POL 345: Quantitative Analysis and Politics

Precept Handout 8

Week 10 (Verzani Chapters 7 and 8: 7.5, 8.6.1-8.6.2)

Remember to complete the entire handout and submit the precept questions to the Blackboard 24 hours before precept. In this handout, we cover the following new materials:

- Using `qnorm()` to calculate the critical value for the α significance level

- Using `pnorm()` to calculate the p-value for hypothesis testing

- Using `prop.test()` to conduct hypothesis testing with proportions under the Central Limit Theorem

- Using `t.test()` to conduct hypothesis tests under the t-distribution

1 Hypothesis Testing under the Central Limit Theorem

- When the sample size is sufficiently large, we can test our hypothesis by conducting a z-test using the standard normal distribution. We conduct a **one-sample z-test** when we want to test whether the mean of the population (from which we have a random sample) is equal to the hypothesized value.

- The *null hypothesis* for a *two-sided test* takes the form $H_0 : \mu = \mu_0$, where μ is the population mean and μ_0 is its hypothesized value. When we know the direction of the difference between μ and μ_0, we conduct a *one-sided test*. The *alternative hypothesis* changes from $H_1 : \mu \neq \mu_0$ (two-sided) to $H_1 : \mu > \mu_0$ or $H_1 : \mu < \mu_0$ (one-sided).

- We calculate the critical value for the α significance level by $\texttt{qnorm}(1 - \alpha/2)$ for a two-sided test, and $\texttt{qnorm}(1 - \alpha)$ or $\texttt{qnorm}(\alpha)$ for a one-sided test, depending on the direction of the prior knowledge.

- We can also calculate the p-value using $\texttt{2 * pnorm(abs(z), lower.tail=FALSE)}$ for a two-sided test and $\texttt{pnorm(z)}$ or $\texttt{pnorm(z, lower.tail=FALSE)}$ for a one-sided test where $\texttt{z}$ is the z-statistic.

- **Example: Public Opinion and the Iraq War.** In 2008, S. S. Gardner published "The Multiple Effects of Casualties on Public Support for War" in the *APSR*. Among other things, Gardner explored the association between gender and support for the war in Iraq. Gardner hypothesized that women were, on average, more likely to believe that invading Iraq was a mistake than they were to support the war. We conduct a one-sample z-test to test whether women were, on average, more likely to feel that it was a mistake to invade Iraq. The dataset includes the variables `female` (binary variable, coded as 1 if female, 0 if male) and `mistake` (binary variable, coded as 1 if respondent felt it was a mistake to invade Iraq and 0 if the respondent supported the invasion).

```
> load("iraq.RData")
> ## Calculate the mean of attitude toward the war among women
> mean.women <- mean(iraq$mistake[iraq$female == 1])
> n.women <- nrow(iraq[iraq$female == 1,])
> ## Normal test for mean: one-sample
> p0 <- 0.3 ## Contrived value for the null
> se <- sqrt(p0 * (1 - p0) / n.women)
> ## Calculate z-value
> z <- (mean.women - p0) / se
> ## Calculate p-value for hypothesis test
> 2 * pnorm(abs(z), lower.tail=FALSE) ## Two-sided p-value

[1] 1.10599e-07

> pnorm(z) ## One-sided p-value

[1] 5.52995e-08
```

- Alternatively, the command `prop.test(x, n, p = NULL, alternative = c("two.sided", "less", "greater"), conf.level = 0.95)` may be used to test the null hypothesis that proportions (probabilities of success) in a group or several groups are the same or equal to a given value. The command takes a vector of counts of successes `x` from a vector of counts of trials `n` with a vector of probabilities for success `p`. We specify the alternative hypothesis as two-sided if the null hypothesis is $H_0 : \mu = \mu_0$. For a one-sided test, we specify the alternative hypothesis as greater than or less than the null value according to prior knowledge.

```
> prop.test(sum(iraq$mistake[iraq$female == 1]),
+           length(iraq$mistake[iraq$female == 1]), p = 0.3,
+           conf.level = 0.95, alternative = "two.sided")

        1-sample proportions test with continuity correction

data:  sum(iraq$mistake[iraq$female == 1]) out of length(iraq$mistake[iraq$female == 1])
X-squared = 27.4665, df = 1, p-value = 1.598e-07
alternative hypothesis: true p is not equal to 0.3
95 percent confidence interval:
 0.109609 0.199855
sample estimates:
       p
0.149425
```

```
> ## for the one-sided
> prop.test(sum(iraq$mistake[iraq$female == 1]),
+           length(iraq$mistake[iraq$female == 1]), p = 0.3,
+           conf.level = 0.95, alternative = "less")

        1-sample proportions test with continuity correction

data:  sum(iraq$mistake[iraq$female == 1]) out of length(iraq$mistake[iraq$female == 1])
X-squared = 27.4665, df = 1, p-value = 7.992e-08
alternative hypothesis: true p is less than 0.3
95 percent confidence interval:
 0.000000 0.191394
sample estimates:
       p
0.149425
```

The results are identical to those found above.

- Setting the argument `correct = TRUE` (default) in the `prop.test()` command improves the approximation of the p-value by adjusting for the discontinuity of the binomial distribution. In other words, setting `correct = FALSE` will give the results identical to those obtained on the previous page.

```
> prop.test(sum(iraq$mistake[iraq$female == 1]),
+           length(iraq$mistake[iraq$female == 1]), p = 0.3,
+           conf.level = 0.95, alternative = "two.sided", correct = FALSE)
```

```
          1-sample proportions test without continuity correction

data:  sum(iraq$mistake[iraq$female == 1]) out of length(iraq$mistake[iraq$female == 1])
X-squared = 28.179, df = 1, p-value = 1.106e-07
alternative hypothesis: true p is not equal to 0.3
95 percent confidence interval:
 0.111274 0.197747
sample estimates:
       p
0.149425


> ## for the one-sided
> prop.test(sum(iraq$mistake[iraq$female == 1]),
+           length(iraq$mistake[iraq$female == 1]), p = 0.3,
+           conf.level = 0.95, alternative = "less", correct = FALSE)

          1-sample proportions test without continuity correction

data:  sum(iraq$mistake[iraq$female == 1]) out of length(iraq$mistake[iraq$female == 1])
X-squared = 28.179, df = 1, p-value = 5.53e-08
alternative hypothesis: true p is less than 0.3
95 percent confidence interval:
 0.000000 0.189311
sample estimates:
       p
0.149425
```

There is overwhelming evidence against the null hypothesis that 30% of women in the population believe entering Iraq was a mistake. That is, were the null hypothesis true, it would be extremely unlikely to observe only 39 of 261 women claim that the Iraq war was a mistake.

2   Hypothesis Testing under Student's t-Distribution

- We conduct a **one-sample t-test** when we want to test whether the mean of the population from which we have a random sample is equal to the hypothesized value. Note that the population distribution is assumed to be normal. In practice, many people use t-test even when the distribution of the data is not known to be normal because this test gives the results more conservative than the test based on the normal approximation.

- To conduct a t-test, we can use the `t.test()` function in **R**. The general syntax for the function is `t.test(x, alternative, mu, conf.level)`, where

  - `x` is the vector of data—we test whether its mean is statistically different from the hypothesized population parameter `mu`;
  - `conf.level` is the test's level of confidence (defaults to `0.95`);

- alternative specifies a direction of the test. Options are `"two.sided"`, `"less"`, or `"greater"` (defaults to `two.sided`). One-sided tests correspond to the situation when we know the direction of the difference (e.g. higher or lower) we expect to observe

- **Example: Civil War and GDP Growth Rate.** A broad range of literature in International Relations addresses the relationship between civil war and GDP. Scholars have sought to understand whether and how engagement in civil wars effects the country's economy. The `growth.RData` dataset includes the following:

  - `war`: A dichotomous variable, taking a value of 1 if the country was engaged in civil war in the *previous* year and a value of 0 if not

  - `growth.rate`: GDP growth over the last 12 months, in inflation-adjusted terms

  - `gdppc`: GDP per capita in constant PPP dollars, $1000's

  We conduct a one-sample t-test to determine whether gdp growth rate is positive among countries that experienced a civil war in the previous year.

```
> load("growth.RData") ## Load data

> t.test(growth$growth.rate[growth$war == 1], alternative="two.sided",
+        mu=0, conf.level=0.95)

        One Sample t-test

data:  growth$growth.rate[growth$war == 1]
t = 1.6671, df = 756, p-value = 0.0959
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.105824  1.298030
sample estimates:
mean of x
 0.596103

> ## one-sided
> t.test(growth$growth.rate[growth$war == 1], alternative="less", mu=0,
+        conf.level=0.95)

        One Sample t-test

data:  growth$growth.rate[growth$war == 1]
t = 1.6671, df = 756, p-value = 0.952
alternative hypothesis: true mean is less than 0
95 percent confidence interval:
    -Inf 1.18496
sample estimates:
mean of x
 0.596103
```

There is reasonable support for the alternative hypothesis that countries engaged in a civil war do not have zero gdp growth, on average, but very little evidence in the directional hypothesis that the growth rate is less than zero. Overall the evidence is not strong enough to reject the null hypothesis at the 95% confidence level, and so we retain the null in this instance.

3   Precept Questions

In preparation for precept, please answer the following questions, which are based on the `success.RData` and the `chechen.RData` data files available on Blackboard. Submit your code following the instructions given in the syllabus.

Above, we examined Gardner's study on the relationship between gender and public opinion regarding the Iraq war. Alternatively, scholars have argued that public opinion regarding wars is driven more by factors other than gender. In 2005, Gelpi, et. al. published "Success Matters: Casualty Sensitivity and the War in Iraq." In this study, the authors investigated whether the public was less supportive of the Iraq war when casualty numbers are high. Using the authors' data, we conduct a hypothesis test on whether Presidential Support was low when total casualties were large. The dataset includes the variables `pres.support` (percentage of population that support the President) and `casualties` (an aggregate measure of casualties from the Iraq War).

1. To begin, calculate the point estimate for Presidential support for the full sample, as well as for a subsetted sample with low casualties (casualties less than 500) and one with high casualties (casualties of 500 or greater). Note that we must first convert Presidential support from percentage points to proportions. Calculate the standard error and 95% confidence intervals.

2. Next, conduct a one-sample two-sided hypothesis test on whether the average level of Presidential support was less than 50% when casualties were high.

3. Finally we return to the Chechen data set (`chechen.RData`) analyzed in Problem Set 1. A description of the data set can be found in the problem set. Begin by calculating a variable `diffattack` indicating the difference in the number of insurgency attacks following and prior to the treatment of indiscriminate Russian fire. Next, conduct the t-test where the null hypothesis is that there is no average difference in attacks among shelled villages between pre-attack and post-attack periods and the alternative hypothesis is that the average attack has increased after shelling. Briefly interpret the results.