Lecture 3

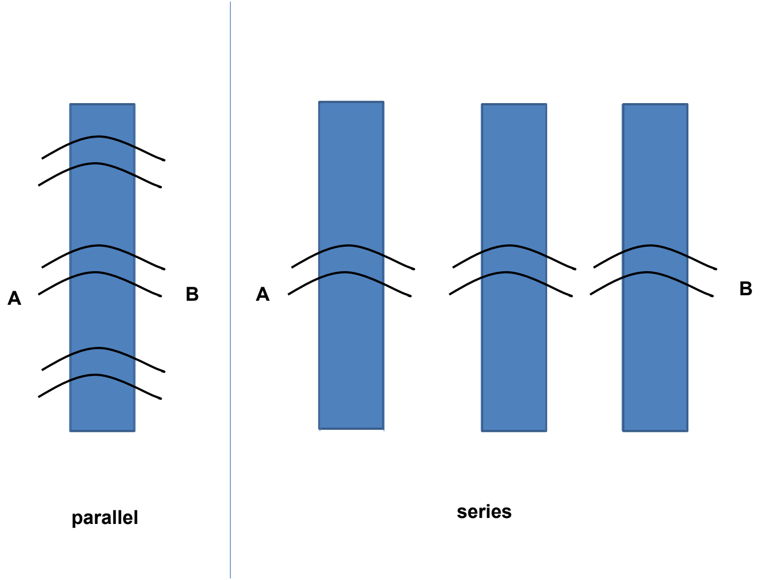
- **Read:** Chapter 1.8-1.10, 2.1-2.2.
- Experiments, Models, and Probabilities
  - Using Independence in Modeling
  - Counting
- Discrete Random Variables
  - Definitions
  - Probability Mass Function

Using Independence in Modeling

Reliability: Complex systems are often simply modeled by assuming that they consist of several *independent* components

- Consider a network with n links.
- For link i, event $W_i =$ link i is operational and is independent of other components
- Probability that a link i is working well: $P[W_i] = p_i$

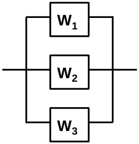
Parallel vs. Series Connections: Bridge Analogy



Using Independence in Modeling (cont.)

- What is probability that the network is reliable for communications?
  - Series connection: $P[\cap_i W_i] = p_1 p_2 \cdots p_n$
    Overall reliability is worse than any of the components

  - Parallel connection:
    $P[\text{system is up}] = 1 - P[\cap_i W_i^c] = 1-(1-p_1)(1-p_2)\cdots(1-p_n)$
    Reliability is better than any of the components, for small p_i, roughly sum

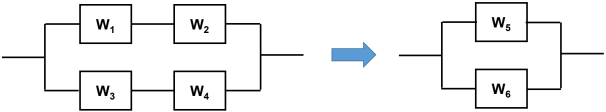


  - Mixture: Same story, can decompose the problem and reduce the series and parallel combinations

Example 1.44

- An operation consists of two redundant parts. The first part has two components in series (W_1 and W_2) and the second part has two components in series (W_3 and W_4). All components succeed with probability p$= 0.9$. Draw a diagram of the operation and calculate the probability that the operation succeeds.

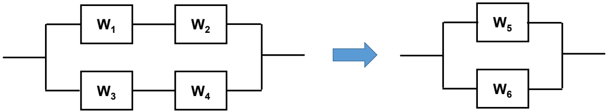
- A diagram of the operation is shown below.



- We can create an equivalent component, W_5, with probability of success p_5 by observing that for the combination of W_1 and W_2,

$$P[W_5] = p_5 = P[W_1 W_2] = p^2 = 0.81$$

Example 1.44 (cont.)



- Similarly, the combination of W_3 and W_4 in series produces an equivalent component, W_6, with probability of success $p_6 = p_5 = 0.81$.
- The entire operation then consists of W_5 and W_6 in parallel which is also shown in the figure.
- The success probability of the operation is

$$P[W] = 1 - (1 - p_5)^2 = 0.964$$

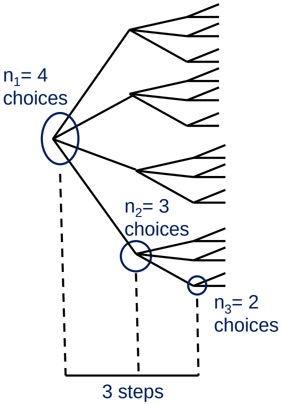
- We could consider the combination of W_5 and W_6 to be an equivalent component W_7 with success probability $p_7 = 0.964$ and then analyze a more complex operation that contains W_7 as a component.

Counting

Motivation: Discrete Uniform Law: Recall that for such a law, all sample points are equally likely, and computing probabilities reduces to just counting, i.e.,

$$P[A] = \frac{\text{number of elements of } A}{\text{total number of sample points}}$$

Fundamental Principle of Counting



- r steps with n_i choices at each step i
- total number of choices $= n_1 \times n_2 \times \ldots \times n_r$

Fundamental Principle of Counting: Examples

- number of license plates with 3 letters followed by 4 digits$=$ $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10$
  - if repetition is prohibited: $26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7$
- number of subsets of a set with n elements$= 2^n$

Permutations

- k-permutation: an ordered sequence of k distinguishable objects
- $(n)_k =$ no. of possible k-permutations of n distinguishable objects:

$$(n)_k = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$$

  This follows from the fundamental counting principle.

- Example: number of words with 4 distinct letters$=$ $\frac{26!}{22!} = 26 \times 25 \times 24 \times 23$

Sampling without Replacement

- Choosing objects from a collection is also called sampling, and the chosen objects are known as a sample.
- A k-permutation is a type of sample obtained by specific rules for selecting objects from the collection.
- In particular, once we choose an object for a k-permutation, we remove the object from the collection, and we cannot choose it again.
- Consequently, this is called sampling without replacement.
- Example: number of license plates with 3 letters followed by 4 digits if repetition is prohibited: $26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 \cdot 7$

Sampling with Replacement

- A second type of sampling occurs when an object can be chosen repeatedly.
- In this case, when we remove the object from the collection, we replace the object with a duplicate.
- This is known as sampling with replacement.
- We want to choose with replacement a sample of k objects out of a collection of n distinguishable objects.
- Sampling with replacement ensures that in each subexperiment needed to choose one of the k objects, there are n possible objects to choose.
- Hence, there must be n^k ways to choose with replacement a sample of k objects.
- Example: number of license plates with 3 letters followed by 4 digits$= 26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10$

Sampling When the Order Does Not Matter

- Both in choosing a k-permutation or in sampling with replacement, different outcomes are distinguished by the order in which we choose objects.
- However, in many practical problems, the order in which the objects are chosen makes no difference.
- For example, in a bridge hand, it does not matter in what order the cards are dealt.
- Suppose there are four objects, A, B, C, and D, and we define an experiment in which the procedure is to choose two objects, arrange them in alphabetical order, and observe the result.
- In this case, to observe AD, we could choose A first or D first or both A and D simultaneously.
- What we are doing is picking a subset of the collection of objects.
- Each subset is called a k-combination.

Combinations

- k-combination:
  - Pick a subset of k out of n objects
  - Order of selection does not matter
  - Each subset is a k-combination
- $\binom{n}{k} =$ no. of possible k-element subsets (i.e., order is not important) that can be obtained out of a set of n distinguishable objects
- **Remark:** In a combination, there is no ordering involved, e.g., 2-permutations of {A, B, C} are AB, AC, BA, CA, BC, CB, while the combinations of 2 out of the 3 letters would be AB, AC, BC

How Many Combinations

- $\binom{n}{k}$ = "n choose k" = no. of possible k-element subsets (i.e., order is not important) that can be obtained out of a set of n distinguishable objects:
- To find $\binom{n}{k}$, we perform the following two subexperiments to assemble a *k-permutation* of n distinguishable objects:
    1. Choose the k items at once (a *k-combination* out of n objects).
    2. Choose an ordering for the k items (a *k-permutation* of the k objects in the *k-combination*).
- The number of outcomes in the combined experiment is $(n)_k$.
- The first subexperiment has $\binom{n}{k}$ possible outcomes (the number we have to derive).
- The second experiment has $(k)_k = k!$ possible outcomes.

$$(n)_k = \binom{n}{k} k! \quad \Rightarrow \quad \binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}$$

How Many Combinations (cont.)

- We encounter $\binom{n}{k}$ in other mathematical studies
- Sometimes it is called a binomial coefficient because it appears (as the coefficient of $x^k y^{n-k}$) in the expansion of the binomial form $(x + y)^n$

Example: Independent Trials and Binomial Probabilities

- n independent coin tosses, $P[H] = p$
- What is the probability of obtaining an n-sequence with k heads?
- $P[HTTHHH] = p^4(1-p)^2$
- $P[sequence] = p^{\#heads} \cdot (1-p)^{\#tails}$

$$
\begin{aligned}
P[k \text{ heads}] &= \sum_{k-\text{head seq.}} P[seq.] \\
&= p^k \cdot (1-p)^{n-k} \cdot (\#\text{ of k-head seqs.}) \\
&= \binom{n}{k} p^k \cdot (1-p)^{n-k}
\end{aligned}
$$

Partitions

How many ways are there to divide a set of n distinct elements into r disjoint sets of $n_1, n_2, ..., n_r$ elements each, with $n_1, n_2, ..., n_r \leq n$?

- A combination of k elements out of n breaks up into two disjoint sets of elements of size k and $n - k$
- Note: $\binom{n}{k} = \#$ of ways of breaking n elements into subsets of size k and $n - k$ each
- Using the fundamental principle of counting, we can answer the above question as

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\ldots\binom{n-n_1-n_2-\ldots-n_{r-1}}{n_r}$$
$$= \frac{n!}{n_1!n_2!\ldots n_r!} = \binom{n}{n_1, n_2, \ldots, n_r}$$

This is called the multinomial coefficient.

Partitions: Example

- Anagrams: How many letter sequences can be obtained by rearranging the word TATTOO?
- Think of an anagram for this word as including 6 slots, which need to be filled with the letters T, A, or O.
- Each solution corresponds to selecting 3 slots for T, 2 slots for O, and 1 slot for A.
- How many ways are there to divide the 6 slots into such a partition?
- Answer: $\frac{6!}{1!2!3!} = 60$
- Note: The letters are not distinct, but the slots are!

Problem 1.8.6

A basketball team has

- 3 pure centers, 4 pure forwards, 4 pure guards
- one swingman who can play either guard or forward.

A pure player can play only the designated position.
How many lineups are there (1 center, 2 forwards, 2 guards)?

Problem 1.8.6 Solution

Three possibilities:

1. swingman plays guard: N_1 lineups
2. swingman plays forward: N_2 lineups
3. swingman does not play: N_3 lineups

$$N = N_1 + N_2 + N_3$$

Problem 1.8.6 Solution (cont.)

We need (1 center, 2 forwards, 2 guards) for each lineup.

center / forward / guard

$$N_1 = \binom{3}{1}\binom{4}{2}\binom{4}{1} = 72 \text{ (swingman plays guard)}$$

$$N_2 = \binom{3}{1}\binom{4}{1}\binom{4}{2} = 72 \text{ (swingman plays forward)}$$

$$N_3 = \binom{3}{1}\binom{4}{2}\binom{4}{2} = 108 \text{ (swingman does not play)}$$

Multiple Outcomes

- Consider n independent trials
- Each having r possible trial outcomes $(s_1, ..., s_r)$
- Such that $P[\{s_k\}] = p_k$

Multiple Outcomes (2)

- Outcome is a sequence:
  - Example: $s_3 s_4 s_3 s_1$

$$\begin{aligned} P[s_3 s_4 s_3 s_1] = p_3 p_4 p_3 p_1 &= p_1 p_3^2 p_4 \\ &= p_1^{n_1} p_2^{n_2} p_3^{n_3} p_4^{n_4} \end{aligned}$$

- where n_i denotes the number of times s_i arose in the sequence
- Probability depends on how many times each outcome occurred

Multiple Outcomes (3)

Let $N_i =$ no. of times s_i occurs. Then,

$$\begin{aligned} P[N_1 = n_1, ..., N_r = n_r] &= M p_1^{n_1} p_2^{n_2} ... p_r^{n_r} \\ &= \binom{n}{n_1, ..., n_r} p_1^{n_1} p_2^{n_2} ... p_r^{n_r} \end{aligned}$$

$$\begin{aligned} M &= \text{Multinomial Coefficient} \\ &= \frac{n!}{n_1! n_2! ... n_r!} \end{aligned}$$

Service Facility Design

- c: service capacity of a facility
- n: # of customers it is assigned
- p: prob. customer requires service
- $N =$ number of customers requiring service
- Given n, p choose c
- Criterion?
  - We establish a probability such that $\mathrm{P}[\mathrm{N} > \mathrm{c}] <$ this probability
  - We choose the smallest c such that $\mathrm{P}[\mathrm{N} > \mathrm{c}]$ is still $<$ this probability

Card Play

- 52-card deck, dealt to 4 players, i.e., 13 cards each
- Find P[each gets an ace]

- Count size of the sample space: Partition the 52 card deck into 4 sets of 13 cards each
$$\binom{52}{13, 13, 13, 13} = \frac{52!}{13!13!13!13!}$$
- Count number of ways of distributing the four aces (One ace in each player's hand)
$$4 \times 3 \times 2$$
- Count number of ways of dealing the remaining 48 cards (Each hand has an additional 12 cards)
$$\binom{48}{12, 12, 12, 12} = \frac{48!}{12!12!12!12!}$$
- So, P[each gets an ace] $=$
$$\frac{\#\text{ hands with one ace each}}{\#\text{ hands in sample space}} = \frac{4 \times 3 \times 2 \cdot \frac{48!}{12!12!12!12!}}{\frac{52!}{13!13!13!13!}}$$

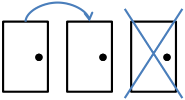
N People

- N people in the class
- Each person can pick heads or tails
- A person will win if only 1 person selects heads

  $\rightarrow$ Ethernet network $\rightarrow$ what is the optimal strategy?

- The probability that a given person wins is the probability that one person picks heads and that the remaining $N-1$ people do not pick heads.
- The probability that a given person selects heads is p; the probability that the remaining people do not select heads is $(1-p)^{N-1}$.
- Therefore, the probability a given person wins is $p(1-p)^{N-1}$.
- Since there are N people, the probability that any one of the N people wins is $Np(1-p)^{N-1}$.

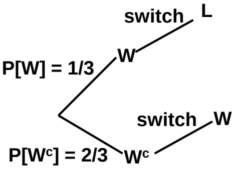
Three Doors



- Three doors, one has a prize behind it
- You pick one, what is the chance you win?
- After you pick your door, I open one of the others and show you there is no prize behind it
- Should you change your choice? (Yes, you should change!)
- If so, how? What is the probability you win?

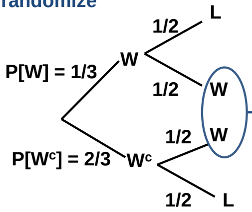
Three Doors (cont.)

W = {first door selected was a winner}



Had a low probability of winning;
by switching you win

If you randomize



(1/3 x 1/2) + (2/3 x 1/2) = 1/2

Birthday Problem

- Suppose class has n students
- What is the probability that **at least two** have a common birthday?
- Assume that for each person, any of the 365 days is equally likely to be their birthday (*Ignore leap years, and assume equal likelihood*)

Birthday Problem (cont.)

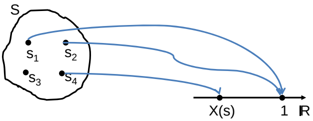
- $A =$ event that at least two have common birthday
- $A^c =$ each birthday is on a different day
- $P[A] = 1 - P[A^c]$, so let us compute $P[A^c] = ?$ (*This is easier.*)
- One person can have a birthday on any one of the 365 days $\rightarrow 365^n$ equally likely outcomes, *so computing probability boils down to counting problem!*
- Number of outcomes with no common birthday $\rightarrow 365 \times 364 \times ...(365 - n + 1)$
- $P[A^c] = \frac{365 \times 364 \times ...(365 - n + 1)}{365^n}$
- e.g., $n = 4 \rightarrow P[A] = 0.016$; $n = 32 \rightarrow P[A] = 0.753$; $n = 56 \rightarrow P[A] = 0.988$

What is a Random Variable?

- A number associated with an experiment, e.g.,
  - Experiment = throw two 4-sided dice
  - Number = the maximum of the two throws
- Different outcomes give different numbers; hence, the name "random"
- **Note:** nothing really random about a function, except that it is a function of outcomes that have probabilities of occurring
- Mathematically: A RV is a function that maps the points of the sample space to real numbers, e.g., $X(s) : S \mapsto \mathbb{R}$
- **Example:** Assign a real number to each outcome
- **Intuition:** what you can measure/observe about an experiment. Dealing with numbers makes analysis simpler. An RV represents a view on what is going on (e.g., the university entrance exam score of a student).
- RVs can be discrete, continuous, or a mixture

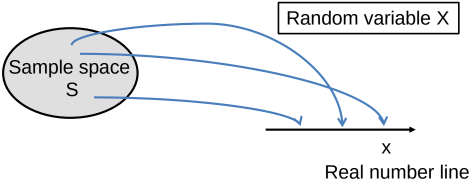
Random Variable as a Function

- sample space S
- an outcome $s \in S$
- an event $A \subset S$
- Probability measure assigns a number between [0,1] to each event
  - $P : A \mapsto P[A]$
  - satisfies three axioms
- random variable: X assigns a real number to each outcome
  - $X : S \mapsto \mathbb{R}$ (not necessarily a 1-1 function)
    $s \mapsto X(s)$

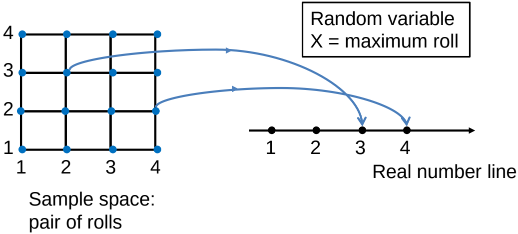


Visualization of a Random Variable

- Rolls of the die

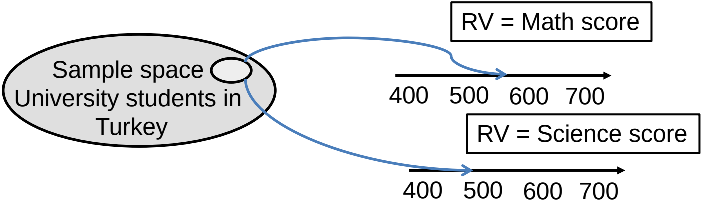


- If the outcome of the experiment is (4, 2) the value of this random variable is 4.

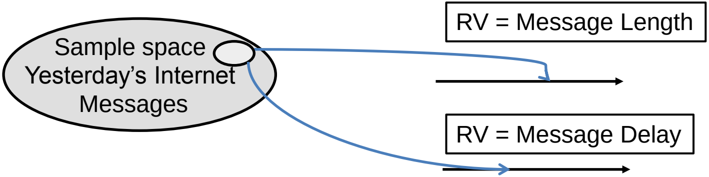


Examples of Random Variables

- Select a university student at random out of a given population: $X =$ math score, $Y =$ combined score



- Select a message from a collection of Internet messages at random: $X =$ length, $Y =$ delay



Random Variables

- **Experiment:** Procedure + Observation
- Observation is a particular outcome
- **Random variable:** Assign a real number to each outcome

  Possibilities: An RV will be denoted by X

  1. RV may be the observation
     e.g., roll of die
  2. RV is a function of the observation
     e.g., {heads, tails}
     X: heads $\rightarrow \pi$
     tails $\rightarrow 0$
  3. RV could be a function of another RV
     e.g., $Y = cos(X)$

Roadmap- Concept about Random Variables

- Is an RV just a function?
- No, there is always an underlying probabilistic model, i.e., (1) sample space, events and (2) probability law. Then,
  - RVs are functions of the outcome of an experiment.
  - A function of an RV is another RV - called derived RV.
  - RVs have certain averages of interest, e.g., mean and variance.
  - RVs can be conditioned on or independent of events or other RVs, i.e., these change the underlying probability law.
  - RVs can be discrete, continuous, or mixed.
- Most of these concepts will follow directly from the underlying probability model.

Discrete Random Variable (RV)

- discrete random variable: RV such that its range S_X is countable
- S_X = range of X (set of possible values X can take)
- S_X is discrete $\Rightarrow$ S_X has a countable number of elements
  e.g., $S_X = \{1, 2, 3, 4, 5, 6\}$ ✓
  $S_X = \mathbb{Z}$ (set of integers) ✓
  $S_X = [0, 1]$ is not a countable set ✗

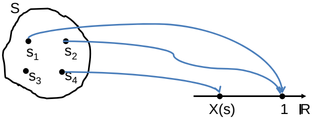
Probability Mass Function (PMF)

- A discrete RV has probability mass function (PMF)

$$\underbrace{p_X(x)}_{PMF} = P[X = x] = P[\underbrace{\{s \in S | X(s) = x\}}_{\text{this is an event, i.e., a subset of } S}]$$

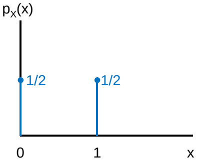
  $(= \text{Prob. of event } \{X = x\},\ x \in S_X\ )$

- Example: $p_X(1) = P[X = 1] = P[\{s_1, s_2\}]$



- Example: Suppose $S_X = \{0, 1\}$ $\qquad p_X(0) = 1/2$

$$p_X(1) = 1/2$$



PMFs and How We Compute Them

Convention: uppercase characters for RVs, lowercase characters for numerical values the RV can take.

To compute PMF $p_X(x)$:

1. Pick an x; collect all samples that give rise to $X = x$
2. Add their probabilities
3. Repeat for all x

Properties of PMFs

- $x \in S_X, \qquad p_X(x) \geq 0$
- $\sum_{x \in S_X} p_X(x) = 1$
- For an event $B \subset S_X$,

$$P[B] = P[x \in B] = \sum_{x \in B} p_X(x)$$

This follows by the additivity axiom because each event {X=x} is disjoint.
(e.g., $B = \{0, 1\}$)