MATH 10

INTRODUCTORY STATISTICS

Tommy Khoo

Your friendly neighbourhood graduated student.

Homework and Stuff

- Deadline extended : hand in a hard copy during Thursday's class.
- I will go through some examples today to help with this homework.

- Office hours is getting crowded. (･´з`･)

- Will try to set up a group tutoring/review session via Email. Weekend?
- Alternatively : I can work through HW4 on Weds office hours.

Homework and Stuff

- Next homework is lighter and has more practice questions on hypothesis testing.
- Homework will be posted by Weds noon.
- Answer key for midterm delayed but will also be posted Weds noon.

Week 7

final exam in 3 weeks

- **Chapter 12 – Test of Means** ← today's lecture

More hypothesis testing. Difference between means.

- **Chapter 13 – Interestingly titled "Power"**

The idea of the "power" of a test.

- **Chapter 14 – (brief introduction to) Regression**

Correction – Binomial Hypothesis Testing

- Emailed out a correction to the slides last week.
- I will tell the grader to give you credit if you used the wrong version of the standard error for homework. Thankfully, I don't think anyone did.
- Recall : sampling distribution of a sample proportion p is approximately Normal with mean π.
- Standard error depends on whether the population proportion π is known.

Correction – Binomial Hypothesis Testing

- Recall : sampling distribution of a sample proportion p is approximately Normal with mean π.
- Standard error depends on whether the population proportion π is known.
- Formula sheet : $\sqrt{\frac{\pi(1-\pi)}{n}}$ if π known. Otherwise, $\sqrt{\frac{p(1-p)}{n}}$.
- Hypothesis testing : we assume the null hypothesis is true, which means π is known and given by the null hypothesis.

Hypothesis Testing – A Review of Main Ideas

- Recall what we learned about Sampling Distributions.

- Suppose we have a large population, and some variable X that we are interested in.
- We take a sample of size n, and calculate the sample mean $\bar{X}$.

- The sampling distribution of the mean tells us the probability of getting certain values of the sample mean. (dist. of $\bar{X}$)

Hypothesis Testing – A Review of Main Ideas

- Suppose we don't know what the true population mean is.
- But we have some kind of idea or hypothesis about its value.

$$H_0 : population\ mean\ is\ \mu \quad , \qquad H_A : not\ \mu$$

Hypothesis Testing – A Review of Main Ideas

- Suppose we don't know what the true population mean is.
- But we have some kind of idea or hypothesis about its value.

$$H_0 : population\ mean\ is\ \mu \quad , \qquad H_A : not\ \mu$$

- Assuming that H_0 is true, then we know what the sampling distribution of the mean is (*estimating the standard error in the t-dist case*).

- We can set a threshold for when we will find H_0 too unlikely and reject it.

Hypothesis Testing – A Review of Main Ideas

$$H_0 : population\ mean\ is\ \mu \quad , \qquad H_A : not\ \mu$$

- We can set a threshold for when we will find H_0 too unlikely and reject it.

- We build this threshold based on our data, which is the sample mean $\bar{X}$, calculated from a sample of size n.

- More precisely, threshold is based on $P(\ sample\ mean\ \geq \bar{X}\ )$ or $P(\ sample\ mean\ \leq \bar{X}\ )$. Let's assume $\bar{X} > \mu$, so we focus on one.

Hypothesis Testing – A Review of Main Ideas

- More precisely, threshold is based on $P(\, sample\ mean\ \geq \bar{X}\,)$ or $P(\, sample\ mean\ \leq \bar{X}\,)$. Let's assume $\bar{X} > \mu$, so we focus on one.

- Lingo / Convention : *Chapter 11, Section 6 – Significant Results*

- Textbook calls this a (one-tailed) probability value or p-value.
- The textbook defines two types: one-tailed and two-tailed p-value.
- Will talk about this later. Just focus on this one-tailed p-value for now.

Hypothesis Testing – A Review of Main Ideas

- More precisely, threshold is based on $P(\,sample\ mean\ \geq \bar{X}\,)$ or $P(\,sample\ mean\ \leq \bar{X}\,)$. Let's assume $\bar{X} > \mu$, so we focus on one.

- Using the sampling distribution, we can figure out what $P(\,sample\ mean\ \geq \bar{X}\,)$ is.

- Recall : probability is area under the curve.
- z- or t-tables give you the wrong area under the curve.
- Need to use $P(A) = 1 - P(\,not\ A)$.

Hypothesis Testing – A Review of Main Ideas

- Using the sampling distribution, we can figure out what $P(\, sample\ mean\ \geq \bar{X}\,)$ is.

- We define a threshold for when H_0 is "too unlikely for us".

- We say that if $P(\, sample\ mean\ \geq \bar{X}\,)$ is "too small", then we "don't believe" in H_0. So, we reject H_0.

- Significant level α determines when this probability is "too small".

Hypothesis Testing – A Review of Main Ideas

- Significant level α determines when this probability is "too small".
- Condition depends on which of the two types of test used.

One-tailed test

$$P(\,sample\ mean\ \geq \bar{X}\,) < \alpha$$

Two-tailed test

$$P(\,sample\ mean\ \geq \bar{X}\,) < \alpha/2$$

Hypothesis Testing – A Review of Main Ideas

One-tailed test : $P(\, sample\ mean \ \geq \bar{X}\,) < \alpha$

Two-tailed test : $P(\, sample\ mean \ \geq \bar{X}\,) < \alpha/2$

- If this condition is met (write it down in the exam!) then we say :

We reject H_0 at α level of significance. "blah blah blah" is *probably* "blah blah blah".

Hypothesis Testing – A Review of Main Ideas

One-tailed test : $P(\,sample\ mean\ \geq \bar{X}\,) < \alpha$

Two-tailed test : $P(\,sample\ mean\ \geq \bar{X}\,) < \alpha/2$

- If this condition is **NOT** met (write it down in the exam!) then we say :

We do not reject H_0 at α level of significance. Inconclusive.

Hypothesis Testing – A Review of Main Ideas

- If this condition is **NOT** met (write it down in the exam!) then we say :

We do not reject H_0 at α level of significance. Inconclusive.

- For exam, you can ask me what conclusion is the question looking for.
- E.g. what does it mean for the vote to be "too close to call".

- There might be very precise questions asking you about the meaning of hypothesis testing in general. E.g. everything is probabilistic. Can always reject or not reject by picking the α.
- Will go through later in this lecture!

Example Related to HW4 Qns 1

- H_0: $\pi = 0.50$, H_A: $\pi \neq 0.50$, is male proportion greater than 0.50?
- Sample proportion $p = 0.60$, sample size $n = 25$.

Example Related to HW4 Qns 1

- H_0: $\pi = 0.50$, H_A: $\pi \neq 0.50$, is male proportion greater than 0.50?
- Sample proportion $p = 0.60$, sample size $n = 25$.

- Assuming null hypothesis is true, population proportion is $\pi = 0.50$.
- Applying Normal approximation to the binomial distribution, the sampling distribution is Normal with mean $\pi = 0.50$ and variance $\pi(1-\pi)/n$.

Example Related to HW4 Qns 1

- H_0: $\pi = 0.50$, H_A: $\pi \neq 0.50$, is male proportion greater than 0.50?
- Sample proportion $p = 0.60$, sample size $n = 25$.

- Assuming null hypothesis is true, population proportion is $\pi = 0.50$.
- Applying Normal approximation to the binomial distribution, the sampling distribution is Normal with mean $\pi = 0.50$ and variance $\pi(1-\pi)/n$.

- Z-statistic/value is $\frac{p - mean}{standard\ error} = \frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{0.1}{0.1} = 1$.

Example Related to HW4 Qns 1

- H_0: $\pi = 0.50$, H_A: $\pi \neq 0.50$, is male proportion greater than 0.50?
- Sample proportion $p = 0.60$, sample size $n = 25$.
- Z-statistic/value is $\frac{p - mean}{standard\ error} = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{0.1}{0.1} = 1.$

- (one-tailed) p-value is :

$$P(\ sample\ prop\ \geq 0.6\ ) = P(\ Z \geq 1\ ) = 1 - P(\ Z < 1\ ) = 1 - 0.8413 = 0.1587.$$

- Rejecting null or not depends on your significance level.

Example Related to HW4 Qns 1

- (one-tailed) p-value is :

$P(\, sample\ prop\ \geq 0.6\,) = P(\, Z \geq 1\,) = 1\ - P(\, Z < 1\,) = 1\ - 0.8413 = 0.1587.$

- Rejecting null or not depends on your significance level α.
- We are comparing areas : is $P(\, sample\ prop\ \geq 0.6\,)$ less than $\alpha/2$?

- Suppose level of significance is $\alpha\ = 0.10$.
- Then, $P(\, sample\ prop\ \geq 0.6\,) = 0.1587 > \frac{0.10}{2} = 0.05$.

Example Related to HW4 Qns 1

- (one-tailed) p-value is :

$P(\, sample\ prop \ \geq 0.6\,) = P(\, Z \geq 1\,) = 1\ - P(\, Z < 1\,) = 1\ - 0.8413 = 0.1587.$

- Suppose level of significance is $\alpha\ = 0.10.$
- Then, $P(\, sample\ prop\ \geq 0.6\,) = 0.1587 > \frac{0.10}{2} = 0.05.$ → write this in exam!!

- Conclusion : → variation of exactly what you should write in exam.

We do not reject the null, at the α level of significance. Results are inconclusive.

Chapter 12, Section 4 – Hypo. Test For Difference Between Means

- The general strategy :

$$z \text{ or } t = \frac{sample\ difference\ - hypothesized\ difference}{standard\ error}$$

- Use z when population variances are given. Sampling distribution is normal.

- Use t when population variances are not given. Sampling distribution is the t-dist.
- Df = (n-1) + (n-1) = 2(n-1).

Chapter 12, Section 4 – Hypo. Test For Difference Between Means

Assumptions for the t-dist case

1. Both populations are normally distributed with the same unknown variance.
2. Both simple random samples are independent and have same size n.

$$\text{MSE} = \frac{S_1^2 + S_2^2}{2}$$

$$\text{Standard Error, SE} = \sqrt{\frac{2\,MSE}{n}} = \sqrt{\frac{S_1^2+S_2^2}{n}}.$$

Break time!! \o/

- No exercise today! Go enjoy your break. ^__^

12 minutes

- Circle is a timer that becomes blue. O_o →

(please ignore if it glitches)

p-value and Bayes Theorem

- (one-tailed) p-value = P(D = data or more extreme | H = null hypothesis is true)

- But $P(\ H \mid D\ )\ =\ \frac{P(\ D \mid H\ )\ P(H)}{P(D)}$.

- We can go further: $P(\ H \mid D\ )\ =\ \frac{P(\ D \mid H\ )P(H)}{P(D \mid H\ )P(H) + P(H \mid D)P(D)}$.

p-value and Bayes Theorem

- (one-tailed) p-value = P(D = data or more extreme | H = null hypothesis is true)
- But $P(\ H \mid D\ ) \ = \ \frac{P(\ D \mid H\ )\, P(H)}{P(D)}$.

- So p-value is NOT the probability that the null hypothesis is true, which is P(H|D).

- Technically, we are using $P(\ D\ |H\ )$ to "guess" whether $P(\ H\ |D\ )$ would be small.
- $P(\ D\ |H\ )$ is probability of data given that null hypothesis is true.
- E.g. if two men committed the murder, what is prob. that DNA matches? → HW3

Chapter 11, Section 10 – Misconceptions

Extremely important for the exams. Might be in final exam. (͡° ͜ʖ ͡°)

- Is the p-value the probability that the null hypothesis is false/true?
- Does a low p-value indicate a large effect?
- If an outcome is not statistically significant, does it mean that the null hypothesis is true?

NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION

Psychology journal bans *P* values

Test for reliability of results 'too easy to pass', say editors.

Chris Woolston

26 February 2015 | Clarified: 09 March 2015

PDF | Rights & Permissions

A controversial statistical test has finally met its end, at least in one journal. Earlier this month, the editors of *Basic and Applied Social Psychology* (*BASP*) announced that the journal would no longer publish papers containing *P* values because the statistics were too often used to support lower-quality research [1].

Authors are still free to submit papers to *BASP* with *P* values and other statistical measures that form part of 'null hypothesis significance testing' (NHST), but the numbers will be removed before publication. Nerisa Dozo, a PhD student in psychology at the University of Queensland in Brisbane, Australia, tweeted:

NATURE | NEWS

Statisticians issue warning over misuse of *P* values

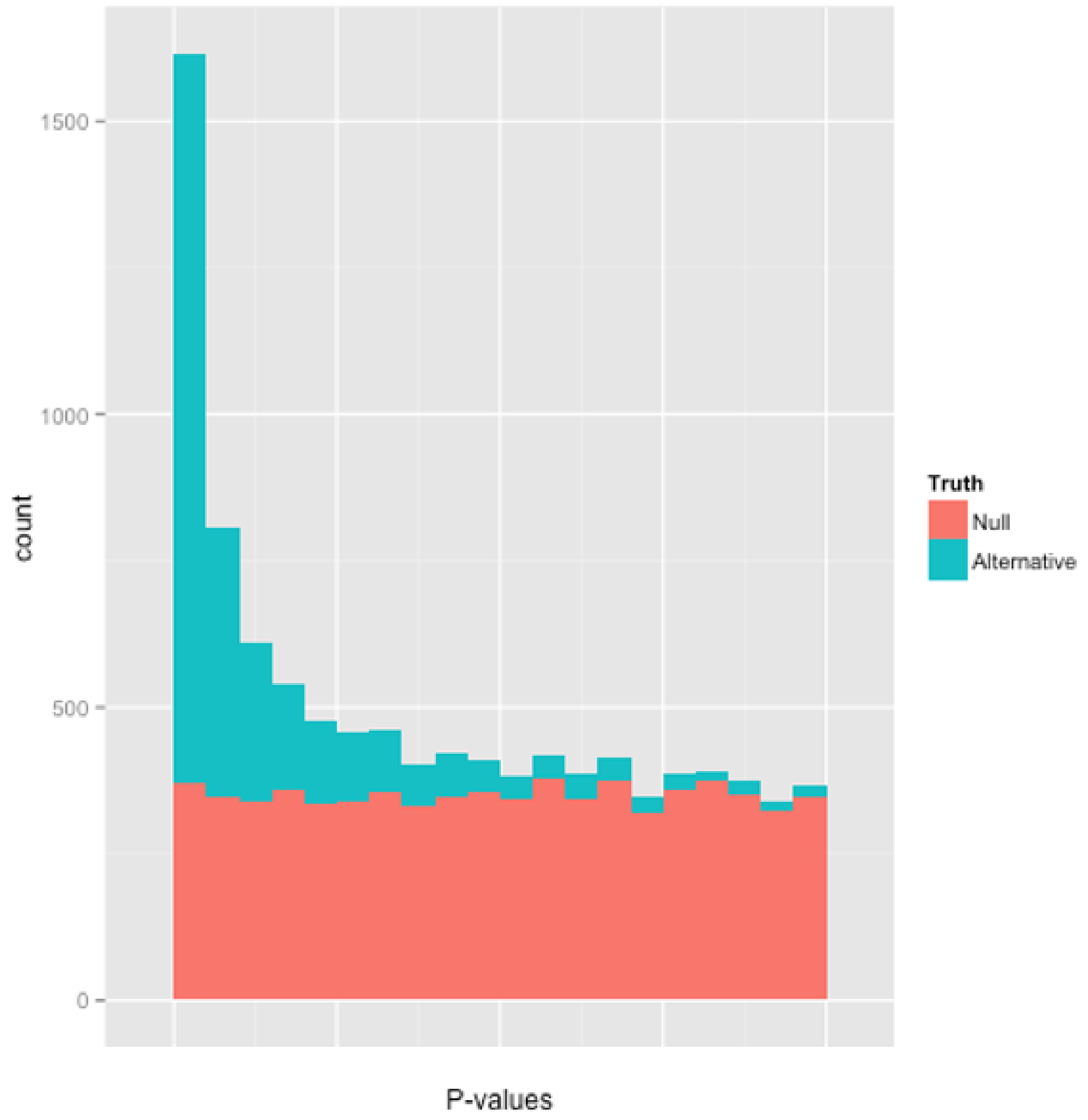
Policy statement aims to halt missteps in the quest for certainty.

Monya Baker

07 March 2016

PDF | Rights & Permissions

Misuse of the *P* value — a common test for judging the strength of scientific evidence — is contributing to the number of research findings that cannot be reproduced, the American Statistical Association (ASA) warns in a statement released today[1]. The group has taken the unusual step of issuing principles to guide use of the *P* value, which it says cannot determine whether a hypothesis is true or whether results are important.



Disclaimer

- P-values from different experiments are *not exactly* compatible.
- The experimental conditions have to be identical.
- Mathematically : sampling distribution will change depending on sample size, and stratified vs simple random sampling etc!
- But if experimental conditions are identical. We should see those distributions.

"The truth is never simple and rarely pure." – Oscar Wilde

p-value : lingo and conventions (not required for exam!)

- Officially, in mathematical statistics, the p-value is defined as the smallest significance level that will cause a rejection of the null.

- You might have also seen this general definition for a two-tailed p-value :

$$2 \cdot min\{\, P(\, sample\ mean\ \geq \bar{X}\,)\,,\, P(\, sample\ mean\ \leq -\bar{X})\,\}$$

p-value : lingo and conventions (not required for exam!)

- You might have also seen this general definition for a two-tailed p-value :

$$2 \cdot min\{\, P(\, sample\ mean\ \geq \bar{X}\,)\,,\, P(\, sample\ mean\ \leq -\bar{X})\,\}$$

- Your textbook defines two versions of the p-value :

  One-tailed p-value $P(\, sample\ mean\ \geq \bar{X}\,)$

  Two-tailed p-value $2 \cdot P(\, sample\ mean\ \geq \bar{X}\,)$

- Your textbook's condition for a two-tailed test is :

$$2 \cdot P(\, sample\ mean\ \geq \bar{X}\,) < \alpha$$

p-value : lingo and conventions (not required for exam!)

- Textbook : Two-tailed p-value $2 \cdot P(\, sample\ mean \ \geq \bar{X}\,)$
- Textbook's condition for a two-tailed test is :

$$2 \cdot P(\, sample\ mean \ \geq \bar{X}\,) < \ \alpha$$

- This is exactly the same as what we have been doing so far :

$$P(\, sample\ mean \ \geq \bar{X}\,) < \ \alpha/2$$

- Both full credit. If you already have $P(\, sample\ mean \ \geq \bar{X}\,)$, no need to multiply it by 2.

Skipped Chapters

- Chapter 12, Section 6, Pairwise Comparisons (Tukey HSD test),
- Chapter 12, Section 7, Specific Comparisons,
- Chapter 12, Section 8 , Correlated Pairs
- Chapter 12, Section 11, Pairwise (Correlated) (Bonferroni correction),

- are not required!

→ Syllabus on the website has been updated.