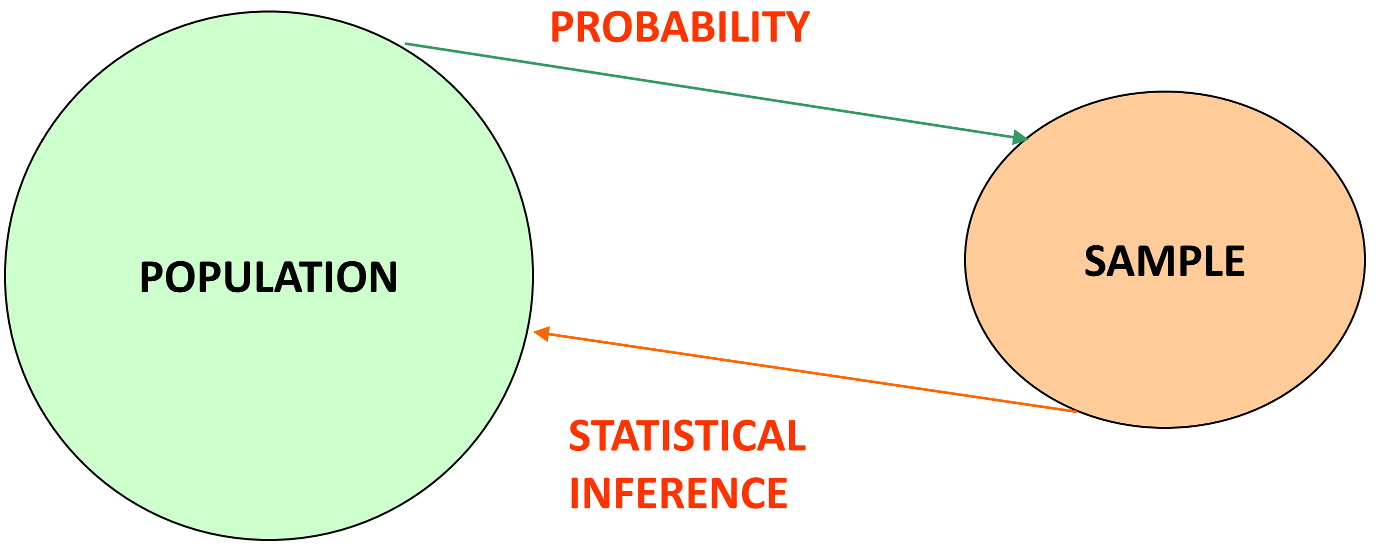
# IAM 530 ELEMENTS OF PROBABILITY AND STATISTICS

## LECTURE 2-INTRODUCTION TO PROBABILITY


PROBABILITY
POPULATION
SAMPLE
STATISTICAL
INFERENCE

# WHY PROBABILITY IS NEEDED?

- Nothing in life is certain. We can quantify the uncertainty using PROBABILITY
- A probability provides a quantitative description of the chances or likelihoods associated with various outcomes.
- It provides a bridge between descriptive and inferential statistics.

# WHAT IS PROBABILITY?

- CLASSICAL INTERPRETATION(Frequency of Occurrence)

  If a random experiment is repeated, the relative frequency for any given outcome is the probability of this outcome.

  **Probability of an event:** Relative frequency of the occurrence of the event in the long run.

  - Example: Probability of observing a head in a fair coin toss is 0.5 (if coin is tossed long enough).

- SUBJECTIVE INTERPRETATION(Indication of Uncertanity)

  The assignment of probabilities to event of interest is subjective.

  - Example: I am guessing there is 50% chance of raining today.

# BASIC CONCEPTS

- Experiment: is the process by which an observation (or measurement) is obtained.
- Random experiment: involves obtaining observations of some kind
- Population: Set of all possible observations.
- Elementary event (simple event): one possible outcome of an experiment
- Event (Compound event): one or more possible outcomes of a random experiment
- Sample space: the set of all outcomes for an experiment.

# EXAMPLES OF A RANDOM EXPERIMENT

| **Experiment** | **Outcomes** |
|---|---|
| Flip a coin | Heads and Tails |
| Record a statistics test marks | Numbers between 0 and 100 |
| Measure the time to assemble a computer | Numbers from zero and above |

SAMPLE SPACE

Countable

Finite number of elements

Uncountable (Continuous )

Infinite number of elements

# EXAMPLES

- Countable sample space examples:
  - Tossing a coin experiment

    $S$ : {Head, Tail}
  - Rolling a dice experiment

    $S$ : {1, 2, 3, 4, 5, 6}
  - Determination of the sex of a newborn child

    $S$ : {girl, boy}
- Uncountable sample space examples:
  - Life time of a light bulb

    $S : [0, \infty)$
  - Closing daily prices of a stock

    $S : [0, \infty)$

# ASSIGNING PROBABILITIES

– Given a sample space S ={$E_1$,$E_2$,…,$E_k$}, the following characteristics for the probability P($E_i$) of the simple event $E_i$ *must hold:*

1. $0 \le P(E_i) \le 1 \quad for \quad each \quad i$

2. $\sum_{i=1}^{k} P(E_i) = 1$

- The probability of an event A is equal to the sum of the probabilities of the simple events contained in A
- If the simple events in an experiment are **equally likely,** you can calculate

$$P(A) = \frac{\text{total outcomes in A}}{\text{total outcomes in S}}$$

EXAMPLE:

An urn contains 11 red balls and 3 white balls. Two balls are drawn. How likely is that both balls are red?

Suppose we model each ball by positive integer 1-11 for red balls, 12-14 for white balls.

S (the sample space): {(1,2),(1,3),....,(13,14)}

- The set S models all possible draws.
- 14*14-14=14*13=182 possibilities

E : event of two red balls.

E={(1,2),(1,3),....,(10,11)}

- Number of elements of E = 11*11-11=110

P: the probability of a finite set is the sum of all it’s elements

$$P(\mathrm{E}) = \sum_{s \in E} P(s) = \sum_{s \in E} \frac{1}{182} = \frac{110}{182}$$

# COUNTING RULES

- While forming the sample space tt some point, we have to stop listing and to use some counting rules.

- Methods to determine how many subsets can be obtained from a set of objects.

# THE m*n RULE

- If an experiment is performed in two stages, with **m** ways to accomplish the first stage and **n** ways to accomplish the second stage, then there are **mn** ways to accomplish the experiment.
- This rule is easily extended to **k** stages, with the number of ways equal to

$$\mathbf{n_1\, n_2\, n_3 \ldots\, n_k}$$

- **Example:** Toss two coins. The total number of simple events is: 2*2=4

- **Example:** Toss three dice. The total number of simple events is: 6*6*6=216

- **Example:** Two balls are drawn from a dish containing two red and two blue balls. The total number of simple events is: 4*3=12

# THE FACTORIAL

- number of ways in which objects can be permuted.

$$n! = n(n-1)(n-2)\ldots 2.1$$

$$0! = 1,\ 1! = 1$$

**Example:** Possible permutations of {1,2,3} are {1,2,3}, {1,3,2}, {3,1,2}, {2,1,3}, {2,3,1}, {3,2,1}. So, there are 3!=6 different permutations.

# PARTITION RULE

- There exists a single set of N distinctly different elements which is partitioned into k sets; the first set containing $n_1$ elements, …, the $k$-th set containing $n_k$ elements. The number of different partitions is

$$\frac{N!}{n_1!n_2!\cdots n_k!} \text{ where } N = n_1 + n_2 + \cdots + n_k.$$

- **Example:** Let’s partition {1,2,3} into two sets; first with 1 element, second with 2 elements.

Partition 1: {1} {2,3}

Partition 2: {2} {1,3}

Partition 3: {3} {1,2}

3!/(1! 2!)=3 different partitions

**Example:** How many different arrangements can be made of the letters “statistics”?

- N=10, n1=3 s, n2=3 t, n3=1 a, n4=2 i, n5=1 c

$$\frac{10!}{3!3!1!2!1!} = 50400$$

# PERMUTATIONS

- Any <u>ordered</u> sequence of r objects taken from a set of n distinct objects is called a **permutation** of size r of the objects.

$$P_r^n = \frac{n!}{(n-r)!}$$

where $n! = n(n-1)(n-2)...(2)(1)$ and $0! \equiv 1.$

- **Example:** How many 3-digit lock combinations can we make from the numbers 1, 2, 3, and 4?

$$P_3^4 = \frac{4!}{1!} = 4(3)(2) = 24$$

# COMBINATION

- Given a set of n distinct objects, any <u>unordered</u> subset of size r of the objects is called a **combination**.

$$C_r^n = \frac{n!}{r!(n-r)!}$$

- **Example:** Three members of a 5-person committee must be chosen to form a subcommittee. How many different subcommittees could be formed?

$$C_3^5 = \frac{5!}{3!(5-3)!} = \frac{5(4)(3)(2)1}{3(2)(1)(2)1} = \frac{5(4)}{(2)1} = 10$$

# COUNTING

| | Number of possible arrangements of size $r$ from $n$ objects | |
|---|---|---|
| | Without Replacement | With Replacement |
| Ordered | $\frac{n!}{(n-r)!}$ | $n^r$ |
| Unordered | $\binom{n}{r}$ | $\binom{n+r-1}{r}$ |

# ELEMENTARY SET OPERATIONS

## INTERSECTION

- The intersection of event A and B is the event that occurs when both A and B occur. It is denoted by $A \cap B$.
- The joint probability of A and B is the probability of the intersection of A and B, which is denoted by $P(A \cap B)$.

## UNION

- The union event of A and B is the event that occurs when either A or B or both occur. It is denoted by $A \cup B$.

$$P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$$

## COMPLEMENT

- The complement of event A (denoted by $A^C$) is the event that occurs when event A does not occur.

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

# PROBABILITY FUNCTION AXIOMS

Let S denote a non-empty or finite or countably infinite set and let $2^S$ denote the set of all subsets of S. A real-valued function P defined on $2^S$ is a *probability function if*

- For any event $E$, $0 \leq P(E) \leq 1$.
- $P(S) = 1$.
- If $E_1, E_2, \ldots$ are pairwise disjoint events, then

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

# THE CALCULUS OF PROBABILITIES

- If P is a probability function and A and B any sets, then

a. $P(B \cap A^C) = P(B) - P(A \cap B)$

b. If $A \subset B$, then $P(A) \leq P(B)$

c. $P(A \cap B) \geq P(A) + P(B) - 1$ (Bonferroni Inequality)

d. $\boldsymbol{P}\left(\bigcup_{\boldsymbol{i}=1}^{\infty} \boldsymbol{A}_{\boldsymbol{i}}\right) \leq \sum_{\boldsymbol{i}=1}^{\infty} \boldsymbol{P}\ \boldsymbol{A}_{\boldsymbol{i}}$ for any sets $A_1, \boldsymbol{A}_2, \cdots$

(Boole's Inequality)

# MUTUALLY EXCLUSIVE EVENTS

- When two events A and B are mutually exclusive or disjoint, , if A and B have no common outcomes.

$$A \cap B = \varnothing \text{ and } P(A \cap B) = 0$$

- The events $A_1, A_2, \ldots$ are pairwise mutually exclusive (disjoint), if

$$A_i \cap A_j = \varnothing \text{ for all } i \neq j.$$

# EQUALLY LIKELY OUTCOMES

- The same probability is assigned to each simple event in the sample space, S.
- Suppose that $S=\{s_1,\dots,s_N\}$ is a finite sample space. If all the outcomes are equally likely, then $P(\{s_i\})=1/N$ for every outcome $s_i$.

# ADDITION RULE

For any two events A and B

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

# COMPLEMENT

- We know that for any event **A:**
  - **$P(A \cap A^C) = 0$**
- Since either **A** or **$A^C$** must occur,

  **$P(A \cup A^C) = 1$**
- so that **$P(A \cup A^C) = P(A) + P(A^C) = 1$**

Then

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

# CONDITIONAL PROBABILITY

The probability that A occurs, given that event B has occurred is called the **conditional probability** of A given B and is defined as

$$\mathrm{P(A|B)} = \frac{P(A \cap B)}{P(B)} \quad if \quad P(B) \neq 0$$

$$0 \leq P(A|B) \leq 1$$

$$P(A|B) = 1 - P(A^C|B), \quad P(A|A) = 1$$

$$P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2|B)$$

Example: A red die and a blue die are thrown.

***A*** = { the red die scores a 6 }

***B*** = { at least one 6 is obtained on the two dice }

$$P(A) = \frac{6}{36} = \frac{1}{6} \text{ and } P(B) = \frac{11}{36}$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{P(A)}{P(B)}$$

$$= \frac{1/6}{11/36} = \frac{6}{11}$$

**Example:** A bowl contains five candies two red and three blue. Randomly select two candies, and define

- A: second candy is red.
- B: first candy is blue.

$$P(A|B) = P(2^{nd}\ \text{red}|1^{st}\ \text{blue}) = 2/4 = 1/2$$

$$P(A|\text{not } B) = P(2^{nd}\ \text{red}|1^{st}\ \text{red}) = 1/4$$

# INDEPENDENT VS. NON-INDEPENDENT EVENTS

- If A and B are **independent**, then

$$P(A \textit{ and } B) = P(A) \times P(B)$$

which means that conditional probability is:

$P(B \mid A) = P(A \text{ and } B) / P(A) = P(A)P(B)/P(A) = P(B)$

- We have a more general multiplication rule for events that are **not independent**:

$$P(A \text{ and } B) = P(B \mid A) \times P(A)$$

- In particular, we would like to know whether they are **independent**, that is, if the probability of one event is **not affected** by the occurrence of the other event.

Two events A and B are said to be **independent** if

- $P(A|B) = P(A)$

and

- $P(B|A) = P(B)$

# Example

Roll two dice

S=all possible pairs ={(1,1),(1,2),…,(6,6)}

- Let A=first roll is 1; B=sum is 7; C=sum is 8 P(A|B)=?; P(A|C)=?

**Solution:**

- P(A|B)=P(A and B)/P(B)

P(B)=P({1,6} or {2,5} or {3,4} or {4,3} or {5,2} or {6,1})

= 6/36=1/6

P(A|B)= P({1,6})/(1/6)=1/6 =P(A)

A and B are independent

- $P(A|C)=P(A \text{ and } C)/P(C)=P(\emptyset)/P(C)=0$

  A and C are disjoint

$P(C)=P(\{2,6\} \text{ or } \{3,5\} \text{ or } \{4,4\} \text{ or } \{5,3\} \text{ or } \{6,2\})$

$= 5/36$

# THE MULTIPLICATIVE RULE FOR INTERSECTIONS

- For any two events, A and B, the probability that both A and B occur is

$$P(A \cap B) = P(A)\ P(B \text{ given that A occurred})$$
$$= P(A)P(B|A)$$

- If the events A and B are independent, then the probability that both A and B occur is

$$P(A \cap B) = P(A)\ P(B)$$

# BAYES' THEOREM

- Suppose you have P(B|A), but need P(A|B).

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} \text{ for } P(B) \neq 0$$

- Can be generalized to more than two events.

# THE LAW OF TOTAL PROBABILITY

Let $B_1$ , $B_2$ , $B_3$ ,..., $B_k$ be mutually exclusive and exhaustive events (that is, one and only one must happen). Then the probability of any event A can be written as

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_k)$$

$$= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \ldots + P(B_k)P(A|B_k)$$

- Suppose that the events $B_1, B_2, B_3, \ldots, B_n$ partition the sample space for some experiment and that $A$ is an event defined on $S$. For any integer, $k$, such that $1 \leq k \leq n$ we have

$$P(B_k \mid A) = \frac{P(A \mid B_k) P(B_k)}{\sum_{j=1}^{n} P(A \mid B_j) P(B_j)}$$

**Example:** You are living in a dorm. One night the fire alarm goes off. How likely is it that there is a fire?

Here H is the event "there is a fire" and

E is the event "the fire alarm goes off."

You want to know P(H|E).

You estimate that all things being equal a fire is unlikely on a given night, setting P(H)=0.001 (roughly one fire in three years).

You know that in a typical semester of about 100 days there are about 3 fire alarms (typically false alarms), so you estimate P(E)=0.03.

- Finally you guess that it is nearly certain someone would set off the alarm if there really were a fire, so you estimate $P(E|H)=0.98$.

- By Bayes' Rule,

  $P(H|E)=P(H)P(E|H)/P(E)=(0.001)(0.98)/(0.03)= 0.033.$

- **Example:**A drug company has designed a test for a disease. Through extensive testing, the company reports that the test produces only 1% false positive results (i.e., a healthy person tests positive) and only 2% false negative results (i.e., a person with the disease tests negative).
- Let P be the event "someone tests positive,"
- N be the event "someone tests negative,"
- H be the event "someone is healthy," and
- D be the event "someone has the disease."
- Then the company is reporting $P(P|H)=0.01$ (or equivalently $P(N|H)=0.99$) and $P(N|D)=0.02$ (or equivalently $P(P|D)=0.98$).

Suppose you test positive for the disease. How likely is it that you in fact have the disease?

It is tempting but incorrect to say 98% since $P(P|D)=0.98$.

But you want to know $P(D|P)$, which may be quite different. It turns out you do not have enough information yet. Oddly enough you must also know $P(D)$, the prevalence of the disease in your population.

Suppose the disease is rare, occurring in only 0.05% of the population. Then applying the second form of Bayes' we get

$$P(D \mid P) = \frac{P(D)P(P \mid D)}{P(D)P(P \mid D) + P(H)P(P \mid H)}$$

$$= \frac{0.0005 * 0.98}{0.0005 * 0.98 + 0.9995 * 0.01} \approx 0.047 = 4.7\%$$