Elementary Statistics – Decision Errors and Power

Dr. Ab Mosca (they/them)

Slides based off slides courtesy of Kaitlyn Cook (https://www.smith.edu/people/kaitlyn-cook)

Plan for Today

- Hypothesis Testing
  - Decision Errors
  - Power

Warm Up: Statistics and CIs

Pieces of a Hypothesis test:

1. ***Two competing and complementary claims about the world***
2. ***Test Statistic***
3. ***Null Distribution***
4. ***P-value***

The ***z-score*** of an observation characterizes the number of standard deviations it falls above or below the mean.

$Z = \frac{x-\mu}{\sigma}$, for a sample proportion, $\hat{p}$, $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$

Practice:

Suppose I rolled a 10 sided die 100 times and recorded the number of times I rolled a 9 (which was 45). I want to know if my die is fair so I do a hypothesis test with these hypotheses: $H_0: my\ die\ is\ fair, H_A: my\ die\ is\ not\ fair$

What is Z? What p-value does Z imply? Should I reject my null hypothesis?

Outcomes of a Hypothesis Test

A hypothesis test has 4 possible outcomes

		Test Conclusion	
		Reject H_0 (we suspect H_0 is false)	Fail to reject H_0 (we suspect H_0 is true)
Truth	H_0 is true	A	B
	H_0 is false (H_A is true)	C	D

In which scenario(s) have we made the correct decision? Why?

Outcomes of a Hypothesis Test

A hypothesis test has 4 possible outcomes

		Test Conclusion	
		Reject H_0 (we suspect H_0 is false)	Fail to reject H_0 (we suspect H_0 is true)
Truth	H_0 is true	A	
	H_0 is false (H_A is true)		D

Decision Errors

A hypothesis test has 4 possible outcomes

		Test Conclusion	
		Reject H_0 (we suspect H_0 is false)	Fail to reject H_0 (we suspect H_0 is true)
Truth	H_0 is true	**Type 1 Error**	
	H_0 is false (H_A is true)		**Type 2 error**

Decision Errors

A hypothesis test has 4 possible outcomes

		Test Conclusion	
		Reject H_0 (we suspect H_0 is false)	Fail to reject H_0 (we suspect H_0 is true)
Truth	H_0 is true	**Type 1 Error**	
	H_0 is false (H_A is true)		**Type 2 error**

- Because we reject H_0 when $p \leq \alpha$, the probability of a Type 1 error is α
- The probability of a Type 2 error is β

Decision Errors

A hypothesis test has 4 possible outcomes

		Test Conclusion	
		Reject H_0 (we suspect H_0 is false)	Fail to reject H_0 (we suspect H_0 is true)
Truth	H_0 is true	**Type 1 Error**	
	H_0 is false (H_A is true)		**Type 2 error**

- Because we reject H_0 when $p \leq \alpha$, the probability of a Type 1 error is α
- The probability of a Type 2 error is β

- The ***power*** of a statistical test is the probability of correctly rejecting a false H_0 $(1 - \beta)$

		Test Conclusion	
		Reject H_0 (we suspect H_0 is false)	Fail to reject H_0 (we suspect H_0 is true)
Truth	H_0 is true	**Type 1 Error**	
	H_0 is false (H_A is true)		**Type 2 error**

Suppose you are the manager of a professional baseball team.

- One of the players on your team consistently had a 0.250 batting average (they hit about one out of every four times at the plate).
- When you sit down to negotiate, this player claims they have been working hard in the off season and improved to become a 0.333 hitter.

Should you offer this player a raise (at the expense of another players salary)?

- What is the parameter of interest in this context?
- If you were to conduct a hypothesis test, what would your null and alternative hypotheses be?
- What would a Type 1 error be in the context of this problem?
- What would a Type 2 error be in the context of this problem?
- Who would be most worried about each type of error (you or the player)?

Practical Significance

Suppose that we want to determine whether a generic formulation of a drug is more harmful (i.e., produces more side effects) than the name brand version, which has a 10% adverse event rate.

$$H_0: p = 0.10$$
$$H_A: p > 0.10$$

Aside: This is called a one-sided test. If $H_A: p \neq 0.10$ we'd be doing a two-sided test.

We conduct a clinical trial and find that the sample proportion of patients who experience a side effect on the generic is $\hat{p} = 0.105$.

Practical Significance

Suppose that we want to determine whether a generic formulation of a drug is more harmful (i.e., produces more side effects) than the name brand version, which has a 10% adverse event rate.

$$H_0: p = 0.10$$
$$H_A: p > 0.10$$

We conduct a clinical trial and find that the sample proportion of patients who experience a side effect on the generic is $\hat{p} = 0.105$.

What is the corresponding z-score if the sample size for the trial is n = 100?

What is the corresponding z-score if the sample size for the trial is n = 50000?