

Due in Tutorial Week of 28 Sep 2010

- 1) [Midterm, Spring 2009] If a random variable X has a normal distribution with mean μ and variance σ^2 then:
 - a) X takes positive values only
 - b) $(X - \mu) / \sigma$ has a standard normal distribution
 - c) $(X - \mu)^2 / \sigma^2$ has a chi-squared distribution with n degrees of freedom
 - d) $(X - \mu) / (s / \sqrt{n})$ has a t distribution with $n-1$ degrees of freedom
 - e) none of the above

- 2) [Midterm, Spring 2009] Suppose you have a random sample of 10 observations from a normal distribution with mean 8 and variance 3. The sample mean (\bar{x}) is 10 and the sample variance is 2. What is the mean and variance of the sampling distribution of \bar{x} ?

- 3) Make a table that shows the pdf and cdf for the sum of 3 coin flips, where coins have a zero on one side and a one on the other side.

- 4) Statistical Snappers:
 - a) Why is a chi-square variable never negative?
 - b) What proportion of the time is a standard normal variable between 0 and 1?
 - c) What proportion of the time is a standard normal variable bigger than 2?

- 5) Suppose you have a random sample of n observations, denoted X_1, X_2, \dots, X_n . Suppose that the population from which the sample is drawn has mean μ and variance σ^2 . As usual, you do **not** know the value of μ or σ^2 .
 - a) Suppose you calculate the sample mean \bar{X} . What is the relationship between \bar{X} and μ ? Be precise.
 - b) What does the sampling variance of \bar{X} measure? Provide an unbiased estimator of it, and **show** that it is unbiased.
 - c) Now suppose that the quantity you **really** want to estimate is the square of the population mean: μ^2 .
 - i) Is $(\bar{X})^2$ an unbiased estimator of μ^2 ? If yes, show it. If not, give an alternative estimator of μ^2 that **is** unbiased (and be sure to **show** that it's unbiased!).
 - ii) Now find another unbiased estimator of μ^2 (yes, there's more than one). Be sure to **show** that your new estimator is unbiased).
 - iii) You now have 2 unbiased estimators of μ^2 . Which is more efficient? Show it, and sketch the sampling distributions of your two estimators.

Part 3: EViews Exercises

The following questions use the EViews workfile pumf2006_greatervancouver.wf1 on posted on the course website. These are the data that I used for your first lecture. The census long form, and the data codebook are also posted. You will need all 3 files to do the assignment.

For each question, provide a short written answer and attach a hardcopy of the EViews output you used to answer the question. When you've finished the assignment, be sure to save a copy of your workfile! We'll use these data again on future assignments.

- 6) Give the sample mean, minimum, maximum, and sample standard deviation of the WAGES variable (this variable measures the respondent's total annual income from wages and salaries in 2005).
 - a) Why are there so many zero's? What does "9999999" mean?
 - b) What does a zero mean? What does "9999999" mean?
 - c) Give the sample mean, minimum, maximum, and sample standard deviation of the WAGES variable excluding all the "9999999"s. Why is it so different? Are there other values of WAGES that should be excluded?
- 7) Create a new variable called WAGES_MISSING that equals 1 if the WAGES variable is missing, and equals 0 otherwise. The sample mean of WAGES_MISSING measures the *proportion* of people for whom the WAGES variable is missing. Report this value.
- 8) Create a new variable called FULLTIME if the variable FPTWK has a value that corresponds with the respondent working mainly full time.
- 9) What is the average number of WAGES given that the respondent worked mainly full time. What about when the respondent did not work mainly full time? (You have just computed 2 conditional means.) How do these values change if you only use data on respondents whose WAGES are not missing?
- 10) Are the incomes of people from wages and salaries (given in the variable WAGES) very spread out? Report the mean and standard deviation of WAGES, and attach a histogram showing its distribution.
- 11) Create a variable COHORT that takes the values 1,2,3,4,5,6 for immigrant permanent residents arriving in the 1950s, 1960s, 1970s, 1980s, 1990s, 2000s, takes the value 7 for non-permanent (temporary) residents, and takes the value 8 for Canadian-born (non-immigrant) people.
 - a) Report the conditional mean of WAGES, for people with WAGE_MISSING=0, for each value of COHORT.
 - b) Do immigrants earn less than Canadian-born people? Give a table that shows this.
- 12) Create a table of means, like in 11) above, that controls for age.
- 13) Statistical Snappers: using Eviews' random number generator, create a variable called MYNORM that is distributed as a standard normal. Create another variable called MYCHI equal to its square.
 - a) What proportion of MYCHI is negative?
 - b) What proportion of MYNORM is between 0 and 1?
 - c) What proportion of MYNORM is bigger than 2?
 - d) Are these numbers different from what you answered in Question 4? Why?