

Unit 8 – Stata for Categorical Data Analysis Homework

Due Wednesday November 11, 2020

Before you Begin
Download from the course website
[**hw8_categorical.dta**](#)

Description of [**hw8_categorical.dta**](#)

This data set contains 1000 observations on 3 variables. It is data from a study of the relationship of coffee consumption and heart attack. There is one covariate, cigarette consumption. Before you begin, familiarize yourself with the data.

Tip - In developing your answers to these questions, be sure to use the **help** command to learn more about the commands that are being introduced and the nifty options that each offers!

- ___ 1. Produce a simple tabulation of the distribution of cigarettes per day (**cigs**). Do this 3 ways, so that you can decide which command you like best!
 - a) Using the command **tab1**
 - b) Using the command **tabulate**
 - c) Using the command **fre**

- ___ 2. Produce a cross-tabulation of coffee consumption (**coffee**) as the row variable and heart attack (**micase**) as the column variable using the command **tab2** . Do this 2 ways
 - a) with row percentages
 - b) with column percentages

- ___ 3. Produce a table showing the proportion in each category of cigarette smoking (**cigs**), together with 95% confidence limits using the command **proportion**.

- ___ 4. Produce a Fisher Exact test of the association of high coffee consumption (**coffee**) with heart attack (**micase**).
 Do this 3 ways, so that you can decide which command you like best!
 - a) Using the command **tab2** with option **exact**
 - b) Using the command **cs**
 - c) Using the command **cc**