ANSWERS CHAPTER 8

THINK IT OVER

think it over

TIO 8.2: The null hypothesis states that the population mean is precisely 20, whereas the alternative states that the population mean can have a range of values, all of which, are greater than 20.

TIO 8.4: If I was callous I would go for the lowest value I could get a way with, say 50%, providing I knew I would not be asked to state the significance level I used. This is why research papers also state the significnance levels used in one form or another.

TIO 8.5: It tells us how the population data, not the sample data, is distributed about the mean of the population which gives an indication of how reliable a model the population mean is

Tio 8.6: One alternative interpretation is to say that on average there are at least 20 sweets in the bag. H_0: $\mu \geq 20$ can always be interpreted to mean 'at least'.

TIO: 8.7: The greater the sample size the smaller the difference between the population mean and the sample mean, since the SE decreases as n increases.

TIO 8.8: A line should be drawn at z=1.645 and assuming a 1 tailed test, the area to the left or right shaded to show the critical region corresponding to an alpha value of 5%.

TIO 8.9: Due to symmetry of the standardised normal curve since the area under a standardised normal distribution is equal to 1. You are interested in the area of the 'tail' which represents the region where you can statistically reject the null hypothesis.

TIO 8.9: we can't hence t-distribution used.

TIO 8.10: the 'boundaries' in some respects are quite arbitrary and using 99% or 0.01% boundary strengthens your case. It is still not definite.

TIO 8.11: The rest of the data is designed to give you some idea as to which sort of statistical testing is appropriate. For example, the kurtosis, skewness will give you some idea as to the shape of the distribution ie. how close it is to a normal distribution for example. Comparing the mean, median and mode will also give you an idea of the shape of the distribution.

TIO 8.12: There is a 82.8% probability that the mean of the population is within the confidence interval. The difference between the SPSS calculated mean and the null hypothesis mean is 0.4. To check this, put in a value of 59.4 as the sample mean (the test value) and you should get the same as in the table below. You will see the probability rise to 100% that the mean of the population falls between the lower and upper bounds.

One-Sample Test

	Test Value = 59.4					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
test result	.000	49	1.000	.00000	-3.6854	3.6854

TIO 8.13: the value of the test statistic goes to 0 since there is no difference between the sample mean and the hypothesised mean. Looking at the formula: $t = \frac{\overline{x} - \mu_0}{s / \sqrt{n}}$ we can see that if $\overline{x}$ is equal to μ_0 (the hypothesised value of the mean, we have 0 in the numerator of the fraction, hence t will equal 0.

EXERCISES

1. (a) eg I need to find out how close the weight of the contents of the packets are to the stated weight
   (b) eg. The stated weight is 50g, customers say it is not 50g
   (c) $H_0{:}\mu = 50$, $H_1{:}\mu \neq 50$
   (d) it is normal or can use the CLT since samples are of sufficient size
   (e) $z = \frac{\overline{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
   (f) 5% is usual but to be confident 1% is better)

2. (a) Take random samples from the production line and calculate their means and standard deviations
   (b) –1.1
   (c) 2 tailed test, z lies between –1.96 and 1.96 therefore accept H_0
   (d) There is no significant evidence, at the 5% level, to suggest that the mean is different from 50g. This does not imply the mean is unchanged, just that from this particular sample, the mean did not fall into the rejection region.

3. (a) 98%: 50±1.165 and at 90% 50±0.823
   (b) The Production Manager doesn't understand hypothesis testing since the statistics will only tell him that if he wants to be 98% confident that the mean of the packets of crisps are 50g then he must expect the weights to vary between 48.835 and 51.165 ie. the higher the value of the confidence, the larger the range of deviations from the mean. It tells him nothing about the actual weights of the packets of crisps.)

4. (a) 2.11: $p = 0.035$
   (b) Since 0.035<0.05 you reject H_0 and conclude there has been a change in the waiting time.
   (c) It doesn't tell you whether waiting times have got longer or shorter only that they have changed.

5. Type 1

6. one way increase α to 0.1 and β to 0.1

7. (a) 271
   (b) Each experiment costs £1000 so therefore 170 experiments can be run. Rearranging $\sqrt{n} = \frac{(z_\alpha + z_\beta)\sigma}{(\mu_0 - \mu_1)}$ to give $\sqrt{170} = \frac{(z_\alpha + z_\beta)2}{0.4}$

$$\left(z_{\alpha}+z_{\beta}\right)=\frac{0.4\times 13}{2}=2.6$$

Therefore must add up to 2.6. At $\alpha=0.1, \beta=0.1$ using the standard normal probability distribution tables to find $z_{0.1}+z_{0.1}=2.56$ which is close to what is required. The consequence of doing this is that the risk of rejecting the null hypothesis when it is true increases to 10% and similarly, accepting the null hypothesis when it is in fact false also increase to 10%.

(c) By increasing α to 0.1means the company is willing to accept evidence to support the null hypothesis at the 10% significance level. This should only affect company sales not patient lives. On the other hand by increasing β to 0.1, the company is saying that they are not too concerned if the drug is effective.