Gov 2000: 4. Sums, Means, and Limit Theorems

Matthew Blackwell

Fall 2016

Where are we? Where are we going?

- Probability: formal way to quantify uncertain outcomes/random variables.
- Last week: how to work with multiple r.v.s at the same time.
- This week: applying those ideas to study large random samples

Large random samples

- In real data, we will have a set of n measurements on a variable:

$$X_1, X_2, \dots, X_n$$

- Or we might have a set of n measurements on two variables:

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$$

- Empirical analyses: sums or means of these n measurements
    - Almost all statistical procedures involve a sum/mean.
    - What are the properties of these sums and means?
    - Can they tell us anything about the distribution of X_i?
- Asymptotics: what can we learn as n gets big?

1/ Sums and Means of Random Variables

Sums and means are random variables

- If X_1 and X_2 are r.v.s, then $X_1 + X_2$ is a r.v.
    - Has a mean $\mathbb{E}[X_1 + X_2]$ and a variance $\mathbb{V}[X_1 + X_2]$
- The sample mean is a function of sums and so it is a r.v. too:

$$\bar{X} = \frac{X_1 + X_2}{2}$$

Distribution of sums/means

	X_1	X_2	$X_1 + X_2$	$\bar{X}$
draw 1	20	71	91	45.5
draw 2	12	66	78	39
draw 3	59	75	134	67
draw 4	3	58	61	30.5
⋮	⋮	⋮	⋮	⋮
			distribution of the sum	distribution of the mean

Independent and identical r.v.s

- We often will work with independent and identically distributed r.v.s, $X_1, \ldots, X_n$
  - Random sample of n respondents on a survey question.
  - Written "i.i.d."
- Independent: $X_i \perp\!\!\!\perp X_j$ for all $i \neq j$
- Identically distributed: $f_{X_i}(x)$ is the same for all i
  - $\mathbb{E}[X_i] = \mu$ for all i
  - $\mathbb{V}[X_i] = \sigma^2$ for all i

Distribution of the sample mean

- Sample mean of i.i.d. r.v.s: $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$
- $\overline{X}_n$ is a random variable, what is its distribution?
  - What is the expectation of this distribution, $\mathbb{E}[\overline{X}_n]$?
  - What is the variance of this distribution, $\mathbb{V}[\overline{X}_n]$?
  - What is the p.d.f. of the distribution?
- How do they relate to the expectation, variance of $X_1, \ldots, X_n$?

Properties of the sample mean

Mean and variance of the sample mean

Suppose that $X_1, \ldots, X_n$ is are i.i.d. r.v.s with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$. Then:

$$\mathbb{E}[\overline{X}_n] = \mu \qquad \mathbb{V}[\overline{X}_n] = \frac{\sigma^2}{n}$$

- Key insights:
  - Sample mean get the right answer on average
  - Variance of $\overline{X}_n$ depends on the variance of X_i and the sample size
  - Not dependent on the (full) distribution of X_i!
- Standard error of the sample mean: $\sqrt{\mathbb{V}[\overline{X}_n]} = \frac{\sigma}{\sqrt{n}}$
- You'll prove both of these facts in this week's HW.

2/ Useful Inequalities

Why inequalities?

- Behavior of r.v.s depend on their distribution, but we often don't know (or don't want to assume) a distribution.
- Today, we'll discuss results for r.v.s with any distribution subject to some restrictions like finite variance.
- Why study these?
  - Build toward massively important results like LLN
  - Inequalities used regularly throughout statistics
  - Gives us some practice with proofs/analytic reasoning

Markov Inequality

Markov Inequality

Suppose that X is r.v. such that $\mathbb{P}(X \geq 0) = 1$. Then, for every real number $t > 0$,

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}.$$

- For instance, if we know that $\mathbb{E}[X] = 1$, then $\mathbb{P}(X \geq 100) \leq 0.01$
- Once we know the mean of a r.v., it limits how much probability can be in the tail.

Markov Inequality Proof

- For discrete X:

$$\mathbb{E}[X] = \sum_{x} x f_X(x) = \sum_{x<t} x f_X(x) + \sum_{x\geq t} x f_X(x)$$

- Because X is nonnegative, $\mathbb{E}[X] \geq \sum_{x\geq t} x f_X(x)$
- Since $x \geq t$, then $\sum_{x\geq t} x f_X(x) \geq \sum_{x\geq t} t f_X(x)$
- But this is just $\sum_{x\geq t} t f_X(x) = t \sum_{x\geq t} f_X(x) = t\mathbb{P}(X \geq t)$
- Implies $\mathbb{E}[X] \geq t\mathbb{P}(X \geq t)$

Chebyshev Inequality

Chebyshev Inequality

Suppose that X is r.v. for which $\mathbb{V}[X] < \infty$. Then, for every real number $t > 0$,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathbb{V}[X]}{t^2}.$$

- The variance places limits on how far an observation can be from its mean.

Proof of Chebyshev

- Let $Y = (X - \mathbb{E}[X])^2$
  - $\rightsquigarrow \mathbb{P}(Y \geq 0) = 1$ (nonnegative)
  - $\mathbb{E}[Y] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{V}[X]$ (definition of variance)
- Note that if $|X - \mathbb{E}[X]| \geq t$ then $Y \geq t^2$ because we just squared both sides.
- Thus, $\mathbb{P}(|X - \mathbb{E}[X]| \geq t) = \mathbb{P}(Y \geq t^2)$
- Apply Markov's inequality:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t) = \mathbb{P}(Y \geq t^2) \leq \frac{\mathbb{E}[Y]}{t^2} = \frac{\mathbb{V}[X]}{t^2}$$

Application: planning a survey

- Suppose we want to estimate the proportion of voters who will vote for Donald Trump, p, from a random sample of size n.
  - $X_1, X_2, \ldots, X_n$ indicating voting intention for Trump for each respondent.
  - By our earlier, calculation, $\mathbb{E}[\overline{X}_n] = p$ and $\mathbb{V}[\overline{X}_n] = \frac{\sigma^2}{n}$
  - Since this is a Bernoulli r.v., we have $\sigma^2 = p(1-p)$
- What does n need to be to have at least 0.95 probability that $\overline{X}_n$ is within 0.02 of the true p?
  - How to guarantee a margin of error of $\pm$ 2 percentage points?

Application: planning a survey

- What does n have to be so that

$$\mathbb{P}(|\overline{X}_n - p| \leq 0.02) \geq 0.95 \iff \mathbb{P}(|\overline{X}_n - p| \geq 0.02) \leq 0.05$$

- Applying Chebyshev:

$$\mathbb{P}(|\overline{X}_n - p| \geq 0.02) \leq \frac{\mathbb{V}[\overline{X}_n]}{0.02^2} = \frac{p(1-p)}{0.0004n}$$

- We don't know $\mathbb{V}[X_i] = p(1-p)$, but:
  - Conservative to use largest possible variance.
  - It can't be bigger than $p(1-p) \leq (1/2) \cdot (1/2) = (1/4)$

$$\mathbb{P}(|\overline{X}_n - p| \geq 0.02) \leq \frac{p(1-p)}{0.0004n} \leq \frac{1}{0.0016n}$$

- We want this probability to be bounded by 0.05 so we need $(1/0.0016n) \leq 0.05$, which gives us $n \geq 12,500$!!

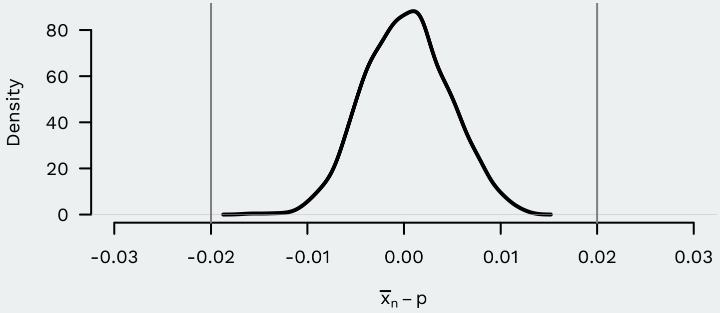
Application: planning a survey

- Do we really need $n \geq 12,500$ to get a margin of error of ± 2 percentage points?
- No! Chebyshev provides a bound that is guaranteed to hold, but actual probabilities are much smaller.
  - We're also using the "worst-case" variance of 0.25.
- Let's simulate 1000 samples of size $n = 12500$ with $p = 0.4$ and show the distribution of the means.
  - What proportion of these are within 0.02 of p?

Application: planning a survey

```
nsims <- 1000
holder <- rep(NA, times = nsims)
for (i in 1:nsims) {
    this.samp <- rbinom(n = 12500, size = 1, prob = 0.4)
    holder[i] <- mean(this.samp)
}
mean(abs(holder - 0.4) > 0.02)
```

```
## [1] 0
```



3/ Law of Large Numbers

Current knowledge

- For i.i.d. r.v.s, $X_1, \ldots, X_n$, with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$ we know that:
  - Expectation is $\mathbb{E}[\overline{X}_n] = \mathbb{E}[X_i] = \mu$
  - Variance is $\mathbb{V}[\overline{X}_n] = \frac{\sigma^2}{n}$ where $\sigma^2 = \mathbb{V}[X_i]$
  - Some bounds on tail probabilities from Chebyshev.
  - None of these rely on a specific distribution for X_i!
- Can we say more about the distribution of the sample mean?
- Yes, but we need to think about how $\overline{X}_n$ changes as n gets big.

Sequence of sample means

- What can we say about the sample mean n gets large?
- Need to think about sequences of sample means with increasing n:

$$
\begin{aligned}
\overline{X}_1 &= X_1 \\
\overline{X}_2 &= (1/2) \cdot (X_1 + X_2) \\
\overline{X}_3 &= (1/3) \cdot (X_1 + X_2 + X_3) \\
\overline{X}_4 &= (1/4) \cdot (X_1 + X_2 + X_3 + X_4) \\
\overline{X}_5 &= (1/5) \cdot (X_1 + X_2 + X_3 + X_4 + X_5) \\
&\vdots \\
\overline{X}_n &= (1/n) \cdot (X_1 + X_2 + X_3 + X_4 + X_5 + \cdots + X_n)
\end{aligned}
$$

- Note: this is a sequence of random variables!

Convergence in Probability

Convergence in probability

A sequence of random variables, $Z_1, Z_2, \ldots$, is said to converge in probability to a value b if for every $\varepsilon > 0$,

$$\mathbb{P}(|Z_n - b| > \varepsilon) \to 0,$$

as $n \to \infty$. We write this $Z_n \overset{p}{\to} b$.

- Basically: probability that Z_n lies outside any (teeny, tiny) interval around b approaches 0 as $n \to \infty$
- Wooldridge writes $\text{plim}(Z_n) = b$ if $Z_n \overset{p}{\to} b$.

Law of large numbers

Theorem: Weak Law of Large Numbers

Let $X_1, \ldots, X_n$ be a an i.i.d. draws from a distribution with mean μ and finite variance σ^2. Let $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Then, $\overline{X}_n \overset{p}{\to} \mu$.

- Intuition: The probability of $\overline{X}_n$ being "far away" from μ goes to 0 as n gets big.
    - The distribution of $\overline{X}_n$ "collapses" on μ
- No assumptions about the distribution of X_i beyond i.i.d. and a finite variance!

LLN proof

- Proof: by Chebyshev and properties of probabilities, we have

$$0 \le \mathbb{P}(|\overline{X}_n - \mu| \ge \varepsilon) \le \frac{\mathbb{V}[\overline{X}_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

- As $n \to \infty$, we know that $\sigma^2/n\varepsilon^2 \to 0$ which by the sandwich theorem implies

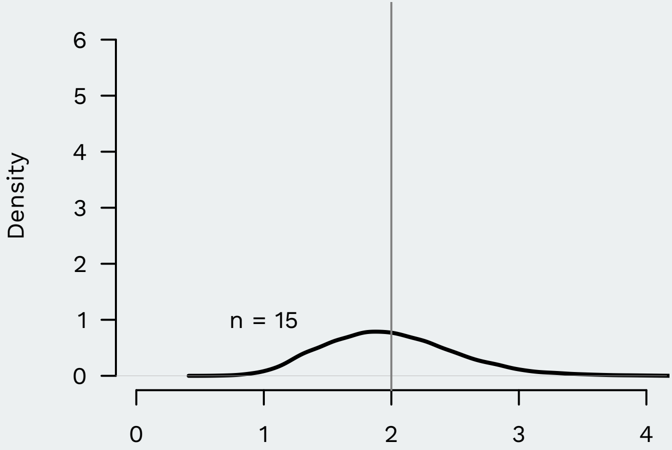
$$\lim_{n\to\infty} \mathbb{P}(|\overline{X}_n - \mu| > \varepsilon) = 0$$

LLN by simulation in R

- Draw different sample sizes from Exponential distribution with rate 0.5
- $\rightsquigarrow \mathbb{E}[X_i] = 2$

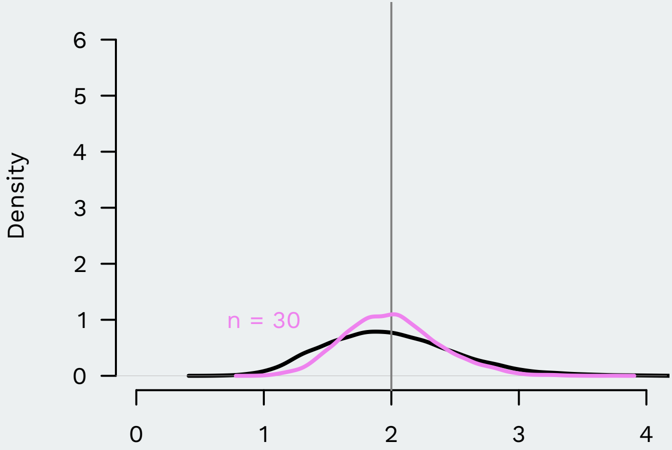
```
nsims <- 10000
holder <- matrix(NA, nrow = nsims, ncol = 6)
for (i in 1:nsims) {
    s5 <- rexp(n = 5, rate = 0.5)
    s15 <- rexp(n = 15, rate = 0.5)
    s30 <- rexp(n = 30, rate = 0.5)
    s100 <- rexp(n = 100, rate = 0.5)
    s1000 <- rexp(n = 1000, rate = 0.5)
    s10000 <- rexp(n = 10000, rate = 0.5)

    holder[i, 1] <- mean(s5)
    holder[i, 2] <- mean(s15)
    holder[i, 3] <- mean(s30)
    holder[i, 4] <- mean(s100)
    holder[i, 5] <- mean(s1000)
    holder[i, 6] <- mean(s10000)
}
```

LLN in action



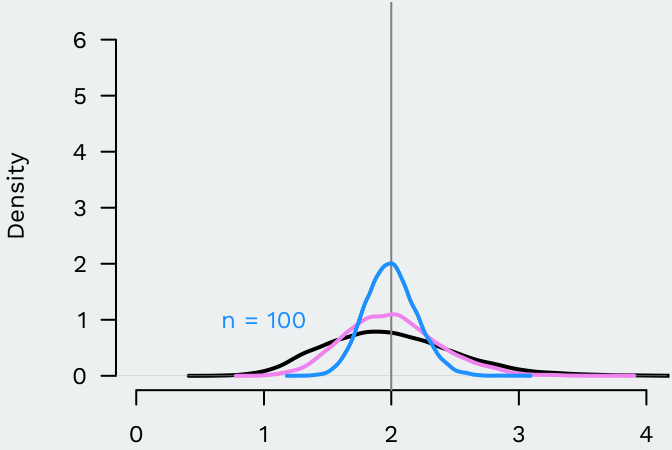
- Distribution of $\overline{X}_{15}$

LLN in action



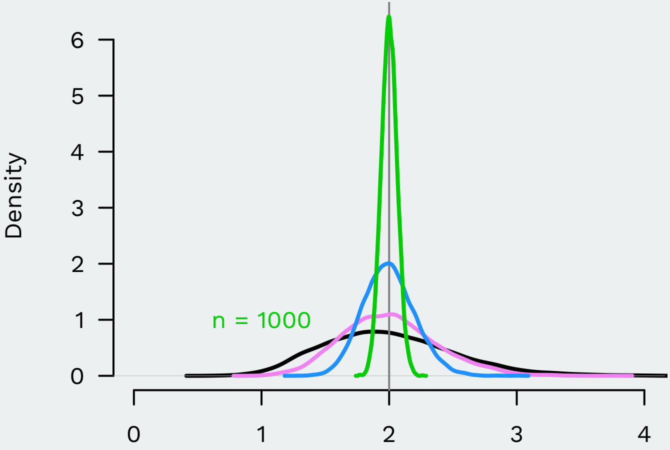
- Distribution of $\overline{X}_{30}$

LLN in action



- Distribution of $\overline{X}_{100}$

LLN in action



- Distribution of $\overline{X}_{1000}$

Properties of convergence in probability

1. if $X_n \stackrel{p}{\to} c$, then $g(X_n) \stackrel{p}{\to} g(c)$ for any continuous function g.
2. if $X_n \stackrel{p}{\to} a$ and $Z_n \stackrel{p}{\to} b$, then
   - $X_n + Z_n \stackrel{p}{\to} a + b$
   - $X_n Z_n \stackrel{p}{\to} ab$
   - $X_n / Z_n \stackrel{p}{\to} a/b$ if $b > 0$

- Thus, by LLN:
  - $\left(\overline{X}_n\right)^2 \stackrel{p}{\to} \mu^2$
  - $\log(\overline{X}_n) \stackrel{p}{\to} \log(\mu)$

4/ Central Limit Theorem

Current knowledge

- For i.i.d. r.v.s, $X_1, \dots, X_n$, with $\mathbb{E}[X_i] = \mu$ and $\mathbb{V}[X_i] = \sigma^2$ we know that:
  - $\mathbb{E}[\overline{X}_n] = \mu$ and $\mathbb{V}[\overline{X}_n] = \frac{\sigma^2}{n}$
  - $\overline{X}_n$ converges to μ as n gets big
  - Chebyshev provides some bounds on probabilities.
  - Still no distributional assumptions about X_i!
- Can we say more?
  - Can we approximate $\Pr(a < \overline{X}_n < b)$?
  - What family of distributions (Binomial, Uniform, Gamma, etc)?
- Again, need to analyze when n is large.

Convergence in Distribution

Convergence in distribution

Let $Z_1, Z_2, \ldots$, be a sequence of r.v.s, and for $n = 1, 2, \ldots$ let $F_n(z)$ be the c.d.f. of Z_n. Then it is said that $Z_1, Z_2, \ldots$ converges in distribution to r.v. W with c.d.f. F_W if

$$\lim_{n \to \infty} F_n(x) = F_W(x),$$

which we write as $Z_n \xrightarrow{d} W$.

- Basically: when n is big, the distribution of Z_n is very similar to the distribution of W
- We use c.d.f.s here to avoid messy details with discrete vs continuous.
- If $X_n \xrightarrow{p} X$, then $X_n \xrightarrow{d} X$

Standardizing an r.v.

- Common to standardize a r.v. by subtracting its expectation and dividing by its standard deviation:

$$Z = \frac{X - \mathbb{E}[X]}{\sqrt{\mathbb{V}[X]}}$$

- Possible to show that for any X, we have (try to prove these to yourself):
  - $\mathbb{E}[Z] = 0$
  - $\mathbb{V}[Z] = 1$
- Sometimes called a z-score.

Central Limit Theorem

Central Limit Theorem

Let $X_1, \ldots, X_n$ be i.i.d. r.v.s from a distribution with mean μ and variance $0 < \sigma^2 < \infty$. Then,

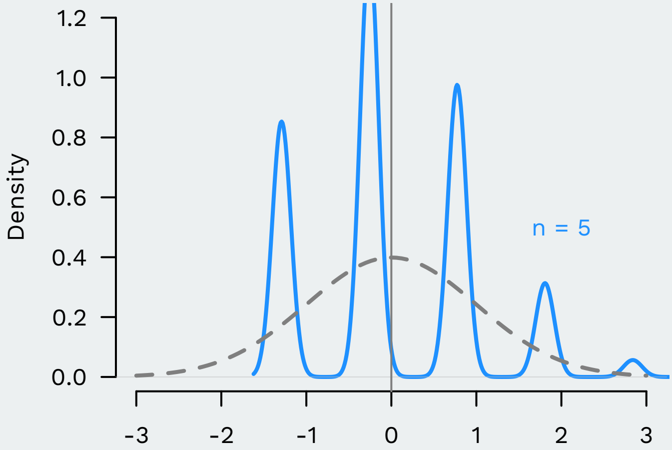
$$\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \xrightarrow{d} N(0, 1).$$

- Distribution free! We don't have to make specific assumptions about the distribution of X_i
- Implies that $\overline{X}_n \sim N(\mu, \sigma^2/n)$
  - $\rightsquigarrow$ easy approximations to probability statements about $\overline{X}_n$ when n is big!

CLT by simulation in R

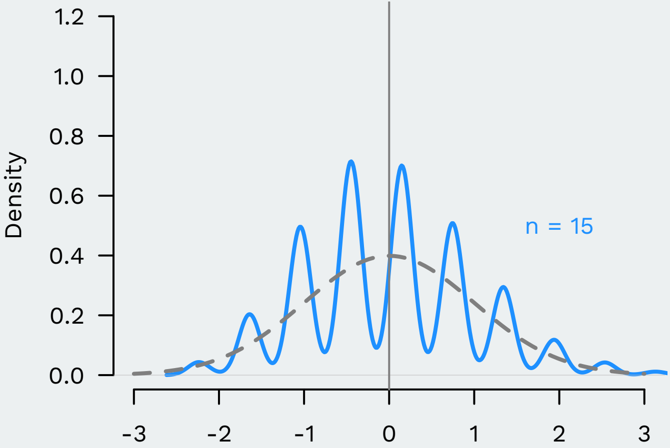
```
set.seed(2138)
nsims <- 10000
holder2 <- matrix(NA, nrow = nsims, ncol = 6)
for (i in 1:nsims) {
    s5 <- rbinom(n = 5, size = 1, prob = 0.25)
    s15 <- rbinom(n = 15, size = 1, prob = 0.25)
    s30 <- rbinom(n = 30, size = 1, prob = 0.25)
    s100 <- rbinom(n = 100, size = 1, prob = 0.25)
    s1000 <- rbinom(n = 1000, size = 1, prob = 0.25)
    s10000 <- rbinom(n = 10000, size = 1, prob = 0.25)

    holder2[i, 1] <- mean(s5)
    holder2[i, 2] <- mean(s15)
    holder2[i, 3] <- mean(s30)
    holder2[i, 4] <- mean(s100)
    holder2[i, 5] <- mean(s1000)
    holder2[i, 6] <- mean(s10000)
}
```

CLT in action



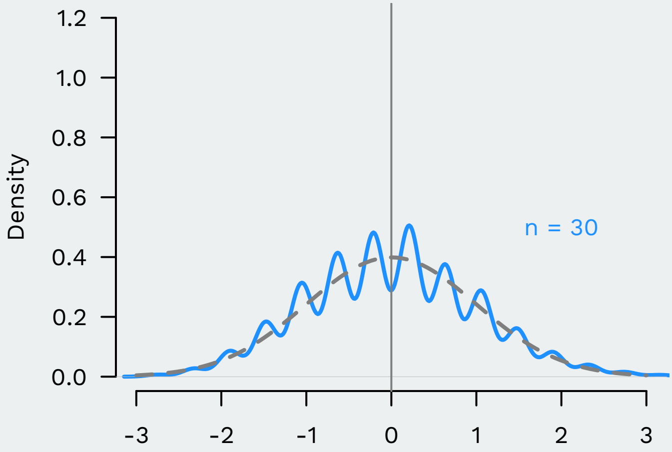
- Distribution of $\frac{\overline{X}_5 - \mu}{\sigma / \sqrt{5}}$

CLT in action



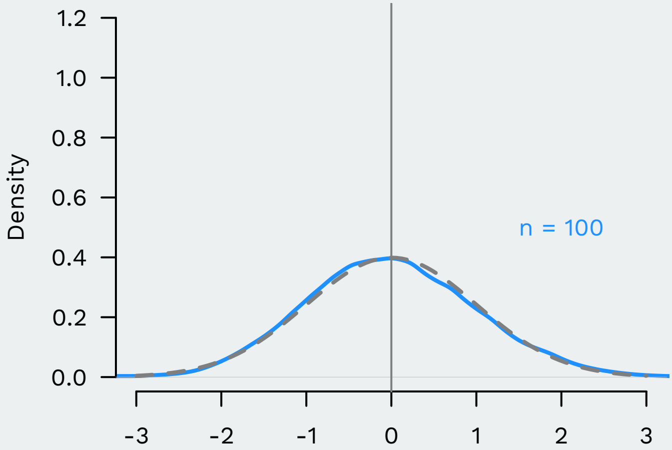
- Distribution of $\frac{\overline{X}_{15}-\mu}{\sigma/\sqrt{15}}$

CLT in action



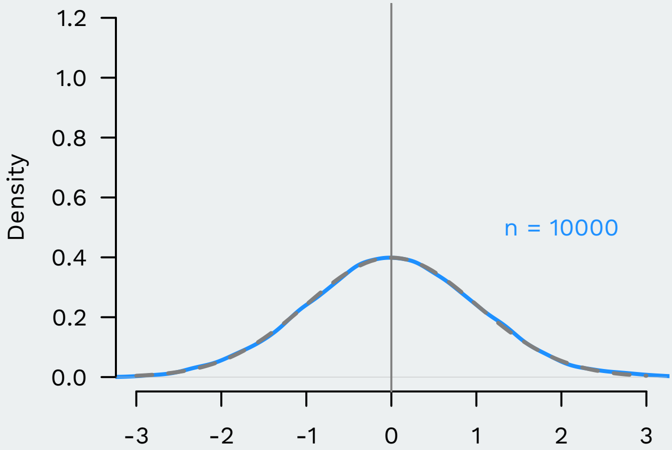
- Distribution of $\frac{\overline{X}_{30} - \mu}{\sigma / \sqrt{30}}$

CLT in action



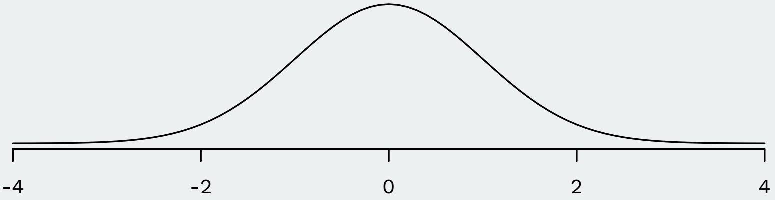
- Distribution of $\frac{\overline{X}_{100} - \mu}{\sigma / \sqrt{100}}$

CLT in action



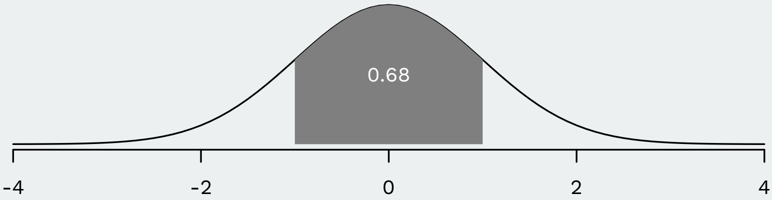
- Distribution of $\frac{\overline{X}_{10000}-\mu}{\sigma/\sqrt{10000}}$

Empirical Rule for the Normal Distribution



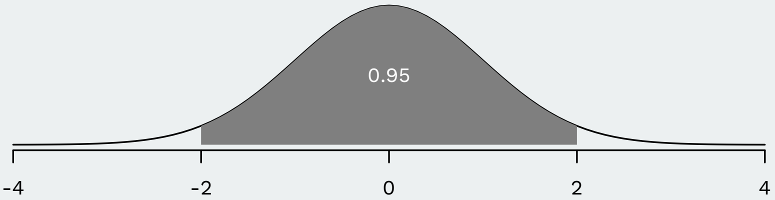
- If $Z \sim N(0, 1)$, then the following are roughly true:

Empirical Rule for the Normal Distribution



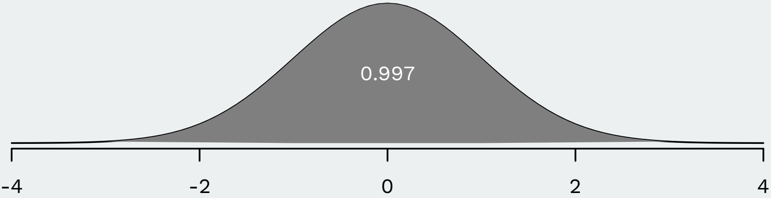
- If $Z \sim N(0, 1)$, then the following are roughly true:
- Roughly 68% of the distribution of Z is between -1 and 1.

Empirical Rule for the Normal Distribution



- If $Z \sim N(0, 1)$, then the following are roughly true:
- Roughly 68% of the distribution of Z is between -1 and 1.
- Roughly 95% of the distribution of Z is between -2 and 2.

Empirical Rule for the Normal Distribution



- If $Z \sim N(0, 1)$, then the following are roughly true:
- Roughly 68% of the distribution of Z is between -1 and 1.
- Roughly 95% of the distribution of Z is between -2 and 2.
- Roughly 99.7% of the distribution of Z is between -3 and 3.

Simulating the empirical rule

- Actual probability of $Z \sim N(0,1)$ between -2 and 2:

```r
pnorm(2) - pnorm(-2)
```

```
## [1] 0.9545
```

- Simulated probability of $\frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}}$ between -2 and 2:
  - $n = 15 \rightsquigarrow$ 0.9683
  - $n = 30 \rightsquigarrow$ 0.9666
  - $n = 100 \rightsquigarrow$ 0.9523
  - $n = 1000 \rightsquigarrow$ 0.9551
  - $n = 10000 \rightsquigarrow$ 0.9546
- Quality of the approximation depends on the underlying distribution of the X_i
  - Obviously if $X_i \sim N(0,1)$ it's going to be perfect with $n = 1$

Slustsky's Theorem

- Let $X_1, X_2, \ldots$ converge in distribution to some r.v. X
- Let $Y_1, Y_2, \ldots$ converge in probability to some number, c
- Slutsky's Theorem gives the following result:
    1. $X_n Y_n$ converges in distribution to cX
    2. $X_n + Y_n$ converges in distribution to $X + c$
- Extremely useful when trying to figure out what the large-sample distribution of an estimator is.

Application: planning a survey

- Trump example: we want the the probability of being within 0.02 from the true p to be 95%.
- $\rightsquigarrow$ we want n such that:

$$\mathbb{P}\left(|\overline{X}_n - p| > 0.02\right) \leq 0.05$$

- By the CLT, if n is large, then

$$\overline{X}_n - p \approx N\left(0, \sigma^2/n\right)$$

- We know $\sigma^2 \leq 1/4$, so to be conservative:
  - $\overline{X}_n - p \approx N\left(0, \frac{1}{4n}\right)$
  - Standardizing $\rightsquigarrow Z = \frac{(\overline{X}_n - p)}{1/\sqrt{4n}} = 2\sqrt{n}(\overline{X}_n - p) \approx N(0,1)$
- Easier to work with standardized r.v.:

$$\mathbb{P}\left(|\overline{X}_n - p| > 0.02\right) \leq 0.05 \iff \mathbb{P}\left(|Z| > 0.02 \times 2\sqrt{n}\right) \leq 0.05$$

Application: planning a survey

- We want:

$$\mathbb{P}(|Z| > 0.04\sqrt{n}) \leq 0.05$$
$$\mathbb{P}(Z < -0.04\sqrt{n}) + \mathbb{P}(Z > 0.04\sqrt{n}) \leq 0.05$$

- The standard normal is symmetric around 0, so:
  - Upper tail probs = lower tail probs
  - $\mathbb{P}(Z < -0.04\sqrt{n}) = \mathbb{P}(Z > 0.04\sqrt{n})$
- Allow us to simplify:

$$2 \times \mathbb{P}(Z < -0.04\sqrt{n}) \leq 0.05$$
$$\mathbb{P}(Z < -0.04\sqrt{n}) \leq 0.025$$

- To solve for n, we need to know q such that $\mathbb{P}(Z \leq q) = 0.025$
  - Inverse of the c.d.f. called the quantile: $q = F^{-1}(0.025)$
  - $q = F^{-1}(p)$ is the (smallest) value of the r.v. such that $\mathbb{P}(X \leq q) = F(q) \geq p$

Application: planning a survey

- We can use the `qnorm()` function in R:

```
qnorm(0.025, mean = 0, sd = 1)
```

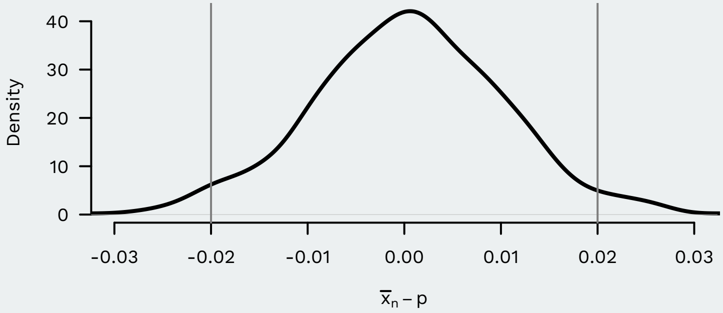
```
## [1] -1.96
```

- if $-0.04\sqrt{n} \leq q$, then $\mathbb{P}\left(Z < -0.04\sqrt{n}\right) \leq 0.025$
- So, we need $-0.04\sqrt{n} \leq -1.96$ or $n > 2401$
- Much lower than the 12,500 from Chebyshev.

Application: planning a survey

```
nsims <- 1000
holder <- rep(NA, times = nsims)
for (i in 1:nsims) {
    this.samp <- rbinom(n = 2401, size = 1, prob = 0.4)
    holder[i] <- mean(this.samp)
}
mean(abs(holder - 0.4) > 0.02)
```

```
## [1] 0.052
```



5/ More Exotic CLTs*

CLT for non-iid r.v.s

- What if we don't have i.i.d. r.v.s? Does the CLT still apply?
- Let $X_1, X_2, \ldots$ be independent (but not identically distributed) with means $\mathbb{E}[X_i] = \mu_i$ and variances $\mathbb{V}[X_i] = \sigma_i^2$.
- Scaled and centered:

$$Y_n = \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu_i}{\left(\sum_{i=1}^n \sigma_i^2\right)^{1/2}}$$

  - No need to divide by n because there are n entries in the sum $\sum_{i=1}^n \mu_i$
- Easy to show that $\mathbb{E}[Y_n] = 0$ and $\mathbb{V}[Y_n] = 1$. Does the CLT apply?

Liapounov CLT

Liapounov CLT

Suppose that the r.v.s $X_1, X_2, \ldots$ are independent and that $\mathbb{E}[|X_i - \mu_i|^3] < \infty$ for $i = 1, 2, \ldots$. Also, suppose that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} \mathbb{E}\left[|X_i - \mu_i|^3\right]}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{3/2}} = 0.$$

Then,

$$Y_n = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{1/2}} \xrightarrow{d} N(0, 1)$$

- Key condition: there isn't one r.v.s in the sequence that is "too big" that could dominate the sum

CLT for dependent sequences

- We have shown the CLT for i.i.d. and for independent r.v.s. What about dependent sequences?
- CLT works for a dependent sequence $X_1, X_2, \ldots$.
  - What does dependent sequence mean? Cov$[X_i, X_j] \neq 0$
- Key condition for dependent CLT: r.v.s aren't "too correlated"
- Overall conditions for CLT to hold: the sum/mean of many, not too correlated, not too big r.v.s

6/ Wrap-up

Limitations of asymptotics

- These results are practically and theoretically very useful.
- But remember that they are approximations
- We don't live in asymptopia—n is always finite.
- Asymptotics often give reasonable answers, but you can check with simulations.

Review

- Sums and means of r.v.s are themselves r.v.s
- Learned about the distribution of the sample mean of i.i.d. r.v.s
  - Expectation $\mathbb{E}[\overline{X}_n] = \mu$
  - Variance $\mathbb{V}[\overline{X}_n] = \sigma^2/n$
  - Converges in probability to true mean (LLN)
  - Converges in distribution to a normal distribution (CLT)
- Ahead: generalizing these ideas to arbitrary estimators of parameters.