

1. (18 pts) A study in February 2008 by the Pew Forum on Religion and Public Life showed that of 997 Virginians sampled, 31% identified themselves as Evangelical Protestants.¹ Nationally the figure is 26%.
 - (a) (12 pts) Test at the 1% level of significance the hypothesis that the proportion of Evangelical Protestants is higher in Virginia than it is nationally.
 - (b) (6 pts) Find a 95% confidence interval estimating the proportion of Virginians who are Evangelical Protestants.
2. (18 pts) A recent study showed a link between the amount of abdominal fat one has when he is in his 40's and whether one suffers from dementia when he is in his 70's. A report in Newsweek Magazine states the following.

The observational study followed some 6,500 people, ages 40 to 45, for more than three decades, starting in the mid-1970s. All participants were given a sagittal abdominal diameter (SAD) measurement, a simple procedure in which a technician uses calipers to measure the distance from the back to the upper abdomen, midway between the upper pelvis and bottom of the ribs. At the end of the study about 1,000 participants, or 16 percent, were diagnosed with dementia. After factoring out known dementia risk factors such as diabetes, stroke, hypertension and high cholesterol, data analysis showed that potbellies were a bigger risk for cognitive impairment than obesity alone.²

The report also states that “Obese participants with large guts were 3.6 times as likely to develop dementia.”

- (a) (2 pts) Clearly this is an observational study. (It says so.) Identify an explanatory variable and the response variable in the study.
- (b) (2 pts) Identify three potential confounding variables for which the researchers controlled.
- (c) (12 pts) Based on the numbers given in the article, suppose we compare a sample of 100 people of normal weight with a sample of 100 people who are obese and have high SAD measurements and find that 30 years later 7 of the normal-weight people suffer from dementia and 25 of the obese, high-SAD people suffer from dementia. Test at the 1% level whether there is a difference in the rates of dementia between the two populations.
- (d) (2 pts) What is the value of the pooled estimate for p ?

¹<http://religions.pewforum.org/maps>

²<http://www.newsweek.com/id/129147>

3. (12 pts) The IRS estimates that the average American spent 26.5 hours filing his taxes for the 2006 tax year.³ The estimate for the 2003 tax year was 25.4 hours. Suppose a sample of 20 taxpayers is taken and is observed to spend an average of 26.5 hours on filing their 2006 taxes, with a standard deviation of 5.6 hours. Test the hypothesis at the 5% level of significance that the average amount of time spent by all taxpayers for 2006 is greater than 25.4 hours. Assume that the distribution of time spent on taxes in 2006 is approximately normally distributed.
4. (22 pts) In the previous problem, suppose that the sample size was 200 instead of 20 and suppose further that the 25.4 average for 2003 was itself based on a sample of 300 taxpayers, and showed a standard deviation of 5.2 hours.
- (12 pts) Using the data from both samples, test at the 5% level of significance the hypothesis that the average time spent filing 2006 taxes was greater than the time spent filing 2003 taxes.
 - (2 pts) What is the value of the pooled estimate for σ ?
 - (6 pts) Estimate the change in the average number of hours spent filing tax returns from 2003 to 2006. Use a 90% confidence interval.
 - (2 pts) What is the value of the margin of error in the confidence interval in part (c)?
5. (4 pts) In the confidence interval of the previous problem,
- If the level of confidence were increased (and all other things were equal), would that increase or decrease the margin of error of the estimate?
 - If the sample size were increased (and all other things were equal), would that increase or decrease the margin of error of the estimate?
6. (12 pts) No doubt you recall the experiment that I did with M&Ms candy on the first day of class. The actual counts from that bag of M&Ms matched pretty closely the expected counts, based on the Mars website.⁴ Ok, so they were telling the truth about milk chocolate M&Ms. But what about the dark chocolate M&Ms? Are they telling the truth about them? Their website says that the color distribution of the dark chocolate M&Ms is

Red	17%
Blue	17%
Yellow	17%
Brown	17%
Orange	16%
Green	16%

³<http://www.cnn.com/2008/LIVING/wayoflife/04/14/tax.burden.ap/index.html>

⁴<http://www.mms.com/us/about/products/>

I just now bought two packages of dark chocolate M&Ms and counted the colors. I found the following:

Red	9
Blue	5
Yellow	27
Brown	23
Orange	32
Green	13

Test the hypothesis at the 1% level of significance that the distribution that I observed for the dark chocolate M&Ms is consistent with the one advertised by Mars on their website.

7. (14 pts) A study was conducted among males age 25 - 29 to determine whether there is a difference among ethnic groups in the percentage who are married. The survey randomly selected 100 White males, 100 Black males, 100 Hispanic males, and 100 Asian males. The following table shows the results.

	White	Black	Hispanic	Asian
Number married	54	38	32	55
Number not married	46	62	68	45

- (a) (12 pts) Test the hypothesis at the 1% level of significance that there is a difference in marriage rates among the different ethnic groups for males age 25 - 29.
- (b) (2 pts) In the table above, fill in the expected counts.