

ECN 102: Analysis of Economics Data

Homework 2

Remy Beauregard

Submission rules: homework must be submitted as *one pdf*; please do not submit multiple files. All submissions should include all Stata code and output produced, any figures generated, and answers to questions. For written questions, work must be shown for answers. Code need not be commented.

All datasets can be downloaded in Stata format from the [AED data page](#).

Question 1: Population mean and variance

Let X denote the number of days it will rain in Davis, CA in July. Suppose $X = 0$ with probability 0.75, $X = 1$ with probability 0.13, $X = 2$ with probability 0.08, and $X = 3$ with probability 0.04 (assume these are all the possible values X can take).

- Obtain the population mean, $\mu = E[X]$, from first principles
- Obtain $\sigma^2 = E[(x - \mu)^2]$ from first principles
- Find the standard deviation of X
- What do these Greek letters μ and σ correspond to, specifically? How do these differ from $\bar{x}$ and s ?

Question 2: Mean and variance of the sample mean

Let $\bar{X}$ be the mean of a random sample of size $n = 400$ from a random variable X that is *not* distributed normally, with mean 45 and variance 100.

- Give the mean of $\bar{X}$
- Give the variance and standard deviation of $\bar{X}$
- Is X normally distributed? Is $\bar{X}$ likely to be normally distributed? Explain.

Question 3: The t distribution

For a given T distribution, the following Stata commands give areas in the **right tail** of the distribution:

- `di ttail(9,1.2)` gives $P[T > 1.2]$ for $T \sim T(9)$ (a t-distribution with $n = 10, df = 9$).
- `di invttail(9,0.025)` gives $t^* : P[T > t^*] = 0.025$ again for $T \sim T(9)$

Now, for $T \sim T(44)$, use these Stata commands to find:

- a) $P[T > 2]$
- b) $P[|T| > 2]$
- c) $P[T < -1.8]$
- d) $P[T < 1.8]$
- e) $P[T > 0]$
- f) $t^* : P[T < -t^*] = 0.025$
- g) $t^* : P[T > t^*] = 0.05$
- h) $t^* : P[|T| > |t^*|] = P[T < -t^* \text{ or } T > t^*] = 0.05$
- i) $t^* : P[T > t^*] = 0.5$

Question 4: T-testing in Stata

Use AED_CALELECTRICITY.DTA from hw1. Suppose we want to test whether the spot price of electricity is different from its one-day ahead forward price.

- a) Describe the variables present in the dataset. Which variable corresponds to the spot price? Which corresponds to the forward price? Is there a variable for the difference between the two?
- b) It is given above that we want to test whether the difference between these two series is different from zero. Is this a one-sided or two-sided test? What would our null and alternate hypotheses be?
- c) To run a ttest on a variable in Stata, we use `ttest [VARNAME]=[NULL VALUE]`. Include the output Stata gives for running this command.
- d) What is the conclusion of our test based on the p-value approach?
- e) What is the conclusion of our test based on the critical value approach? (You may need to compute an additional quantity to answer this)

- f) What is the conclusion of our test based on the given confidence interval?
- g) If we had instead wanted to test whether the difference between the spot price and forward price was *greater than zero*, what would our hypotheses have been? What would our conclusion have been?
- h) Why does Stata provide three p-values and alternate hypotheses in its output for `ttest` but only one t-statistic?