Lecture 24-25 Worksheet

Chrysafis Vogiatzis

Every worksheet will work as follows.

1. You will be entered into a Zoom breakout session with other students in the class.
2. Read through the worksheet, discussing any questions with the other participants in your breakout session.
   - You can call me using the "Ask for help" button.
   - Keep in mind that I will be going through all rooms during the session so it might take me a while to get to you.
3. Answer each question (preferably in the order provided) to the best of your knowledge.
4. While collaboration between students in a breakout session is highly encouraged and expected, each student has to submit their own version.
5. You will have 24 hours (see Compass) to submit your work.

Worksheet 1: Formulating hypotheses

Formulating statistical hypotheses can be pretty difficult at first. That said, there are a few rules we may follow:

- State the null hypothesis as an equality.
- State the alternative hypothesis as either a two-sided inequality ($\neq$) or a one-sided inequality ($\leq, \geq$) depending on the hypothesis being tested.

Let us practice that in the next few problems.

Problem 1: The proportion of students with an internship

The University of Illinois at Urbana-Champaign (UIUC) is interested in finding how many students have internships lined up for next summer. UIUC believes 50% of them have secured internships. Formulate the statistical hypothesis that the true percentage is indeed equal to 50%.

Answer to Problem 1.

H_0 :

H_1 :

Problem 2: The proportion of students with an internship: second take

What would you change if you only cared about whether the true proportion is 50%, or if it is less than 50%?

Answer to Problem 2.

H_0 :

H_1 :

Problem 3: The α and the β errors

Assume you ask a sample of UIUC students of size n through an online survey. In the survey, x of them will reply that they have indeed already secured an internship. Define (in a sentence or two) what α (Type I error) and what β (Type II error) are in this setup.

Answer to Problem 3.

Worksheet 2: World Tourism Organization

The World Tourism Organization (WTO) needs your help! Due to COVID-19 and the effects it's had in tourism and travel, the WTO is preparing a new advertising campaign to encourage travel within the United States for US residents, after COVID-19 has been placed under control.

The WTO has certain ideas about United States residents and their travel patterns, but no idea whether they are right or not! Help them by performing a full scale hypothesis test.

Problem 4: Formulating a hypothesis about a proportion

The WTO claims that a fraction $p = 0.55$ of the population has visited 7 or more states. Is that the case or maybe a different fraction of the population has visited 7 or more states? Formulate this proportion statistical hypothesis test.

Answer to Problem 4.

Problem 5: Collecting a sample

Our plan is as follows. We will ask the $n = 90$ students in the class and help WTO reject or fail to reject their null hypothesis of $p = 0.55$. Let x be the number of students who have indeed been to 7 or more states. What is the observed proportion $\hat{p} = \frac{x}{n}$ distributed as?

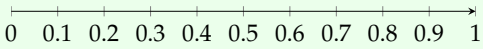
Answer to Problem 5.

$\hat{p} \sim$

Problem 6: Drawing

Get your artistic selves out! Based on your answer in Problem 5, draw the distribution here.

Answer to Problem 6.



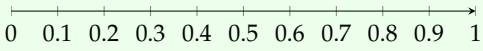
Problem 7: Adding a significance level

Let us select a significance level of 95% $\implies$ $\alpha = 5\%$. Mark the acceptance region for $\hat{p}$ by redrawing your plot with marks at

$$p_0 \pm z_{\alpha/2} \cdot \sqrt{\frac{p_0\,(1-p_0)}{n}},$$

where $p_0 = 0.55$.

Answer to Problem 7.



Problem 8: Bringing this together

Call me in your breakout room to get your personalized sample! [1] Based on the sample you were given (or that you got online), should you reject or fail to reject the null hypothesis that the true proportion is $p = 0.55$?

[1] If you are not in a breakout room when doing this, follow this link to get your observed proportion.

Answer to Problem 8.

Worksheet 3: Extensions

Problem 9: The β error for overestimating

The WTO is worried that they may be overestimating the true proportion. What is the power of the test $(1 - \beta)$ should the true proportion be 0.3?

Answer to Problem 9.

Problem 10: The β error for underestimating

What if they are underestimating? What is the power of the test $(1 - \beta)$ should the true proportion be 0.8?

Answer to Problem 10.

Problem 11: P-values

Before we head out, let us check something interesting. Given a specific observed proportion $\hat{p}$, what is the maximum value of α such that $\hat{p}$ leads to a rejection? First, try it for a specific value. What is the maximum value of α such that $\hat{p} = 0.67$ leads to rejection? You may assume that we are still dealing with the original hypothesis test of

$$H_0 : p = 0.55.$$
$$H_1 : p \neq 0.55.$$

Answer to Problem 11.

Problem 12: P-values

How would you calculate this in the general case? We call the maximum value of α that leads to rejection the P-value. How can a P-value be calculated in general?

Answer to Problem 12.