

Lecture 2

Quantitative Political Science

Prof. Bisbee

Vanderbilt University

Lecture Date: 2023/08/29

Slides Updated: 2023-08-28

Agenda

1. The normal distribution
2. Probability theory
3. Calculating probability
4. Bayes' Rule

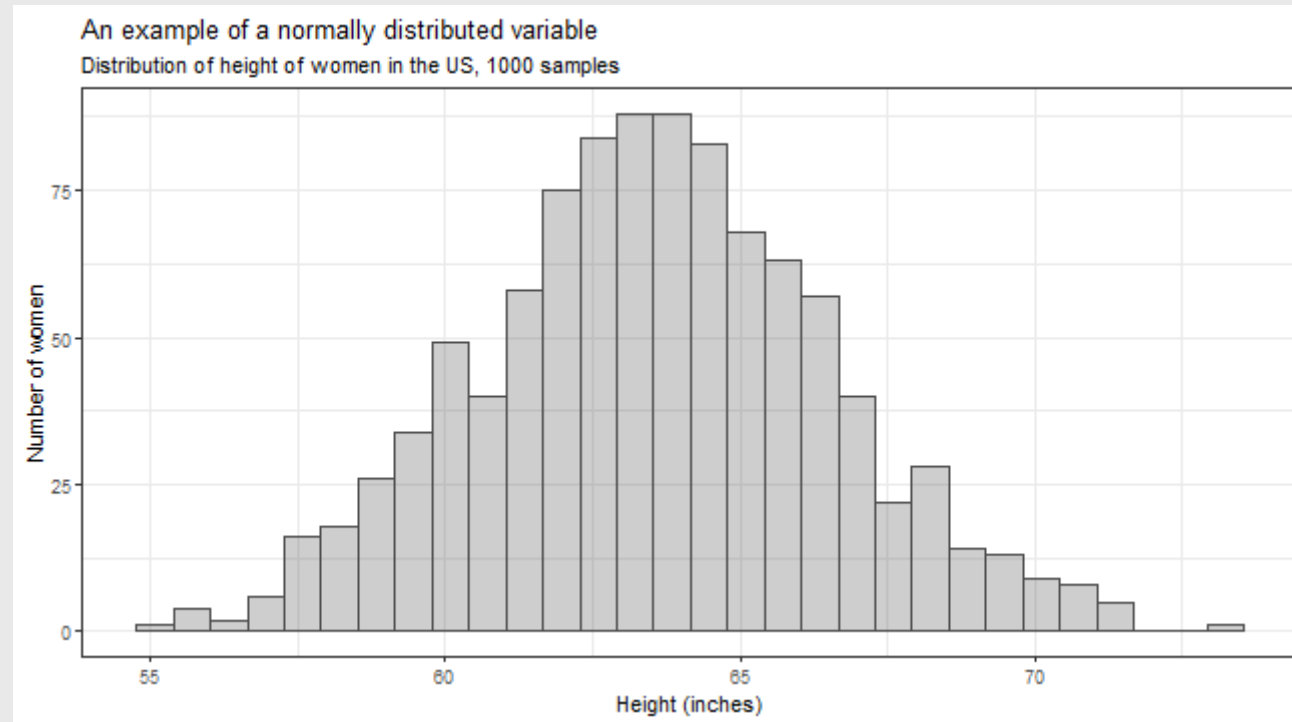
The Normal Distribution

- Bell shaped?
 - Unimodal
 - Symmetric
- Describes many empirical distributions we observe

```
require(tidyverse)
x <- rnorm(n = 1000, mean = 63.5, sd = 3)

p <- data.frame(x = x) %>%
  ggplot(aes(x = x)) +
  geom_histogram(alpha = .3, color = 'grey30') +
  labs(x = 'Height (inches)',
       y = 'Number of women',
       title = 'An example of a normally distributed variable',
       subtitle = 'Distribution of height of women in the US, 1000 samples') +
  theme_bw()
```

The Normal Distribution

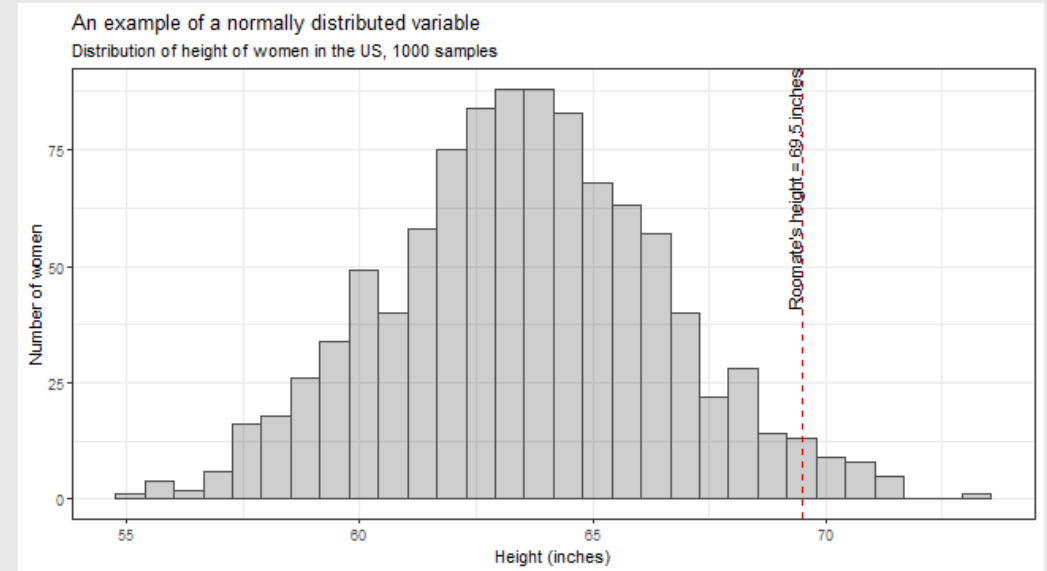


The Normal Distribution

- Your cousin's best friend's new roommate is 69.5 inches tall
 - She is taller than what proportion of women in the American population?
- Tchebysheff's Theorem:
 - $\bar{y} \pm s$ contains 68% of observations
 - $\bar{y} \pm 2s$ contains 95% of observations
 - $\bar{y} \pm 3s$ contains basically all of the observations

The Normal Distribution

```
p <- data.frame(x = x) %>%
  ggplot(aes(x = x)) +
  geom_histogram(alpha = .3,color =
'grey30') +
  labs(x = 'Height (inches)',
    y = 'Number of women',
    title = 'An example of a normally
distributed variable',
    subtitle = 'Distribution of height of
women in the US, 1000 samples') +
  geom_vline(xintercept = 69.5,color =
'red',linetype = 'dashed') +
  annotate(geom = 'text',x = 69.5,y =
Inf,label = "Roomate's height = 69.5
inches",
    angle = 90,hjust = 1,vjust =
-.15) +
  theme_bw()
```



The Normal Distribution

- Can calculate it manually

```
est <- mean(x < 69.5)
```

```
## [1] 0.972
```

- Why is it not exact?

```
est <- mean(rnorm(n = 10000, mean = 63.5, sd = 3) < 69.5)
```

```
## [1] 0.978
```

- Preview of the **Law of Large Numbers**

```
est <- mean(sapply(1:1000, function(x)  
mean(rnorm(n = 1000, mean = 63.5, sd = 3) <  
69.5))))
```

```
## [1] 0.977394
```

The Normal Distribution

- The **mean** of independently drawn observations from **any** populations tends toward the Normal Distribution
- This creates lots of Normal Distributions!
 - **However**, do NOT make the mistake of assuming all distributions are normal
- We haven't talked about **samples** versus **populations**
 - `rnorm()` samples from the normal distribution
- Taking data as given, resisted making **inferences** to a population
 - That's the next big step, but we need to first get familiar with **Probability Theory**

Probability Theory: The Logic

- Familiar with moving from populations to samples
 - Given a 52 card deck with 13 spades, what is the probability of drawing a spade at random from a perfectly shuffled deck?
 - **Population:** 52 card deck with 13 spades
 - **Sample:** draw a card a random
 - $1/4$: we know the distribution of spades in the population, and we are precise about sampling process

Probability Theory; The Logic

- Run it in reverse!
- We know the sample
 - Central tendency (\bar{y})
 - Dispersion (s)
- Make assumptions about sampling process to make **good** guesses about the population's central tendency (μ) and dispersion (σ^2)
- Learn the tools of the **population-to-sample** process to understand the **sample-to-population** process

A Probabilistic Model for an Experiment

- **Experiments** describe the *process by which an observation is made*
- An **observation** is a quantity of interest:
 - The price of a stock
 - The number of subjects who choose "A" instead of "B"
 - The proportion of respondents who approve of Biden's job as president
- Experiments produce outcomes called **events**

Example of an Experiment

- Election in a country with 101 voters; majority rule; everyone votes for either A or B
 - Possible **events**?
 - a. Candidate A wins
 - b. Candidate B wins
 - c. Candidate A wins with 76 votes
 - d. Candidate B wins with 56 votes
 - e. Candidate A wins in a "landslide" (≥ 67 votes)
 - Others?
- Note that a. can be decomposed into several other events
 - a.1 Candidate A wins with 51 votes
 - a.2 Candidate A wins with 52 votes

Example of an Experiment

- Types of events
 - **Compound:** a., b., and e.
 - **Simple:** c. and d.
- **Set theory**
 - Associate a distinct point -- a **sample** point -- with each simple event
 - Refer to simple event i as E_i
- **Sample space** S : set of all possible sample points
- How many in this example?
- $S = \{E_1, E_2, \dots, E_{102}\}$

Set Theory

- S consists of countable (finite) number of sample points
 - **Discrete sample space**
 - Sample points are mutually exclusive
- Compound events are sets of sample points
 - "A wins by a landslide" occurs if one of E_{68}, \dots, E_{102} occurs
 - $A = \{E_{68}, \dots, E_{102}\}$
- Thus an **event** is a collection of sample points
 - A subset of S

Probabilities in Set Theory

- Assign a number $P(A)$ to each event A in S
 - $P(A) \forall A \in S$
- Three axioms of probability give structure to this assignment
 1. $P(A) \geq 0$
 2. $P(S) = 1$
 3. $P(A_1 \cup A_2 \cup A_3 \cup \dots A_n) = \sum_{i=1}^n P(A_i)$
- With these axioms in hand, $P(A)$ is defined as the **probabilities of A**

Calculating Probability of Event

- **Sample point** method
 - Define experiment and sample space S
 - Assign reasonable probabilities to sample points s.t. $P(E_i) \geq 0$ and $\sum_i P(E_i) = 1$
 - For $A = E_{i_1} \cup E_{i_2} \cup \dots E_{i_J}$, calculate $P(A) = \sum_j^J P(E_{i_j})$
- **Event composition** method
 - Decompose and compose event A into **unions** and **intersections** of events with conveniently calculated probabilities

Event Composition Method

- Need four tools
 1. Definitions of **conditional probability** and **independence**
 2. **Multiplicative** and **additive** laws
 3. Probability of an event and its **complement**
 4. The **law of total probability** and **Bayes' Rule**
- Before this, recall that
 - $A \cap B$: compound event where *both* A and B happen (AND)
 - $A \cup B$: compound event where *either* A or B happens, or both (OR)

1. Definitions

- **Conditional probability:** $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- **Independence:** For two events A and B , if *any one* of the following conditions holds:
 - $P(A|B) = P(A)$
 - $P(B|A) = P(B)$
 - $P(A \cap B) = P(A)P(B)$
- then A and B are *independent* events
- If none hold, then A and B are dependent events

2. Laws

- **Multiplicative law** (intersection)

- $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$ **OR**
- $P(A \cap B) = P(A)P(B)$ if A and B are independent events
- For three events $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)$
- For many events $P(A_1 \cap A_2 \cap \dots A_k) = P(A_1)P(A_2|A_1) \dots P(A_k|A_1 \cap A_2 \cap \dots A_{k-1})$

- **Additive law** (union)

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- If A and B are mutually exclusive, $P(A \cap B) = 0$

Aside on a simple proof

- Denote "not A " as A^c (sometimes written $\sim A$ or, in WMS, \bar{A})
- **Complementary events** means that $P(A) = 1 - P(A^c)$
- Proof:
 - $S = A \cup A^c$ (given $A \subset S$)
 - $P(S) = P(A) + P(A^c)$ (by Axiom 3, since A and A^c are mutually exclusive)
 - $1 = P(A) + P(A^c)$ (by Axiom 2)
 - $P(A) = 1 - P(A^c)$ ■

Example of event composition method

- Randomly assign 16 students into 3 teams of 6, 5, and 5
- 11 of the students are male
- What is the probability that the team of six students (call this "Team 1") is entirely male?

1. Define event as "the first six assigned students are male"

2. Decompose event into simpler events A through F

- A : First student picked is male
- B : Second student picked is male
- ...
- F : Sixth student picked is male

3. Event occurs iff A through F all occur, meaning $A \cap B \cap C \cap D \cap E \cap F$.

Example contd

- Thus we want $P(A \cap B \cap C \cap D \cap E \cap F)$
- How to solve?
- **Multiplicative Law**
 - $P(A_1 \cap A_2 \cap \dots A_k) = P(A_1)P(A_2|A_1) \dots P(A_k|A_1 \cap A_2 \cap \dots A_{k-1})$
 - $P(A_1 \cap A_2 \cap \dots A_6) = P(A_1)P(A_2|A_1) \dots P(A_6|A_1 \cap A_2 \cap \dots A_5)$

Example contd

- Step through it
 - What is the probability that the first student picked is male? $\frac{11}{16}$
 - $P(A_1 \cap A_2 \cap \dots A_6) = \frac{11}{16} P(A_2|A_1) \dots P(A_6|A_1 \cap A_2 \cap \dots A_5)$
 - What is the probability that the second student picked is male? $\frac{10}{15}$
 - $P(A_1 \cap A_2 \cap \dots A_6) = \frac{11}{16} \cdot \frac{10}{15} P(A_3|A_1 \cap A_2) \dots P(A_6|A_1 \cap A_2 \cap \dots A_5)$
- And so on
 - $P(A_1 \cap A_2 \cap \dots A_6) = \frac{11}{16} \cdot \frac{10}{15} \cdot \frac{9}{14} \cdot \frac{8}{13} \cdot \frac{7}{12} \cdot \frac{6}{11} = \frac{3}{52}$

Factor notation

- Note that the product of the numerators is $11 \cdot 10 \cdot \dots \cdot 2 \cdot 1$ except we remove the final five multiplicands
- Same as $\frac{11 \cdot 10 \cdot \dots \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$ which is $\frac{11!}{5!}$
- Similarly, denominator can be written as $\frac{10!}{16!}$
- Thus $P(A_1 \cap A_2 \cap \dots A_6) = \frac{11!}{5!} \cdot \frac{10!}{16!}$

Law of Total Probability

- In a discrete sample space S , we can define S as the union of k mutually exclusive events
 - $S = B_1 \cup B_2 \cup \dots B_k$, where
 - $B_i \cap B_j = \emptyset, \forall i \neq j$
- Note that a collection of sets $\{B_1, B_2, \dots B_k\}$ is a **partition** of S iff
 - their union is S and
 - they are themselves mutually exclusive

Law of Total Probability

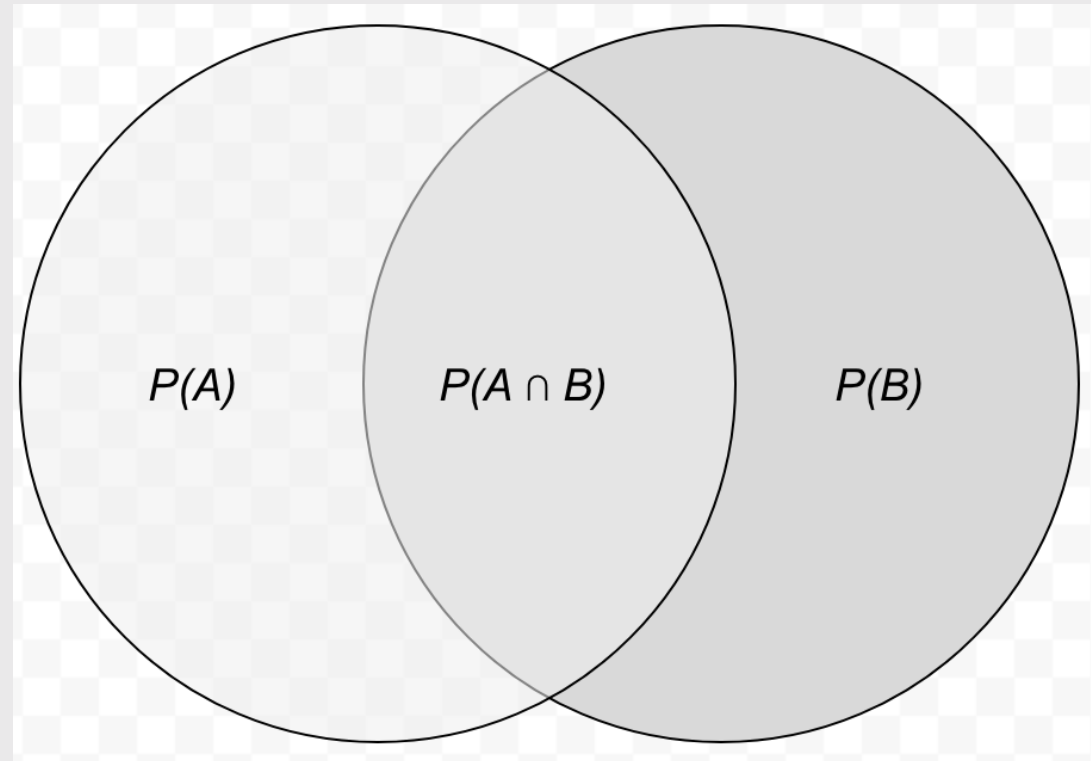
- If A is a subset of S , it can be **decomposed** as the union of its intersections with each of the partitions of S as follows:
 - $A = (A \cap B_1) \cup (A \cap B_2) \cup \dots (A \cap B_k)$
- If $\{B_1, B_2, \dots B_k\}$ is a partition of S such that $P(B_i) > 0 \forall i = 1, 2, \dots k$
- then the **Law of Total Probability** states:
 - $P(A) = \sum_{i=1}^k P(A|B_i)P(B_i)$

Bayes' Rule

- $P(B_j|A) = \frac{P(A \cap B_j)}{P(A)}$
- Substitute the definition into the numerator and the Law of Total Probability into the denominator
- $P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^k P(A|B_i)P(B_i)}$

Bayes' Rule

- Simpler example with two sets
- $P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$



Bayes' Rule

- Example time!
- WMS 2.124, pg. 73: "A population of voters contains 40% Republicans and 60% Democrats. It is reported that 30% of the Republicans and 70% of the Democrats favor an election issue. A person chosen at random from this population is found to favor the issue in question. Find the conditional probability that this person is a Democrat."
- WMS 2.125, pg. 73: "A diagnostic test for a disease is such that it (correctly) detects the disease in 90% of the individuals who actually have the disease. Also, if a person does not have the disease, the test will report that he or she does not have it with probability .9. Only 1% of the population has the disease in question. If a person is chosen at random from the population and the diagnostic test indicates that she has the disease, what is the conditional probability that she does, in fact, have the disease? Are you surprised by the answer? Would you call this diagnostic test reliable?"