Econometrics – Problem Set #3

Junwoo Yang

April 3, 2021

#3.2 Let Y be a Bernoulli random variable with success probability $P(Y = 1) = p$, and let $Y_1, \cdots, Y_n$ be i.i.d. draws from this distribution. Let $\hat{p}$ be the fraction of successes (1s) in this sample.

(a) Show that $\hat{p} = \overline{Y}$.

Proof. $\hat{p} = \dfrac{\#\text{success}}{n} = \dfrac{\sum_{i=1}^n Y_i}{n} = \overline{Y}$. □

(b) Show that $\hat{p}$ is an unbiased estimator of p.

Proof. $E(\hat{p}) = E\left(\dfrac{\sum_{i=1}^n Y_i}{n}\right) = \dfrac{1}{n}\displaystyle\sum_{i=1}^n E(Y_i) = \dfrac{1}{n}\displaystyle\sum_{i=1}^n p = p.$ □

(c) Show that $\mathrm{Var}(\hat{p}) = p(1-p)/n$.

Proof. $\mathrm{Var}(\hat{p}) = \mathrm{Var}\left(\dfrac{\sum_{i=1}^n Y_i}{n}\right) = \dfrac{1}{n^2}\displaystyle\sum_{i=1}^n \mathrm{Var}(Y_i) = \dfrac{1}{n^2}\displaystyle\sum_{i=1}^n p(1-p) = \dfrac{p(1-p)}{n}.$ □

#3.3 In a poll of 500 likely voters, 270 responded that they would vote for the candidate from the democratic party, while 230 responded that they would vote for the candidate from the republican party. Let p denote the fraction of all likely voters who preferred the democratic candidate at the time of the poll, and let $\hat{p}$ be the fraction of survey respondents who preferred the democratic candidate.

(a) Use the poll results to estimate p.

Proof. $\hat{p} = \frac{270}{500} = 0.54.$ □

(b) Use the estimator of the variance of $\hat{p}$, $\hat{p}(1-\hat{p})/n$, to calculate the standard error of your estimator.

Proof. We need to be careful about notation.

$$\begin{aligned}
\mathrm{Var}(\hat{p}) &= \sigma_{\hat{p}}^2 = \frac{p(1-p)}{n} \\
\widehat{\mathrm{Var}(\hat{p})} &= \hat{\sigma}_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n} = \frac{0.54(1-0.54)}{500} = 0.0004968 \\
SE(\hat{p}) &= \sqrt{\widehat{\mathrm{Var}(\hat{p})}} = \sqrt{0.0004968} = 0.02228901074.
\end{aligned}$$

□

(c) What is the p-value for the test of $H_0\colon p = 0.5$ vs. $H_1: p \neq 0.5$?

Proof. By using t-statistic and CLT, we get

$$t^{act} = \frac{\hat{p} - 0.5}{SE(\hat{p})} = \frac{0.54 - 0.5}{0.02228901074} = 1.794606$$

$$p\text{-value} = 2\Phi(-|t^{act}|) = 2\Phi(-1.794606) = 0.0727165$$ □

(d) What is the p-value for the test of $H_0: p = 0.5$ vs. $H_1: p > 0.5$?

Proof. p-value $= 1 - \Phi(t^{act}) = 1 - \Phi(1.794606) = 0.03635825$. □

(e) Why do the results from (c) and (d) differ?

Proof. (c) is a two-sided test, which has a wider rejection region than one-sided test of (d) because the null hypothesis is rejected not only when p is greater than 0.5 but also when p is less than 0.5. Therefore, the p-value is twice as different. □

(f) Did the poll contain statistically significant evidence that the democratic candidate was ahead of the republican candidate at the time of the poll? Explain.

Proof. The answer depends on how to set up the alternative hypothesis. If H_1 indicates that $p \neq 0.5$ like (c), H_0 is accepted because the p-value is greater than 0.05. However, if H_1 indicates that $p > 0.5$ like (d), then H_0 can be rejected at $\alpha = 0.05$ significance level because the p-value is less than 0.05. Also, the judgment depends on what level of α is set. If the hypotheses are tested at $\alpha = 0.1$ significance level, both tests can reject $H_0: p = 0.5$ because both p-values are less than 0.1. □

#**3.4** Using the data in Exercise 3.3:

(a) Construct a 95% confidence interval for p.

Proof. Let $z_{0.025}$ be 97.5^{th} percentile of $N(0,1)$. That means $P(Z < z_{0.025}) = 0.975$ where $Z \sim N(0,1)$. Then by using CLT, t-statistic can be approximated by standard normal distribution:

$$\begin{aligned}0.95 &= P\left(\left|\frac{\hat{p} - p}{SE(\hat{p})}\right| < z_{0.025}\right) = P\Big(p \in (\hat{p} - z_{0.025}SE(\hat{p}), \hat{p} + z_{0.025}SE(\hat{p}))\Big) \\ &= P\Big(p \in (0.54 - 1.959964 \times 0.02228901074, 0.54 + 1.959964 \times 0.02228901074)\Big) \\ &= P\Big(p \in (0.4963143, 0.5836857)\Big)\end{aligned}$$

Hence the 95% C.I. for p is $(0.4963143, 0.5836857)$. □

(b) Construct a 99% confidence interval for p.

Proof. Similarly, the 99% C.I. for p is

$$(\hat{p} - z_{0.005}SE(\hat{p}), \hat{p} + z_{0.005}SE(\hat{p})) = (0.4825873, 0.5974127).$$ □

(c) Why is the interval in (b) wider than the interval in (a)?

Proof. The interval in (b) is wider because the significance level is lower. □

(d) Without doing any additional calculations, test the hypothesis $H_0: p = 0.50$ vs $H_1: p \neq 0.50$ at the 5% significance level.

Proof. Since $0.05 \in (0.4963143, 0.5836857)$ which is 95% C.I. for p, we cannot reject H_0 at $\alpha = 0.05$ significance level. □

#3.5 (optional) A survey of 1000 registered voters is conducted, and the voters are asked to choose between candidate A and candidate B. Let p denote the fraction of voters in the population who prefer candidate A, and let $\hat{p}$ denote the fraction of voters in the sample who prefer candidate A.

(a) You are interested in the competing hypotheses $H_0\colon p = 0.4$ vs. $H_1: p \neq 0.4$. Suppose that you decide to reject H_0 if $|\hat{p} - 0.4| > 0.01$.

(i) What is the size of this test?

Proof. The probability $K(p)$ of rejecting $H_0\colon p = 0.4$ is called power function of the test. At the value p_1 of the parameter, $K(p_1)$ is the power at p_1. The size of the test is the probability that the test actually incorrectly rejects H_0 when H_0 is true, which is exactly the probability of a Type I error. That is, $K(0.4) = \alpha$ is called significance level of the test. Note that $\hat{p} \sim (p, \frac{p(1-p)}{1000})$.

$$\begin{aligned} K(0.4) = \alpha &= P(\text{reject } H_0 \,|\, H_0 \text{ is true}) \\ &= P(|\hat{p} - 0.4| > 0.01 \,|\, p = 0.4) \\ &= P\Big(\Big|\frac{\hat{p} - 0.4}{\sqrt{p(1-p)/1000}}\Big| > \frac{0.01}{\sqrt{p(1-p)/1000}}\Big) \\ &= 1 - P\Big(-0.6454972 < \Big|\frac{\hat{p} - 0.4}{\sqrt{p(1-p)/1000}}\Big| < 0.6454972\Big) \\ &\approx 1 - \Phi(0.6454972) + \Phi(-0.6454972) \\ &= 0.518605 \end{aligned}$$

where approximation due to CLT. The rejection rule we used needs to be adjusted because it causes a large probability of type I error. □

(ii) Compute the power of this test if $p = 0.45$.

Proof. The power of at $p = 0.45$, $K(0.45)$, is the probability of making the correct decision (namely, rejecting $H_0\colon p = 0.4$ when $p = 0.45$). Hence, we are pleased that here it is large.

$$\begin{aligned} K(0.45) &= P(\text{reject } H_0 \,|\, p = 0.45) \\ &= P(|\hat{p} - 0.4| > 0.01 \,|\, p = 0.45) \\ &= 1 - P(-0.01 < \hat{p} - 0.4 < 0.01 \,|\, p = 0.45) \\ &= 1 - P\Big(\frac{-0.06}{\sqrt{p(1-p)/1000}} < \frac{\hat{p} - 0.45}{\sqrt{p(1-p)/1000}} < \frac{-0.04}{\sqrt{p(1-p)/1000}} \,\Big|\, p = 0.45\Big) \\ &= 1 - P\Big(-3.81385 < \frac{\hat{p} - 0.45}{\sqrt{p(1-p)/1000}} < -2.542567\Big) \\ &\approx 1 - \Phi(-2.542567) + \Phi(-3.81385) \\ &= 0.9945663 \end{aligned}$$

where approximation due to CLT. □

(b) In the survey, $\hat{p} = 0.44$.

(i) Test $H_0\colon p = 0.4$ vs $H_1: p \neq 0.4$ using a 10% significance level.

Proof. The same here uses CLT. Note that $SE(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{1000}} = 0.01569713$.

$$\begin{aligned} t^{act} &= \frac{\hat{p} - 0.4}{SE(\hat{p})} = \frac{0.44 - 0.4}{\sqrt{0.44 \times 0.56/1000}} = 2.548236 \\ p\text{-value} &= 2\Phi(-|t^{act}|) = 0.01082692 \end{aligned}$$

Since p-value is lower than 0.1, H_0 is rejected at 10% significance level. □

(ii) Test $H_0: p = 0.4$ vs $H_1: p < 0.4$ using a 10% significance level.

Proof. Since p-value $= \Phi(t^{act}) = 0.9945865$ is grater than 0.1, H_0 can't rejected at 10% significance level. □

(iii) Construct a 90% confidence interval for p.

Proof. $(\hat{p} - z_{0.05}SE(\hat{p}), \hat{p} + z_{0.05}SE(\hat{p})) = (0.4141805, 0.4658195)$. □

(iv) Construct a 99% confidence interval for p.

Proof. $(\hat{p} - z_{0.005}SE(\hat{p}), \hat{p} + z_{0.005}SE(\hat{p})) = (0.3995669, 0.4804331)$. □

(v) Construct a 60% confidence interval for p.

Proof. $(\hat{p} - z_{0.2}SE(\hat{p}), \hat{p} + z_{0.2}SE(\hat{p})) = (0.426789, 0.453211)$. □

(c) Suppose that the survey is carried out 30 times, using independently selected voters in each survey. For each of these 30 surveys, a 90% confidence interval for p is constructed.

(i) What is the probability that the true value of p is contained in all 30 of these confidence intervals?

Proof. Each of CIs either does or does not contain parameter p (the fraction in population). It is not probability concept. We can't say like that probability of a 90% CI contains true value of p is 0.9. However, if many such intervals were calculated, about 90% of them should contain p. Therefore, the probability that all 30 CIs contain p is wrong approach. On the contrary, in Bayesian statistics, parameter p is not fixed and can move. Thus a 90% CI means that it contains p with probability 0.9. In this case, probability that all CIs contain p is $0.9^{30} = 0.04239116$. □

(ii) How many of these confidence intervals do you expect to contain the true value of p?

Proof. It should be interpreted that 27 (90% of 30) CIs would contain p but 3 (10% of 30) CIs wouldn't. □

(d) In survey jargon, the margin of error is $1.96 \times SE(\hat{p})$; that is, it is half the length of the 95% confidence interval. Suppose you want to design a survey that has a margin of error of at most 0.5%. That is, you want $P(|t^{act}| > 0.005) < 0.05$.[1] How large should n be if the survey uses simple random sampling?

Proof. We want half length of 95% CI to be less than 0.005. That is, $z_{0.025}SE(\hat{p}) < 0.005$. We get

$$z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < 0.005$$

solving for n

$$n > \frac{\hat{p}(1-\hat{p})z_{0.025}^2}{0.005^2}.$$

Note that $\hat{p}(1-\hat{p})$ has maxima of 0.25 when $\hat{p} = 0.5$. Therefore, if n is large enough to satisfy

$$n > \frac{0.25 \times 1.959964^2}{0.005^2} = 38414.59,$$

we obtain the desired result for all $\hat{p}$. Therefore, we need at least 38415 samples. □

[1] The our textbook actually says that $P(|\hat{p} - p| > 0.005 \leq 0.005)$, but I think it's wrong.

#3.6 Let $Y_1, \cdots, Y_n$ be i.i.d. draws from a distribution with mean μ. A test of $H_0\colon \mu = 10$ vs. $H_1 : \mu \neq 10$ using the usual t-statistic yields a p-value of 0.07.

(a) Does the 90% confidence interval contain $\mu = 10$? Explain.

Proof. Since p-value is less than 0.1, $H_0\colon \mu = 10$ is rejected at $\alpha = 0.1$. Therefore, This 90% C.I. doesn't contain $\mu = 10$. □

(b) Can you determine if $\mu = 8$ is contained in the 95% confidence interval? Explain.

Proof. To decide whether $H_0\colon \mu = 8$ is rejected, it requires t-statistic for $\mu = 8$, which means information about sample mean $\hat{\mu}$ or standard error $SE(\hat{\mu})$ is needed. However, we only know p-value under $\mu = 10$. If either of $\hat{\mu}$ or $SE(\hat{\mu})$ is given, we can calculate p-value under $\mu = 8$ by using CLT, then we can decide whether 95% C.I. contains $\mu = 8$. □

#3.9 (optional) Suppose that a plant manufactures integrated circuits with a mean life of 1000 hours and a standard deviation of 100 hours. An inventor claims to have developed an improved process that produces integrated circuits with a longer mean life and the same standard deviation. The plant manager randomly selects 50 integrated circuits produced by the process. She says that she will believe the inventor's claim if the sample mean life of the integrated circuits is greater than 1100 hours; otherwise, she will conclude that the new process is no better than the old process. Let μ denote the mean of the new process. Consider the null and alternative hypotheses $H_0\colon \mu = 1000$ vs. $H_1 : \mu > 1000$.

(a) What is the size of the plant manager's testing procedure?

Proof. Let Y be random variable representing life of integrated circuits produced by new process. In situation where $SE(\overline{Y})$ is not given, we assume that $\mathrm{Var}(Y) = 100$ to calculate the size of test. Note that the size of test should be approximated using t-statistic and CLT because $\mathrm{Var}(Y)$ isn't observable. Under null distribution, $Y \sim (1000, 100^2)$ and $\overline{Y} \sim (1000, \frac{100^2}{50})$.

$$\begin{aligned} K(1000) &= P(\overline{Y} > 1100 \,|\, \mu = 1000) \\ &= P\Big(\frac{\overline{Y} - 1000}{100/\sqrt{50}} > \frac{1100 - 1000}{100/\sqrt{50}}\Big) \\ &\approx 1 - \Phi(7.071068) \\ &= 7.687184 \times 10^{-13} \end{aligned}$$

where approximation due to CLT. □

(b) Suppose the new process is in fact better and has a mean integrated circuit life of 1150 hours. What is the power of the plant manager's testing procedure?

Proof. Similarly, we also approximate power of test using z-statistic and CLT.

$$\begin{aligned} K(1150) &= P(\overline{Y} > 1100 \,|\, \mu = 1150) \\ &= P\Big(\frac{\overline{Y} - 1150}{100/\sqrt{50}} > \frac{1100 - 1150}{100/\sqrt{50}}\Big) \\ &\approx 1 - \Phi(-3.535534) \\ &= 0.9997965 \end{aligned}$$

which means that the probability of type II error, β, is close to zero. □

(c) What testing procedure should be plant manager use if she wants the size of her test to be 1%?

Proof. Because α should be larger than (a), the rejection region is expected to be generous.

$$0.01 = K(1000) = P\Big(\frac{\overline{Y} - 1000}{100/\sqrt{50}} > z_{0.01}\Big) = P\Big(\overline{Y} > z_{0.01}\frac{100}{\sqrt{50}} + 1000\Big) = P(\overline{Y} > 1032.9).$$

Hence, in order for the size of test to be 0.01, she should conclude that the new process is better than the old one when $\overline{Y} > 1032.9$. □

#3.12 To investigate possible gender discrimination in a British firm, a sample of 120 men and 150 women with similar job descriptions are selected at random. A summary of the resulting monthly salaries follows:

	Average salary ($\overline{Y}$)	Standard deviation (s_Y)	n
Men	£8200	£450	120
Women	£7900	£520	150

(a) What do these data suggest about wage differences in the firm? Do they represent statistically significant evidence that average wages of men and women are different? (To answer this question, first, state the null and alternative hypotheses; second, compute the relevant t-statistic; third, compute the p-value associated with the t-statistic; and, finally, use the p-value to answer the question.)

Proof. The null and alternative hypotheses for wage differences are

$$H_0 : \mu_1 - \mu_2 = 0 \quad \text{vs.} \quad H_1 : \mu_1 - \mu_2 \neq 0$$

where subscript 1 and 2 represent men and women respectively. The t-statistic is

$$t^{act} = \frac{\overline{Y}_1 - \overline{Y}_2 - 0}{SE(\overline{Y}_1 - \overline{Y}_2)} = \frac{8200 - 7900}{\sqrt{\frac{450^2}{120} + \frac{520^2}{150}}} = 5.078064.$$

The p-value of two-sided test is

$$p\text{-value} = 2\Phi(-|t^{act}|) = 2\Phi(-5.078064) = 3.813003 \times 10^{-7}.$$

Because p-value is very low, we can reject H_0 at very low significance levels. Therefore, there are statistically significant wage differences between men and women. □

(b) Do these data suggest that the firm is guilty of gender discrimination in its compensation policies? Explain.

Proof. The one-sided test, i.e., $H_1 : \mu_1 - \mu_2 > 0$, still results in a very low level of p-value. It can be said that men's wages are superior to women's. But this does not imply gender discrimination. There are many factors that determine wage levels, such as education level and work ability. The above data simply shows information about wage differences between men and women without any consideration for these variables. That is, what this data suggests is a necessary condition for claiming gender discrimination, but not sufficient. □

#3.16 Assume that grades on a standardized test are known to have a mean of 500 for students in Europe. The test is administered to 600 randomly selected students in Ukraine; in this sample, the mean is 508, and the standard deviation (s) is 75.

(a) Construct a 95% confidence interval for the average test score for Ukrainian students.

Proof. The 95% CI for population mean of Ukrainian student which is denoted μ, is

$$\begin{aligned}(\overline{Y} - z_{0.025}SE(\overline{Y}),\ \overline{Y} + z_{0.025}SE(\overline{Y})) &= \Big(508 - z_{0.025}\frac{75}{\sqrt{600}},\ 508 + z_{0.025}\frac{75}{\sqrt{600}}\Big) \\ &= (501.9989,\ 514.0011).\end{aligned}$$

□

(b) Is there statistically significant evidence that Ukrainian students perform differently than other students in Europe?

Proof. Since the 95% CI does not contain $\mu = 500$, $H_0\colon \mu = 500$ can rejected at 5% significance level. Thus Ukrainian students perform differently than other students in Europe. □

(c) Another 500 students are selected at random from Ukraine. They are given a 3-hour preparation course before the test is administered. Their average test score is 514 with a standard deviation of 65.

(i) Construct a 95% confidence interval for the change in average test score associated with the prep course.

Proof. Let the subscripts 1 and 2 prep and not prep respectively.

$$\begin{aligned}
&\left(\overline{Y}_1 - \overline{Y}_2 - z_{0.025} SE(\overline{Y}_1 - \overline{Y}_2),\ \overline{Y}_1 - \overline{Y}_2 - z_{0.025} SE(\overline{Y}_1 - \overline{Y}_2)\right) \\
&= \left(514 - 508 - z_{0.025}\sqrt{\frac{65^2}{500} + \frac{75^2}{600}},\ 514 - 508 + z_{0.025}\sqrt{\frac{65^2}{500} + \frac{75^2}{600}}\right) \\
&= (-2.274902,\ 14.2749).
\end{aligned}$$

□

(ii) Is there statistically significant evidence that the prep course helped?

Proof. Since the 95% CI for difference contain $\mu_1 - \mu_2 = 0$, H_0 cannot rejected at $\alpha = 0.05$ significance level. Therefore, preparation course is not statistically helpful. □

(d) The original 600 students are given the prep course and then are asked to take the test a second time. The average change in their test scores is 7 points, and the standard deviation of the change is 40 points.

(i) Construct a 95% confidence interval for the change in average test scores.

Proof. Suppose X represents the difference in average scores between the first tests and second tests that taken after preparation course.

$$\begin{aligned}
(\overline{X} - z_{0.025} SE(\overline{X}),\ \overline{X} + z_{0.025} SE(\overline{X}) &= \left(7 - 1.959964\frac{40}{\sqrt{600}},\ 7 + 1.959964\frac{40}{\sqrt{600}}\right) \\
&= (3.799392,\ 10.20061).
\end{aligned}$$

□

(ii) Is there statistically significant evidence that students will perform better on their second attempt, after taking the prep course?

Proof. The above 95% CI does not contain $\mu_X = 0$, we say that test score will improve after taking the preparation course. □

(iii) Students may have performed better in their second attempt because of the prep course or because they gained test-taking experience in their first attempt. Describe an experiment that would quantify these two effects.

Proof. Suppose there are n students who took a test once. Take a preparation course for about half of the students. Then let all n students take the second test. Test the mean difference between the first and second tests for each group of students who have been taught and who have not. If a significant mean difference is identified only for the group of students who have been taught, there is no effect from the experience of taking the test. □

#E3.2 A consumer is given the chance to buy a baseball card for \$1, but he declines the trade. If the consumer is now given the baseball card, will he be willing to sell it for \$1? Standard consumer theory suggests yes, but behavioral economists have found that ownership tends to increase the value of goods to consumers. That is, the consumer may hold out for some amount more than \$1 (for example, \$1.20) when selling the card, even though he was willing to pay only some amount less than \$1 (for example, \$0.88) when buying it. Behavioral economists call this phenomenon the endowment effect. John List investigated the endowment effect in a randomized experiment involving sports memorabilia traders at a sports-card show. Traders were randomly given one of two sports collectibles, say good A or good B, that had approximately equal market value. Those receiving good A were then given the option of trading good A for good B with the experimenter; those receiving good B were given the option of trading good B for good A with the experimenter.

(a) (i) Suppose that, absent any endowment effect, all the subjects prefer good A to good B. What fraction of the experiment's subjects would you expect to trade the good that they were given for the other good?

Proof. Let X be a Bernoulli random variable representing the good A or good B traders receive. Since they receive good at random, $P(X = A) = P(X = B) = \frac{1}{2}$. Since all subjects prefer good A, all subjects who receive good B will exercise their options.

$$\begin{aligned} E(\text{trade}) &= P(X = A)E(\text{prefer } B|X = A) + P(X = B)E(\text{prefer } A|X = B) \\ &= \frac{1}{2} \times 0 + \frac{1}{2} \times 1 = \frac{1}{2}. \end{aligned}$$ □

(ii) Suppose that, absent any endowment effect, 50% of the subjects prefer good A to good B, and the other 50% prefer good B to good A. What fraction of the subjects would you expect to trade the good they were given for the other good?

Proof. Similarly,

$$\begin{aligned} E(\text{trade}) &= P(X = A)E(\text{prefer } B|X = A) + P(X = B)E(\text{prefer } A|X = B) \\ &= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}. \end{aligned}$$ □

(iii) Suppose that, absent any endowment effect, X% of the subjects prefer good A to good B, and the other $(100 - X)$% prefer good B to good A. Show that you would expect 50% of the subjects to trade the good they were given for the other good.

Proof. Similarly,

$$\begin{aligned} E(\text{trade}) &= P(X = A)E(\text{prefer } B|X = A) + P(X = B)E(\text{prefer } A|X = B) \\ &= \frac{1}{2} \times (100 - X)\% + \frac{1}{2} \times X\% = 50\%. \end{aligned}$$ □

(b) Using the sports-card data, what fraction of the subjects traded the good they were given? Is the fraction significantly different from 50%? Is there evidence of an endowment effect?

Proof. First of all, import and look over the data.

```
# Data Import
rm(list=ls())
library(readxl)
cards <- read_excel('Sportscards.xlsx')

# Data Skimming
head(cards)
```

```
## # A tibble: 6 x 9
##   goodb dealer trades_p_m years_trade income  male education   age trade
##   <dbl>  <dbl>      <dbl>       <dbl>  <dbl> <dbl>     <dbl> <dbl> <dbl>
## 1     1      1         70          12      6     1         4    37     1
## 2     1      1         40           2      5     1         5    40     1
## 3     1      1         35          10      7     1         5    29     0
## 4     0      1         33           3      5     1         4    42     1
## 5     0      1         32          10      4     1         3    25     0
## 6     0      1         30          10      3     1         6    24     1
```

```r
summary(cards)
```

```
##      goodb           dealer      trades_p_m     years_trade         income     
##  Min.   :0.000   Min.   :0.0   Min.   : 0.00   Min.   : 0.000   Min.   :1.000  
##  1st Qu.:0.000   1st Qu.:0.0   1st Qu.: 3.00   1st Qu.: 3.000   1st Qu.:2.000  
##  Median :1.000   Median :0.5   Median : 8.00   Median : 7.000   Median :5.000  
##  Mean   :0.527   Mean   :0.5   Mean   :10.24   Mean   : 8.655   Mean   :4.149  
##  3rd Qu.:1.000   3rd Qu.:1.0   3rd Qu.:15.00   3rd Qu.:11.250   3rd Qu.:6.000  
##  Max.   :1.000   Max.   :1.0   Max.   :70.00   Max.   :60.000   Max.   :8.000  
##       male          education          age            trade       
##  Min.   :0.0000   Min.   :1.000   Min.   :10.00   Min.   :0.0000  
##  1st Qu.:1.0000   1st Qu.:2.000   1st Qu.:25.00   1st Qu.:0.0000  
##  Median :1.0000   Median :4.000   Median :34.00   Median :0.0000  
##  Mean   :0.8986   Mean   :3.628   Mean   :34.69   Mean   :0.3378  
##  3rd Qu.:1.0000   3rd Qu.:5.000   3rd Qu.:43.00   3rd Qu.:1.0000  
##  Max.   :1.0000   Max.   :6.000   Max.   :76.00   Max.   :1.0000  
```

Now we test hypothesis $H_0 : \mu = 0.5$ vs. $H_1 : \mu \neq 0.5$.

```r
n <- nrow(cards) # the number of samples
Ybar_act <- mean(cards$trade) # actual realized value of sample mean
s_Y <- sd(cards$trade) # sample variance
t_act <- (Ybar_act - 0.5)/(s_Y/sqrt(n)) # actual t-statistic
pvalue <- 2*pnorm(-abs(t_act)) # p-value for two sided-test
cat(n,Ybar_act,s_Y,t_act,pvalue)
```

```
## 148 0.3378378 0.474579 -4.156922 3.225641e-05
```

Because $n = 148 > 30$, standard normal distribution can be used as a good approximation to t-statistic by CLT. Thus p-value for two-sided test:

$$2\Phi\left(-\left|\frac{\overline{Y}^{act} - 0.5}{s_Y/\sqrt{n}}\right|\right) = 2\Phi\left(-\left|\frac{0.3378378 - 0.5}{0.474579/\sqrt{148}}\right|\right) = 2\Phi(-4.1569219) = 3.2256415 \times 10^{-5}.$$

The result is that we can reject null hypothesis that $\mu = 0.5$ at a very low significance level. Hence, The fraction in population differs significantly from 50%, so there exists an endowment effect. □

(c) Some have argued that the endowment effect may be present but that it is likely to disappear as traders gain more trading experience. Half of the experimental subjects were dealers, and the other half were nondealers. Dealers have more experience than nondealers. Repeat (b) for dealers and nondealers. Is there a significant difference in their behavior? Is the evidence consistent with the hypothesis that the endowment effect disappears as traders gain more experience?

Proof. First, make sure that the dealers actually trade more than the nondealers. Note that the proportion of dealer is actually half.

```r
library(dplyr)
dealers <- cards %>%
  group_by(dealer) %>%
  summarise(mean(trades_p_m), sd(trades_p_m), mean(years_trade), sd(years_trade), n()) %>%
  data.frame()
print(dealers)

##   dealer mean.trades_p_m. sd.trades_p_m. mean.years_trade. sd.years_trade. n..
## 1      0         5.662162       6.424007          6.945946        9.370811  74
## 2      1        14.824324      11.029671         10.364865        6.747043  74

# Does the dealers actually have a higher number of trades per month than the nondealers?
Ybar_m_diff <- dealers[1,2]-dealers[2,2]
SE_m_diff <- sqrt((dealers[1,3])^2/dealers[1,6] + (dealers[2,3])^2/dealers[2,6])
t_m_diff <- (Ybar_m_diff - 0)/SE_m_diff
p_m_diff <- 2*pnorm(-abs(t_m_diff))

# Does the dealers actually have a higher number of years traded than the nondealers?
Ybar_y_diff <- dealers[1,4]-dealers[2,4]
SE_y_diff <- sqrt((dealers[1,5])^2/dealers[1,6] + (dealers[2,5])^2/dealers[2,6])
t_y_diff <- (Ybar_y_diff - 0)/SE_y_diff
p_y_diff <- 2*pnorm(-abs(t_y_diff))

cat(p_m_diff,p_y_diff)

## 6.6237e-10 0.01086459
```

Since both p-values are less than 0.05, we reject the null hypotheses that there is no difference in trade frequency at $\alpha = 0.05$ significance level. Thus the statement that dealers have more experience than nondealers is acceptable. Now we test following three pairs of hypotheses:

(1) $H_0 \colon \mu_{\text{Dealers}} = 0.5$ vs. $H_1 \colon \mu_{\text{Dealers}} \neq 0.5$

(2) $H_0 \colon \mu_{\text{Nondealers}} = 0.5$ vs. $H_1 \colon \mu_{\text{Nondealers}} \neq 0.5$

(3) $H_0 \colon \mu_{\text{Dealers}} - \mu_{\text{Nondealers}} = 0$ vs. $H_1 \colon \mu_{\text{Dealers}} - \mu_{\text{Nondealers}} \neq 0$.

```r
traded <- cards %>%
  group_by(dealer) %>%
  summarise(mean(trade), sd(trade), n()) %>%
  data.frame()

# Test for (1) and (2)
traded[,5] <- (traded[,2] - 0.5)/(traded[,3]/sqrt(traded[,4]))
traded[,6] <- 2*pnorm(-abs(traded[,5]))
colnames(traded) <- c("Dealer", "Ybar_act", "s_Y", "n", "t_act", "pvalue")
print(traded)

##   Dealer  Ybar_act       s_Y  n      t_act       pvalue
## 1      0 0.2297297 0.4235304 74 -5.4894589 4.031669e-08
## 2      1 0.4459459 0.5004626 74 -0.9291215 3.528261e-01

# Test for (3)
Ybar_diff <- traded[1,2]-traded[2,2]
SE_diff <- sqrt((traded[1,3])^2/traded[1,4] + (traded[2,3])^2/traded[2,4])
t_diff <- (Ybar_diff - 0)/SE_diff
p_diff <- 2*pnorm(-abs(t_diff))
cat(Ybar_diff,SE_diff,t_diff,p_diff)

## -0.2162162 0.07621456 -2.836941 0.004554803
```

Based on the above results, we can conclude that the fraction of the dealers traded the good they were given to whole dealer group in population does not differ from 50%, while for nondealers it is significantly different from 50%. In addition, the population means for each group of dealers and nondealers are significantly different. Namely, we accept H_0, H_1, and H_1 in order. Note that all tests can be applied at $\alpha = 0.005$ significance level.

The results of these three tests consistently show that there is an endowment effect only in the nondealer group. Therefore, The hypothesis that the endowment effect disappears as traders gain more experience is reasonable. □