Econ 140 - Spring 2016
Section 2

GSI: Fenella Carpena

January 28, 2016

In this section note, we will review the normal distribution, the sample mean, the law of large numbers, and the central limit theorem (second half of Stock & Watson, Chapter 2). Please be advised that this note is not intended to be a comprehensive review of lecture or the textbook, since there is a lot more material than we have time to cover. However, I have tried to focus on the concepts which I believe are necessary to be successful in our class.

1 The Normal Distribution

The normal distribution is the most widely used distribution in statistics. It is really important for this course because we will rely heavily on it for many topics.

- A normal random variable is a continuous r.v. that can take on any value. We say that X has a **normal distribution** with expected value μ and variance σ^2, written as $X \sim \mathcal{N}(\mu, \sigma^2)$. The normal pdf is symmetric about it mean, μ.
- A special case of the normal distribution occurs when the mean is 0 and the variance is 1. This is called the **standard normal distribution.**
  - If the r.v. Z has $\mathcal{N}(0,1)$ distribution, then we say Z is a **standard normal random variable.**
  - *Notation: The letter Z is usually used to denote a standard normal r.v.*
  - The pdf of the standard normal r.v. is denoted $\phi(z)$.
  - The cdf of the normal distribution is denoted $\Phi(z)$. It is obtained as the area under ϕ, to the left of z. Further, $\Phi(z) = P(Z \leq z)$.
  - The values of the probabilities $\Phi(z)$ of a standard normal r.v. are tabulated in the back of your textbook (see *Appendix Table 1*). It is very important to understand how this table works, since we will be using it over the course of the semester.
- **Important property:** Any normal r.v. X can be transformed into a standard normal r.v.
  - Mathematically, if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $(X - \mu)/\sigma \sim \mathcal{N}(0, 1)$.
  - The term $(X - \mu)/\sigma$ is also sometimes called the Z-statistic.
  - The above property means that to turn X into a normal r.v., we need to subtract its mean μ and divide by the sd σ.
  - This property is important because it allows us to find the probabilities of *any* normal r.v., by first transforming it to a standard normal and then using the *Appendix Table 1* at the back of the textbook.
- **Important formulas to remember when working with a normal random variable Z:**
  1. $P(Z > z) = 1 - P(Z \leq z) = 1 - \Phi(z)$

2. $\Phi(-z) = P(Z < -z) = P(Z > z)$. This holds because of symmetry of the normal normal distribution.
3. $P(|Z| > z) = P(Z > z) + P(Z < -z) = 2P(Z > z) = 2P(Z < -z) = 2\Phi(-z)$ Again, this holds because of symmetry of the normal distribution.
4. For any constants a and b with $a < b$, $P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$

2 Sample Average (a.k.a. Sample Mean)

- We denote the sample mean as $\overline{X}$. It is the average of n individual measurements from a population, where n is our sample size. Mathematically, we state this formally as $\overline{X} = \frac{X_1+X_2+\ldots+X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
- If the X_i's (that is, $X_1, \ldots, X_n$) are independent and identically distributed (i.i.d.), this means the following.
  - *Independent*: Each of the X_i's are independent from each other. For example, X_1 and X_5 are independent, X_2 and X_3 are independent, etc. As a consequence of independence, the covariance of any pair of X_i's is zero. For example, $cov(X_1, X_5) = 0$, $cov(X_2, X_3) = 0$.
  - *Identically Distributed*: All individual measurement X_i have the same distribution. In particular, this means that all X_i's have the same mean μ_X and the same standard deviation σ_X.
- Simple random sampling produces n random observations $X_1, \ldots, X_n$ that are i.i.d.
- The sample average $\overline{X}$ is a random variable and has a sampling distribution, because the value of the sample mean is different every time we take a different sample. In other words, there is only population but many possible samples, so the value of the sample statistic we obtain would differ from sample to sample. This variability is called **sampling variation**.
- Important formulas to know about $\overline{X}$ (make sure you know how to derive these, see equation 2.44 and 2.45 of the textbook):
  - $E(\overline{X}) = \mu_X$
  - $var(\overline{X}) = \frac{\sigma_X^2}{n}$
  - $sd(\overline{X}) = \frac{\sigma_X}{\sqrt{n}}$

3 Law of Large Numbers (LLN)

- The **Law of Large Numbers (LLN)** states that if $X_1, \ldots, X_n$ are i.i.d. with $E(X_i) = \mu_X$ and if large outliers are unlikely, then $\overline{X} \xrightarrow{p} \mu_X$.
- What does $\overline{X} \xrightarrow{p} \mu_X$ mean? This is "convergence in probability." We say that $\overline{X}$ converges in probability to μ_X if for any constant c, the probability that $\overline{X}$ is in the range $(\mu_X - c)$ to $(\mu_X + c)$ becomes close to 1 as n increases.
- Intuitively, the LLN says that as the sample size n increases, with very high probability (approaching 1), the sample average $\overline{X}$ will be arbitrarily close to the expected value μ_X

4 Central Limit Theorem (CLT)

- In general, we do not know what the distribution of the sample mean $\overline{X}$ is.
- However, the **Central Limit Theorem (CLT)** tells us that if (1) $X_1, \ldots, X_n$ are i.i.d. with $E(X_i) = \mu_X$ and $var(X_i) = \sigma_X^2$, (2) $var(X_i) = \sigma_X^2$ is finite, i.e. $var(X_i) = \sigma_X^2 < \infty$, and (3) the sample size n is sufficiently large, then $\overline{X}$ is approximately normally distributed with mean μ_X and variance σ_X^2/n. That is,

$$\overline{X} \sim \mathcal{N}\left(\mu_X, \frac{\sigma_X^2}{n}\right)$$