Answers (Problem set 8)

Online S520

1. *Trosset exercise 9.6.6*

   Let X_i denote the time that light bulb i burns before failing. The company claims that $\mu = E(X_i) \geq 800$, whereas the dissatisfied customers believe that $\mu < 800$. Before the consumer organization will allocate financial resources to countering the company's advertising campaign, it needs to be convinced that their claim is indeed exaggerated. Accordingly, it will test the null hypothesis $H_0 : \mu \geq \mu_0 = 800$ against the alternative hypothesis $H_1 : \mu < \mu_0 = 800$.

   The test statistic is

   $$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = (745.1 - 800)/(238/\sqrt{100}) = -2.307.$$

   Using a standard normal distribution, we can compute the P-value (or significance probability) (refer to the formula on Trosset p.217 Test (c))

   $$p = P_{\mu_0}(T_n \leq -2.307) \approx \Phi(-2.307) = P(Z \leq -2.307) \quad \text{(Refer to Trosset p. 215)}$$

   ```
   > pnorm(-2.307)
   [1] 0.01052741
   ```

   This is below the significance level of 0.05, so we reject the null hypothesis; there is strong evidence against the company. (You could also argue for a t-test here, in which case the P-value would be 0.012 — not substantively different.)

2. (Trosset exercise 9.6.7.)

   Let X_i denote the information units for Picture A minus the information units for Picture B elicited from patient i. The average numbers of information units elicited using A and B are identical if $\mu = E(X_i) = 0$. Accordingly, it will test the null hypothesis $H_0 : \mu = \mu_0 = 0$ against the alternative hypothesis $H_1 : \mu \neq \mu_0 = 0$.

   The test statistic is

   $$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = (-0.1833 - 0)/(5.18633/\sqrt{60}) = -0.27.$$

   Using a standard normal distribution, we can compute the P-value (or significance probability) as below (refer to the formula on Trosset p.217 Test (a))

$$\begin{aligned} p &= P_{\mu_0}(|T_n| \geq |-0.27|) \\ &= 1 - P_{\mu_0}(|T_n| < |-0.27|) \\ &= 1 - P_{\mu_0}(-0.27 < T_n < 0.27) \\ &= 1 - P_{\mu_0}(T_n < 0.27) + P_{\mu_0}(T_n < -0.27) \\ &\approx 1 - \Phi(0.27) + \Phi(-0.27) \quad \text{(Refer to Trosset p. 215)} \\ &= 1 - P(Z < 0.27) + P(Z < -0.27) \end{aligned}$$

```
> 1 - pnorm(0.27) + pnorm(-0.27)
[1] 0.7871603
```

Or use the results on Lecture Notes pages 6 and 7:
The two-tailed P-value is:

```
t = (-0.1833 - 0)/(5.18633/sqrt(60))
2 * (1 - pnorm(abs(t)))

[1] 0.7842652
```

(If use a t-distribution,

```
> t = (-0.1833 - 0)/(5.18633/sqrt(60))
> 2 * (1 - pt(abs(t), 59))
[1] 0.7852214
```

This is not a small probability (much bigger than $\alpha = 0.05$), so we do not reject H_0. The data is compatible with the null hypothesis that describing Picture A in the morning is equivalent to describing Picture B in the afternoon.

3. *A researcher wanted to see if live reggae music improved students' math test scores.He selected a sample of 61 students from his university and gave them a math test. The students then studied for two and a half hours while an acoustic reggae band played quietly. The students then took another math test of the same difficulty.*

   The researcher's variable was the change in test score for each student — a positive change meant the student did better, while a negative change meant a student did worse. He found that the changes had an approximately normal distribution with mean 6.5 and standard deviation 12.

   (a) *Assuming the sample is random, test the hypothesis that the population average is positive.*

   The experimental unit is a student. The students are sampled from one population (some university.) The measurements we are using are the changes in test scores (after reggae minus before reggae.) Let μ be the population mean — i.e. the mean change in test score if the treatment was given to *all* students at the university. The instructions give a direct for the test, so set up one-sided hypotheses:

$$\begin{aligned} H_0 &: \mu \leq 0 \\ H_1 &: \mu > 0. \end{aligned}$$

Now find the test statistic and P-value:

```
> t.stat = (6.5 - 0) / (12 / sqrt(61))
> 1 - pnorm(t.stat)
[1] 1.165593e-05
```

Alternatively, use the t-distribution:

```
> 1 - pt(t.stat, df=60)
[1] 4.04772e-05
```

Either way, the P-value is tiny (less than 10^{-4}): If the null hypothesis model were true, we'd see a sample average as big as 6.5 less than 1 in 10,000 times. So it's reasonable to reject the null hypothesis and conclude that there is a strong evidence that live reggae music does improve students' math test score.

(b) *Does the study provide convincing evidence that live reggae music improves students' math test scores? Explain why or why not.*

Nope. This was not a well-controlled experiment — we don't know *what* is improving the test scores. Is it the reggae? The extra practice? Was the second test really of the same difficulty? Were students more relaxed the second time for non-reggae-related reasons? We can't tell.

4. *Every semester, a professor gives his undergraduate statistics classes a test for psychic powers. The test consists of guessing which side of a screen a picture will appear on: left or right. In one trial of the test, a student has to guess "Left" or "Right"; then R's random number generator will randomly choose one side of the screen to display a picture of a star. The student repeats the process for a total of 20 trials.*

   (a) *In words, the null hypothesis for a particular student is that they don't have psychic powers and they're just randomly guessing. The alternative hypothesis is that they do have psychic powers and in the long run, they can do better than randomly guessing. Let p be the probability the student guesses correctly on any particular trial. Write down mathematical null and alternative hypotheses in terms of p.*

$$\begin{aligned} H_0 &: p \leq 0.5 \\ H_1 &: p > 0.5 \end{aligned}$$

   (b) *Suppose the null hypothesis is true for a particular student. Let Y be the number of times the student guesses correctly. Then if the null is true, Y has a Binomial$(20, 0.5)$ distribution. Which of the following is true?*

      i. *If the student guesses 13 right out of 20, the significance probability (P-value) is the probability under the null of guessing 13 or more out of 20.*

ii. *If the student guesses 13 right out of 20, the significance probability (P-value) is the probability under the null of guessing 13 or fewer out of 20.*

(i) is true. We want a small probability to be evidence in favor of psychic powers, i.e. the higher the score, the smaller the P-value. This happens with (i); the opposite happens with (ii).

(c) *Suppose a student guesses 13 right out of 20. What is the P-value? Do you think 13 out of 20 is intriguing evidence that the student has psychic powers?*

The P-value is `1 - pbinom(12, 20, 0.5)` $= 0.13$. This isn't really evidence of psychic powers – a score that's in the top 13% is not especially impressive.

(d) *Now suppose a student guesses 19 right out of 20. What is the P-value? Do you think 19 out of 20 is intriguing evidence that the student has psychic powers?*

The P-value is `1 - pbinom(18, 20, 0.5)` $= 0.00002$. You might a priori reject the idea of psychic powers, in which case there's no result of this test that would convince you the student had psychic powers. Otherwise, this is impressive. Of course, the sample size is only 20, so we would want to give the student a more thorough follow-up test before claiming they really did have psychic powers.

5. *Recently, the number of airline companies that offer in-flight Wi-Fi service to passengers has increased. However, it is estimated that only 10% of the passengers who have Wi-Fi available to them are willing to pay for it. Suppose Gogo, the largest provider of airline Wi-Fi service, would like to test this hypothesis by randomly sampling 125 passengers and asking them if they would be willing to pay $4.95 for 90 minutes of onboard Internet access. Suppose that 20 passengers indicated they would use this service. What is the significance probability? Based on this sample, can we conclude that the proportion of airline passengers willing to pay for onboard Wi-Fi service is different than 10%? Use $\alpha = 5\%$. **(Use Binomial probability)***

Solutions: This is a two-tailed test of the null $H_0 : p = 0.1$. The P-value is `2 * (1-pbinom(19, 125, 0.1))` $= 0.047 < \alpha = 0.05$. So there is enough evidence to reject the null hypothesis and conclude that the proportion of airline passengers passengers willing to pay for onboard Wi-Fi service is different than 10%, at the significance level 5%.