

# MSDS 6372: Homework 1

Ben Goodwin

1/12/2021

#Question 1, Chapter 23 Problem 18

“Does consuming nine small snack-sized meals each day, rather than consuming three meals reduce LDL cholesterol levels?” We are interested if nibbling can result in a drop of 6% or more in the mean cholesterol level.

Steps to design a study:

- 1) State the objective:

We are interested in measuring mean difference in LDL cholesterol levels between those who nibble and those who don't.

- 2) Determine the scope of inference: Will this be a randomized experiment or an observational study?  
This will be a randomized experiment with groups random chosen to nibble and to eat regular meals

What experimental or sampling units will be used? The sampling units will be cholesterol levels

What are the populations of interest? The populations of interest are groups of people who either nibble or eat three meals per day.

- 3) Understand the system under study: The system under study are mean LDL cholesterol levels.
- 4) Decide how to measure a response: We are concerned with a difference in mean LDL cholesterol level.
- 5) List factors that can affect the response: Design factors:

Factors to vary (treatments and controls) Nibbling or three meal diet

Factors to fix

Confounding Factors: Factors to control by design (blocking) Lifestyle

Factors to control by analysis (covariates) Groups with entirely different eating patterns Actual eating patterns

Factors to control by randomization Grouping

- 6) Plan the conduct of the experiment (time line): This study should be able to be accomplished in three months, this is an adequate amount of time for LDL levels to change in the general population based solely on diet.
- 7) Outline the statistical analysis: The responses will be in terms of LDL levels. We have both before and after treatment measurements for both groups. We will use a 2 sample paired t-test to conduct the experiment.
- 8) Determine sample size:

```

PROC POWER;
ONECORR SIDES = 1
ALPHA = 0.05
CORR = 0.5
NTOTAL = .
POWER = 0.8;
RUN;

```

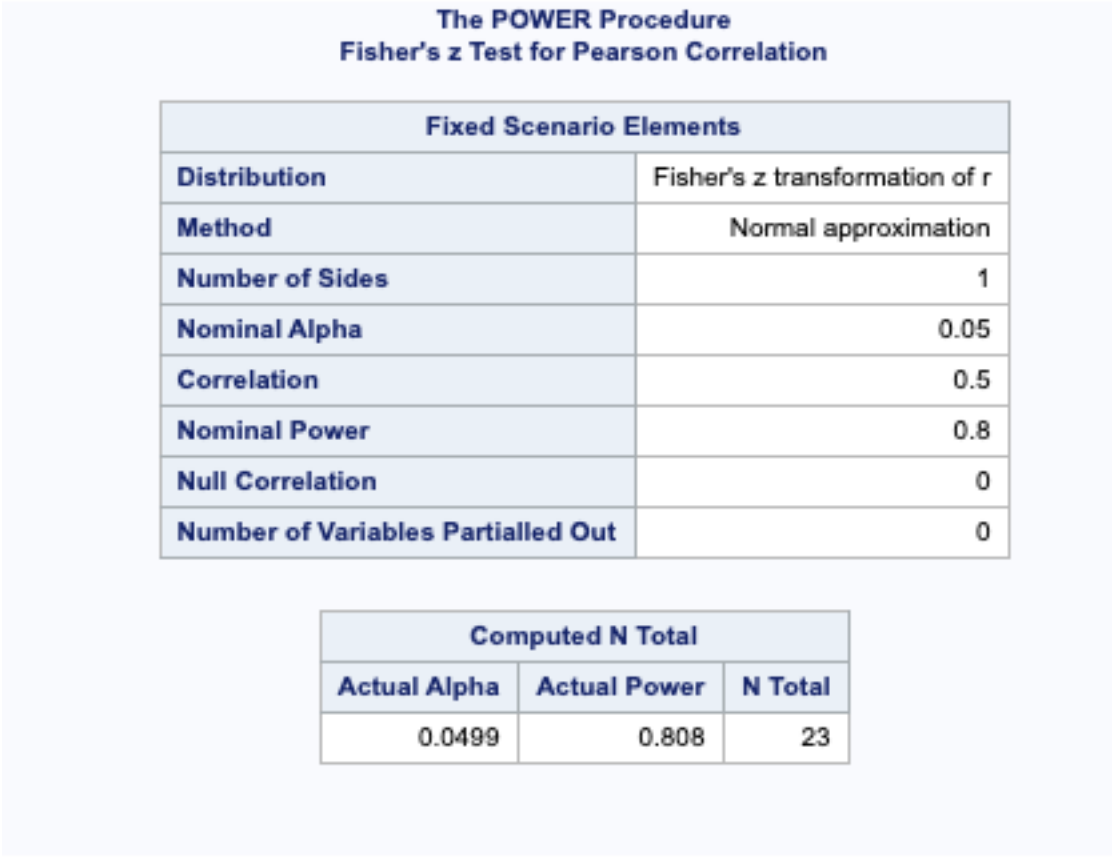


Figure 1: unchanged image

Formula from 709

$$n = \frac{[t_{df}(1-\alpha/2)]^2 S_e^2}{(Practically\ significant\ difference)^2} (C_1^2 + C_2^2 + \dots + C_k^2) = 23$$

##Question 2

Search

**Ben Goodwin**

Competitions  
Novice

Joined a month ago · last seen in the past day

Home Competitions (1) Datasets Notebooks Discussion Organizations Followers Notifications ... **Edit Profile**

Account (User ID 6337431)

User Name  
**bensharn**  
 Your username cannot be changed.

Display Name  
**Ben Goodwin**  
 The name users will see publicly. We recommend using your full name.

Email Address  
**bgoodwin@smu.edu**

Figure 2: unchanged image

198...
 

Ben Goodwin

0.66267

1

~10s

Your First Entry

Welcome to the leaderboard!

Your score represents your submission's accuracy. For example, a score of 0.7 in this competition indicates you predicted Titanic survival correctly for 70% of people.

What next? You've got a few options:
 

- 🧠 Learn skills that can improve your score in [our Intro to Machine Learning course by Dan Becker](#).
- 🔍 Check out [the discussion forum](#) to find lots of tutorials and insights from other competitors.
- 🏆 Find a new challenge by entering one of our [open, active competitions](#) or searching our [public datasets](#).

198...	JunkeyMonkeyBaby		0.66028	1	2mo
198...	landrunner		0.66028	1	2mo
198...	Kyohei Higashi		0.66028	1	2mo

Figure 3: unchanged image