ISTA 370: Homework 6

Spring, 2015
Due Sunday, December 6th 3rd, 11:59 PM
50 points total

This week's homework is due before midnight on December 6th. It is about testing hypotheses. If you are not clear on how hypothesis testing works, read pp. 105 - 117 of the textbook. In any case, read the slides from the last two lecture ("Testing Hypotheses about Means") on two-sample tests and read Sections 4.4 and 4.5, that is, pp. 117 - 132.

Question 1 Don't use R for this question. Work it out with pencil and paper and possibly a calculator (or even in your head!) – but not using R's built-in functions; the goal here is to understand the components of the Z test calculation. The formula for a two-sample Z test is:

$$Z = \frac{\left((\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)\right)}{\sqrt{\left(\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}\right)}}$$

The $N(0,1)$ quantiles $\alpha = 0.05, 0.025, 0.01$ are $\pm 1.645, \pm 1.96, \pm 2.33$, respectively.

(a) (10 points) The the average height of athletes on the track team is believed to be the same as the average heights of athletes on the football team. The standard deviation of heights of each team is 6 inches. The track and football teams each have a total of 36 members, and the difference in the average heights is 2 inches.

    i. What is H_0?

    ii. Is this a one-tailed or two-tailed test?

    iii. What is the critical value for a two-tailed test at $\alpha = 0.05$?

    iv. What is the critical value for a one-tailed test at $\alpha = 0.025$?

    v. What is the standard error of $\mu_1 - \mu_2$?

    vi. Can you reject H_0 at any of the α levels?

(b) (10 points) Drug A is thought to reduce blood pressure by five points more than Drug B; the standard deviation associated with each drug is 2.0. In a two-sample test, with $N_A = N_B = 100$, Drug A reduced blood pressure by 27 points and Drug B reduced blood pressure by 21 points.

    i. What is H_0?

    ii. Is this a one-tailed or two-tailed test?

    iii. What is the standard error of $\mu_1 - \mu_2$?

    iv. What is the value of Z?

    v. Can you reject H_0?

Question 2 (15 points) Load the Longley dataset available in `R`. This is in the `datasets` package. It shows relationships between economic variables. Find the correlation (for this you can use `R`'s `cor` function) between GNP and the number of people unemployed, and test whether it is significantly different from zero using Fisher's r to Z test (discussed on pages 130-132 in the text). Is this a one- or two-tailed test? What is the standard error? What is Z? Can you reject H_0 at the $\alpha = 0.05$ level?

Question 3 (15 points total) On pages 125-130 of the book the t test is described (along with its two-sample and paired-sample forms). We use the t test instead of the z test when we don't know the population standard deviation, σ, and have to estimate it ($s = \frac{1}{N-1}\sum_i (x_i - \mu)^2$). Consider two small samples of test scores:

$$A = 17, 29, 14, 21, 29, 4, 10$$
$$B = 19, 31, 18, 27, 28, 6, 11$$

(a) (3 points) Figure out how to run a two-sample t test on these data in `R`.

(b) (3 points) Figure out how to run a paired sample t test on these data in `R`.

(c) (6 points) Explain the difference in results (i.e., what the two tests are doing differently and why this leads to different results).

(d) (3 point) Explain why the first test has more degrees of freedom.