STA 291
Lecture 6

- **Randomness** and **Probability**

- Abstract /mathematical but necessary because this is the mathematical theory underlying all statistical inference
- Fundamental concepts that are very important to understanding *Sampling Distribution, Confidence Interval,* and *P-Value*
- Our goal is to learn the rules involved with assigning probabilities to events, and related calculations after assign probability.

Probability: Basic Terminology

- **Events: --** where probability live.

  usually denoted by A, B, C, etc. (We say probability of event A etc)

- **Sample space** is the largest event. S

- **Sample Space:** The collection of all possible outcomes of an *experiment*.

- **Event:** A specific collection of outcomes.

- **Simple Event:** An event consisting of exactly one outcome.

- Each event is given (assign) a probability: a number between 0 and 1, inclusive. Denoted by P(A)

- Could assign probability using a table (if we can list all events with their probabilities)

- Otherwise, (with too many events) we need a rule of assign probability

Sample Spaces, and Events

Examples of experiment/sample space:

1. Flip a coin: {head, tail}
2. Flip a coin 3 times: {HHH, HHT, ….TTT}
3. Roll a die: {1,2,3,4,5,6}
4. Draw and record opinion of a SRS of size 50 from a population
5. Time your commuting time in a weekday morning.
6. Score of A football game between two chosen teams
7. Give AIDS patient a treatment and record how long he/she lives.

- There are often more than one way to assign probability to a sample space. (could be infinitely many ways)
- Statistical techniques we will learn later help identify which one “reflecting the truth”.
- In Chap. 5 here we do not identify which probability is more “true” but just learn the consequences of a (legitimate) probability assignment.

legitimate probability

- P(A) has to be a number between 0 and 1, inclusive (rule 1)
- The probability of the sample space S must be 1: P(S) =1 (rule 2)

- When we are able to list all simple events,

O_i The summation over i for $P(O_i)$ must be 1

$$\Sigma P(O_i) = 1$$

P(A) could be computed by sum of $P(O_i)$ over those O's inside A.

Assigning Probabilities to Events

Example:

- – equally likely outcomes (classical approach)
- – relative frequency (using historical data)
- -- Subjective

- Monty Hall show: chance of winning is 1/3 or 1/2 ?
- Both are legitimate, but only one is true.

Equally Likely Approach

- The equally likely outcomes approach usually relies on symmetry/geometry to assign probabilities to events.
- Suppose that an experiment has only n *(distinct)* outcomes. The equally likely approach to probability assigns a probability of $1/n$ to each of the outcomes.
- Further, if an event A is made up of m *simple* outcomes, then $Prob(A) = P(A) = m/n$.

- Notice the notation: event is denoted by capital letter   A
- The probability assigned to this event is denoted by P(A) or Pr(A) or Prob(A)

- We say P(win) = 0.5   or P(win) = 0.65 etc.

Assign Probability: Equally Likely Approach

- Examples:

1. Roll a fair die, the sample space is

{ 1, 2, 3, 4, 5, 6 }

There are 6 outcomes in sample space, each one get a probability of 1/6.

– The probability of getting “5” is 1/6.

-- The probability of getting {“3” or above} is 4/6.

This does not mean that whenever you roll the die 6 times, you definitely get exactly one “5”

- One free throw by John Wall:
- Outcomes: {hit, miss}
- Two free throw: S= { hit-hit, hit-miss, miss-hit, miss-miss}

- Flip a fair coin twice: sample space =

  { (H, H), (H, T), (T, H), (T, T) }

There are four pairs. Each one gets ¼ probability.

Let A ={ two heads }, P(A) = ¼

Let A = {at least a head}, P(A) = ¾

- For a population of size 30, using SRS, we select two individuals.
- There are 435 different choices. i.e. There are 435 outcomes in the sample space.
- Each one has probability 1/435 = 0.00229

- How do I know there are 435 different choices?
- 435= (30x29)/2

- The first interview has 30 choices, the second has 29. (no repeats)
- But {John, Jane} is the same as {Jane, John}. This is the reason of dividing by 2

- Method of counting: number of ways to chose m out of n possible. (result not ordered)

- Lottery game “Pick 3”. You pick 3 digits, each range from 0 to 9
- There are 1000 different choices. Sample space has 1000 outcomes in it.

- To win the game “straight” you have to match all 3 number in the correct order.
- There is only one way (one outcome) to win. The probability is 1/1000 = 0.001

- For “straight”, a play cost you $1

- Do you still want to play?

- To win the game “front pair”, you need to match the first 2 out of 3 numbers.

- there are 9 ways to win “front pair”, [the 10 th would be the straight win, not counted as front pair] the probability is 9/1000. We write

$$P(\text{front pair}) = 9/1000$$

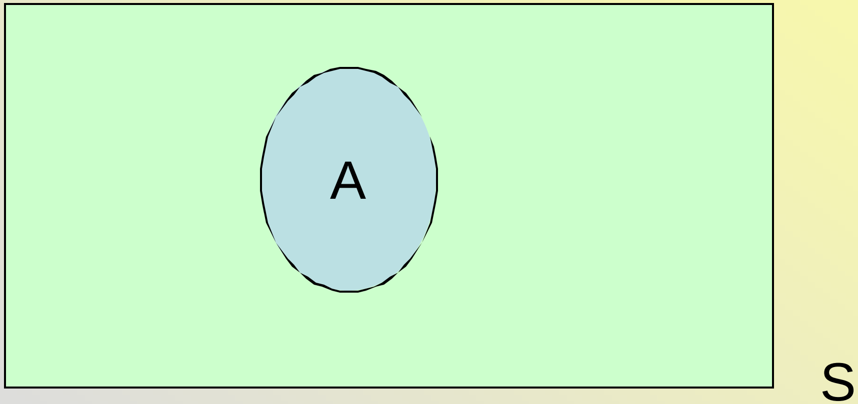
- Rule 3 P(A) = 1- P(A-complement)

$$P(A) = 1 - P(A^C)$$

Complement

- Let *A* denote an event.
- **The complement of an event** ***A*****:** All the outcomes in the sample space *S* that **do not belong** to the event *A*. The complement of *A* is denoted by A^C = green

A = blue



Law of Complements:

$$P(A) = 1 - P(A^c)$$

Homework 3

- The next online homework assignment will be available later today.
- Due Feb 11, (at 11 PM)

Attendance Survey Question 10

- On a 4”x6” index card
  - Please write down your name and section number
  - Today’s Question: Have You ever played

Pick 3 of Lotto Kentucky

Powerball

None of above