Stat 20: Discussion at section

B. M. Bolstad, bolstad@stat.berkeley.edu

Oct 29, 2003

One sample normal model

In the one sample normal model we have a sample $Y_1, \ldots, Y_n$ of independent random variables from a common normal distribution with mean μ and standard deviation σ (both parameters are unknown). We estimate μ using the sample mean

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$$

and the standard deviation σ using the sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_i - \bar{Y}\right)}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2}{n-1}}$$

Hypothesis tests

We typically carry out hypothesis tests about the mean in particular the tests are

$$H_0 : \mu = \mu_0 \text{ vs } H_A : \mu \neq \mu_0$$

$$H_0 : \mu \leq \mu_0 \text{ vs } H_A : \mu > \mu_0$$

$$H_0 : \mu \geq \mu_0 \text{ vs } H_A : \mu < \mu_0$$

where μ_0 is a known constant value. Assuming the null hypothesis is true then the t-statistic

$$t = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}}$$

has t distribution with $n-1$ degrees of freedom. Note that denominator is the standard error of $\bar{Y}$ ie SE $\left(\bar{Y}\right) = \frac{s}{\sqrt{n}}$

A 100C% confidence interval for μ is given by

$$\bar{Y} \pm t^* \frac{s}{\sqrt{n}}$$

where t^* is given by the value satisfying

$$P\left(-t^* < T < t^*\right) = C$$

where T has t-distribution with $n-1$ degrees of freedom.

Question 1

A random sample of 10 one bedroom apartments (for a particular community) has the following monthly rents

$$500, 650, 600, 505, 450, 550, 515, 495, 650, 395$$

Does this data give good evidence that the mean rent is above \$500? Carry out an appropriate test and then give a 95% confidence interval.

Answer:

Let μ be the mean monthly rent for the community. The appropriate hypothesis test will be

$$H_0 : \mu \leq 500 \text{ vs } H_A : \mu > 500$$

First lets compute the summary statistics (sample mean, standard deviation)

$$\sum_{i=1}^{10} Y_i = 500 + 650 + \cdots + 395 = 5310$$

$$\sum_{i=1}^{10} Y_i^2 = 500^2 + 650^2 + \cdots + 395^2 = 2881300$$

Therefore

$$\bar{Y} = \frac{\sum_{i=1}^{10} Y_i}{n} = \frac{5310}{10} = 531$$

and

$$s = \sqrt{\frac{\sum_{i=1}^{10} Y_i^2 - n\bar{Y}^2}{n-1}} = \sqrt{\frac{2881300 - 10(531)^2}{10-1}} = 82.79$$

Now we compute the t statistic.

$$t = \frac{531 - 500}{82.79/\sqrt{10}} = 1.184$$

note that the t-statistic has the t distribution with df 9 (under the null hypothesis).

We may put bounds on the P-value for this test statistic using the t-table. In particular notice that

$$1.100 < 1.184 < 1.383$$

and so

$$0.15 > P(T > 1.184) > 0.10$$

Thus we can say that $0.10 <$ P-value < 0.15 and so we cannot reject the null hypothesis. That is we have no information to conclude that the mean is really above 500.

A 95% confidence interval will be given by

$$531 \pm 2.262 \left(\frac{82.79}{\sqrt{10}} \right)$$

and so the 95% confidence interval is $(471.78, 590.22)$. Again we would note that 500 is contained in the interval.