

# SST115 - Hypothesis Testing Worksheet

## Question 1 – Conceptual

**Part A:** What different types of questions do CIs and Hypothesis testing answer?

**Part B:** Describe the difference between the Null hypothesis and the Alternate hypothesis. Which one are we trying to show is true?

**Part C:** What is a Null distribution?

**Part D:** What does the p-value tell us for a hypothesis test?

---

## Question 2 – Identify the Parameters

For each of the following situations, state whether the parameter of interest is a mean or a proportion.

**Part A:** A poll shows that 64% of Americans personally worry a great deal about federal spending and the budget deficit.

**Part B:** A survey reports that local TV news has shown a 17% increase in revenue within a two year period while newspaper revenues decreased by 6.4% during this time period.

**Part C:** In a survey, high school and college students are asked whether they use geolocation services on their smart phones.

**Part D:** In a survey, smart phone users are asked whether they use a web-based taxi service.

**Part E:** In a survey, smart phone users are asked how many times they used a web-based taxi service over the last year.

---

### Question 3 – Writing Hypotheses

Write the null and alternative hypotheses in words and then symbols for each of the following situations.

**Part A** New York is known as “the city that never sleeps”. A random sample of 25 New Yorkers were asked how much sleep they get per night. Do these data provide convincing evidence that New Yorkers on average sleep less than 8 hours a night?

**Part B** Employers at a firm are worried about the effect of March Madness, a basketball championship held each spring in the US, on employee productivity. They estimate that on a regular business day employees spend on average 15 minutes of company time checking personal email, making personal phone calls, etc. They also collect data on how much company time employees spend on such non- business activities during March Madness. They want to determine if these data provide convincing evidence that employee productivity decreases during March Madness.

## P-values for Hypothesis Tests for a proportion

To get p-values, we are going to use the `pnorm()` function. We will do so with the following steps. Keep in mind what the null distribution looks like, since it gives us the value of the mean and standard deviation we will put into `pnorm()`.

$$\hat{p} \sim N(p_0, \sqrt{\frac{p_0(1-p_0)}{n}})$$

- 1) write down the null and alternate hypothesis
- 2) the sign in  $H_A$  will determine how we use `pnorm()`
- 3) calculate  $\hat{p}$
- 4) throw everything in `pnorm()`

**$H_A$ : parameter > ‘hypothesized value’ (right-tail test)** We will put the value of  $\hat{p}$  into the `pnorm()` function but we will need to specify `lower.tail=FALSE` because R by default finds ‘less than’ probabilities and we want a ‘greater than’ probability.

```
pnorm(p_hat, mean=p0, sd = sqrt(p0*(1-p0)/n), lower.tail=FALSE)
```

## Question 4

Write down the conditions for performing a hypothesis test for a single proportion and also what the null distribution looks like when these conditions are met.

## Question 5 – Monday Breakups

The following is an example from Dr. Laura Ziegler's intro statistics class at ISU.

Facebook is a social networking website. One piece of data that members of Facebook often report is their relationship status: single, in a relationship, married, it's complicated, etc.

With the help of Lee Byron of Facebook, David McCandless examined changes in peoples' relationship status, in particular, breakups. A plot of the results showed that there were repeated peaks on Mondays. Based on this initial examination of data, McCandless speculated that breakups are reported at higher frequency on Mondays.

To test this research hypothesis, McCandless collected a random sample of 75 breakups reported on Facebook within the last year. Of these sampled breakups, 20 occurred on a Monday.

**Research Question:** Are people more likely to break up on Mondays than other days of the week?

**Part A:** Describe the parameter in context and provide the appropriate symbol.

**Part B:** What is the value of the statistic?

**Part C:** Is this an observational study or an experiment?

**Part D:** What types of inferences we will be able to draw? Causal claims? Generalizations?

**Part E:** State the null and alternate hypothesis using the notation in the slides.

**Part F:** Are the conditions met to conduct a hypothesis test using the Normal distribution? State the conditions and whether they are met or not met.

**Part G:** Find the values of the mean and standard deviation for the null distribution (show work).

**Part H:** Find the p-value using `pnorm()`.

**Part I:** Write up a conclusion to answer the research question.