Data Science for Everyone

Week 9: Hypothesis testing, more coding

Angela Lai

New York University

Outline

- Logistics
- Hypothesis Testing
- Questions?

Logistics

- All grades up to and including the midterm should now be finalized
- Lab 6 out, due at 8 p.m. ET on April 8
- Usual submission process but remember that you'll get your feedback via email

Logistics

- Lecture materials and solved demo notebooks can be accessed at **this link**
- If you want feedback on your project data set, fill out the Google form in the email I sent out last week by the end of Monday, April 6

Concept Review: Distributions

Probability distribution: a mathematical function that provides the probabilities of occurrence of different possible outcomes in an experiment

Empirical distribution: distribution of observed data, such as data in random samples.

Concept Review: Distributions

We often do not know the true, underlying probability distribution behind our data, and instead only have access to the empirical distribution.

If I knew the underlying probability distribution of all the data I'm interested in, life would be a lot simpler.

Hypothesis Testing

We have two competing hypotheses and want to decide between them. How?

Hypothesis Testing

We have two competing hypotheses and want to decide between them. How?

At a high level:

1. Get observed data.
2. Formulate your null hypothesis (H_0) and your alternative hypothesis (H_a).
3. Decide on how you will pick between the two. (What's your decision rule/significance level?)
4. Based on the results, either reject OR fail to reject the null hypothesis.

Hypothesis Testing

Generally:

Null hypothesis, H_0: A claim about some statistic. Usually a general statement or default position that there is no relationship between two measured phenomena or no association among groups. (source)

Alternative hypothesis, H_a: A belief that counters the null hypothesis. H_a doesn't really make a claim of its own. Its primary objective is to reject the null hypothesis. (source)

Hypothesis Testing

In more technical terms:

Null hypothesis: The hypothesis under which we can simulate data. Under this model, if the data look different from what the null hypothesis predicts, the difference is due to nothing but chance.

Alternative hypothesis: Something other than chance made the actual data differ from the predictions from the model under the null hypothesis.

Hypothesis Testing: Fair Coin

We have a coin. Suppose we want to test its fairness.

What's the **null hypothesis**? The **alternative hypothesis**?

Hypothesis Testing: Fair Coin

In this case:

Null hypothesis: The coin is fair. That is, the results are like draws made at random with replacement and in this case, the probability we get heads or tails is equal. That is, $p(H) = p(T) = 1/2$.

Alternative hypothesis: The coin is not fair.

Hypothesis Testing: Fair Coin

What does this mean in simpler terms?

The null hypothesis is basically a guess about the true probability distribution. In this case, we hypothesize $p(head) = p(tails) = 1/2$. The alternative hypothesis is $p(head) \neq p(tail)$: the coin is not fair.

Hypothesis Testing

What steps do we take?

Hypothesis Testing

What steps do we take?

Set a **significance level**: the probability of the study rejecting the null hypothesis, given that the null hypothesis were assumed to be true.

Use the null hypothesis to simulate the distribution of the test statistic.

Then, look at where our **observed test statistic** falls on that distribution.

If our p-value is below the significance level we set, reject the null hypothesis. Otherwise, we fail to reject the null!

p-value: probability of getting our observed test statistic given that the null hypothesis was true.

Hypothesis Testing: Fair Coin

Observed data: I flipped the coin 10 times and got heads 9 times. Let my observed test statistic be $t = 9$. Set a significance level of $\alpha = 0.05$.

I assume the null hypothesis is true and, again and again for a large number of times, flip an imaginary fair coin 10 times. By doing this, I get a simulated distribution of the number of heads I'd expect to see if my coin is fair.

Now, looking at this simulated distribution, how likely it is that I would get my test statistic or something more extreme? (How often did I get 9 or 10 heads?) This is my p-value.

If my p-value is really low–particularly, if it's below the significance level I set–reject the null.

Hypothesis Testing

We do not accept the null hypothesis! We can only fail to reject it.

	Null is true	Null is false
Fail to reject null	Correct inference	Type II Error
Reject null	Type I Error	Correct inference

Questions?

Any questions?