Chapter 6: Large Random Samples

Sections

- 6.1: Introduction
- 6.2: The Law of Large Numbers
    - Skip p. 356-358
- 6.3: The Central Limit Theorem
    - Skip p. 366-368
- Skip 6.4: The correction for continuity

Remember: The Midterm is October 25th in class (same room)

- It will cover Chapters 1 - 7
- The exam will be closed book but you may bring a “cheat sheet”
    - One sheet of letter-sized paper. You can write on both sides and no restrictions on what can be on there. I will provide the normal table.
- Homework 7, due Oct. 18 will be from the last part of Chapter 7 and you will get that back on Oct. 23, i.e. before the midterm.

Introduction

- Intuitively we expect the average of many i.i.d. random variables to be close to their mean
- For example: Let $X_1, X_2, X_3, \ldots$ be a random sample from a $N(\mu, \sigma^2)$ distribution and let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. We can show that for any constant c
$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| \leq c) = 1$$
- The Law of large numbers gives a mathematical foundation to this for more distributions
- The Central Limit Theorem gives an approximate probability distribution for how close the sample average is to the mean

Inequalities

Theorem 6.2.1: Markov Inequality

Let X be a non-negative random variable, i.e. $P(X \geq 0) = 1$. Then for any constant $t > 0$

$$P(X \geq t) \leq \frac{E(X)}{t}$$

- Gives a bound to how much probability can be at large values

Theorem 6.2.2: Chebychev Inequality

Let X be a random variable and suppose $\mathrm{Var}(X)$ exists. Then for any constant $t > 0$

$$P(|X - E(X)| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}$$

- Gives a bound to how far away X is from its mean, and relates it to the variance

Example - Using the Chebychev inequality

- Let X be a continuous random variable with mean μ and variance σ^2.
- By using $t = k\sigma$ in the Chebychev inequality we get

$$P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{(k\sigma)^2} = \frac{1}{k^2}$$

which can also be written as

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

In other words: No matter what distribution X has:

  - There is at least 75% chance that X is within 2σ from its mean (set $k = 2$)
  - There is at least 88.9% chance that X is within 3σ from its mean (set $k = 3$)

The sample mean

The *sample mean* is defined as $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$

Theorem 6.2.3: Mean and variance of $\overline{X}$

Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and variance σ^2. Then

$$E(\overline{X}_n) = \mu \quad \text{and} \quad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$$

- That is, the variance of the average is smaller than for a single random variable
- Using Chebychev's inequality we get (for any distribution)

$$P(|\overline{X}_n - \mu| \geq t) \leq \frac{\sigma^2}{nt^2}$$

The weak Law of Large Numbers

Def: Convergence in probability

A sequence of random variables, $Z_1, Z_2, Z_3, \ldots$ is said to *converge to b in probability* if for every number $\epsilon > 0$

$$\lim_{n\to\infty} P\left(|Z_n - b| < \epsilon\right) = 1$$

This is often written as

$$Z_n \xrightarrow{P} b$$

Theorem 6.2.4: (Weak) Law of Large Numbers (LLN)

Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and a finite variance. Then

$$\overline{X}_n \xrightarrow{P} \mu$$

Histogram as an approximation to a pdf

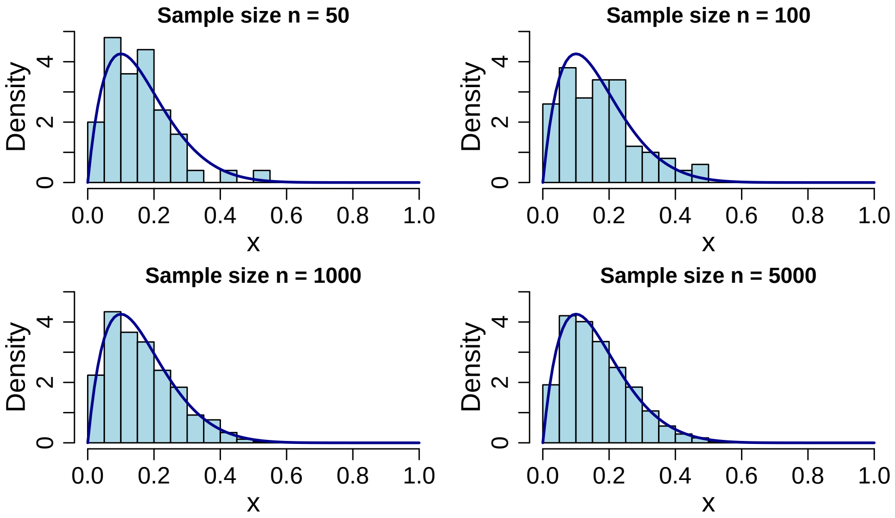
Theorem 6.2.6: Histograms

Let $X_1, X_2, X_3, \ldots$ be a sequence of i.i.d. random variables.
Let $c_1 < c_2$ be two constants.
Define $Y_i = 1$ if $c_1 \le X_i < c_2$ and $Y_i = 0$ otherwise.
Then $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$ is the proportion of X_i's that lie in the interval $[c_1, c_2)$ and

$$\overline{Y}_n \xrightarrow{P} P(c_1 \le X_1 < c_2)$$

- This means that the area of a bar in a histogram converges to the probability of that interval
- I.e. the histogram is an approximation to the pdf.

Example: Random samples from the Beta distribution



Convergence in distribution

Def: Convergence in distribution

Let $X_1, X_2, X_3, \ldots$ be a sequence of random variables. Let F_n be the cdf for X_n for all n and let F^* also be a cdf. We then say that the sequence *converges in distribution to F^** if

$$\lim_{n \to \infty} F_n(x) = F^*(x)$$

for all x for which F^* is continuous. F^* is called the *Asymptotic distribution of X_n*

The Central Limit Theorem

Theorem 6.3.1: Central Limit Theorem (CLT)

Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and finite variance σ^2. Then for each fixed number x

$$\lim_{n\to\infty} P\left(\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq x\right) = \Phi(x)$$

where $\Phi(x)$ is the standard normal cdf.

That is,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma}$$

converges in distribution to the standard normal distribution

Example: Sample mean of Binomials

Say $X_1, X_2, X_3, \ldots$ are i.i.d. Binomial with parameters k and p.
Then $\mu = E(X_i) = kp$ and $\sigma^2 = \mathrm{Var}(X) = kp(1-p)$.
For large n the distribution of

$$\frac{\sqrt{n}(\overline{X} - kp)}{\sqrt{kp(1-p)}} \quad \text{is approximately } N(0,1)$$

In other words, the distribution of the sample mean $\overline{X}_n$ is approximately

$$N\left(kp, \frac{kp(1-p)}{n}\right)$$

Say $k = 10$, $p = 0.2$ and $n = 25$.
Then the distribution of $\overline{X}_{25}$ is approximately

$$N\left(2, \frac{10 \times 0.2 \times 0.8}{25} = 0.064\right)$$

Example: Sample mean of Binomials – continued

The distribution of the sample mean $\overline{X}_n$ is approximately

$$N\left(kp, \frac{kp(1-p)}{n}\right)$$

For $k = 10$, $p = 0.2$ and $n = 25$, the distribution of $\overline{X}_{25}$ is approximately $N(2, 0.064)$

Then we can calculate for example

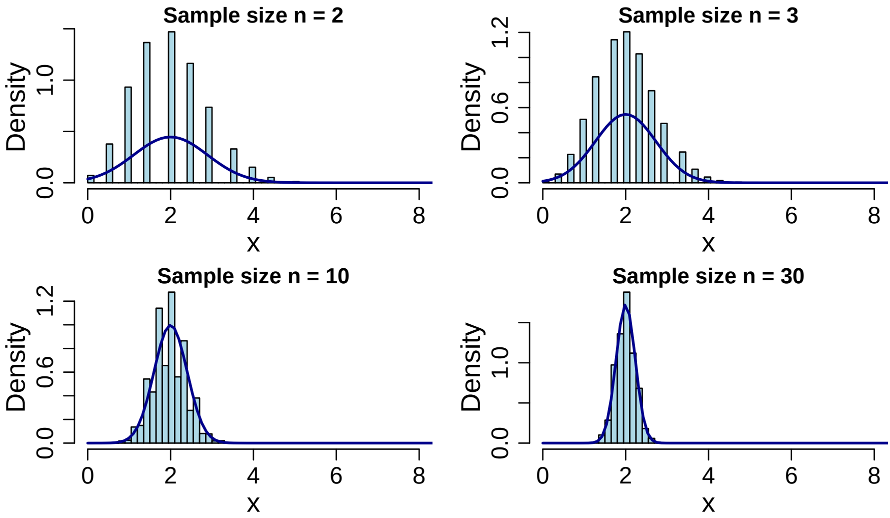
$$P(\overline{X}_{25} \leq 1.5)$$

We can also calculate the minimum number n so that

$$P(|\overline{X}_n - \mu| < 0.1) \geq 0.9$$

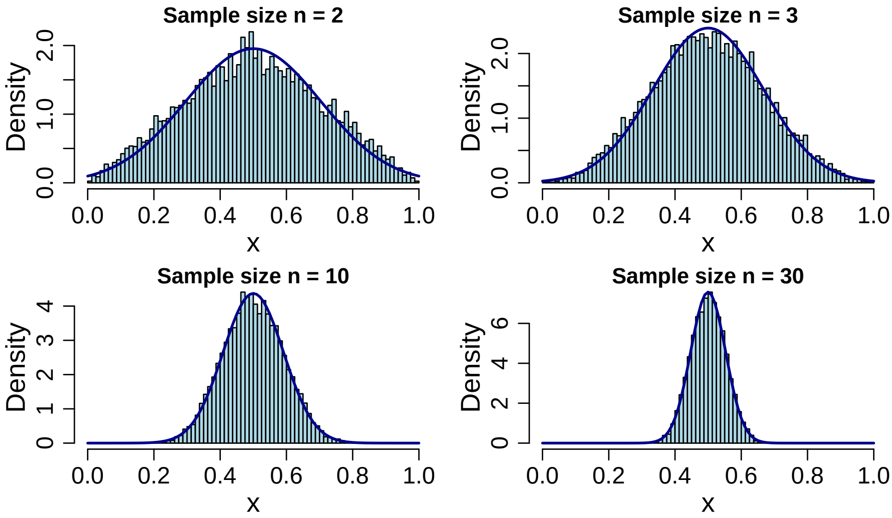
Example: Sampling from a $\text{Binomial}(10, 0.2)$ distr.

Histograms of 10,000 sample means and the normal approx. for different *n*



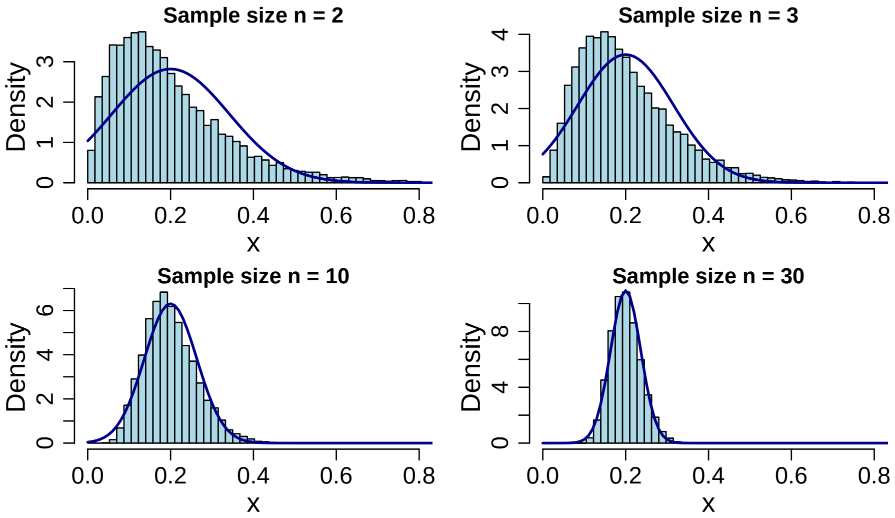
Example: Sampling from a $\mathrm{Uniform}(0, 1)$ distribution

Histograms of 10,000 sample means and the normal approx. for different *n*



Example: Sampling from an $\mathrm{Expo}(5)$ distribution

Histograms of 10,000 sample means and the normal approx. for different *n*



Example: Empty bottle?

- Suppose that people attending a party pour drinks from a bottle containing 63 ounces of a certain liquid.
- Suppose also that the expected size of each dring is 2 ounces and the standard deviation is 1/2 ounce and that all drinks are poured independently.
- Determine the probability that the bottle will not be empty after 36 drinks have been poured.

Delta method

Theorem: 6.3.2: Delta Method

Let $Y_1, Y_2, \dots$ be a sequence of random variables. Suppose

$$a_n(Y_n - \theta) \text{ converges in distribution to } F^*(x)$$

where $F^*(x)$ is a continuous distribution and $a_1, a_2, \dots$ is a sequence of numbers such that $\lim_{n\to\infty} a_n = \infty$.
Let $g(x)$ be a function with a continuous derivative and $g'(\theta) \neq 0$. Then

$$\frac{a_n(g(Y_n) - g(\theta))}{g'(\theta)} \text{ converges in distribution to } F^*(x)$$

Example: Binomial again

In our Binomial example we found that

$$\frac{\sqrt{n}(\overline{X}_n - 2)}{\sqrt{0.064}} \text{ converges in distribution to } N(0,1)$$

Find the asymptotic distribution of $\log(\overline{X}_n)$

END OF CHAPTER 6