STAT 400 Lecture 2

Chapter 1.1 Review

The discipline of statistics

- Statistics find patterns and get rid of noise.
- Statistics deals with the ***collection and analysis of data.***
  - Collection of data: Experimental Design
  - Analysis of data: Statistical Inference.

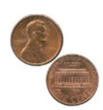
Sample & Population

- Population: The entire group of individuals that we want information about (e.g. the height of all the students in this classroom)

- Sample: The collection of the observations that are obtained from finite number of repeated trials. (e.g. measured the height of the 10 students in this room and record their height)

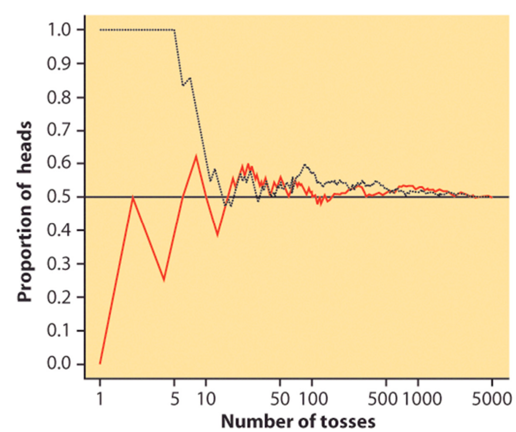
- Statistical Inference

Randomness and probability



- A experiment is **random** if individual outcomes cannot be predicted with certainty before the experiment is performed.
- For a random experiment, the fraction of times a certain outcome in a long run can be determined.
- The **probability** of any outcome of a random phenomenon can be defined as the proportion of times the outcome would occur in a very long series of repetitions.

Example: Coin Toss



Probability of heads is 0.5
= proportion of times you get
heads in many repeated trials

Probability models

Probability models mathematically describe the outcome of random processes. They consist of two parts:

1) ***S* = Sample Space:** This is a set, or list, of all possible outcomes of a random process.

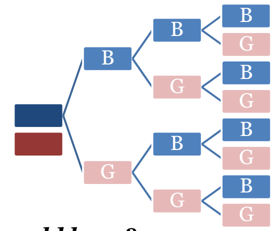
2) A **probability** for each possible event in the sample space S.

Example: Probability Model for a Coin Toss
S = {Head, Tail}
Probability of heads = 0.5
Probability of tails = 0.5

Sample space

- Important: It's the question that determines the sample space.

A. A couple wants 3 children.
What are the possible sequences of boys (B) and girls (G)?



BBB
BBG
GGB
BGG
GBB
GBG
GGB
GGG
=S

Note: S has 2^3=8 elements.

B. A couple wants 3 children.
What is the number of girls they could have?
- S = {0,1,2,3}

C. A researcher designs a new maze for lab rats. What are the possible outcomes for the time the rats will take to finish the maze (in minutes)?
- S = (0, ∞] = (all numbers > 0)

Random variable

- Mathematically, we use a variable to represent the possible outcomes of a random experiment. The variable is called ***random variable***.
- A random variable is usually denoted by soma capital letter, such as X, Y, Z
- A random variable can only take numerical values with each representing a possible outcome in the sample space.

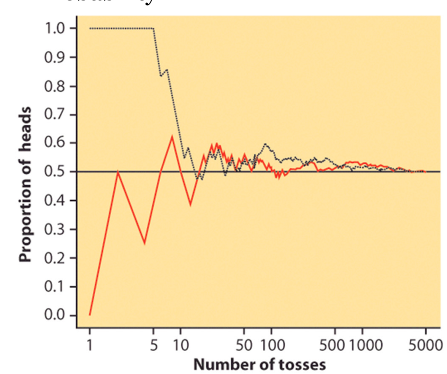
- *Example: Probability Model for a Coin Toss*
- *S* = {Head, Tail}
- Probability of heads = 0.5
- Probability of tails = 0.5

Let X=1 if the outcome is head
X=0 if the outcome is tail
We have
X=1 with Prob=0.5
X=0 with Prob=0.5

How to get the probability?

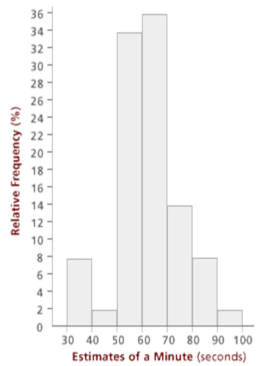
- Sample ⟶ Population
- Relative Frequency ⟶ Probability

- If certain outcome has occurred *f* times in these *n* trials; the *f* is called the frequency of the outcome and *f/n* is called the relative frequency of the outcome.
- *f/n* is not stable when *n* is small and is stable when *n* is big.



Relative Frequency Histogram

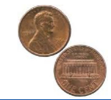
- The value of a random variable X is on the x-axis
- The y-axis shows the relative frequency of each possible outcome
- **Each outcome is present as a bar.**



How long is a minute?

Probability Mass Function

- As n increases, the relative frequency is closer and closer to the probability.



Number of tosses	Number of heads	Relative frequency to get heads
4	1	0.25
100	56	0.56
1000	510	0.510
10000	4988	0.4988

- The *probability* of an event *A*, denoted by *P*(*A*), can be considered as the long run relative frequency of the event *A*.

Chapter 1.2 Properties of Probability

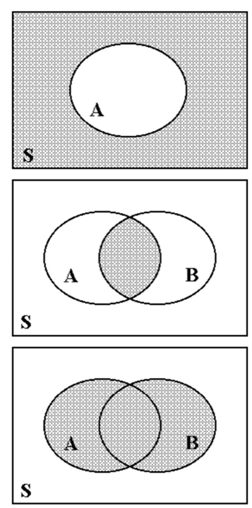
Event: An outcome or a set of outcomes of a random phenomenon, i.e. a subset of the sample space.

Event *A* occurs if we observe an outcome that is a member of the set *A*.

> Example: Toss a coin three times.
> *S* = ______________________.
> Let A be the event that we get exactly two tails.
> Then *A* = ______________________.
> Let C be the event that we get a head in the second toss.
> Then *C* = ______________________.

Null event: denoted by ∅.

Venn Diagrams



Complement of A

A'

(not A, $\overline{A}$, A^c)

contains all elements that are **not** in A

Intersection of A and B

$A \cap B$

(A and B, A B)

contains all elements that are in A **and** in B

Union of A and B

$A \cup B$

(A or B)

contains all elements that are either in A **or** in B or both

A and *B* are **mutually exclusive** (or disjoint) if $A \cap B = \varnothing$.

A and *B* are **mutually exhaustive** if $A \cup B = S$.

So, A and A' are mutually exclusive and exhaustive.

Example :Pick a person in this class at random

This is an Experiment.

Sample space: S = { all people in class }.

Let event A ="person is male" and event B ="person travelled by bike today". Suppose I pick a male who did not travel by bike. Say whether the following events have occurred:

- $A' \cup B$ _______.
- $(A \cap B)'$ _______.

Basic Probability Rules

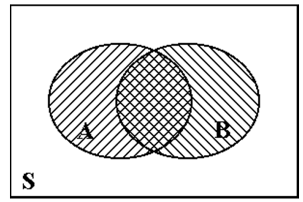
Theorem 1. $P(A') = 1 - P(A)$.

Theorem 2. $P(\varnothing) = 0$.

Theorem 3. If $A \subset B$, then $P(A) \le P(B)$.

Theorem 4. For any event A, $P(A) \le 1$.

!	For any event A, $0 \le P(A) \le 1$	!	$P(S) = 1$, where S is the sample space.	!



Theorem 5. (Addition Rule)

If A and B are any two events, then
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
If events A and B are mutually exclusive, then
$P(A \cup B) = P(A) + P(B)$

Basic Probability Rules

Theorem 6.

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Proof: Exercise.

Example: Suppose the die is loaded so that the probability of an outcome is proportional to the outcome, i.e.
$P(1) = p$, $P(2) = 2p$, $P(3) = 3p$, $P(4) = 4p$, $P(5) = 5p$, $P(6) = 6p$.
i)Find the value of p that would make this a valid probability model.
P(S)= $p + 2p + 3p + 4p + 5p + 6p = 21p = 1 \rightarrow p = 1/21$
ii) Let A = { the outcome is odd}, B = { the outcome is less than 5},
Find the probabilities P(A), P(B), P(A′), P(A ∩ B), P(A ∪ B).
Let X be the outcome.
P(A) = $9p$ = 9/21 , P(A ′) = 1 - P(A)= 12/21,
P(B) = 10 p = 10/21,
P(A ∩B)= P(X=1 or 3)= 4 p =4/21,
P(A ∪ B) = P(A) + P(B) - P(A ∩B)= 15/21.

Assigning Probabilities

- **Finite number of outcomes** : If the sample space S is finite, the probability of an event is the sum of the probabilities of the distinct outcomes making up the event.
- **Equally likely outcomes** : If a random phenomenon has k equally likely outcomes, each individual outcome has probability $\frac{1}{k}$.

Example: A card is drawn from an ordinary pack of 52 playing cards. What is the probability that the card is
- a seven?
- a spade?
- a club or a diamond?
- a club or a king?

Example: A couple want 3 children.

- What are the arrangements (ordered sequences) of boys (B) and girls (G)?
- Genetics tells us that the probability that a baby is a boy or a girl is the same, 0.5.
  → Sample space: {BBB, BBG, BGB, GBB, GGB, GBG, BGG, GGG}
  → All eight outcomes in the sample space are equally likely.
  → The probability of each is thus 1/8.

- What are the numbers (X) of girls they could have?
- The same genetic laws apply. We can use the probabilities above to calculate the probability for each possible number of girls.
  → Sample space {0, 1, 2, 3}
  → $P(X = 0) = P(\text{BBB}) = 1/8$
  → $P(X = 1) = P(\text{BBG or BGB or GBB}) = P(\text{BBG}) + P(\text{BGB}) + P(\text{GBB}) = 3/8$
  → $P(X = 2) = ?$
  → $P(X = 3) = ?$