PSET 12 - Due: Sunday, August 4

---

1. The Avenal State Prison in California has thousands of files on former inmates, an estimated 20 million pages of documents. A proposal has been made to archive many of these old files to compact discs. The state will release money for this project only if prison officials present evidence to suggest the true mean age of all files is greater than 10 years.

   (a) Write the null and alternative hypotheses about μ, the mean age of all the files, that the prison officials would want to test.

   (b) For the hypotheses in part (a), describe the Type I and Type II errors in the context of the problem.

   (c) From the warden's perspective, which error is more serious? Why?

   (d) From a state senator's point of view, which error is more serious? Why?

---

a.)

$$H_0 : \ \mu = 10$$
$$H_a : \ \mu > 10.$$

b.)

$$\text{Type I error}: \ \alpha = P(\text{Rejecting } H_0 \text{ when } H_0 \text{ is true})$$
$$= P(\bar{X} > 10 \text{ when } \mu = 10).$$

In other words, it is the probability of the state releasing money for the project based on incorrect evidence that the mean age of the files is greater than 10 years

$$\text{Type II error}: \ \beta = P(\text{Not rejecting } H_0 \text{ when } H_a \text{ is true})$$
$$= P(\bar{X} < 10 \text{ when } \mu > 10).$$

In other words, a type II error is the probability of not releasing money for the project when in reality they should. When incorrect evidence shows that the mean is not greater than 10 when it is.

c.) A type II error is more serious. Not releasing funds when in reality they should is more detrimental to the project than releasing funds when in reality they don't need to. A type II error in this case would prevent the archiving project from being funded and executed.

d.) A type I error would be more serious in this case due to the fact that money would be unnecessarily spent when it doesn't need to be.

2. Let μ denote the true average radioactivity level (in picocuries per liter). The value 5 pCi/L is considered the dividing line between safe and unsafe water.

  (a) (Showing the water is unsafe.) Suppose that we tested $H_0 : \mu = 5$ versus $H_a : \mu > 5$. In the context of the story describe a (i) Type I error and a (ii) Type II error.

  (b) (Showing the water is safe.) Suppose that we tested $H_0 : \mu = 5$ versus $H_a : \mu < 5$. In the context of the story describe a (i) Type I error and a (ii) Type II error.

  (c) Which set of hypotheses would you recommend testing? Explain your answer.

a.)

$$\text{Type I error}: \; \alpha = P(\text{Rejecting } H_0 \text{ when } H_0 \text{ is true})$$
$$= P(\bar{X} > 5 \text{ when } \mu = 5).$$

A Type I error is the probability of false evidence showing that the mean radioactivity level is greater than 5 pCi/L when in reality it is not. Leading to a false positive.

$$\text{Type II error}: \; \alpha = P(\text{Not rejecting } H_0 \text{ when } H_a \text{ is true})$$
$$= P(\bar{X} \leqslant 5 \text{ when } \mu > 5).$$

A Type II error is the probability of false evidence showing that the mean radioactivity level is less than 5 pCi/L when in reality it is greater. Leading to a false negative

b.)

$$\text{Type I error}: \; \alpha = P(\text{Rejecting } H_0 \text{ when } H_0 \text{ is true})$$
$$= P(\bar{X} < 5 \text{ when } \mu = 5).$$

A Type I error is the probability of false evidence showing that the mean radioactivity level is less than 5 pCi/L when in reality it is not. Leading to a false positive.

$$\text{Type II error}: \; \alpha = P(\text{Not rejecting } H_0 \text{ when } H_a \text{ is true})$$
$$= P(\bar{X} \geqslant 5 \text{ when } \mu < 5).$$

A Type II error is the probability of false evidence showing that the mean radioactivity level is greater than 5 pCi/L when in reality it is less. Leading to a false negative

c.) Since we can control the probability of a type I error and keep it small, it would be wise to choose the test in which the type II error leads to a more ideal error. In this case the more ideal error would be claiming the water is unsafe when in reality it is safe, rather than the type II error being the case where we claim the water is safe when it is actually unsafe. Since the first test $H_0 : \mu = 5$ versus $H_a : \mu < 5$ leads to the favorable type II error, this is the test we should preform.

3. Lightbulbs of a certain type are advertised as having an average lifetime of 750 hours. The price of these bulbs is very favorable and so a potential customer has decided to go ahead with the purchase unless it can be conclusively demonstrated that the true average lifetime is smaller than what is advertised. A random sample of 20 bulbs was selected and the lifetime of each was recorded. Suppose that the sample mean was 738.4 hours with a sample standard deviation of 41.2 hours. Does the sample provide evidence that the true mean lifetime is less than 750? Assume lifetimes vary according to a normal distribution and test using $\alpha = 0.10$.

We have $n = 20$, $\bar{x} = 738.4$, $s = 41.2$, and $\alpha = 0.1$. Furthermore

$$H_0 : \ \mu = 750$$
$$H_a : \ \mu < 750.$$

For a left-tailed test with significance level $\alpha = 0.1$, we have $-t_{0.1,19} = -1.328$ Thus, the rejection area is $t < -1.28$.

Given the population distribution is normal, the population standard deviation is unknown, and the sample size is small, we will use the t statistic

$$t = \frac{738.4 - 750}{\frac{41.2}{\sqrt{20}}}$$
$$= -1.2591.$$

Since $t > -t_{0.1,19}$, we do not reject the null hypothesis. There is insufficient evidence to suggest the true mean lifetime of the lightbulbs is less than 750 at the $\alpha = 0.1$ significance level.

4. In an investigation of the toxin produced by a certain poisonous snake, a researcher prepared 26 different vials, each containing 1 gram of the toxin, and then determined the amount of antitoxin needed to neutralize the toxin. The sample average amount of antitoxin necessary was found to be 1.93 mg, and the sample standard deviation was 0.42 mg. Previous research had indicated that the true average neutralizing amount was 1.75 mg/g of toxin.

(a) Does the new data contradict the previous research by showing the true mean differs from 1.75? Test using $\alpha = 0.05$.

(b) Does the validity of the analysis in part (a) depend on any assumption about the population distribution of neutralizing amount? Explain.

a.) With $n = 26$, $\bar{x} = 1.93$, $s = 0.42$, and $\alpha = 0.05$

$$H_0 : \ \mu = 1.75$$
$$H_a : \ \mu \neq 1.75.$$

Since the sample size is small, we will use the t statistic

For a two-tailed test with level of significance $\alpha = 0.05$, we have $t_{\frac{0.05}{2},25} = 2.060$. Thus the area of rejection will be $|t| > 2.060$. Assuming the null hypothesis is true, the test statistic is

$$t = \frac{1.93 - 1.75}{\frac{0.42}{\sqrt{26}}}$$
$$= 2.1853.$$

Since $t > t_{0.05,25}$, we reject the null hypothesis. There is sufficient evidence to conclude the true mean neutralizing amount differs from 1.75 at the $\alpha = 0.05$ significance level.

b.) Yes, since the sample size is small, $n < 30$, we need to assume the underlying population distribution is normal in order to use the t distribution

5. For each of the following find (i) bounds for the p value of the hypothesis test and (ii) compare the p value to the given level of significance and give a decision regarding the test of hypotheses.

(a) $H_0 : \mu = 25$ versus $H_A : \mu > 25$ $T = 1.84$ ($n = 15$) $\alpha = 0.01$

(b) $H_0 : \mu = 100$ versus $H_A : \mu \neq 100$ $T = 2.90$ ($n = 20$) $\alpha = 0.05$

a.) Looking at the t table, for 14 degrees of freedom, we see that the t value 1.84 is between 1.761 and 2.145. Which corresponds to α values 0.05 and 0.025 respectively. Thus

$$0.05 < p < 0.025.$$

Since all values in this interval are greater than the significance level $\alpha = 0.01$, we fail to reject the null hypothesis.

b.) Again using the t table, we find

$$0.005 < 2p < 0.001$$
$$\implies 0.0025 < p < 0.0005.$$

Since all values in this interval are less than $\alpha = 0.05$, we reject the null hypothesis.

6. A manufacturer of nickel-hydrogen batteries randomly selects 100 nickel plates for test cells, cycles them a specified number of times, and determines that 16 of the plates have blistered. Does this provide enough evidence to conclude that more than 10% of all such plates would blister under such circumstances? Test using $\alpha = 0.05$.

For $n = 100$, $\hat{p} = \frac{16}{100} = 0.16$, and $\alpha = 0.05$

$$H_0 : \ p = 0.10$$
$$H_a : \ p > 0.10$$

.

For a right-tailed test with $\alpha = 0.05$, we have $Z_\alpha = Z_{0.05} = 1.645$. Thus, the area of rejection is $Z > 1.645$. Under the assumption that the null hypothesis is true, the test statistic is

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$
$$= \frac{0.16 - 0.1}{\sqrt{\frac{0.1 \cdot 0.9}{100}}}$$
$$= 2.$$

Since $Z > Z_{0.05}$. We reject the null hypothesis, there is sufficient evidence to suggest more than 10% of plates would blister under the circumstances at the $\alpha = 0.05$ significance level.

7. A recent survey indicated that 54% of employed Canadians admit that they have taken a sick day from work when they really were not sick. A random sample of 400 employed Canadians from Quebec was selected and 204 admitted that they faked an illness.

   (a) Is there sufficient evidence to conclude that the proportion of employed Canadians in Quebec who have faked an illness is different from that of the general Canadian population? Test using $\alpha = 0.05$.

   (b) (i) What conditions must be checked so that the test in part (a) is valid? (ii) Check the conditions and indicate whether or not they are satisfied.

a.) We have $n = 400$, $\hat{p} = \frac{204}{400} = 0.51$, and $\alpha = 0.05$

$$H_0 : \ p = 0.54$$
$$H_a : \ p \neq 0.54.$$

For a two tailed-test with $\alpha = 0.05$, we have $Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$. Thus, the rejection region is $|z| > 1.96$. Under the assumption that the claim in the null hypothesis is true, we have the test statistic

$$Z = \frac{0.51 - 0.54}{\sqrt{\frac{0.54(0.46)}{400}}}$$
$$= -1.2039.$$

Since $Z > -1.96$ and $Z < 1.96$, the test statistic does not lie in the rejection region and we do not reject the null hypothesis. There is insufficient evidence to suggest that the proportion of employed Canadians in Quebec who have faked an illness is different from that of the general Canadian population at the $\alpha = 0.05$ significance level.

b.) The conditions that must be checked are

   1. $np_0 \geqslant 10$

   2. $n(1 - p_0) \geqslant 10$

Checking these conditions, we see

   1. $np_0 = 400(0.54) = 216 \geqslant 10$

   2. $n(1 - p_0) = 400(0.46) = 184 \geqslant 10$

Thus, the conditions are satisfied and the test is valid.