DATA 557: Applied Statistics and Experimental Design
Homework 2

Abhishek Saini

February 2022

Question 1

A manufacturing process is run at a temperature of 60 deg C. The manufacturer would like to know if increasing the temperature would yield an increase in output. Increasing the temperature would be more expensive, so an increase would only be used in future if it increased output. It seems unlikely that increasing the temperature would decrease output and, even if it did, there would be no value in having that information. An experiment was performed to assess the effect of temperature on the output of a manufacturing process. For this experiment, temperatures of 60 or 75 degrees C were randomly assigned to process runs. It was desired to gather more information about output at the new temperature so temperatures were randomly assigned to process runs at a ratio of 2 to 1 (2 runs at temperature 75 for every 1 at temperature 60). The process output was recorded from each run. The variables in the data set are:
run: Run number temp: Temperature output: Process output

1.1. Perform the large-sample Z-test to compare mean output for the two temperatures. Give the value of the test statistic and the p-value for the test.

The test statistic is given as

$$Z = \frac{\bar{X_A} - \bar{X_B}}{\sqrt{s_A^2/n_A + s_B^2/n_B}} = 2.551155$$

The p-value is 0.005368331.

```
> temp_exp = read.csv("temperature_experiment.csv")
> temp_60 = temp_exp[temp_exp$temp == '60',]$output
> temp_75 = temp_exp[temp_exp$temp == '75',]$output
> sd_err = sqrt(var(temp_60)/length(temp_60) + var(temp_75)/length(temp_75))
>
> Z = (mean(temp_75) - mean(temp_60))/(sd_err);Z
```

```
[1] 2.551155
>
> p_val = 1-pnorm(Z) ;p_val
[1] 0.005368331
```

1.2. Do you reject the null hypothesis at a significance level of 0.05?

Yes, since the p value is smaller than 0.05 we reject the null hypothesis

1.3. State the null hypothesis for the test.

The null hypothesis of the test is that the sample means of the two populations are equal.

$$H_0 : \mu_A = \mu_B$$

1.4. Perform the unequal-variance (Welch) t-test to compare mean output in the two temperature groups. Report the test statistic and the p-value for the test.

The test statistic is 2.5512 and the p-value is 0.008531

```
> with(temp_exp,
+ t.test(output[temp == '75'], output[temp == '60'], alternative = "greater", var.equal=F))
```

```
        Welch Two Sample t-test

data:  output[temp == "75"] and output[temp == "60"]
t = 2.5512, df = 25.633, p-value = 0.008531
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 6.032275      Inf
sample estimates:
mean of x mean of y
  1019.46   1001.24
```

1.5. Perform the equal-variance t-test to compare mean output in the two temperature groups. Report the test statistic and the p-value for the test.

The test statistic is 1.9248 and the p-value is 0.03224

```
> with(temp_exp,
+ t.test(output[temp == '75'], output[temp == '60'], alternative = "greater", var.equal=T))
```

```
    Two Sample t-test
data:  output[temp == "75"] and output[temp == "60"]
t = 1.9248, df = 28, p-value = 0.03224
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 2.116827      Inf
sample estimates:
mean of x mean of y
  1019.46   1001.24
```

1.6. Which of the three tests do you think is most valid for this experiment? Why?

The unequal variance (Welch) t-test is the most valid for this experiment. Since the number of samples are small, it is better to favour t-test over the large sample z-test. There's a big difference in the sample variance of the two samples which suggests we should go for the unequal variance (Welch) t-test. Also, equal variance is a strong assumption which will may not always be true.

1.7. Calculate a 95% confidence interval for the difference between mean output using the large-sample method.

The 95% confidence interval is given as $(6.47268, \infty)$

```
> z.05 = qnorm(0.95)
> lower = mean(temp_75) - mean(temp_60) - z.05*sd_err
> lower
[1] 6.47268
```

1.8. Calculate a 95% confidence interval for the difference between mean output using a method that corresponds to the Welch test.

The 95% confidence interval is $(6.032275, \infty)$

1.9. Calculate a 95% confidence interval for the difference between mean output using a method that corresponds to the equal-variance t-test.

The 95% confidence interval is $(2.116827, \infty)$

1.10. Apart from any effect on the mean output, do the results of the experiment suggest a disadvantage of the higher temperature?

The higher variability in output could be a potential disadvantage of using a higher temperature as you could get extreme results occasionally which may not be desirable.

Question 2

The data are from an experiment to compare 4 processing methods for manufacturing steel ball bearings. The 4 process methods were run for one day and a random sample of 1% of the ball bearings from the day was taken from each of the 4 methods. Because the processes produce ball bearings at different rates the sample sizes were not the same for the 4 methods. Each sampled ball bearing had its weight measured to the nearest 0.1 g and the number of surface defects was counted. The variables in the data set are:
Sample: sample number Method: A, B, C, or D Defects: number of defects Weight: weight in g

2.1. The target weight for the ball bearings is 10 g. For each of the 4 methods it is desired to test the null hypothesis that the mean weight is equal to 10. What test should be used?

We can use the 1-sample t-test for each of the 4 methods.

2.2. Give the p-values for the tests for each method. Include your R code for this question.

The p value for method A is 0.9367
The p value for method B is 0.6216
The p value for method C is 0.1421
The p value for method D is 0.04371

```r
defects = read.csv("defects.csv")

weight_A = defects[defects$Method == 'A',]$Weight
weight_B = defects[defects$Method == 'B',]$Weight
weight_C = defects[defects$Method == 'C',]$Weight
weight_D = defects[defects$Method == 'D',]$Weight

t.test(weight_A, mu=10)
t.test(weight_B, mu=10)
t.test(weight_C, mu=10)
t.test(weight_D, mu=10)
```

2.3. Apply a Bonferroni correction to your results from the previous question to account for inflation of type I error rate due to multiple testing. How does the Bonferroni correction change your conclusions? In particular, do you have evidence to reject the null hypothesis that the mean weight for all 4 methods is equal to 10, at significance level 0.05?

Bonferroni correction implies we have to reduce the type I error probability by a factor of 4 = 0.05/4 = 0.0125. With this correction, we fail to reject any of the null hypothesis as the lowest p-value from previous section is 0.04371.

2.4. It is is desired to compare mean weights of the 4 methods. This is to be done first by performing pairwise comparisons of mean weight for the different methods. What test should be used for these comparisons?

To test using pairwise comparison, we do the 2-sample unequal variance (Welch) t-test.

2.5. Report the p-values from all pairwise comparisons. Include your R code for this question.

The p-value on comparing method A with method B is 0.6815
The p-value on comparing method A with method C is 0.2221
The p-value on comparing method A with method D is 0.06482
The p-value on comparing method B with method C is 0.1308
The p-value on comparing method B with method D is 0.03977
The p-value on comparing method C with method D is 0.3918

```r
t.test(weight_A, weight_B)
t.test(weight_A, weight_C)
t.test(weight_A, weight_D)
t.test(weight_B, weight_C)
t.test(weight_B, weight_D)
t.test(weight_C, weight_D)
```

2.6. Apply a Bonferroni correction to your results of the previous question to account for inflation of type I error rate due to multiple testing. What conclusion would you draw from these results? Would you reject the null hypothesis of no difference between any pair of means among the 4 methods, at significance level 0.05?

With Bonferroni correction for 6 hypothesis tests we divide the type I error rate by $6 = 0.05/6 = 0.00833$. With this type I error correction, we fail to reject any of the null hypothesis since the lowest p-value is 0.03977.

2.7. Compare the mean weights for the 4 methods using ANOVA. State the F-statistic and the p-value for the F-test. Include your R code for this question.

The F-statistic is 2.617 and the p-value is 0.0515.

```
> summary(aov(defects$Weight ~ factor(defects$Method), data=defects))
                        Df Sum Sq Mean Sq F value Pr(>F)
factor(defects$Method)   3   1.29  0.4299   2.617 0.0515 .
Residuals              261  42.87  0.1643
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2.8 What do you conclude from the ANOVA?

Since the p-value is not less than 0.05, we fail to reject the null hypothesis of no difference between the group means.

2.9 How does your conclusion from ANOVA compare to the conclusion from the pairwise comparisons?

We fail to reject the null with BOTH ANOVA and with the pairwise comparison using Bonferroni's correction.