized statistic.
3. A note on 1-sided and 2-sided tests.
4. Predicting faces example.

No class Mon Jan16, MLK day.

http://www.stat.ucla.edu/~frederic/13/W23 .

1. Heart Transplant Example.

Example 1.3

Heart Transplants

- The *British Medical Journal* (2004) reported that heart transplants at St. George's Hospital in London had been suspended after a spike in the mortality rate
- Of the last 10 heart transplants, 80% had resulted in deaths within 30 days
- This mortality rate was over five times the national average.
- The researchers used 15% as a reasonable value for comparison.

Heart Transplants

- Does a heart transplant patient at St. George's have a higher probability of dying than the national rate of 0.15?
- Observational units
  - The last 10 heart transplantations
- Variable
  - If the patient died or not
- Parameter
  - The actual probability of a death after a heart transplant operation at St. George's

Heart Transplants

- **Null hypothesis:** Death rate at St. George's is the same as the national rate (0.15).
- **Alternative hypothesis:** Death rate at St. George's is higher than the national rate.

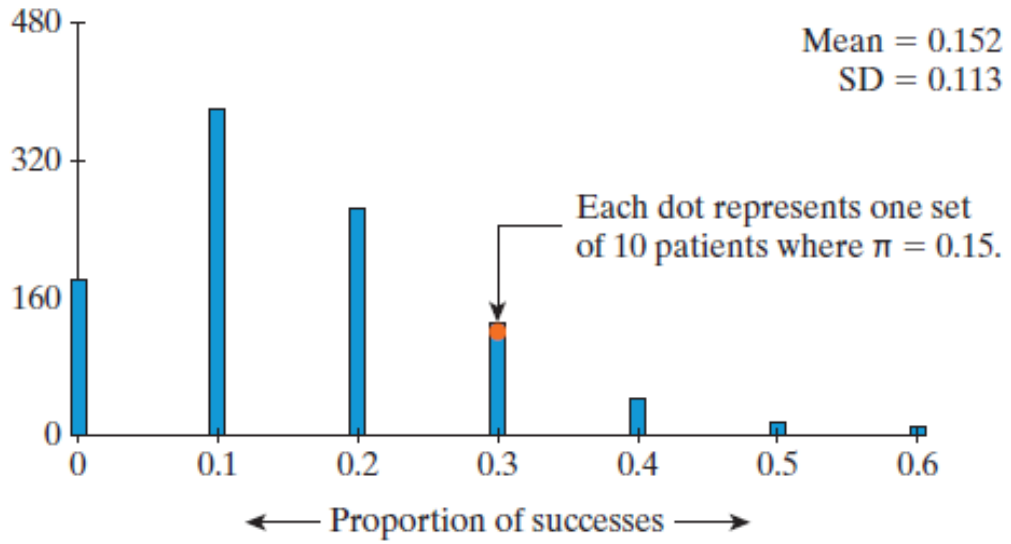
- **H_0:** $\pi = 0.15$ **H_a:** $\pi > 0.15$

- Our **statistic** is 8 out of 10 ($\hat{p} = 0.8$)

Heart Transplants

Simulation

- Null distribution of 1000 repetitions of drawing samples of 10 "patients" where the probability of death is equal to 0.15.



What is the p-value?

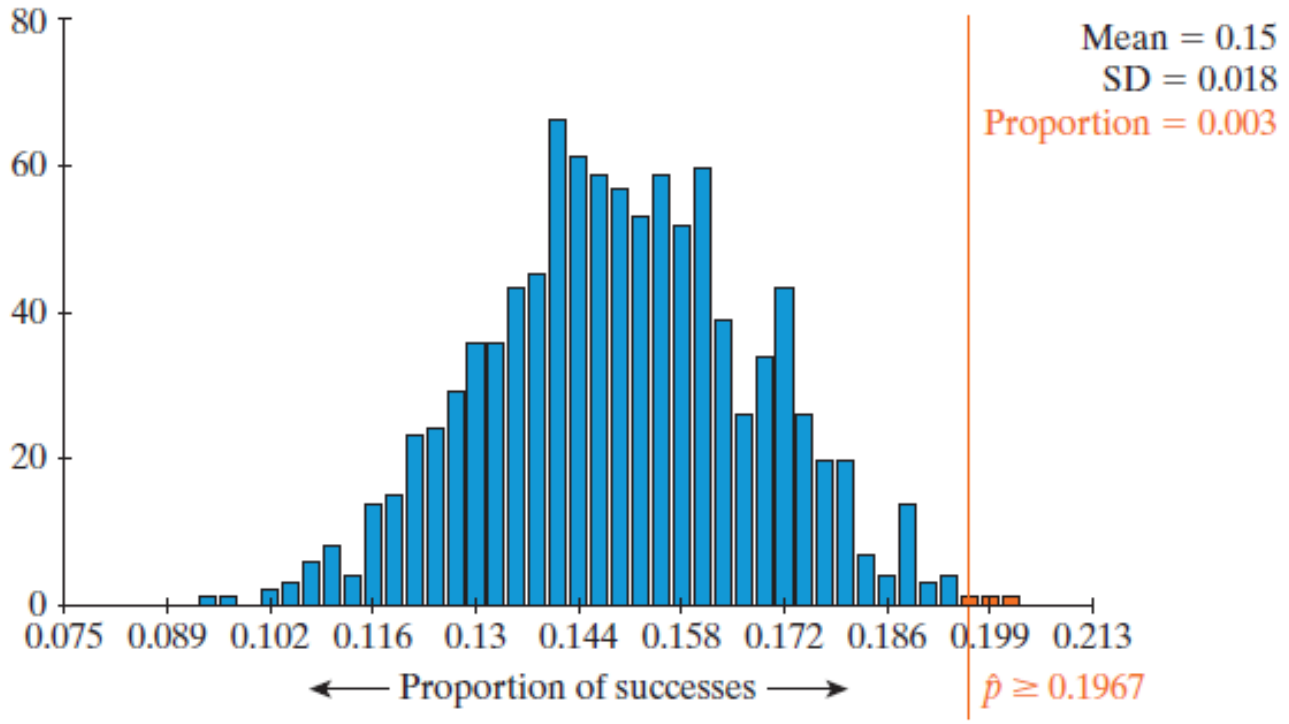
Heart Transplants

Strength of Evidence

- Our p-value is 0, so we have very strong evidence against the null hypothesis.
- Even with this strong evidence, it would be nice to have more data.
- Researchers examined the previous 361 heart transplantations at St. George's and found that 71 died within 30 days.
- Our new statistic, $\hat{p}$, is 71/361 ≈ 0.1967

Heart Transplants

- Here is a null distribution and p-value based on the new statistic.



Heart Transplants

- The p-value was about 0.003
- We still have very strong evidence against the null hypothesis, but not quite as strong as the first case

- Another way to measure strength of evidence is to ***standardize*** the observed statistic

Heart Transplants

- The p-value was about 0.003
- We still have very strong evidence against the null hypothesis, but not quite as strong as the first case

- Another way to measure strength of evidence is to ***standardize*** the observed statistic

2. The Standardized Statistic

- The ***standardized statistic*** is the number of standard deviations our sample statistic is above the mean of the null distribution (or below the mean if it is negative).

- $z = \dfrac{\textit{statistic} - \textit{mean of null distribution}}{\textit{standard deviation of null distribution}}$

- The sd of the null distribution is the *standard error.*
- For a single proportion, we will use the symbol *z* for standardized statistic.
- In the formula above, for the mean, we should use the long-term proportion (probability) given in the null hypothesis. If you do simulations, the mean of the simulated statistics should be close to this.

The Standardized Statistic

- Here are the standardized statistics for our two studies.

$$z = \frac{0.80 - 0.15}{0.113} = 5.75 \qquad z = \frac{0.197 - 0.15}{0.018} = 2.61$$

- In the first, our observed statistic was 5.75 standard deviations above the mean.
- In the second, our observed statistic was 2.61 standard deviations above the mean.
- Both of these are very strong, but we have stronger evidence against the null in the first.

Guidelines for strength of evidence

- If a standardized statistic is below -2 or above 2, we have strong evidence against the null.

Standardized Statistic	**Evidence Against Null**
between -1.5 and 1.5	not much
below -1.5 or above 1.5	moderate
below -2 or above 2	strong
below -3 or above 3	very strong

3. A quick note on 1-sided versus 2-sided tests.

- On my exams, I will tell you explicitly whether to do a 1 or 2 sided test.
- On hw problems, you might have to decide whether to do a 1-sided or 2-sided test.
- With the hw, if in the problem you are given that you are only looking for evidence in one direction as evidence against the null hypothesis, then you do a 1-sided test. If you are looking for *any* difference in proportions as evidence against the null hypothesis, then do a 2-sided test.

Two-Sided Tests

- The change to the alternative hypothesis affects how we compute the p-value.
- Remember that the p-value is the probability (assuming the null hypothesis is true) of obtaining a proportion that is equal to or **more extreme** than the observed statistic
- In a *two-sided test*, **more extreme** goes in both directions.

4. What impacts p-values and strength of evidence? Faces example,

Section 1.4.

Predicting Elections from Faces

Predicting Elections

- Do voters make judgments about candidates based on facial appearances?
- More specifically, can you predict an election by choosing the candidate whose face is more competent-looking?
- Participants were shown two candidates and asked who has the more competent-looking face.