

15.075 Exam 3

Instructor: Cynthia Rudin

TA: Dimitrios Bisias

November 22, 2011

Grading is based on demonstration of conceptual understanding, so you need to show all of your work.

Problem 1

A company makes high-definition televisions and does not like to have defective pixels. Historically, the mean number of defective pixels in a TV is 20. An MIT engineer is hired to make better TV's that have fewer defective pixels. After her first week of work she claims that she can significantly improve the current method. To check her claim you try her new method on 100 new televisions. The average number of defective pixels in those 100 TV's is 19.1. Assume that the new method doesn't change the standard deviation of defective pixels, which has always been 4.

1. Test if the new method is significantly better than the old one at the $\alpha = 0.05$ level.
2. Using the new method, assume that the mean number of defective pixels is actually 19. What is the chance that your test from part 1 will conclude that the new method is statistically more effective?
3. How many televisions will you have to check so that the test you did in part 1 will conclude that the new method is effective, with 95% probability? Please assume again that the mean number of defective pixels is actually 19.

Solution

1. The hypotheses are:

$$H_0 : \mu = 20$$

$$H_1 : \mu < 20$$

We perform a z-test. Under the null hypothesis, the z-statistic is approximately standard normal. The observed value of the z-statistic is:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{19.1 - 20}{4/10} = -2.25$$

Since this value is less than $-z_\alpha = -z_{0.05} = -1.645$, the null hypothesis is rejected at the $\alpha = 0.05$ level.

2. The question asks for the power of the test at 19. The rejection region is $z < -1.645$, or $\bar{x} < 20 - 1.645 \frac{4}{\sqrt{100}} = 19.34$. The power of the test is

$$\begin{aligned}\pi(19) &= P(\text{Test rejects } H_0 | \mu = 19) = P(\bar{x} < 19.34 | \mu = 19) \\ &= P\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < \frac{19.34 - 19}{4/\sqrt{100}}\right) = 0.804.\end{aligned}$$

Using Eq.(7.8), we get the same result:

$$\pi(\mu) = \pi(19) = \Phi\left(-z_\alpha + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma}\right) = \Phi\left(-1.645 + \frac{(20 - 19)\sqrt{100}}{4}\right) = 0.804.$$

3. The rejection region for a sample of n tires is $\bar{x} < 20 - 1.645 \frac{4}{\sqrt{n}}$. We want n such that

$$P\left(\bar{x} < 20 - 1.645 \frac{4}{\sqrt{n}} \mid \mu = 19\right) \geq 0.95.$$

Since $P(Z < 1.645) = 0.95$, we have

$$1.645 \leq \frac{20 - 1.645 \frac{4}{\sqrt{n}} - 19}{4/\sqrt{n}} \quad \rightarrow \quad 1.645 \leq \frac{1\sqrt{n}}{4} - 1.645.$$

Then $n \geq \left(\frac{(1.645+1.645) \cdot 4}{1}\right)^2 = 173.19$, so 174 televisions should be tested for the power to be at least 0.95.

Using Eq.(7.10), we get the same result:

$$n \geq \left[\frac{(z_\alpha + z_\beta)\sigma}{\mu_0 - \mu}\right]^2 = \left[\frac{(z_{0.05} + z_{1-0.95})4}{20 - 19}\right]^2 = \left[\frac{(1.645 + 1.645)4}{20 - 19}\right]^2 = 173.19.$$

so again 174 televisions should be tested for the power to be at least 0.95.

Problem 2

Two methods of memorizing words are to be compared. You choose two groups of 5 people, where the first person in the first group has the same characteristics as the first person in the second group (they have the same educational level, age, etc.). Same thing about the second person in each group - they are also similar to each other in terms of education, age, etc. Same thing for the third, fourth and fifth people from each group. The first group is assigned to the first method of memorization and the second group to the other method. The number of words recalled in a memory test after a week's training with these two methods is shown below.

	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5
Method 1	25	30	22	27	29
Method 2	21	20	23	18	17

Test the hypothesis that the first method is better than the second method at the 0.05 level. You may assume normality of the data.

Solution

We will perform a paired t-test. The hypotheses are:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

We have: $d_1 = 25 - 21 = 4$, $d_2 = 30 - 20 = 10$, $d_3 = 22 - 23 = -1$, $d_4 = 27 - 18 = 9$, $d_5 = 29 - 17 = 12$.
It is: $\bar{d} = 6.8$ and

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d(i) - \bar{d})^2}{n - 1}} = 5.26$$

Under the null hypothesis, the t-statistic

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

has a t_4 distribution. Since

$$t = \frac{6.8}{5.26 / \sqrt{5}} = 2.89 > t_{4,0.05} = 2.132$$

we conclude that the first method is significantly better than the second method.

Problem 3

After your MIT graduation you decide to go to Monte Carlo for vacation. You visit a casino and decide to gamble; specifically, you want to play roulette. Before you start betting, you watch 500 roulette games at the casino, and you find that red is hit 260 times. Can you determine whether the roulette is fair? In case you don't know how roulette is played in Monte Carlo, it involves spinning a wheel that has many slots around it. There is 1 gold slot, 18 red slots, and 18 black slots.

1. Set up the hypotheses.
2. Calculate the P-value. Can you reject null hypothesis at the 0.05 level?
3. Find a 95% CI for the proportion of red results.

Solution

1. We let p be the probability of a red slot on a single spin, which is $\frac{18}{37} = 0.4865$ if the roulette is fair. The alternative hypothesis for this test is two-sided because we just want to check if p is different from $\frac{18}{37}$. The hypotheses are:

$$H_0 : p = p_0 = 0.4865$$

$$H_1 : p \neq p_0 = 0.4865$$

2. Since $n = 500$, we do a large-sample test. We are given the total number of red hits is $Y = 260$, so the sample proportion is

$$\hat{p} = \frac{Y}{n} = \frac{260}{500} = 0.52.$$

Under the null hypothesis, \hat{p} is approximately normally distributed with mean $p_0 = \frac{18}{37} = 0.4865$ and variance $\frac{p_0 q_0}{n} = \frac{0.4865 \cdot 0.5135}{500} = 0.000499$. Thus the following test statistic is approximately standard normal:

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}.$$

The observed value of the test statistic is

$$z = \frac{0.52 - 0.4865}{\sqrt{\frac{0.4865 \cdot 0.5135}{500}}} = \frac{0.0335}{0.0224} = 1.50.$$

The p -value is $2 \cdot P(Z > 1.5) = 0.134$. So no, we cannot reject H_0 at the 0.05 level because $0.134 > 0.05$.

3. To compute a confidence interval, we estimate the variance of \hat{p} by $\frac{\hat{p}\hat{q}}{n}$. A 95% CI for the proportion of red hits is

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} = 0.52 \pm 1.96 \sqrt{\frac{0.52 \cdot 0.48}{500}} = [0.4762, 0.5638].$$

Notice that 0.4865 is in this interval, so we do not reject H_0 .

Problem 4

The weight of an object is measured using an electronic scale that reports the true weight plus a random fluctuation that is normally distributed with zero mean. The manufacturing company of the electronic scale claims that the standard deviation of the fluctuation is 2 milligrams. Assume that the fluctuations are independent. The company measures the weight of one object 8 times, and observes these values:

100.9, 98.2, 101.5, 102.2, 105.1, 99.4, 93.6, 97.5

1. You believe that the standard deviation of the fluctuation is greater than what the company claims. Is there statistically significant evidence for your belief at $\alpha = 5\%$ level?
2. Compute a 95% confidence interval for the variance of the fluctuation.

Solution

1. We perform a chi-square test with hypotheses:

$$H_0 : \sigma = 2$$

$$H_1 : \sigma > 2$$

When H_0 is true, the statistic

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

has a χ^2_7 distribution. The observed value of the chi-square statistic is:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = 21.1$$

The P-value is $P(\chi^2_7 > 21.1) = 0.0036$ and therefore there is statistically significant evidence for our belief at $\alpha = 5\%$ level.

2. A 95% two-sided confidence interval for the variance of the fluctuation is given by:

$$\left[\frac{(n-1)s^2}{\chi^2_{7,0.025}}, \frac{(n-1)s^2}{\chi^2_{7,0.975}} \right] = [5.27, 49.94]$$

Problem 5

You are an investment manager of an asset management company and you want to evaluate the performance of two hedge funds. For this you record their past performance for the last 10 years. The returns for both of the hedge funds in different years are independent and normally distributed with unknown mean and variance. The mean return for Hedge Fund 1, averaged over the 10 years, is $\bar{X} = 22.15\%$ and the standard deviation calculated from those 10 measurements is 2.79%. For Hedge Fund 2 the mean return, averaged over the 10 years, is $\bar{Y} = 20.6\%$ and the standard deviation calculated from those 10 measurements is 2.46%.

1. Test equality of the variances for the two hedge funds using $\alpha = 0.1$.
2. Test for the equality of the means of the two hedge funds' performances, using $\alpha = 0.05$ with a suitable test. Would you recommend using a t-test calculated from a pooled variance or the method for unequal variances? If you rejected the test in (1) then you may assume the variances are unequal, otherwise you may assume that the variances are equal.

Solution

1. We perform an F test with hypotheses:

$$\begin{aligned}H_0 : \sigma_1^2 &= \sigma_2^2 \\H_1 : \sigma_1^2 &\neq \sigma_2^2\end{aligned}$$

The F -statistic is

$$F = \frac{s_1^2}{s_2^2} = 1.29$$

Under H_0 this statistic has $F_{9,9}$ distribution. From the table in the book we see that this is less than the critical value 3.18, therefore we do not reject the null hypothesis.

2. The hypotheses are:

$$\begin{aligned}H_0 : \mu_1 &= \mu_2 \\H_1 : \mu_1 &\neq \mu_2\end{aligned}$$

Since we do not reject the hypothesis that $\sigma_1^2 = \sigma_2^2$, we compute the pooled variance.

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{9 \cdot 2.79^2 + 9 \cdot 2.46^2}{9 + 9} = 6.92.$$

Under the null hypothesis that the mean returns of the two hedge funds are equal, the t-statistic:

$$t = \frac{\bar{X} - \bar{Y}}{S \cdot \sqrt{\frac{1}{10} + \frac{1}{10}}}$$

has a $t_{10+10-2} = t_{18}$ distribution. Its observed value is:

$$t = \frac{\bar{x} - \bar{y}}{s \cdot \sqrt{\frac{1}{10} + \frac{1}{10}}} = 1.33$$

Since $|t|$ is less than $t_{18,0.025} = 2.1$, we do not reject the null hypothesis at $\alpha = 5\%$ level.

MIT OpenCourseWare
<http://ocw.mit.edu>

15.075J / ESD.07J Statistical Thinking and Data Analysis
Fall 2011

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.