Quiz I Review
Probabilistic Systems Analysis
6.041SC

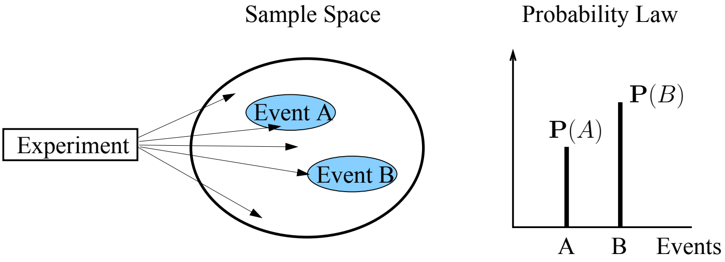
Massachusetts Institute of Technology

Quiz Information

- Content: Chapters 1-2, Lecture 1-7, Recitations 1-7, Psets 1-4, Tutorials 1-3

A Probabilistic Model:

- **Sample Space**: The set of all possible outcomes of an experiment.
- **Probability Law**: An assignment of a nonnegative number **P**(E) to each event E.



Probability Axioms

Given a sample space Ω:

1. **Nonnegativity:** $\mathbf{P}(A) \geq 0$ for each event A
2. **Additivity:** If A and B are disjoint events, then

$$\mathbf{P}(A \cup B) = P(A) + P(B)$$

   If $A_1, A_2, \ldots,$ is a sequence of disjoint events,

$$\mathbf{P}(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$$

3. **Normalization** $\mathbf{P}(\Omega) = 1$

Properties of Probability Laws

Given events A, B and C:

1. If $A \subset B$, then $\mathbf{P}(A) \leq \mathbf{P}(B)$
2. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$
3. $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$
4. $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(A^c \cap B) + \mathbf{P}(A^c \cap B^c \cap C)$

Discrete Models

- **Discrete Probability Law:** If Ω is finite, then each event $A \subseteq \Omega$ can be expressed as

$$A = \{s_1, s_2, \ldots, s_n\} \qquad s_i \in \Omega$$

Therefore the probability of the event A is given as

$$\mathbf{P}(A) = \mathbf{P}(s_1) + \mathbf{P}(s_2) + \cdots + \mathbf{P}(s_n)$$

- **Discrete Uniform Probability Law:** If all outcomes are equally likely,

$$\mathbf{P}(A) = \frac{|A|}{|\Omega|}$$

Conditional Probability

- Given an event B with $\mathbf{P}(B) > 0$, the conditional probability of an event $A \subseteq \Omega$ is given as

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

- $\mathbf{P}(A|B)$ is a valid probability law on Ω, satisfying
  1. $\mathbf{P}(A|B) \geq 0$
  2. $\mathbf{P}(\Omega|B) = 1$
  3. $\mathbf{P}(A_1 \cup A_2 \cup \cdots |B) = \mathbf{P}(A_1|B) + \mathbf{P}(A_2|B) + \cdots$, where $\{A_i\}_i$ is a set of disjoint events
- $\mathbf{P}(A|B)$ can also be viewed as a probability law on the restricted universe B.

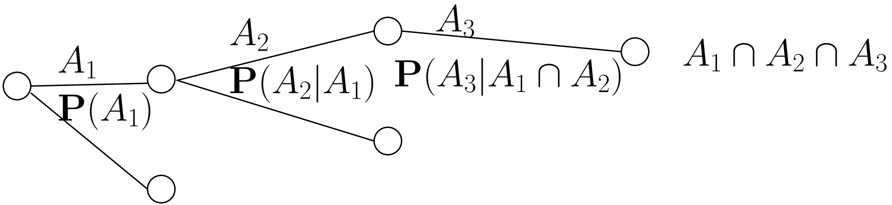
Multiplication Rule

- Let $A_1, \ldots, A_n$ be a set of events such that

$$\mathbf{P}\left(\bigcap_{i=1}^{n-1} A_i\right) > 0.$$

Then the joint probability of all events is

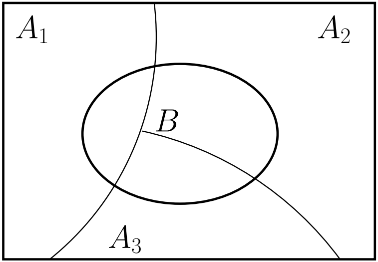
$$\mathbf{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2) \cdots \mathbf{P}(A_n| \bigcap_{i=1}^{n-1} A_i)$$

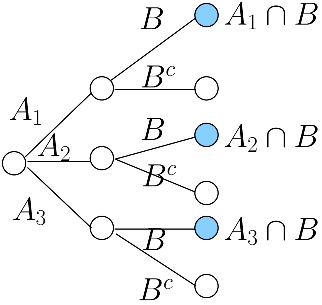


Total Probability Theorem

Let $A_1, \ldots, A_n$ be disjoint events that partition Ω. If $\mathbf{P}(A_i) > 0$ for each i, then for any event B,

$$\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbf{P}(B|A_i)\mathbf{P}(A_i)$$

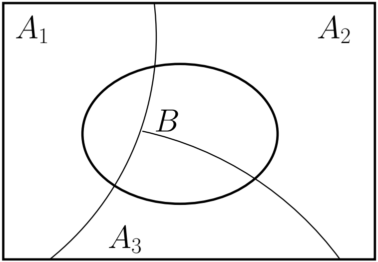


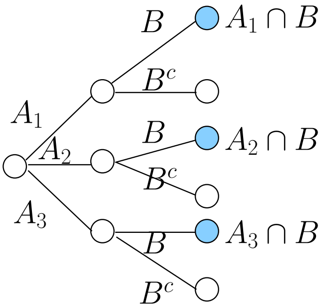


Bayes Rule

Given a finite partition $A_1, \ldots, A_n$ of Ω with $\mathbf{P}(A_i) > 0$, then for each event B with $\mathbf{P}(B) > 0$

$$\mathbf{P}(A_i|B) = \frac{\mathbf{P}(B|A_i)\mathbf{P}(A_i)}{\mathbf{P}(B)} = \frac{\mathbf{P}(B|A_i)\mathbf{P}(A_i)}{\sum_{i=1}^{n} \mathbf{P}(B|A_i)\mathbf{P}(A_i)}$$





Independence of Events

- Events A and B are **independent** if and only if

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$$

  or

$$\mathbf{P}(A|B) = \mathbf{P}(A) \quad \text{if } \mathbf{P}(B) > 0$$

- Events A and B are **conditionally independent** given an event C if and only if

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

  or

$$\mathbf{P}(A|B \cap C) = \mathbf{P}(A|C) \quad \text{if } \mathbf{P}(B \cap C) > 0$$

- Independence $\not\Leftrightarrow$ Conditional Independence.

Independence of a Set of Events

- The events $A_1, \ldots, A_n$ are **pairwise independent** if for each $i \neq j$

$$\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$$

- The events $A_1, \ldots, A_n$ are **independent** if

$$\mathbf{P}\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} \mathbf{P}(A_i) \quad \forall\ S \subseteq \{1, 2, \ldots, n\}$$

- Pairwise independence $\not\Rightarrow$ independence, but independence $\Rightarrow$ pairwise independence.

Counting Techniques

- **Basic Counting Principle:** For an m-stage process with n_i choices at stage i,
$$\# \text{ Choices} = n_1 n_2 \cdots n_m$$
- **Permutations:** k-length sequences drawn from n distinct items without replacement (order is important):
$$\# \text{ Sequences} = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$
- **Combinations:** Sets with k elements drawn from n distinct items (order within sets is not important):
$$\# \text{ Sets} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Counting Techniques-contd

- **Partitions**: The number of ways to partition an n-element set into r disjoint subsets, with n_k elements in the k^{th} subset:

$$\binom{n}{n_1, n_2, \ldots, n_r} = \binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_r-1}{n_r}$$
$$= \frac{n!}{n_1!n_2!, \cdots, n_r!}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$
$$\sum_{i=1}^{r} n_i = n$$

Discrete Random Variables

- A **random variable** is a real-valued function defined on the sample space:

$$X : \Omega \to R$$

- The notation $\{X = x\}$ denotes an event:

$$\{X = x\} = \{\omega \in \Omega | X(\omega) = x\} \subseteq \Omega$$

- The **probability mass function (PMF)** for the random variable X assigns a probability to each event $\{X = x\}$:

$$p_X(x) = \mathbf{P}(\{X = x\}) = \mathbf{P}\big(\{\omega \in \Omega | X(\omega) = x\}\big)$$

PMF Properties

- Let X be a random variable and S a countable subset of the real line
- The axioms of probability hold:
  1. $p_X(x) \geq 0$
  2. $\mathbf{P}(X \in S) = \sum_{x \in S} p_X(x)$
  3. $\sum_x p_X(x) = 1$
- If g is a real-valued function, then $Y = g(X)$ is a random variable:

$$\omega \xrightarrow{X} X(\omega) \xrightarrow{g} g(X(\omega)) = Y(\omega)$$

with PMF

$$p_Y(y) = \sum_{x | g(x) = y} P_X(x)$$

Expectation

Given a random variable X with PMF $p_X(x)$:

- $\mathbf{E}[X] = \sum_x x p_X(x)$
- Given a derived random variable $Y = g(X)$:

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x) = \sum_y y p_Y(y) = E[Y]$$

$$\mathbf{E}[X^n] = \sum_x x^n p_X(x)$$

- **Linearity** of Expectation: $\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$.

Variance

The expected value of a derived random variable $g(X)$ is

$$\mathbf{E}[g(X)] = \sum_x g(x) p_X(x)$$

The variance of X is calculated as

- $var(X) = \mathbf{E}[(X - \mathbf{E}[X])^2] = \sum_x (x - \mathbf{E}[X])^2 p_X(x)$
- $var(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2$
- $var(aX + b) = a^2 var(X)$

Note that $var(x) \geq 0$

Multiple Random Variables

Let X and Y denote random variables defined on a sample space Ω.

- The **joint PMF** of X and Y is denoted by

$$p_{X,Y}(x,y) = \mathbf{P}\big(\{X = x\} \cap \{Y = y\}\big)$$

- The **marginal PMFs** of X and Y are given respectively as

$$p_X(x) = \sum_y p_{X,Y}(x,y)$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y)$$

Functions of Multiple Random Variables

Let $Z = g(X, Y)$ be a function of two random variables

- **PMF:**

$$p_Z(z) = \sum_{(x,y)|g(x,y)=z} p_{X,Y}(x, y)$$

- **Expectation:**

$$\mathbf{E}[Z] = \sum_{x,y} g(x, y) p_{X,Y}(x, y)$$

- **Linearity:** Suppose $g(X, Y) = aX + bY + c$.

$$\mathbf{E}[g(X, Y)] = a\mathbf{E}[X] + b\mathbf{E}[Y] + c$$

Conditioned Random Variables

- Conditioning X on an event A with $\mathbf{P}(A) > 0$ results in the PMF:

$$p_{X|A}(x) = \mathbf{P}\big(\{X = x\}|A\big) = \frac{\mathbf{P}\big(\{X = x\} \cap A\big)}{\mathbf{P}(A)}$$

- Conditioning X on the event $Y = y$ with $\mathbf{P}_Y(y) > 0$ results in the PMF:

$$p_{X|Y}(x|y) = \frac{\mathbf{P}\big(\{X = x\} \cap \{Y = y\}\big)}{\mathbf{P}\big(\{Y = y\}\big)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

Conditioned RV - contd

- Multiplication Rule:

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y)$$

- Total Probability Theorem:

$$p_X(x) = \sum_{i=1}^{n} \mathbf{P}(A_i)p_{X|A_i}(x)$$
$$p_X(x) = \sum_{y} p_{X|Y}(x|y)p_Y(y)$$

Conditional Expectation

Let X and Y be random variables on a sample space Ω.

- Given an event A with $\mathbf{P}(A) > 0$

$$\mathbf{E}[X|A] = \sum_x x p_{X|A}(x)$$

- If $P_Y(y) > 0$, then

$$\mathbf{E}[X|\{Y = y\}] = \sum_x x p_{X|Y}(x|y)$$

- **Total Expectation Theorem:** Let $A_1, \ldots, A_n$ be a partition of Ω. If $\mathbf{P}(A_i) > 0 \ \forall i$, then

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

Independence

Let X and Y be random variables defined on Ω and let A be an event with $\mathbf{P}(A) > 0$.

- X is independent of A if either of the following hold:

$$p_{X|A}(x) = p_X(x) \ \forall x$$
$$p_{X,A}(x) = p_X(x)\mathbf{P}(A) \ \forall x$$

- X and Y are independent if either of the following hold:

$$p_{X|Y}(x|y) = p_X(x) \ \forall x \forall y$$
$$p_{X,Y}(x,y) = p_X(x)p_Y(y) \ \forall x \forall y$$

Independence

If X and Y are independent, then the following hold:

- If g and h are real-valued functions, then $g(X)$ and $h(Y)$ are independent.
- $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$ (inverse is not true)
- $var(X + Y) = var(X) + var(Y)$

Given independent random variables $X_1, \ldots, X_n$,

$$var(X_1 + X_2 + \cdots + X_n) = var(X_1) + var(X_2) + \cdots + var(X_n)$$

Some Discrete Distributions

	X	$p_X(k)$	$\mathbf{E}[X]$	$var(X)$
Bernoulli	$\begin{cases} 1 & \text{success} \\ 0 & \text{failure} \end{cases}$	$\begin{cases} p & k=1 \\ 1-p & k=0 \end{cases}$	p	$p(1-p)$
Binomial	Number of successes in n Bernoulli trials	$\binom{n}{k} p^k (1-p)^{n-k}$ $k = 0, 1, \ldots, n$	np	np(1-p)
Geometric	Number of trials until first success	$(1-p)^{k-1} p$ $k = 1, 2, \ldots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Uniform	An integer in the interval [a,b]	$\begin{cases} \frac{1}{b-a+1} & k = a, \ldots, b \\ 0 & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)(b-a+2)}{12}$