

Introduction to Computational Statistics INSH 5301

Homework 06

02/10/2020

PLEASE copy and paste the whole question number and text into submission so I can grade easily.

When I grade easier, you might get better grade!

For all problems, please show all your work, using R or latex where appropriate.

	18-29	30-44	45-59	60+
Democrat	86	72	73	71
Independent	52	51	55	54
Republican	61	74	70	73

1. You conduct an exit poll after an election, with the shown above:

1.a. Based on the exit poll results, is age independent of Party ID or not? Conduct a chi-squared test by hand, showing each step in readably-formatted latex.

1.b. Verify your results using R to conduct the test.

2.a. Now test for independence using ANOVA (an F test). Your three groups are Democrats, Independents, and Republicans. The average age for a Democrat is 43.3, for an Independent it's 44.6, and for a Republican it's 45.1. The standard deviations of each are D: 9.1, I: 9.2, R: 9.2. The overall mean age is 44.2. Do the F test by hand, again showing each step.

2.b. Check your results in R using simulated data. Generate a simulated dataset by creating three vectors: Democrats, Republicans, and Independents. Each vector should be a list of ages, each with a length equal to the number of Democrats, Independents, and Republicans in the table above, and the appropriate mean and sd based on 2.a (use `rnorm` to generate the vectors). Combine all three into a single dataframe with two variables: age, and a factor that specifies D, I, or R. Then conduct an F test using R's `aov` function on that data and compare the results to 2a. Do your results match 2a? If not, why not?