Hypothesis Testing

Math 122

Coin Flips

I flip a coin 100 times and see only 33 Hs.

Is the coin fair?

Why or why not?

Coin Flips

I flip a coin 100 times and see only 43 Hs.

Is the coin fair?

Why or why not?

Rare Event Rule

If an assumption H_0 implies that the probability of an observed event is exceptionally small, then that assumption is probably false.

Rare Event Rule

Suppose that an assumption H_0 implies that the probability of a certain event is 0.0001.

If we observe that event happening then either

1. H_0 is correct and we have observed a highly unlikely event, or
2. H_0 is false.

It is more likely that H_0 is false.

Ingredients of a Hypothesis Test (Informal)

- Symbols (Parameters: μ, p, n, x)
- Claim to be tested (in terms of the parameter)
- Working assumption (Called the Null hypothesis - H_0)
- Alternative to the working assumption (Called the Alternative Hypothesis - H_1 the opposite of H_0)
- Probability (P-value)
- Conclusion

Claim

$$p = a \quad p \neq a$$
$$p \leq a \quad p > a$$
$$p \geq a \quad p < a$$

Null Hypothesis (Working Assumption) Denoted H_0

$$p = a$$
$$p \leq a$$
$$p \geq a$$

Alternative Hypothesis Denoted H_1

- Opposite of H_0

The opposite of this	is this
$p = a$	$p \neq a$
$p \leq a$	$p > a$
$p \geq a$	$p < a$

- At the end of the test, we

  either believe H_1 or H_0

P-value

- This is the probability of getting results at least as extreme as the observed value.
- This is a measure of consistency between our observations and H_0.
- If P is **large**, then our observations are consistent with H_0.
- If P is **small**, the our observations are inconsistent with H_0, and we reject H_0.

P-value and formal conclusion

- If **P** $\leq$ **α** then H_0 is not consistent with the observations.
  - Reject H_0 and support H_1.

- If **P > α** then H_0 is consistent with the observations.
  - Do not reject H_0 and do not support H_1.

Final Conclusion

- If your **claim is H_0** then your conclusion will be
  - There is enough sample evidence to **reject** the claim.
  - There is **not** enough sample evidence to **reject** the claim.
- If your **claim is H_1** then your conclusion will be
  - The evidence **supports** the claim.
  - The sample evidence does not **support** the claim.

In any hypothesis test:

- Define H_0 (null hypothesis) and H_1 (alternative)
- Collect data.

- Assume H_0 and calculate the probability P of seeing data as extreme as our data.

- If P is small, the observations are inconsistent with H_0. Reject H_0. Support H_1.
- If P is large, the observations are consistent with H_0. Do not reject H_0. Do not support H_1.

Steps in a Hypothesis Test

1. Define parameters.
2. State the claim in symbols.
3. State the opposite of the claim.
4. Determine H_0 and H_1.
5. Select a significance level α.
6. Decide on a distribution and test statistic.
7. Find a P-value.
8. Decide on a formal conclusion.
9. Restate the formal conclusion referring to the original claim.

Types of Claims We Will Test

(A whirlwind overview of the next few weeks)

All depend on some variant of the Central Limit Theorem

One Proportion

- Use
  - one proportion z-test or
  - 1-PropZTest

- **Example**: In a random sample of 87 Grey Forest Glow Worms, 51 were striped. Use this data to test the claim that **most Grey Forest Glow Worms are striped.**

Two Proportions

- Use
  - two proportion z-test or
  - 2-PropZTest

- **Example**: Among 107 male Grey Forest Glow Worms, 72 were striped. Among 96 female GFGWs, 37 were striped. Use this data to test the claim that **the proportion of male GFGWs which are striped is greater than the proportion of female GFGWs which are striped**.

One Mean

- Use
  - t-test or
  - T-Test

- **Example**: A sample of adult GFGWs had the lengths listed below (in inches). Use this data to test the claim that the mean length of an adult GFGW is not 1 inch.

- 0.75, 0,.82, 0.97, 0.99, 1.05, 1.17, 1.28, 1.35

Two Independent Means

- Use
  - two mean t-test or
  - 2-SampTTest

- **Example**: A random sample of 37 GFGWs had an average length of 1.11in with a standard deviation 0.09in. A random sample of 43 blue glow worms had an average length of 1.15in with a standard deviation of 0.12in. Test the claim that **these two types of worms have the same average length**.

Matched Pairs

- Use
  - matched pairs or
  - (a more complex process)

- **Example**: Below are listed the weights of several GFGWs (in oz) along with the weight of food eaten in one day by the same worm. Test the claim that a GFGW on average eats more than its body weight in a day.

Weight	0.09	0.09	0.11	0.12	0.15	0.15	0.16
Eaten	0.07	0.10	0.12	0.13	0.16	0.17	0.16

Linear Correlation

- Use
  - correlation or
  - LinRegTTest
- H_0 is always that there is no linear correlation.
- H_1 is always that there is linear correlation.
- **Example**: Below are listed the weights of several GFGWs (in oz) along with the weight of food eaten in one day by the same worm. Test the claim that **there is a linear correlation between a GFGW's weight and how much the worm eats in a day.**

Weight	0.09	0.09	0.11	0.12	0.15	0.15	0.16
Eaten	0.07	0.10	0.12	0.13	0.16	0.17	0.16

χ^2 Goodness of Fit

- Use
  - chi-squared GOF or
  - $\chi^2 - \text{GOF}$
- H_0 is always that the observed frequencies match the expected values.
- H_1 is always that the observed frequencies do not match the expected values.
- **Example**: A candy company claims that 25% of its candy is brown, 35% is red, and 40% is blue. A bag of candy contained the numbers of each color in the table below. **Test the claim that the actual frequencies of colors match the claimed distribution.**

Brown	Red	Blue
21	33	37

Contingency Table

- Use
  - contingency table or
  - $\chi^2 - \text{GOF}$

- H_0 is always that the rows and columns are independent.
- H_1 is always that the rows and columns are dependent.
- **Example**: The genders and dietary preferences of a sample of GFGWs were observed with the results below. Test the claim that **dietary preference is dependent on gender.**

	Male	Female
Leaves	32	57
Bark	48	41

-

ANOVA

- Use
  - ANOVA or
  - ANOVA(
- H_0 is always that the samples come from populations with equal means.
- H_1 is always that the samples do not come from populations with equal means.
- **Example**: The table below lists the lengths of samples of three colors of glow worms. Test the claim that **all three colors have the same mean length.**
-

Grey	0.9	0.9	1.11	1.12	1.17	1.18	1.18
Blue	0.78	0.82	0.90	0.99	0.99	1.10	
Green	1.0	1.0	1.1	1.1	1.17	1.2	1.25