Probability Theory

October 12, 2022

Statistical Inference

- Next several weeks focused on statistical inference
- Remember—we are interested in making inferences about a population, but only have data from a sample
  - How confident can we be that a relationship in the sample means a relationship in the population?
- Today—the basics of probability theory & sampling distributions

Difference of Means

Is R married?	Summary of Hours per day watching TV Mean	Std. Dev.	Freq.
No	3.2383268	2.5831093	1028
Yes	2.611691	1.8652487	958
Total	2.9360524	2.2864047	1986

Difference of Proportions

Voted Dem in '00 & '04	Respondent's sex Male	Female	Total
0	405 61.46	482 53.08	887 56.60
1	254 38.54	426 46.92	680 43.40
Total	659 100.00	908 100.00	1,567 100.00

Cross-tab

Party ID: 3 cats	4 quantiles of income06 1	2	3	4	Total
Democrat	487 48.22	564 48.70	410 42.75	269 39.73	1,730 45.48
Independent	283 28.02	242 20.90	178 18.56	85 12.56	788 20.72
Republican	240 23.76	352 30.40	371 38.69	323 47.71	1,286 33.81
Total	1,010 100.00	1,158 100.00	959 100.00	677 100.00	3,804 100.00

Randomness

- Applies to the outcomes of a response variable
- Possible outcomes are known, but it is uncertain which outcome will be realized in any given trial
- Examples:
  - Rolling Dice
  - Spinning a Wheel
  - Tossing a Coin
  - Drawing Cards
  - Random Number Generator

Repeated Trials

- While individual outcomes are difficult to predict, predictable patterns emerge with a large number of observations
  - With random outcomes, the proportion of outcomes is difficult to predict in the short run, but *they become very predictable in the long run*

- Law of Large Numbers—we will talk more about this
  - Note: There is no such thing as a "Law of Small Numbers"
    - e.g. "Chance is lumpy"

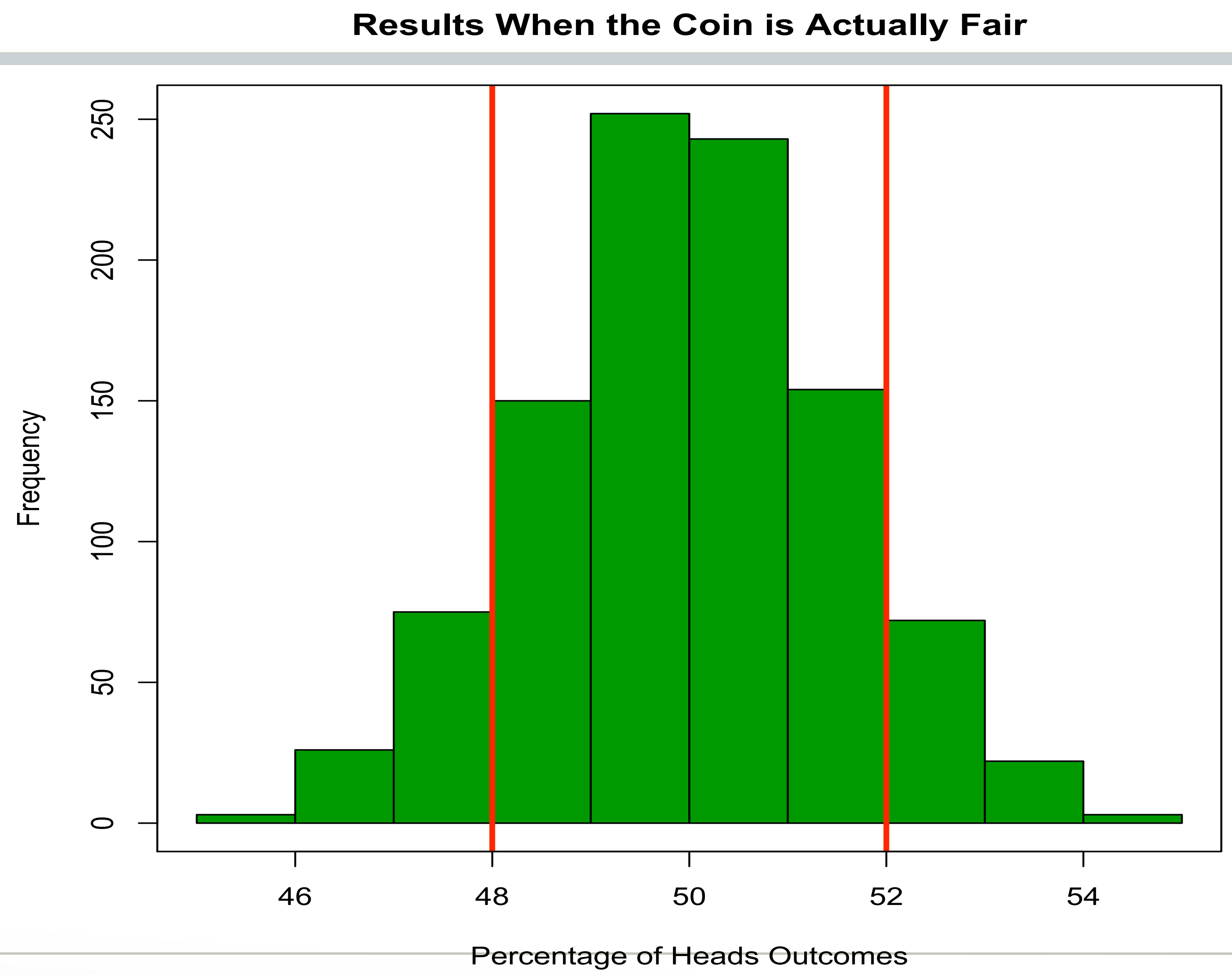
Independence

- Trials are **independent** if the outcome of any one trial is not affected by the outcome of any other trial
  - Example—Coin Toss
  - Not Independent—Drawing Cards (without replacement)
- The **probability** of an outcome is the (expected) proportion of times that the outcome would occur in the long run
  - Example—Coin Tosses
  - Example—Roulette Wheel

A “Fair” Coin & Sampling Variability

- Flip a coin 1000 times and record the percentage of outcomes that come up “Heads”
  - This gives you a statistic from a single trial (i.e., a single sample)
- Then, repeat this process for 1000 trials
- Make a histogram of the variability
  - …the frequency of trials with each (percentage-heads) outcome
- What should we expect for the probability of flipping a “Heads”?
  - …given independent trials, what would a random (i.e., “fair”) coin flip look like in the long-run?

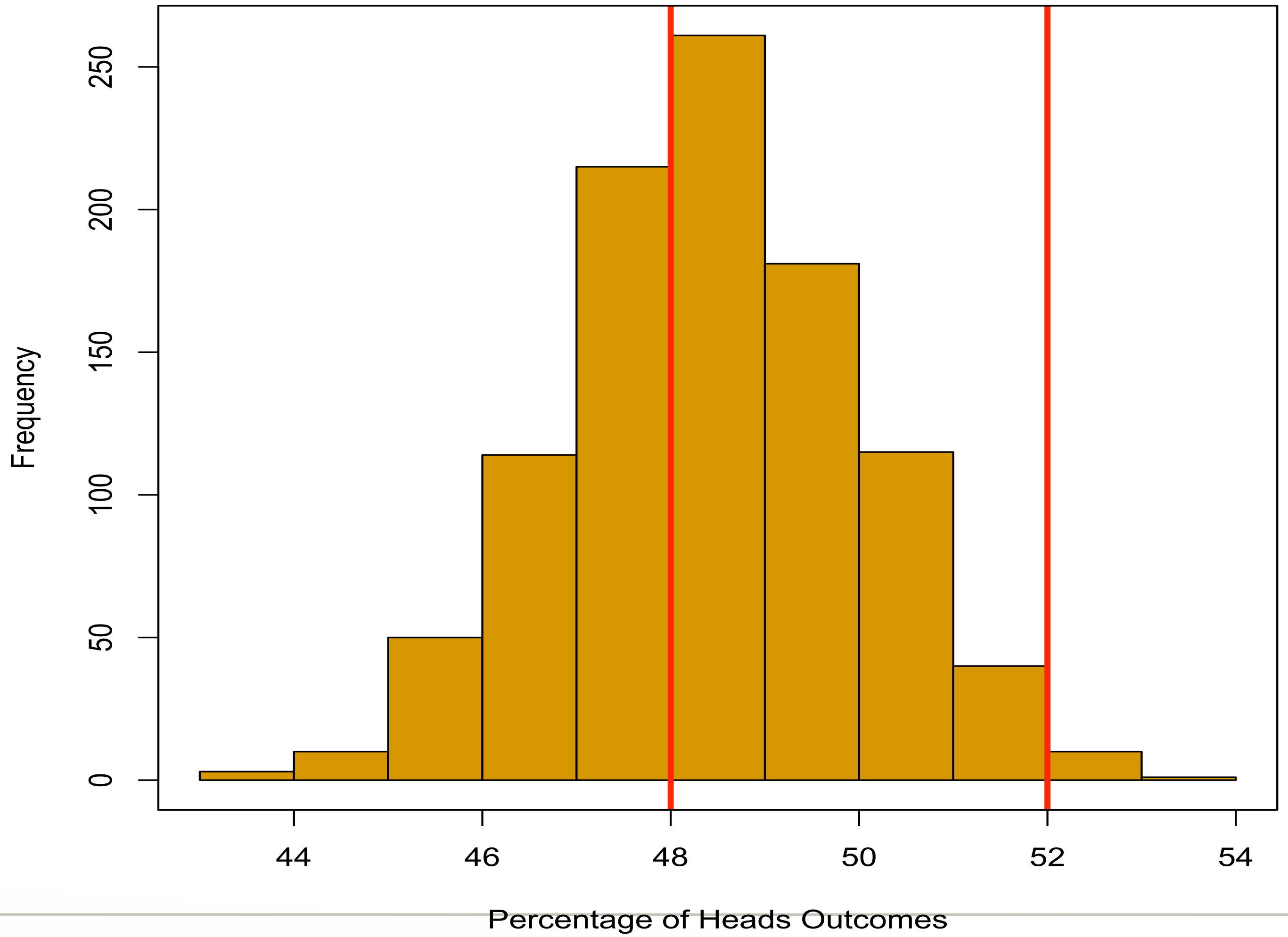
A "Fair" Coin



Results When the Coin is Actually Fair
Frequency
0
50
100
150
200
250
46
48
50
52
54
Percentage of Heads Outcomes

An "Unfair" Coin

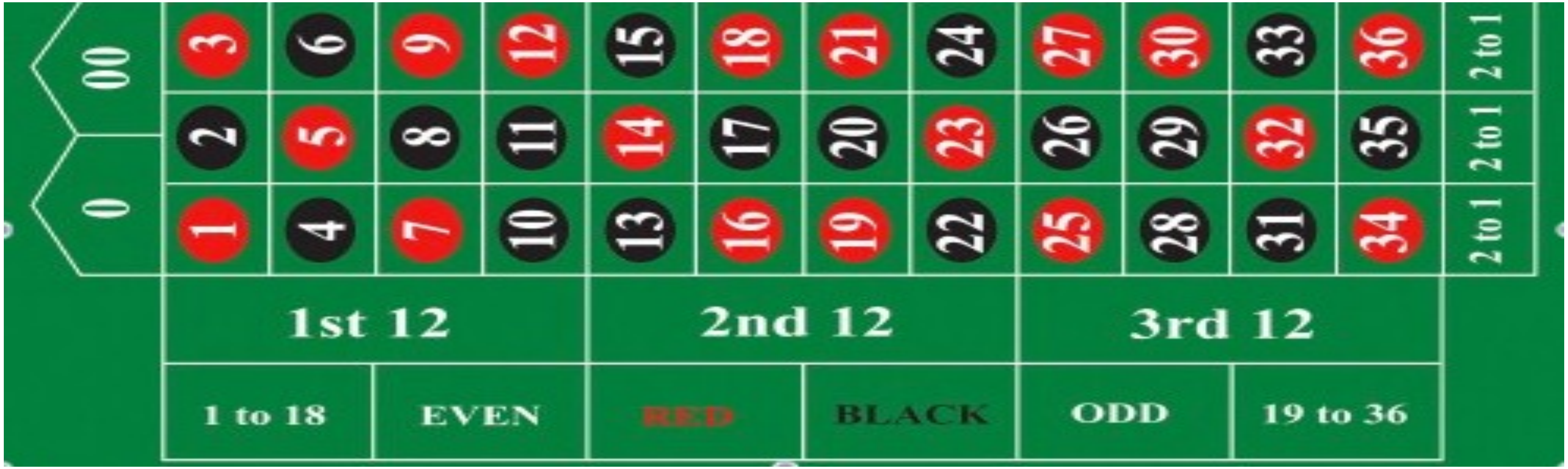
Results When the Coin Slightly Favors Tails



Roulette Wheel

American style double zero roulette wheel.
Photo courtesy www.abbiati.it

Roulette Betting



Probability Model

- A probability model is a mathematical representation of a random phenomenon.
  - It is defined by its **sample space**, **events** within the sample space, and **probabilities** associated with each event
- The **sample space** is the set of all possible outcomes
- Die: {1,2,3,4,5,6}
- Coin Toss Twice: {(H,H),(H,T),(T,H),(T,T)}
- Example: Tossing a Coin Four Times

Events

- An **event** is a subset of the sample space—this may represent multiple possible individual outcomes
  - Example: Tossing coin four times, getting exactly three heads
- Each outcome occurs with probability p in the set of [0,1]
  - All individual (event) probabilities sum to 1
- Probability of event A, P(A) is calculated by adding the probabilities of the individual outcomes
- P(A)=number of outcomes in event A/number of total outcomes in the sample space
  - Example: probability of getting at least three heads on four coin tosses

Complements

- Consists of all outcomes in the sample space that are not A
  - Is denoted by A^c
  - $P(A^C)=1-P(A)$
- Example:
  - Probability of Roulette Wheel Ending up on 3: 1/38
  - Complement: Probability it does not end up on 3: 37/38
- Not an example of complement: Probability you win tic-tac-toe, probability that opponent wins
- Complement: Probability you win tic-tac-toe, probability you do not win tic-tac-toe

Example of Complement

- Canada has two official languages, English and French. Choose a Canadian at random and ask, "What is your mother tongue?" Here is the distribution of responses, combining many separate languages from the broad Asian/ Pacific region:

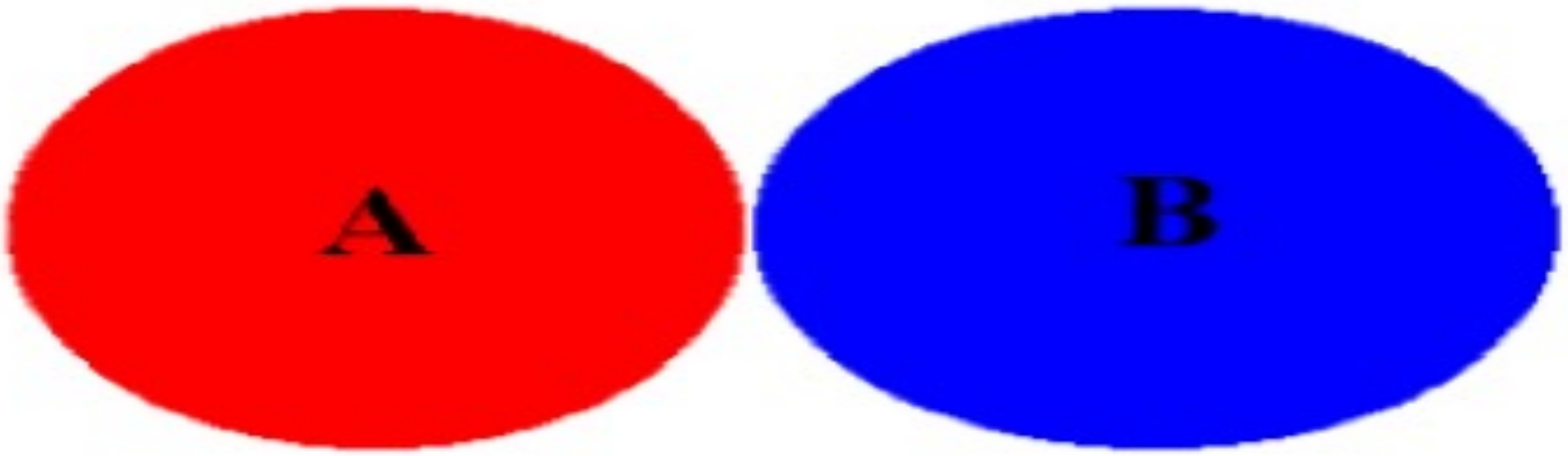
Language	English	French	Asian/Pacific	Other
Probability	?	0.23	0.07	0.11

a) What probability should replace "?" in the distribution?
b) What is the probability that a Canadian's mother tongue is not English?

Intersection

- The intersection of A and B consists of the outcomes that are in both A and B
  - Two events, A and B, are disjoint if they do not have any common outcomes

Disjoint P(A or B)



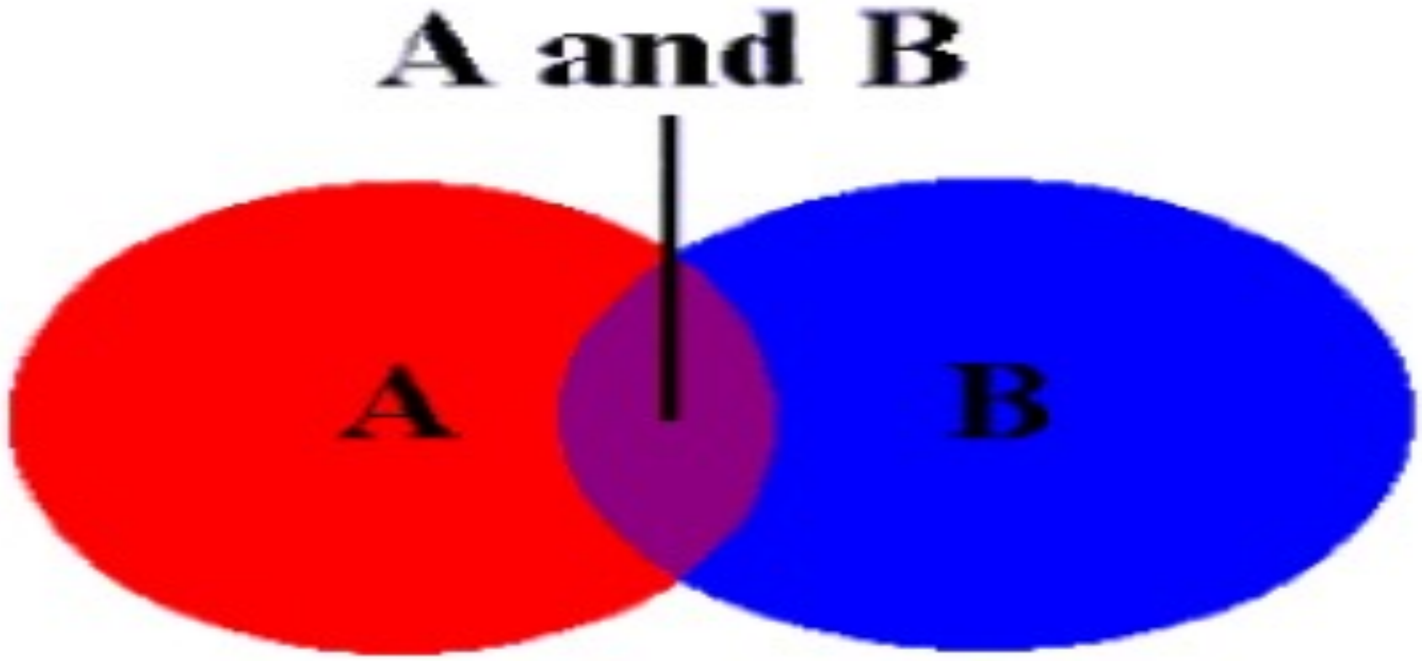
Intersection

- The intersection of A and B consists of the outcomes that are in both A and B
  - Two events, A and B, are disjoint if they do not have any common outcomes
- For the intersection of two independent events, A & B:
  - P(A & B)=P(A)*P(B)
    - This is the multiplication rule
- Example: Probability that when rolling a fair dice twice, I will get a number greater than four each time
- Example: Probability that when rolling a fair dice twice, I will get a number less than five each time

Union

- The union is any collection of events in which at least one of the collection occurs
- Disjoint events:
  - P(A or B)=P(A)+P(B)
- Not Disjoint events
  - P(A or B)=P(A)+P(B)-P(A and B)
- These are really the same equation
- In a roulette game, I place a bet on 3, 7, 15, and 22. What is the probability that I win on at least one of these bets?
- If I roll a fair dice, what is the probability that I get a number that is either less than 5 or even?

Union & Intersection P(A and B)



Joint Probability

- For the intersection of two independent events, A and B:
  - P(A & B)=P(A)*P(B)
    - But, can only use this if events are independent
- For the union of two events:
  - P(A or B)=P(A)+P(B)-P(A and B)
    - Can use this if either disjoint or not disjoint

Conditional Probability

- For events A and B, the conditional probability of event A, given that B has occurred, is:
  - $P(A|B)=P(A \text{ and } B)/P(B)$
- Multiplicative Rule:
  - $P(A \,\&\, B)=P(A|B)*P(B)$
  - $P(A \,\&\, B)=P(B|A)*P(A)$
- Checking for Independence:
  - A & B are independent if:
    - $P(A|B)=P(A)$
    - $P(B|A)=P(B)$
    - $P(A \text{ and } B)=P(A)*P(B)$

Working with Conditional Probabilities

Highest Education	Total population	Employed
Did not finish high school	28,021	11,552
High school but no college	59,844	36,249
Some college but no degree	46,777	32,429
College graduate	51,568	39,250

a) You know that someone is employed. What is the conditional probability that he or she is a college graduate?

b) You know that a second person is a college graduate. What is the conditional probability that he or she is employed?

Working with Conditional Probabilities

- Assume in State A, 10% of the population are government employees and, of government employees, 15% are corrupt. 95% of corrupt officials have a spouse whose salary is ten times higher than their own while, in the rest of the population, only 3% have a spouse who earns ten times as much. What is the probability that a person who has a spouse who earns ten times as much as they do will be a corrupt official?

Random Variables

- X is a random variable
  - A **random variable** is a variable whose value is a numerical outcome of a random phenomenon
- x is a particular value of the variable
- x is in the set of X
- The **probability distribution** of a random variable specifies its possible values and their probabilities

Random Variables

- Two types of random variables:
  - Discrete
    - X has a finite number of possible values
    - Probability distribution of X lists values and their probabilities
    - Can determine probability of event by summing probabilities of individual outcomes
      - P(A or B)=P(A)+P(B)-P(A and B)
  - Continuous
    - X can take any value in an interval of numbers
    - Probability distribution of X is described by a density curve
    - Probability of event is area under the density curve and above the values of X that make up the event
- Normal distribution one type of continuous probability distribution

Mean of a random variable

- Mean of a probability distribution (μ)—the long-run average outcome
- May be interested in mean of several random variables
  - μ_x
  - μ_y
- The mean of a random variable is a weighted average in which each outcome is weighted by its probability
  - Multiply each possible value by its probability, then add all the products
  - Example—average pay-off from betting $10 on "3" on roulette wheel over many repetitions (casino payout is 35:1)
    - What is the expected casino advantage on this bet (in the long run)?

The Law of Large Numbers

- Draw independent observations at random from any population with a finite mean μ
- As the number of observations drawn increases, the mean of the observed values ($\bar{x}$) in the sample approaches the mean of the population (μ)
  - That's why the house always wins
- What is "large?"
  - Depends on variability of outcomes
- Remember: There is no law of small numbers

Getting to Statistical Inference

- Remember, we talked about how random sampling gives us “unbiased” estimates of value in population
- Still have to worry about random sampling error
- Logic of probability theory developed here will allow us to infer from value in sample to population
- Next week, we will discuss sampling distributions and statistical significance