**STATISTICS FOR ENGINEERS**
**ASSIGNMENT VIII**
**ANSWERS**

## Problem 8.1

The main production line at Bakugan Industries was recently altered, with a new machine (the SR-90) introduced that is supposed to produce the watches much more cheaply than before. However, one of your line workers comes to you saying that the number of defective watches has increased since the SR-90 was introduced. For the next two weeks, you count the number of defective watches produced. At the end of those two weeks, you compare the number of defective watches to the two weeks prior to the alteration. The table below gives the number of defective watches per shift. Does the data support the observation of the line worker?

$\vdots$ (data omitted)

*Solution:* First, note that this is a *one-tailed* test (we are asked to test if the defective watches have increased). Second, note that these are independent samples (the same watch faces are not being measured). Finally, note that this will be a t-test.

With that, we need to pull six pieces of information from the table: $n_B$, $\bar{B}$, and $S_B^2$ for before, and $n_A$, $\bar{A}$, and $S_A^2$ for after. As there are definitely two *different* samples, we perform an independent samples t-test for comparing the two means. The statistics are $n_B = 10$, $\bar{B} = 134.2$, $S_B^2 = 483.3$, $n_A = 12$, $\bar{A} = 158.5$, and $S_A^2 = 923.4$. Performing the test (either by hand or using a computer), we find $t = -2.0488$, $\nu = 16.395$, and $p = 0.9716$, which means that the data does support the line worker's contention (at the $\alpha = 0.05$ level). $\diamond$

## Problem 8.2

You are an analyst for an acne drug trial. ... First, do those who received the drug have statistically fewer facial blemishes after taking the drug for two weeks? Next, do those who did not receive the drug have statistically fewer facial blemishes after two weeks? Finally, did those who took the drug statistically improve more than those who did not take the drug?

⋮ (data omitted)

*Solution:* This question asks for two difference of means tests and one independent sample t-test. All three hypotheses are directional, so these will be three one-tailed tests.

The first question looks at the differences among those who did take the drug ($n = 5$, $\bar{D}_1 = -4.8$, $S_1^2 = 5.2$, $t = 4.2099$, $\nu = 7.244$, and $p = 0.001846$).

The second question looks at the differences among those who did *not* take the drug ($n = 5$, $\bar{D}_2 = -0.6$, $S_2^2 = 4.3$, $t = 0.3136$, $\nu = 7.054$, and $p = 0.3814$).

The first test tells us that the participants who took the drug had statistically fewer facial blemishes after two weeks ($D_1 = -4.8$, $p = 0.0018$). The second test tells us that those who did not take the drug did not have statistically fewer facial blemishes after two weeks ($D_2 = -0.6$, $p = 0.38$).

If you actually did the one-sample test on the differences, your answers will be slightly different (due to the test forcing the variances to be equal), but your conclusions will be the same.

The final question asked if the differences between these two samples is statistically significant. To do this, we use an independent samples t-test. Using the appropriate formula (or using our computer), and using the statistics in the previous part of this problem, we get the the difference in reactions between the two groups is statistically significant at the $\alpha = 0.05$ level ($t = -3.047$, $\nu = 7.929$, $p = 0.008031$). ◇

## Problem 8.3

.... Before the workshop, of the 305 children participants, 28.2% did not have an evacuation plan in place, while 36.6% did (the remainder were uncertain). After the program, 22.4% of the children still did not have an evacuation plan in place, while 49.0% did (the remainder were still uncertain). Was the difference in the proportion of children having an evacuation plan in place significantly different from those who did not have an evacuation plan, before the workshop? If 53.4% of the participants were female, and if we assume that having an evacuation plan is independent of gender, how many of the participants were female and had no evacuation plan before the workshop?

*Solution:* The first part is a direct application of the formula for proportions (two-tailed), with $n = 305$, $p_X = 0.282$, and $p_Y = 0.366$. Note that this is a two-tailed test.

$$
\begin{aligned}
z_{p/2} &= \frac{p_X - p_Y}{\sqrt{\frac{p_X(1-p_X)}{n} + \frac{p_Y(1-p_Y)}{n}}} \\
&= \frac{0.282 - 0.366}{\sqrt{\frac{0.282(1-0.282)}{305} + \frac{0.366(1-0.366)}{503}}} \\
&= \frac{-0.084}{\sqrt{0.00142}} = \frac{-0.084}{0.0377} = -2.2255
\end{aligned}
$$

Alternatively, you could have used the tilde version. However, since that is a small-sample correction, it is not needed here. Substituting and solving, we get $z_{p/2} = -2.2255$. Thus, the p-value is $p \approx 2 \times 0.0129 = 0.0258$. As such, we can reject the null hypothesis that the two proportions are statistically different and conclude, at the $\alpha = 0.05$ level, that the two proportions are different.

The second question tests if you remember the effects of independence. The probability of having no plan before the workshop is 0.282. The probability of being a female before the workshop is 0.534. As these two are independent, we know the probability of their intersection is the product of their probabilities, which is 0.1509. Thus, as there are 305 participants, we know that about 46 of the participants are females who did not have an evacuation plan in place before the workshop. $\diamond$

## Problem 8.4

Boxes of nails contain 100 nails each. A sample of 10 boxes is drawn, and each box of nails is weighed. The average weight is 1.500kg, with a standard deviation of 5.000g. Assume that the weight of the box is zero. If we let $\mu_{box}$ denote the mean weight of a box of nails, what is the 95% confidence interval for $\mu_{box}$? What is the 95% confidence interval for the average weight of a nail, $\mu_{nail}$?

*Solution:* This is a straight-forward application of confidence intervals. As $n = 10$, we use the t-distribution with $\nu = 9$.

$$1500 \pm 2.26\sqrt{\frac{25}{10}} = 1500 \pm 3.57$$

Thus, our confidence interval is $\mu_{box} \in (1496.4, 1503.6)$ grams, where the symbol '$\in$' means 'is in'.

The solution for the average nail is quite similar. As the weight of the nails is (assumed) normally distributed, and as the weight of the nail is a linear multiple of the weight of 100 nails, we know the average nail weight is $\mu_{nail} = 15$g, with a standard deviation of $s_{nail} = 0.05$. (Note that the standard deviation behaves nicely. If we had used the variance, we would have to divide the variance by the square of 100.) Thus, using the typical formula, we have the 95% confidence interval for the average weight of a nail as $\mu_{nail} \in (14.964, 15.036)$ grams.

Of course, since the transformation from kg to g is linear, we could have just divided our original confidence interval endpoints by 100 and gotten the same answer.

As we are still just measuring 10 things, $n = 10$, and we must use a t-test. $\diamond$

## Problem 8.5

A random sample of a dozen grades from a recent examination (not yours) are provided in the table below. Find the arithmetic, geometric, and harmonic mean of the sample. Find the variance, standard deviation, and interquartile range of the sample. Find the 95% confidence interval for the population mean. The professor states that the class average for the test was 75%. Does this sample support his assertion?

$\vdots$ (data omitted)

*Solution:* Using our calculator or our computer, we find $AM = 77.17$, $GM = 74.98$, and $HM = 69.99$. We also find $s^2 = 415.61$, $s = 20.39$, and $IQR = 20.5$. The 95% confidence interval for the population mean is

$$95\ \%\ \text{CI for}\ \mu = \quad 77.17 \pm 2.201\sqrt{415.61/12} = 77.17 \pm 12.95$$

Thus, $\mu \in (64.21, 90.12)$. Remember that this is based on the t-test, since there are only 12 scores in our sample.

Finally, if the professor said the class average was 75%, who are we to correct him? Alright, as the asserted value, 75, was within the 95% confidence interval, then the data support his contention (at the $\alpha = 0.05$ level). $\diamond$