SMA 6304 / MIT 2.853 / MIT 2.854
Manufacturing Systems
Lecture 3: Probability

Lecturer: Stanley B. Gershwin

Slides by Larry Wein and Stan Gershwin

Trick Question

I flip a coin 100 times, and it shows heads every time.

Question: What is the probability that it will show heads on the next flip?

Probability and Statistics

Probability $\neq$ *Statistics*

Probability: mathematical theory that describes uncertainty.

Statistics: set of techniques for extracting useful information from data.

Interpretations of probability

Frequency

The probability that the outcome of an experiment is A *is* $P(A)$

if the experiment is performed a large number of times and the fraction of times that the observed outcome is A is $P(A)$.

Interpretations of probability

Parallel universes

The probability that the outcome of an experiment is A *is* $P(A)$

if the experiment is performed in each parallel universe and the fraction of universes in which the observed outcome is A is $P(A)$.

State of belief

The probability that the outcome of an experiment is A *is* $P(A)$

if that is the opinion of an observer *before* the experiment is performed.

Interpretations of probability

Abstract measure

The probability that the outcome of an experiment is A *is* $P(A)$

if $P()$ satisfies a set of conditions.

Axioms of probability

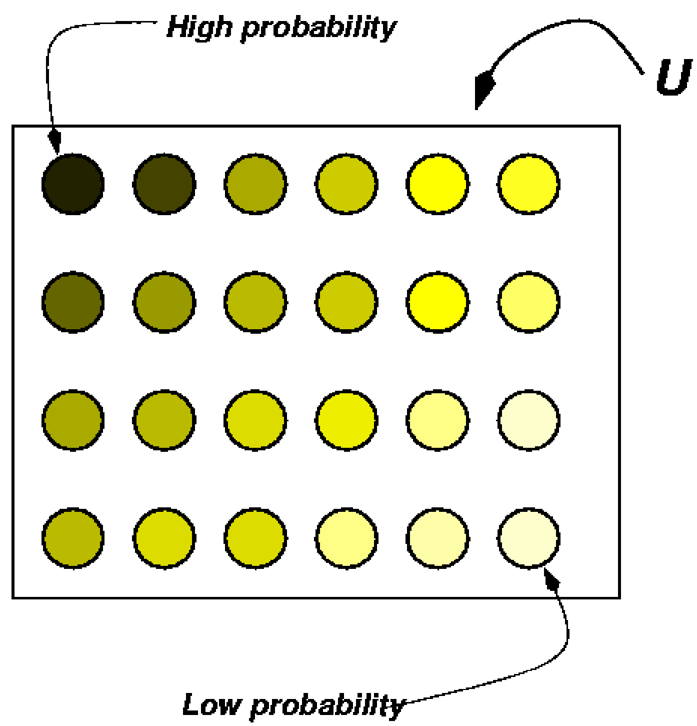
Let U be a set of *samples*. Let E_1, E_2, ... be subsets of U. Let ϕ be the *null set* (the set that has no elements).

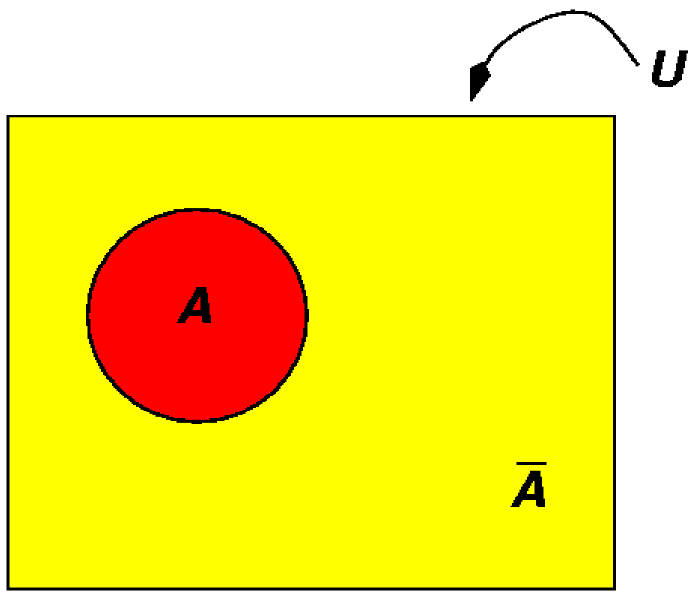
- $0 \leq P(E_i) \leq 1$
- $P(U) = 1$
- $P(\phi) = 0$
- If $E_i \cap E_j = \phi$, then $P(E_i \cup E_j) = P(E_i) + P(E_j)$

- Subsets of U are called *events.*

- $P(E)$ is the *probability* of E.

Discrete Sample Space



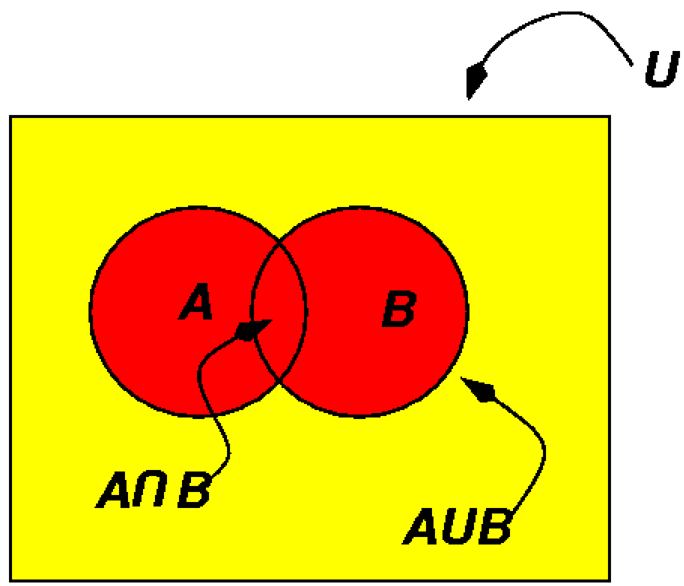


$$P(\bar{A}) = 1 - P(A)$$

Probability Basics

Set Theory

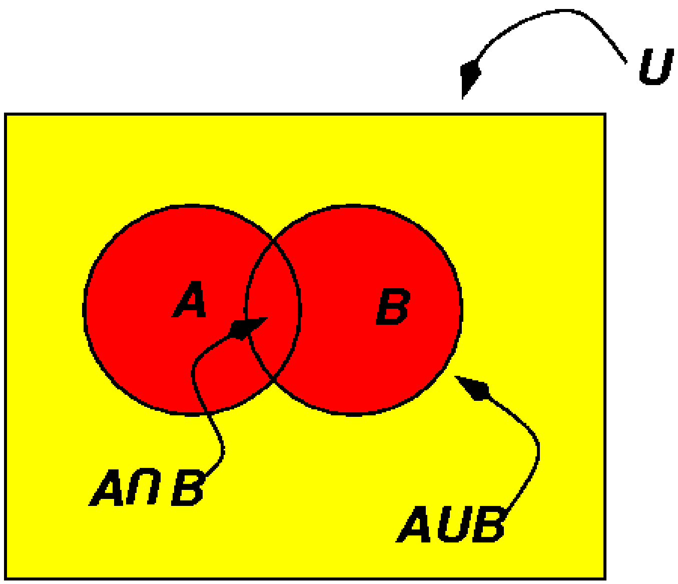
Venn diagrams



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

Conditional Probability

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



We can also write $P(A \cap B) = P(A|B)P(B)$.

Throw a die.

- A is the event of getting an odd number (1, 3, 5).
- B is the event of getting a number less than or equal to 3 (1, 2, 3).

Then $P(A) = P(B) = 1/2$ and
$P(A \cap B) = P(1,3) = 1/3$.

Also, $P(A|B) = P(A \cap B)/P(B) = 2/3$.

Law of Total Probability

- Let $B = C \cup D$ and assume $C \cap D = \phi$. We have $P(A|C) = \dfrac{P(A \cap C)}{P(C)}$ and $P(A|D) = \dfrac{P(A \cap D)}{P(D)}$.

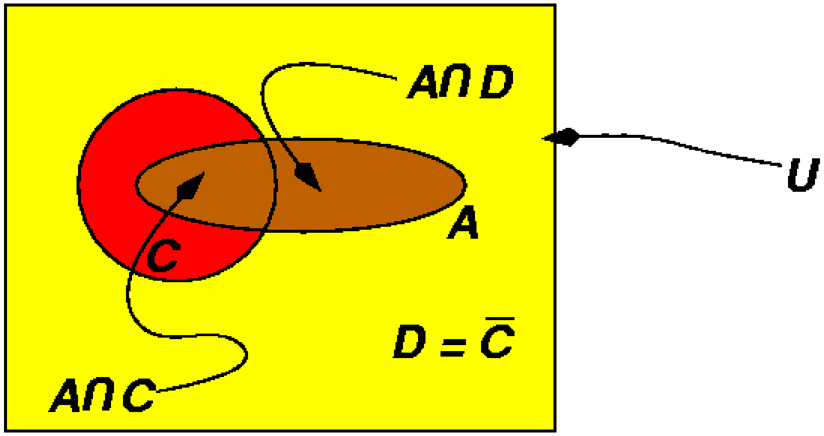
- Also $P(A \cap B) = P(A \cap (C \cup D)) =$
  $P(A \cap C) + P(A \cap D) - P(A \cap (C \cap D)) =$
  $P(A \cap C) + P(A \cap D)$

- Since $P(A|B)$ prob $(B) = P(A|C)P(C) + P(A|D)P(D)$ we have $P(A|B) = P(A|C)P(C|B) + P(A|D)P(D|B)$.

Law of Total Probability

An important case is when $C \cup D = B = U$, so that $A \cap B = A$. Then

$$P(A) = P(A \cap C) + P(A \cap D) =$$
$$P(A|C)P(C) + P(A|D)P(D).$$



Let V be a vector space. Then a *random variable* X is a mapping (a function) from U to V.

If $\omega \in U$ and $x = X(\omega) \in V$, then X is a random variable.

A <u>random variable</u> is a function that assigns a numerical value to each possible outcome of a probability experiment

Example 1: X = **the number that comes up on the roll of a dice**

Example 2: X = **the number of heads in 3 flips of a coin**

Outcome	Probability	Value of X
(h,h,h)	.125	X = 3
(h,h,t)	.125	X = 2
(h,t,h)	.125	X = 2
(h,t,t)	.125	X = 1
(t,h,h)	.125	X = 2
(t,h,t)	.125	X = 1
(t,t,h)	.125	X = 1
(t,t,t)	.125	X = 0

Example

Pseudo-random number generator

A *pseudo-random number generator* is a set of random variables $X_1(\omega), X_2(\omega), \ldots$ where there is a function F such that

$$X_{n+1}(\omega) = F(X_n(\omega))$$

and F is such that the sequence of $X_n(\omega)$ satisfies certain conditions.

Pseudo-random number generators are used extensively in *simulation*.

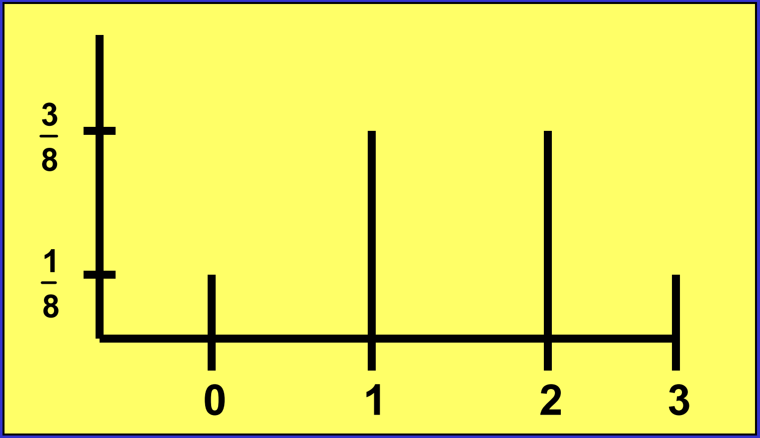
Random variables are either

- **discrete:** can only assume a finite set of values
- **continuous:** can take on any value within some interval of real numbers

Today: discrete random variables
Next Class: continuous random variables

The probability distribution of a discrete random variable X is the possible values $x_1, \ldots, x_n$ and corresponding probabilities $p_i, \ldots, p_n$ where $\sum_{i=1}^{n} p_i = 1$

Example 3:



Number of heads in 3 flips

Summary Measures

Mean =

E [X] = expected value of X

= mean of X

= the average outcome

= μ_X

E [X] = $\sum_{i=1}^{n} x_i P(X = x_i)$

Var (x) = variance of x

= σ_x^2

= $E\left[(x - \mu_x)^2\right]$

= average squared deviation from the mean

= $\sum_{i=1}^{n} P(X = x_i)(x_i - \mu_x)^2$

Example 3: X = number of heads in 3 flips of a coin

$$E(X) = \sum_i x_i P(X = x_i) = 1(\tfrac{3}{8}) + 2(\tfrac{3}{8}) + 3(\tfrac{1}{8}) = 1.5$$

$$\sigma^2(x) = \sum_i P(X = x_i)(x_i - \mu_x)^2 =$$

$$= \frac{1}{8}(0-1.5)^2 + \frac{3}{8}(1-1.5)^2 + \frac{3}{8}(2-1.5)^2 + \frac{1}{8}(3-1.5)^2 = \frac{3}{4}$$

Variance is in "squared" units, so we take square root

$\sigma(x) = \sqrt{\sigma^2(x)} =$ standard deviation of x

The bigger $\sigma \Rightarrow$ the more spread out the distribution

$\Rightarrow$ the more uncertainty in the random variable

Linear Functions of a Random Variable

Y=aX+b (a and b are known numbers)

$$E[f(X)]=\sum_i p_i f(x_i)$$

$$E(Y)=\sum_i p_i(ax_i+b)=a\sum_i p_i x_i+b\sum_i p_i$$

$$=aE(X)+b$$

$$Var(Y)=E(aX+b-[a\mu_x+b])^2$$

$$=E(a[X-\mu_x])^2$$

$$=\sum_i p_i(a[x_i-\mu_x])^2$$

$$=a^2\sum_i p_i(x_i-\mu_x)^2$$

$$=a^2Var(x)$$

Binomial Distribution

Each trial is

- a success with probability P
- a failure with probability 1- P

X= number of successes in **n** independent trials

$$P(X = x) = \frac{n!}{x!(n-x)!} p^{x} (1-p)^{n-x} \qquad x = 0, \ldots, n$$

$$E(x) = np \qquad Var(x) = np(1-p)$$

Example: **X** = number of heads in 3 tosses of a coin

$$n = 3 \quad P = \frac{1}{2}$$

$$P(X = x) = \frac{3!}{x!(3-x)!} \left(\frac{1}{2}\right)^{3}$$

$$P(X = 0) = P(X = 3) = \frac{1}{8}$$

$$P(X = 1) = P(X = 2) = \frac{3}{8}$$

Manufacturing application: X=number of defective parts in a lot

Two Random Variables

$$X = x_i \quad \text{and} \quad Y = y_i \quad \text{with probability} \quad P_i$$

COV(x,y)= covariance of X and Y

$$= E\left[\left(X - \mu_x\right)\left(Y - \mu_y\right)\right]$$

$$= \sum_{i=1}^{n} p_i \left(x_i - \mu_x\right)\left(y_i - \mu_y\right)$$

if X and Y are independent then COV (X,Y) = 0

if X and Y tend to vary in the same direction, then COV (X,Y) > 0

if X and Y tend to vary in the opposite direction, then COV (X,Y) < 0

$$E\left(aX + bY\right) = aE\left(X\right) + bE\left(Y\right)$$

VAR(X+Y)= VAR(X) + VAR(Y) + 2COV(X,Y)

if X and Y are independent, then *VAR(X+Y)= VAR(X)+VAR(Y)*

CORR(X,Y) = correlation of X and Y

$$= \frac{COV(X,Y)}{\sigma_x \sigma_y} \in \left[-1, 1\right]$$

is “NORMALIZED COVARIANCE”