STAT2912: Statistical Tests

Semester 2	Statistical Tests (Advanced)	2011

Tutorial 2

1. The hourly wages in a particular industry are normally distributed with mean \$21.6 and standard deviation \$2.5. A company in this industry employs 40 workers, paying them an average of \$20.3 per hour. Can this company be accused of paying substandard wages?

   (a) State the null hypothesis and alternative hypothesis.

   (b) Calculate an appropriate test statistic, the p-value and then state your conclusion.

   (c) What is the rejection region based on the average hourly wage for an $\alpha = 0.01$ level test?

2. Let $X_1, X_2, \ldots, X_{25}$ be a sample from a normal distribution with unknown mean μ and known variance $\sigma^2 = 25$. We wish to test the hypothesis $H_0 : \mu = 7$ vs $H_1 : \mu \neq 7$.

   (a) Calculate the rejection region based on $\bar{X}$ for the z-test at level $\alpha = 0.05$.

   (b) For the z-test in (a), find the power at each of the following alternative values for μ:

   $$\mu_a = 5,\ 6,\ 6.5,\ 7,\ 7.5,\ 8,\ 9.$$

   (c) Sketch a graph of the power function by using the calculations in (b).

3. Suppose we have 25 observations from an exponential distribution with density function:

   $$f(x) = \frac{1}{\beta} e^{-x/\beta}, \text{ if } x \geq 0; \quad 0, \text{ otherwise.}$$

   (Check that $\mathrm{E}X = \beta$ and $\mathrm{Var}(X) = \beta^2$.)

   A possible test of $H_0 : \beta = 1$ vs $H_1 : \beta > 1$ is

   Reject H_0 if $\bar{x} \geq c$.

   Find the approximate value of c so that the test has significance level $\alpha = 0.05$ (Hint: use the central limit theorem in Lecture 1.2).

4. Do the exercise at the end of Lecture 2.2.

Computer Exercises week 2

1. The survival times in days of 72 guinea pigs after they were injected with tubercle bacilli are stored in the vector object `survival`. (Data from Bjerkedal, T. (1960), *Amer. J. Hygiene.*)

   (a) Set up the screen to take a 2×2 array of plots by using `par(mfrow=c(2,2))`.

   (b) Rename the dataset as `su`.

   (c) Obtain a box-plot representation of the data `su`. Is the data set skewed? Too skewed? Obtain a normal qq-plot for this data. What does this plot indicate?

   (d) Calculate the mean and variance of the sample, and then apply them to test the hypothesis that

   $$H_0: \ \mu = 160 \quad \text{vs} \quad H_1 : \mu < 160, \tag{1}$$

   where μ is the mean survival time. Do you need to assume the survival time is normal to get a good approximation for the p-value?

   (e) Use `t.test` to test the hypothesis (1). What assumption is involved in using `t.test`?

   (f) Compare the p-value in (d) with that in (e). What is your conclusion?

   (g) Find the critical value of the large sample test in (d) at the significance level $\alpha = 0.05$. Is the large sample test significant at the level $\alpha = 0.05$ (using your calculations in (d))?

   (h) Create a new data.frame called `su1`, which is same as `su` except the fourth observation is 2000.

   (i) Repeat questions (c) and (e) by using data `su1`. Comment the difference.