

STAT W4201 001, Homework 3

Brian Weinstein (bmw2148)

Feb 17, 2016

Code is attached here and also posted at <https://github.com/BrianWeinstein/advanced-data-analysis>. Where relevant, code snippets and output are included in-line.

Problem 1: Ramsey 4.30

Problem 2: Ramsey 4.32

Problem 3: Ramsey 5.19

Problem 4:

Consider the Bumpuss data in Chapter 2, compute the power of the two-sided two sample t-test of size 0.05 (i.e., reject the null hypothesis if the absolute value the t-statistic is greater than or equal to 2), under the alternative that $\mu_x - \mu_y = \bar{x} - \bar{y} = 0.01$ and $\sigma = s_p = 0.0214$.

Problem 5:

Show that the two-sided two sample t-test is equivalent to the anova F-test, if the number of groups is two.

For $I = 2$ groups, the F-statistic is given by

$$\text{F-statistic} = \frac{SS_B / [(n-1) - (n-I)]}{SS_W / (n-I)},$$

where n_1 and n_2 are the sizes of samples 1 and 2, respectively, $n = n_1 + n_2$ is the total sample size, SS_B is the “between groups” sum of squared residuals, and SS_W is the “within groups” sum of squared residuals.

Simplifying, we find

$$\text{F-statistic} = \frac{SS_B / (I-1)}{SS_W / (n-I)} = \frac{SS_B / (2-1)}{SS_W / (n-2)} = \frac{SS_B / 1}{SS_W / (n-2)}.$$

If the observations from group 1 are $\sim N(\mu_1, \sigma^2)$ and the observations from group 2 are $\sim N(\mu_2, \sigma^2)$, we know that

$$\text{F-statistic} \sim F_{1, n-2}, \text{ which is equivalent to } t_{n-2}^2.$$

i.e., an F distribution with a numerator degrees of freedom of 1 and a denominator degrees of freedom of $n-2$ is equivalent to the square of a t distribution with $n-2$ degrees of freedom.

Finish
problem
5

Problem 6:

Consider X_1, \dots, X_{10} are i.i.d. $N(0, \sigma^2)$, Y_1, \dots, Y_{10} are i.i.d. $N(\mu, \sigma^2)$ and hypothesis testing:

$$H_0 : \mu = 0$$

$$H_A : \mu \neq 0.$$

Compute the power of a two sided two sample t -test of size 0.05 when $\sigma^2 = 1$ and $\mu = 0.1, 0.5, 1$, and 2. Plot the power as a function of μ . Then, increase the sample size in each group to 20 and draw the power function in the same plot as that of the sample size 10.

Problem 7:

Under the setting of the previous problem, show that, under the null hypothesis, the p -value follows the uniform distribution on the interval $[0, 1]$ and perform simulations to confirm it.

Todo list

Finish problem 5 1