Week 7 Quiz - Hypothesis Testing - Solutions

FIT5197 teaching team

Note you will need to use the z- and t- tables in the unit [Formula Sheet (https://lms.monash.edu/mod/resource/view.php?id=7439150)](https://lms.monash.edu/mod/resource/view.php?id=7439150) to answer the following questions.

Question 1

1,500 men followed the Atkin's diet for a month. A random sample of 29 men gained an average of 6.7 pounds. Test the hypothesis that the average weight gain per man for the month was over 5 pounds. The standard deviation for the population of the 1500 men is known to be 7.1.

Answer 1

First we are given a random sample of $n = 29$ men from a population of 1500 men. We are also given the the sample mean $\hat{\mu} = 6.7$, told the standard deviation of the population is known to be $\sigma = 7.1$, and asked to test the hypothesis that average weight gain is greater than $\mu_0 = 5$ pounds. Given that no further information is provided we assume we are dealing with a Gaussian distribution for the average weight gain in our population. This implies we need to apply a one-sided (because the hypothesis involves the words ``greater than'') z-test (because σ is known and we assumed a Gaussian distribution).

First we need to define the null and alternative hypothesis. The hypothesis we are asked to test is that the average weight gain of the population μ is greater than $\mu_0 = 5$ pounds. Since there is no equality in this expression we choose to set it as the the alternative hypothesis based on the formulations for hypothesis tests taught in the lectures. Thus we define the null and alternative hypotheses as follows:

$$H_0 : \mu \leq 5$$
$$H_A : \mu > 5$$

In this situation we treat standardised differences

$$z_\mu = \frac{\bar{\mu} - \mu_0}{\sigma/\sqrt{n}}$$

that are large and positive as evidence against the null because the other direction will just support the null. So the p-value is the probability of seeing a z-score at least as large as z_μ, i.e.,

$$p = p(Z > z_\mu) = 1 - p(Z < z_\mu)$$

where $Z \sim N(0, 1)$.

Now if at this point it is not clear to you why we are using a one-sided z-test you can use the unit Formula Sheet to help you figure out which test to apply depending on the situation.

Looking up the unit Formula Sheet (https://lms.monash.edu/mod/resource/view.php?id=7439150) and going to section 7 Hypothesis Tests we can see the following text:

given an arbitrary {\bf test statistic} x with CDF $P(X)$ (i.e. x could be z or t), then the p-value is given by

$$p = \begin{cases} 2P(-|x|) & \text{if null hypothesis is equality} \\ 1 - P(x) & \text{if null hypothesis involves } \leq \\ P(x) & \text{if null hypothesis involves } \geq \end{cases}$$

This tells us that given our null hypothesis involves $\leq$, we need to determine our p-value using $p = 1 - P(x)$. Now, if for some reason we are still not sure if we need to use the z or t statistic we can look further down the formula sheet and see the following text: assume dataset of count n with mean $\bar{X}$ and sample variance S^2:

assumptions	**null-hypo.**	**test statistic**
Gaussian, σ^2 known	μ_0	$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$
Gaussian, σ^2 unknown	μ_0	$t_{n-1} = \frac{\bar{X}-\mu_0}{S/\sqrt{n}}$

We use this table because we only have one sample and one sample mean (there is another table in the formula sheet when you have two samples and two sample means). So we know we are assuming the average weight gain for the population follows a Gaussian distribution and we know σ^2 for the population is known and that we are testing a null hypothesis involving the mean and it's relationship to a specific value μ_0. Looking at the table above this tells us to use the z-statistic. Now since we also know from the formula sheet that we should determine our p-value using $p = 1 - P(x)$, for the case of the z-statistic this means our p-value is $p = 1 - P(z)$, where $P(z)$ is the value of the cdf for the value z. To work out the p-value we first need our z-score and then we need to look up the z-table. We have:

$$\begin{aligned} z_\mu &= \frac{\bar{\mu} - \mu_0}{\sigma/\sqrt{n}} \\ &= \frac{6.7 - 5}{7.1/\sqrt{29}} \\ &= 1.289. \end{aligned}$$

Considering our z-table uses z-scores with two decimal places we can see our z-score rounds to $z_\mu = 1.29$ Looking up the z score in the z-table in the formula sheet we see the corresponding p-value is $P(z) = 0.9015$. This probability is the area under the pdf to the left of your z-score. For this one-sided one-sample z-test, you want the area in the right tail under the pdf, so subtract from 1, i.e. Our p-value is $p = 1 - P(z) = 1 - 0.9015 = 0.0985$. Applying the typical significance threshold $\alpha = 0.05$ we see that the p-value is greater than this so we cannot reject the null hypothesis that $\mu \leq 5$ and so we cannot conclude that $\mu > 5$, i.e. we cannot conclude that the average weight gain is greater than $\mu_0 = 5$ pounds.

Question 2

A premium golf ball production line must produce all of its balls to 1.615 ounces in order to get the top rating (and therefore the top dollar). Samples are drawn hourly and checked. If the production line gets out of sync with a statistical significance of more than 1%, it must be shut down and repaired. This hour's sample of 18 balls has a mean of 1.611 ounces and a standard deviation of 0.065 ounces. Do you shut down the line?

Answer 2

In this question we are given that the sample mean is $\hat{\mu} = 1.611$ ounces, the sample standard deviation is $\hat{\sigma} = 0.065$ and the sample size is $n = 18$. Since we are given the sample standard deviation and not the population standard deviation we have to apply the t-test, not the z-test, as the true standard deviation of the population is unknown. This time we are told our statistical significance level is 1%, i.e. $\alpha = 0.01$. We also note that the target weight is $\mu_0 = 1.615$ ounces and so we should apply a two-sided one-sample t-test to see if the produced ball weights in the latest hour deviate from the desired ball weight. We can also note that given the limited information we have to assume a Gaussian distribution for the population ball weight. We want to apply a two-sided test because we want to test the null hypothesis that our population mean ball weight μ is equal to $\mu_0 = 1.615$, so we define our null and alternative hypotheses to be:

$$H_0 : \mu = 1.615$$
$$H_A : \mu \neq 1.615$$

In this situation we treat standardised differences

$$t_{\hat{\mu}} = \frac{\bar{\mu} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

that are large and either positive or negative as evidence against the null because in between will just support the null. So the p-value is the probability of seeing a t-score at least as large as $|t_{\hat{\mu}}|$ or at most as small as $-|t_{\hat{\mu}}|$, i.e.,

$$p = 2p(T < -|t_{\hat{\mu}}|)$$

where $t_{\hat{\mu}} \sim T(n-1)$.

If we are still clueless, looking up our [Formula Sheet (https://lms.monash.edu/mod/resource/view.php?id=7439150)](https://lms.monash.edu/mod/resource/view.php?id=7439150) we see that for a two-sided test where the null hypothesis involves equality that our p-value should be $2P(-|x|)$. Looking at the table for single sample, single sample mean tests we also see that because we assumed a Gaussian distribution for the population ball weight, the true standard deviation is unknown (we only have the sample standard deviation) and our hypothesis involves comparison of the mean against μ_0, we should apply the t-test. So our p-value should be $p = 2P(-|x|) = 2p(T < -|t_{\hat{\mu}}|) = 2P(-|t_{\hat{\mu}}|)$. First we compute our t-score:

$$\begin{aligned} t_{\hat{\mu}} &= \frac{\bar{\mu} - \mu_0}{\hat{\sigma}/\sqrt{n}} \\ &= \frac{1.611 - 1.615}{0.065/\sqrt{18}} \\ &= -0.261 \end{aligned}$$

Now when dealing with t-tests and solving them using printed t-tables for the sake of efficiency the procedure is slightly different to z-tests and we don't usually worry about the p-value. Instead we focus on the t-score required to achieve a specific significance level and check if the absolute value of our t-score is less than or greater than, the required t-score. If the absolute value of our t-score is greater than the required t-score, we conclude statistical significance and reject the null hypothesis.

Now we note that we have the following degrees of freedom, $df = n - 1 = 18 - 1 = 17$. Looking up the $df = 17$ row and the two-sided significance level $\alpha = 0.01$ in the t-table in the formula sheet, the critical t-score is 2.892. Since the absolute value of our t-score $|t_{\hat{\mu}}| = | - 0.261| = 0.261$ is less than the critical t-score of 2.898, we conclude that we cannot reject the null hypothesis that $\mu = 1.615$ and so we decide not to switch off the production line.

Note that in practice we are better off using the greater numerical accuracy of R to compute our p-values or critical statistic values for the t- and z- statistics rather than using old fashioned tables, but in the exam you need to use the t-table and z-table from the formula sheet.

Question 3

There is some variability in the amount of phenobarbital in each capsule sold by a manufacturer. However, the manufacturer claims that the mean value is 20.0 mg. To test this, a sample of 25 pills yielded a sample mean of 19.7 with a sample standard deviation of 1.3. What inference would you draw from these data? In particular, are the data strong enough evidence to discredit the claim of the manufacturer? Use the 5 percent level of significance.

Answer 3

State the hypothesis:

$$H_0 : \mu = 20$$

$$H_A : \mu \neq 20$$

We use fewer than 30 observations in our sample and we don't have population standard deviation. So our p-value should be $p = 2P(-|x|) = 2p(T < -|t_{\hat{\mu}}|) = 2P(-|t_{\hat{\mu}}|)$. First we compute our t-score:

$$\begin{aligned} t_{\hat{\mu}} &= \frac{\bar{\mu} - \mu_0}{\hat{\sigma}/\sqrt{n}} \\ &= \frac{19.7 - 20}{1.3/\sqrt{25}} \\ &\approx -1.154 \end{aligned}$$

If you put into R `2*pt((19.7-20)/(1.3/sqrt(25)),25)` you will get ≈ 0.26 and make the conclusion that the evidence is not strong enough to reject the null hypothesis

Question 4

Twenty years ago, entering male high school students of Central High could do an average of 24 pushups in 60 seconds. To see whether this remains true today, a random sample of 36 freshmen was chosen. If their average was 22.5 with a sample standard deviation of 3.1, can we conclude that the mean is no longer equal to 24? Use the 5 percent level of significance.

Answer 4

State the hypothesis:

$$H_0 : \mu = 24$$
$$H_A : \mu \neq 24$$

we don't have population standard deviation. So our p-value should be calculated using t-distribution

$$\begin{aligned} t_{\hat{\mu}} &= \frac{\bar{\mu} - \mu_0}{\hat{\sigma}/\sqrt{n}} \\ &= \frac{22.5 - 24}{3.1/\sqrt{36}} \\ &\approx -1.9 \end{aligned}$$

If you put into R `2*pt((22.5-24)/(3.1/sqrt(36)),36)` you will get ≈ 0.0063 and make the conclusion that the evidence is strong enough to reject the null hypothesis that students can still do 24 push-ups in 60 seconds

In [8]:
```
2*pt((22.5-24)/(3.1/sqrt(36)),36)
```

0.00627140571562647

R code hackers brain-melting challenge

Solve this problem using calculations in R and the relevant built in cdf function.

An economist was curious to see if women were not satisfied with their jobs. A random sample of 25 women gave an average job satisfaction score of 46 out of 100. Given that the population standard deviation of female job satisfaction scores is 5, test the hypothesis that the average female job satisfaction score is less than or equal to 50 out of 100.

Answer

We'll treat the RV X for the job satisfaction scores to be normally distrubuted since scores can be thought of as a continuous variable. Let μ_x denote our estimate of the average job satisfaction score of women. The population standard deviation is known so we just use the standard z-statistic.

We will test:

$$H_0 : \mu_x \leq 50$$
$$H_A : \mu_x > 50$$

This will be a one-sided test.

In [8]:
```
#Calculate based on the formula
mu_x <- 46
mu <- 50
sigma <- 5
count_x <- 25
z_value <- (mu_x-mu)/(sigma/sqrt(count_x))
pval = 1 - pnorm((z_value))
result = ifelse(pval > 0.05,"we have weak/no evidence against the null", ifelse(pval<0.01,
                                                                                 "we have strong evidence against
                                                                                 "we have moderate evidence agair

cat("The p-value is:", pval, "\n")
cat("so,", result)
```

```
The p-value is: 0.9999683
so, we have weak/no evidence against the null
```

Since we have weak evidence against the null hypothesis we cannot reject it.