# Math 11: Discussion 9

TA: Thu Nguyen
Tue. 05/26/2020

**Recap of discussion 9:**

1. Hypothesis testing
2. 1-sample vs. 2-sample test for mean
3. Decision rules

## Contents



## 9.1 Upcoming assignments

| Assignments | Chapters | Deadlines |
|---|---|---|
| Quiz | Ch. 19 | Tue. 05/26 |
| Homework | Ch. 20 | Mon. 06/01 |
| Quiz | Ch. 20 | Tue. 06/02 |
| Homework | Ch. 21 | Wed. 06/03 |
| Quiz | Ch. 21 | Thu. 06/04 |
| Lab 7 | | Fri. 05/15 |

Key concepts (not exhaustive):

1. *Hypothesis testing:* proportion and mean
2. *Mean:* 1-sample and 2-sample
3. *Error types:* type 1 ($\alpha$), and type 2 ($\beta$)

## 9.2 Hypothesis Testing

***Motivation:*** Same type of problem (inferring the population from the available samples) but we are now looking from a different perspective: we want to test if our current belief about the population is true, or more precisely, if the observed sample statistics is statistically significantly different from the belief.

We divide into 4 main steps:

1. State the null hypothesis, $H_0$, and the alternative hypothesis, $H_1$
2. Calculate the test statistics, $\hat{p}$ if qualitative data and $\bar{x}$ if quantitative
3. Calculate $p$-value given the test statistics
4. Give a conclusion: either (1) Reject $H_0$, or (2) Fail to Reject $H_0$

**Remark:** In principle, we have the choices to compare either the test statistics or the induced $p$-value, both will return the same answer. Depending on the context, 1 approach may be more straight-forward.

We consider the problem of hypothesis testing in 3 different contexts:

1. *qualitative data:* test if a proportion $p = p_0$ where $p_0$ is the current belief;
2. *quantitative data:*
   (a) 1-sample $t$-test: for if the mean $\mu_X = \mu_0$, where $\mu_0$ is the current belief;
   (b) 2-sample $t$-test: for if there are differences between the means of 2 different populations in question, $\mu_X - \mu_Y = 0$;

### 9.2.1 Qualitative data - Proportion

While the null hypothesis $H_0$ will mostly be an equality, there are 3 options for the alternative hypothesis $H_1$. We use the example of qualitative data, testing for proportion $p$, the other 2 cases are similar with the appropriate test statistics.

$$\begin{aligned} H_0 &: \quad p = p_0 \\ H_1 &: \quad p > p_0; \quad \text{or} \quad p < p_0; \quad \text{or} \quad p \neq p_0 \end{aligned}$$

**Example 1.** *Let's fill out the decision table in the case of testing for proportion below:*

| $H_1$ | *Conclusion 1: Reject* $H_0$ | *Conclusion 2: Fail to Reject* $H_0$ |
|---|---|---|
| $H_1 : p > p_0$ | $\mathbb{P}(Z \geq z^*) \leq \alpha$ | $\mathbb{P}(Z \geq z^*) > \alpha$ |
| $H_1 : p < p_0$ | $\mathbb{P}(Z \geq z^*) \geq 1 - \alpha$ | $\mathbb{P}(Z \geq z^*) < 1 - \alpha$ |
| $H_1 : p \neq p_0$ | $\mathbb{P}(\lvert Z\rvert \geq z^*) \leq \frac{\alpha}{2}$ | $\mathbb{P}(\lvert Z\rvert \geq z^*) > \frac{\alpha}{2}$ |

**Example 2.** *(Adaptation of Fall '17, Question 9)*

*It is currently believed that among all $12^{th}$ grade students, 26% of them can be classified as proficient in Math. A local school district wants to investigate if their students are better at Math than the national average. They randomly sample 600 students and find that 185 of the students are considered proficient.*

*Assume a confidence level of $\alpha = .05$, determine if this is indeed the case.*

*We will follow this outline:*

*(a) Write the hypothesis.*

*(b) Calculate the test statistics.*

*(c) Calculate the corresponding p value.*

*(d) Give a conclusion and give some meaning based on the context.*

*(e) Interpret the p value in part (c).*

*(f) If they sample 1,000 students instead, would that change the probability of making type 1 error? If yes by how much?*

---

*Solution:*

We first recognize that this question concerns proportion. From the question, we collect some important information: let $p_0$ be the current believed proportion:

$$p_0 = .26; \qquad n = 600; \qquad \hat{p} = \frac{185}{600} = .3083$$

(a) Hypothesis:

$$H_0 : p = .26; \qquad H_A : p > .26$$

(b) Test statistics:

$$z^* = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{.3083 - .26}{\sqrt{\frac{.26(.74)}{600}}} = 2.70$$

(c) Corresponding $p$ value:

$$\mathbb{P}(Z \geq z^*) = \mathbb{P}(Z \geq 2.7) = .0035$$

(d) Conclusion:

$$p = .0035 < .05 = \alpha \quad \Longrightarrow \quad \text{Reject } H_0$$

*Meaning:* at the confidence level of $\alpha = .05$ there appears evidence that the local students are more proficient at Math than the national average.

(e) *Interpretation of the p value:* the probability of observing an event as extreme as the event observed assuming that $H_0$ is true, ie. assuming that 26% of the local students are proficient at Math.

(f) Since the probability of making type 1 error equals $\alpha$, which is independent of the sample size, changing the sample size will not affect the probability.

### 9.2.2 Quantitative data - Mean

**Example 3.** *Let's see how the same data set can be tested for both* 1*-sample or* 2*-sample tests. Consider the following data set of the number of hours spent studying at UCSD:*

| *Regular week* | 30 | 52 | 28 | 15 | 32 |
|---|---|---|---|---|---|
| *Finals week* | 32 | 46 | 34 | 20 | 32 |

*Give examples of* 1*-sample and* 2*-sample tests.*

*Solution:*

(a) 1-sample $t$-test: it is believed that students spend on average 30 hours per week on studying, however UCSD students study more than that, we want to know if this is true.

(b) 2-sample $t$-test: it is believed that students spend on average the same number of hours per week on studying regardless of if they are having finals, however UCSD students tend to study more than during finals week, we want to know if this is true.

Figure 1 summarizes of the different test statistics and the corresponding distributions.

| Data type | Parameter | Test statistics | Distribution |
|---|---|---|---|
| *Proportion* | $\bar{p} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{x_i=1}$ | $z^* = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ | $\mathcal{N}(0,1)$ |
| 1-*sample* $t$-*test* | $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ | $t^* = \frac{\bar{x}-\mu_X}{s/\sqrt{n}}$ | approximately $T_{n-1}$ |
| 2-*sample* $t$-*test* | $d = \bar{x} - \bar{y}$ | $t^* = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{s_X^2}{n_X}+\frac{s_Y^2}{n_Y}}}$ | $T_{\min\{n_X-1, n_Y-1\}}$ |

Table 1: Summary of different test statistics and the corresponding distributions.

**Example 4.** *(Adaptation of Fall '17, Question 9)*

*It is desired to know the know the fitness level of $K - 12$ students. One way is through the 1 mile run time. At the moment, the national average run time among 12 year-olds is 8.75 minutes. A local school randomly choose 40 students and find that their average is 8.95 minutes with a standard deviation of .8 minutes.*

*Assume a confidence level of $\alpha = .1$,*

*(a) Determine if their students' average is different from the national average.*

*(b) Determine if their students' fitness average is worse than the national fitness average.*

*Solution:*

We first recognize that this question concerns 1-sample mean. From the question, we collect some important information: let $\bar{x}$ be the observed mean and $s$ the observed standard deviation:

$$\mu = 8.75 \qquad n = 40; \qquad \bar{x} = 8.95; \qquad s = .8;$$

(a) We first write the hypothesis:

$$H_0 : \mu = 8.75; \qquad H_A : \mu \neq 8.75$$

We next calculate the test statistics:

$$t^* = \frac{\bar{x} - \mu}{{}^{s}/\sqrt{n}} = \frac{8.95 - 8.75}{{}^{.8}/\sqrt{40}} = 1.58$$

Given $n = 40$, we are working with $t$-distribution with df $= 39$. Since our $t$-table does not return $p$ value, we have to instead look-up the threshold $t_\alpha$. In our table, the df closest is df $= 40$, so we are going to use this:

$$\text{2-tailed } T_{40} \text{ with } \alpha = .1 : \quad t_\alpha = 1.684$$

We observe that

$$t^* = 1.58 < 1.740 = t_\alpha \qquad \Longrightarrow \quad \text{Fail to reject } H_0$$

We conclude that at the confidence level of $\alpha = .1$ there does not appear evidence that their students' average is different from the national average.

(b) We rewrite the hypothesis:

$$H_0 : \mu = 8.75; \qquad H_A : \mu > 8.75$$

Since this is a 1-tailed test:

$$\text{1-tailed } T_{40} \text{ with } \alpha = .1 : \quad t_\alpha = 1.303$$

We observe that

$$t^* = 1.58 > 1.303 = t_\alpha \qquad \Longrightarrow \quad \text{Reject } H_0$$

We conclude that at the confidence level of $\alpha = .1$ there appears evidence that their students' average is higher from the national average, and hence the fitness is worse than the national average.

**Example 5.** *(Adaptation of Fall '17, Question 6)*

*Let us now investigate if a new drug is effective at helping people lose weight. Suppose we have carried out our experiment noting the weight loss and found the following:*

| *Treatment* | $n$ | *Mean* | *Standard deviation* |
|---|---|---|---|
| *Drug* | 22 | 2.4 | 4.9 |
| *Placebo* | 18 | .05 | 2.3 |

*Assume a confidence level of $\alpha = .05$, determine if the drug is indeed effective.*

*Solution:*

We first recognize that this question concerns 2-sample mean. From the question, we collect some important information: let $X, Y$ be the statistics for drug and placebo respectively:

$$\bar{x} = 2.4; \quad s_x = 4.9; \quad n_x = 22; \qquad \text{and} \qquad \bar{y} = .05; \quad s_y = 2.3; \quad n_y = 18;$$

Let $d$ be the observed difference:

$$d = \bar{x} - \bar{y} = 2.35$$

We first write the hypothesis:

$$H_0 : d = 0; \qquad H_A : d > 0$$

We next calculate the test statistics:

$$t^* = \frac{d}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} = \frac{2.35}{\sqrt{\frac{4.9^2}{22} + \frac{2.3^2}{18}}} = 1.997$$

Given $n_x = 22, n_y = 18$, we are working with $t$-distribution with df $= 17$. Since our $t$-table does not return $p$ value, we have to instead look-up the threshold $t_\alpha$:

$$\text{1-tailed } T_{17} \text{ with } \alpha = .05 : \quad t_\alpha = 1.740$$

We observe that

$$t^* = 1.997 > 1.740 = t_\alpha \quad \Longrightarrow \quad \text{Reject } H_0$$

We conclude that at the confidence level of $\alpha = .05$ there appears evidence that the drug is effective at helping people lose weight.