# Week 11 Quiz

**Problem 1**

Two samples of sizes 25 and 35 are independently drawn from two normal populations, where the unknown variances are assumed to be equal. The number of degrees of freedom for the equal-variances t-test statistic is:
a. 58
b. 60
c. 62
d. 57
e. 68

**Answer: A**

**Problem 2**
Which of the following statements is correct?

a. the pooled-variances t-test is used whenever the population standard deviations can be assumed to be equal, regardless of the sample size.
b. the unequal-variances t-test is used whenever the population standard deviations are unknown and cannot assumed to be equal.
c. the z-test can be used as a close approximation to the unequal-variances t-test when the population standard deviations are not assumed to be equal but both sample sizes are large (typically greater than 30)
d. all of the above statements are true.

**Answer: D**

**Problem 3.**
Two independent samples from populations that are normally distributed produced the following statistics: for sample 1 the sample size was 40, the sample mean was 34.2 and the sample standard deviation was 12.6. For sample 2, the sample size was 32, the sample mean was 49.1 and the sample standard deviation was 19.4. Assume that population variances are equal. Given a significance level of 5%, at what approximate value of t should you reject the null hypothesis that states that the two population means are equal, in favor of the two-sided alternative?

a. ±2.32
b. ±1.285
c. ±1.68
d. ±1.96
e. ±2.13

**Answer D.**

**Problem 4.**

The owner of Bun & Run Hamburgers wishes to compare the sales per day at different locations. The mean number of hamburgers sold for 10 randomly selected days at Northside was 83.55 with a sample standard deviation of 10.50. For a randomly selected 12 days at Southside, the mean number of hamburgers sold was 69.54 with a sample standard deviation of 14.25. We wish to test whether there is a difference in the mean number of hamburgers sold at the two locations using a 5% significance level. What is the value of the test statistic in this case?

a. 1.84
b. 0.24
c. 2.57
d. 1.71
e. 2.20

**Answer C:**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(10-1)110.25 + (12-1)203.06}{10+12-2} = 161.296$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{(83.55 - 69.54) - 0}{\sqrt{161.296/10 + 161.296/12}} = 2.576$$

**Problem 5**

A researcher randomly sampled 30 graduates of an MBA program and recorded data concerning their starting salaries. The sample comprised 18 women whose average starting salary is R48000, and 12 men whose average starting salary is R55000. It is known that the sample standard deviations of starting salaries for women and men are R11500 and R13000 respectively. The researcher was attempting to show that female MBA graduates have significantly lower average starting salaries than male MBA graduates. What is the value of the test statistic in this case?

a. -1.55
b. -2.16
c. -0.86
d. -1.40
e. -1.68

**Answer A.**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(18-1)132250000 + (12-1)169000000}{18+12-2} = 146687500$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{(48000 - 55000) - 0}{\sqrt{146687500/18 + 146687500/12}} = -1.551$$

**Problem 6**

It is known that the *sample variances* of final exam marks for first-year statistics students at a particular South African university are 45.3 for female students and 52.1 for male students. Samples of 27 female and 31 male first-year statistics students from the university are selected and the sample exam marks are calculated. For females, the sample mean mark is 52.3 % and for males, the sample mean mark is 55.4 %. If we wish to test whether females have, on average, higher exam marks than males, what would the test statistic value of the hypothesis test in this case be?

a. 1.04
b. 1.58
c. 0.49
d. -1.15
e. -1.68

**Answer E.**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(27-1)45.3+(31-1)52.1}{27+31-2} = 48.943$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}} = \frac{(52.3-55.4)-0}{\sqrt{48.943/27+48.943/31}} = -1.683$$

**Problem 7**

A new Grade 9 maths curriculum is to be tested. Several students in a Grade 9 maths course at a particular high school were randomly divided into two groups. The experimental group received teaching according to the new curriculum while the control group did not. All the students were given a test of computational skill (out of a total of 5 points) after the course and the results were as follows: for the experimental group, the sample size was 28, the sample mean was 1.99 and the sample variance was 85. For the control group, the sample size was 31, the sample mean was 3.54, and the sample variance was 80. We wish to test at the 5 % level of significance whether there is any difference in the mean test scores. We assume that the underlying population variances are equal. What is the value of the test statistic for this hypothesis test?

a. -1.03
b. 0.44
c. -0.80
d. -0.66
e. 0.13

**Answer: D.**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(28-1)85+(31-1)80}{28+31-2} = 82.368$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}} = \frac{(1.99-3.54)-0}{\sqrt{82.368/28+82.368/31}} = -0.655$$

**Problem 8.**

A comparison of the price-earnings (P/E) ratio for the top and bottom 100 companies on the JSE is being prepared. A financial advisor randomly sampled each group to determine whether there was any difference in the P/E ratios of the two groups of companies. Assume unequal population variances but that the populations are normally distributed. For 6 top 100 companies, the average P/E ratio was 21.03 with a sample variance of 128.17. For 8 bottom 100 companies, the average P/E ratio was 10.67 with a sample variance of 125.15. The test is conducted at the 5 % level of significance. What would be the value of the test statistic for the hypothesis test in this case?

a. 1.34
b. 0.74
c. 1.67
d. 1.72
e. 2.25

**Answer: D.**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(6-1)128.17 + (8-1)125.15}{6+8-2} = 126.408$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{(21.03 - 10.6) - 0}{\sqrt{126.408/6 + 126.408/8}} = 1.718$$

**Problem 9.**

A sociologist wants to test if there is any difference between the mean income of farm workers in two different areas. In area A a random sample of 41 workers yielded a mean income of R21000 with a standard deviation of R9010. In area B a random sample of 9 workers yielded a mean income of R15078 and a standard deviation of R5624. If we assume that the population variances are not equal, what would be the value of the test statistic for the hypothesis test in this case?

a. 0.21
b. 1.25
c. 0.67
d. 1.88
e. 2.56

**Answer: D.**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{(41-1)81180100 + (9-1)31629376}{41+9-2} = 72921646$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} = \frac{(21000 - 15078) - 0}{\sqrt{72921646/41 + 72921646/9}} = 1.884$$

**Problem 10**

The national electricity supplier claims that switching off the hot water cylinder at night does not result in saving electricity. In order to test this claim a newspaper reporter obtains the cooperation of 16 house owners with similar houses and salaries. Eight of the selected owners switch their cylinders off at night. The consumption of electricity in each house over a period of 30 days is measured; the units are kWh (kilowatt-hours). For households that switched off their hot water cylinders, average consumption over the 30 days was 680kWh with a variance of 425kWh2. For those that did not switch off their hot water cylinders, average electricity consumption was 700kWh with a variance of 300kWh2. If we wish to test the assumption of equal population variances at the 5 % level of significance, what is the value of the test statistic?

a. -1.50
b. -2.10
c. -1.29
e. -1.95

**Answer: B.**

**Step 3: Evaluate the test statistic.**

We first find the pooled sample variance in the following

$$s^2_{pool} = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2} = \frac{(8-1)425+(8-1)300}{8+8-2} = 362.5$$

The test statistic is defined to be:

$$TS = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{s_1^2/n_1+s_2^2/n_2}} = \frac{(680-700)-0}{\sqrt{362.5/8+362.5/8}} = -2.101$$

**Problem 11**

A teacher is interested in performing a hypothesis test to compare the mean math score of the girls and the mean math score of the boys. She randomly selects 10 girls from the class and then randomly selects 10 boys. She arranges the girls' names alphabetically and uses this list to assign each girl a number between 1 and 10. She does the same thing for the boys.

A) Paired t-test. Since the boys and girls are in the same class and are hence dependent samples, they can be linked.
B) 1-sample t-test. The teacher should compare the sample mean for the girls against the population mean for the boys.
C) Two-sample t-test. There is no natural pairing between the two populations.
D) Paired t-test. Since there are 10 boys and 10 girls, we can link the two samples.

**Answer: C**

**Problem 12**

The first 10 students who arrived for the Friday lecture filled out a questionnaire on their attitudes toward the instructor. The first 10 who were late for the lecture were spotted and afterward filled out the same questionnaire. The appropriate design for testing the significance of the difference between the means is related paired-sample t-test.

B) independent 2-sample t-test.
C) one-sample t-test.
D) one-sample z-test.

**Answer: B.**

# Summary of Weekly Quiz 11

## 1. Five Number Summary :

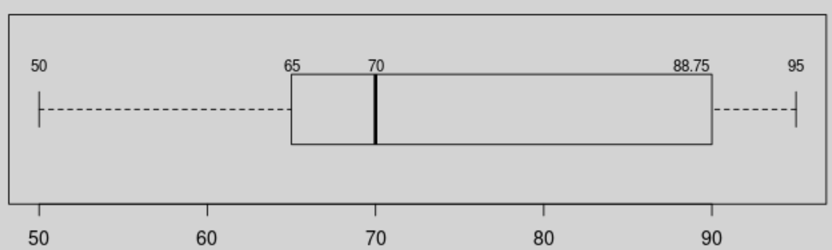
The five-number summary is use used to describe the shape of the distribution of a given numerical data. It consists of five numbers: minimum data value, first quartile, median, the third quartile, and the maximum data value.

The five-number summary of this given data set is:

| stats | value |
|---|---|
| Min. | 50.00 |
| 1st Qu. | 65.00 |
| Median | 70.00 |
| 3rd Qu. | 88.75 |
| Max. | 95.00 |

## 2. Boxplot :

The boxplot is a geometric representation of the five-number summary. The boxplot of the given data set is given below.




The class boundary is: 49.9, 59.9,69.9,79.9,89.9,99.9

| cut.data.freq | Freq | midpts | rel.freq | cum.freq | rel.cum.freq |
|---|---|---|---|---|---|
| [5e+01,6e+01] | 2 | 54.90 | 0.05 | 2 | 0.05 |
| (6e+01,7e+01] | 9 | 64.90 | 0.24 | 11 | 0.29 |
| (7e+01,8e+01] | 11 | 74.90 | 0.29 | 22 | 0.58 |
| (8e+01,9e+01] | 6 | 84.90 | 0.16 | 28 | 0.74 |
| (9e+01,1e+02] | 10 | 94.90 | 0.26 | 38 | 1.00 |