CENG 551 Probability and Stochastic Processes for Engineers

Week # 3
Counting & Random Variables

Dr. Deniz Ozdemir
Deniz.ozdemir@yasar.edu.tr
Web: dozdemir.yasar.edu.tr

Plan for the Session

- More on Independence
- Counting

- Random variables
- Probability mass function (pmf)
  - Binomial Random Variable
- Expectation
  - Example

2

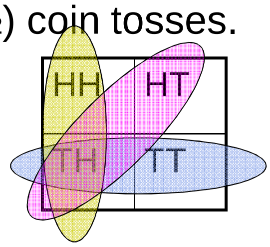
Independence

- Simply put $P(A|B) = P(A)$
- This implies that

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

Conditioning may affect independence

- Assume A and B are independent:
- If we are told that C occurred, are A and B MAY NOT BE independent.
- Two independent fair (p=½) coin tosses.



  - Event *A*: First toss is H
  - Event *B*: Second toss is H
  - Event *C*: The two outcomes are different.

Independence of a Collection of Events

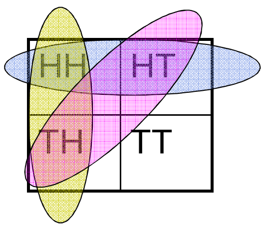
- Intuitive definition:
  - Information about some of the events tells us nothing about probabilities related to remaining events.
  - Example: $P(A_1 \cup (A_2 \cap A_3) | A_4 \cap A_5) = P(A_1 \cup (A_2 \cap A_3))$

- Mathematical definition:
  - For any distinct $i, j, \ldots, q$

$$P(A_i \cap A_j \cap \ldots \cap A_q) = P(A_i)\, P(A_j) \ldots P(A_i)$$

Independence vs. Pairwise Independence

- Example 1 revisited
  - Two independent fair (p=½) coin tosses.
  - Event *A*: First toss is H
  - Event *B*: Second toss is H
  - Event *C*: The two outcomes are different.
  - **P**(*A*) = **P**(*B*) = **P**(*C*) =1/2



- P($A \cap C$) =
- P(C|A $\cap$ B) =

- Pairwise independence **does not** imply independence.

True or False?

- If $P(A|B) = P(A)$, then $P(B|A^c) = P(B)$.
- If 10 out 10 independent fair coin tosses resulted in tails, the events “first toss was tails” and “10th toss was tails” are independent.
- If 5 out 10 independent fair coin tosses resulted in tails, the events “first toss was tails” and “10th toss was tails” are independent.
- If the events $A_1,\ldots,A_n$ form a partition of the sample space, and if B, C are some other events, then

$$P(B \mid C) = \sum_i P(A_i \mid C) P(B \mid A_i)$$

Discrete Uniform Law

- Let all sample points be equally likely
- Then,

$$P(A) = \frac{\text{number of elements of } A}{\text{total number of sample points}}$$

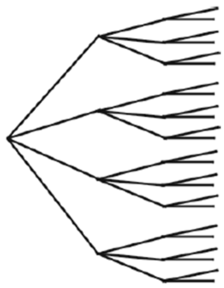
- Just count …

Basic Counting Principle

- r steps
- n_i choices at step i
- Number of choices is:

$$n_1 n_2 \ldots n_r$$

- Number of license plates with 3 letters a 4 digits =
- …if repetition is prohibited =
- **Permutations**: Number of ways of ordering n elements is =



Example

- Probability that six rolls of a six-sided die all give different numbers?
  - Number of outcomes that make the event happen =
  - Number of elements in the sample space=
  - Answer=

Combinations

- $\binom{n}{k}$: number of k-element subsets of a given n element set.
- Two ways of constructing an ordered sequence of k **distinct items:**
  - Choose the k items one at a time: $n(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$ choices
  - Choose k items, then order them (k! possible orders) $\binom{n}{k} k! = \frac{n!}{(n-k)!}$ $\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}$

Combinations: Some Properties

- $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r} \quad 1 \le r \le n$
- $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$
- Identity: $\sum_{k=0}^{n} \binom{n}{k} =$

Basic Counting Principle

Number of Possible Arrangements of Size r from n Objects:

	Without Replacement	With Replacement
Ordered:	$\frac{n!}{(n-r)!}$	n^r
Unordered:	$\binom{n}{r}$	$\binom{n+r-1}{r}$

Binomial Probabilities

- n independent coin tosses
  - $P(\text{H}) = p$
  - $P(\text{HTTHH}) =$
  - $P(\text{sequence}) = p^{\#\text{ of heads}} (1-p)^{\#\text{ of tails}}$
  - $P(k \text{ heads}) =$

$= \sum_i k\text{-head seq}_i.\ P(\text{seq}_i)$

$= (\#\ k\text{-head seq})\ p^k (1-p)^{n-k}$

$= \binom{n}{k} p^k (1-p)^{n-k}$

Birthday Problem

- At a gathering of s randomly chosen students what is the probability that at least 2 will have the same birthday?

P(at least 2 have same birthday)=1-P(all s students have different birthdays).

Assume 365 days in a year. Think of students' birthdays as a sample of these 365 days.

The total number of possible outcomes is:

$N=365^s$ (ordered, with replacement)

The number of ways that s students can have different birthdays is

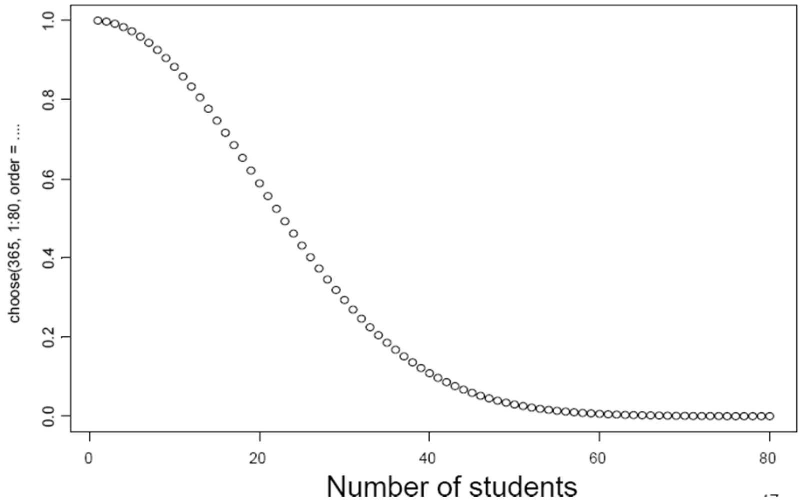
$M=364!/(365-s)!$ (ordered, without replacement)

P(all s students have different birthdays) =

M / N.

Birthday Problem

Probability that all students have different birthdays



Coin Tossing Problem

- Event B: 3 out of 10 tosses were "heads".
  - What is the (conditional) probability that the first 2 tosses were heads, given that B occurred?
- All outcomes in conditioning set B are equally likely:
  - Probability: $p^3(1-p)^7$
  - Conditional probability law is uniform.
- Number of outcomes in B:
- Out of the outcomes in B, how many start with HH?

Partitions

- 52-card deck, dealt to 4 players.
- Find P(each gets an ace)
- Count size of the sample space (possible combination of "hands")
- Count number of ways of distributing the four aces: $4 \cdot 3 \cdot 2$
- Count number of ways of dealing the remaining 48 cards
- Answer:

Another Counting Rule

- The number of ways of classifying n items into k groups with r^i in group i, $r^1+r^2+\ldots+r^k=n$, is:
- $n! / (r^1! \, r^2! \, r^3! \ldots r^k!)$

Hogwart's Problem

- How many ways are there to assign 100 incoming students to the 4 houses at Hogwarts?

~1.6(10^{57})

Random Variables

- Before this we talked about "Probabilities" of events and sets of events where in many cases we hand selected the set of fine grain events that made up an event whose probability we were seeking. Now we move onto another more interesting way to group this point: using a function $p_x(x)$ to ascribe values to every point in a sample space (discrete or continuous)
- One example might be the number of heads in 3 tosses of a coin.

Random Variables

- An assignment of a value (number) to every possible outcome.
- Mathematically: A function from the sample space to the real numbers:
  - –Discrete or Continuous
- Can have several random variables defined on the same sample space

Random Variables

- A random variable (r.v.) associates a unique numerical value with each outcome in the sample space.
- Notation:
  - Random Variable X
  - Experimental Value x
- Example: 1 coin toss. Define X:
  - $X = \begin{cases} 1 \text{ if } x = \text{H} \\ 0 \text{ if } x = \text{T} \end{cases}$ i.e. $X(\text{H}) = 1$, $X(\text{T}) = 0$
- Example: $Y = g(X)$

Random Variables

- Temperature in DC on Feb 1.
- Length of queue at Migros
- Amount of water in a “tall Americano”
- The number of points Efes score in a game they win
- The number of words in your emails

Discrete Random Variables

- **Discrete random variables:** number of possible values is finite or countably infinite: $x_1, x_2, x_3, x_4, x_5, x_6, \ldots$
- *Probability mass function (p.m.f.)*
  - $f(x) = P(X = x)$ (Sum over all possible values =1 always)
- *Cumulative distribution function (c.d.f)*
  - $F(x) = P\ (X \leq x) = \Sigma_{k \leq x} f(k)$

Probability Mass Function

- ("probability law",
- "probability distribution")
- Notation: $$p_X(x) = \mathrm{P}(X = x)$$
- **Example**: X= number of coin tosses until first head
  - Assume independent tosses, $\mathrm{P}(H) = p > 0$
  - $p_X(k) = P(X = k)$

    $= P(TT\ldots TH)$

    $=(1\text{-}p)^{k\text{-}1}p, \qquad k=1,2,..$

Geometric probability distribution With parameter p

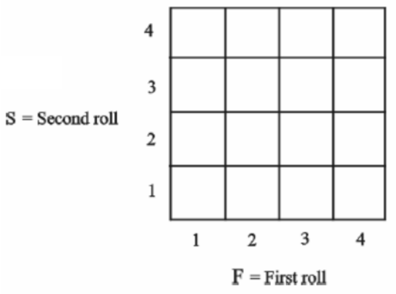
How to compute a pmf $p_X(x)$

- Collect all possible outcomes for which X is equal to x: $\{ \omega \in \text{Sample Space} \mid X(\omega)=x \}$
- Add their probabilities.
- Repeat for all x.

- **Example**: Two independent throws of a fair tetrahedral die:
  - F: outcome of first throw
  - S: outcome of second throw

  $L = min\ (F, S)$

  $p_L(2) =$



Probability Mass Function

- $p_x(x_0)$ probability that the experimental value of a random variable X obtained on a performance of the experiment is equal to x_0
- The same story value of Probability Mass Function (pmf). It can extend up to more dimensions which then allows for conditional pmfs

$\sum_i p_x(x_0)$ $\qquad 0 <= p_x(x_0) <= 1$

Binomial pmf with p and n

- X: number of heads in n independent coin tosses
  - $P(H) = p > 0$
- Let $n = 4$
- $p_X(2) =$

$$= 6p^2(1\text{-}p)^4$$
$$= \binom{4}{2} p^2(1\text{-}p)^4$$

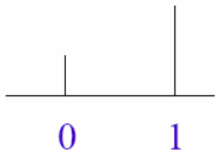
- In general:

$$p_X(k) = \binom{n}{k} p^k(1\text{-}p)^{n\text{-}k}, \quad k = 1,2,..,n$$

Selected Discrete Distributions:

- Single Coin flip P(heads/success) = p
  - *Probability of success*

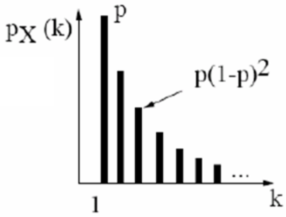
$$f(x) = P(X = x) = \begin{cases} p & \text{if } x = 1 \text{ (success)} \\ 1 - p & \text{if } x = 0 \text{ (failure)} \end{cases}$$



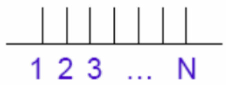
  - *"Bernoulli Distribution"*
- Multiple coin flips
  - x successes out of n trials
  - $f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 1,\ldots, n$
  - *"Binomial distribution"*

Selected Discrete Distributions:

- Number of failures before the first success
  - $f(x)=P(X=x) = (1-p)^{x-1}p, \quad x=1,2,..$
  - *“Geometric Distribution”*



- N equally likely events
  - $f(x)=P(X=x) = 1/N, \quad x=1,2,..$
  - *“Uniform distribution”*



Selected Discrete Distributions:

- *Hypergeometric distribution*
  - Drawing balls from the box without replacing the balls
- *Poisson Distribution*
  - number of occurrences of a rare event
- *Multinomial Distribution*
  - more than two outcomes
- *Negative Binomial Distribution*
  - number of trials to get r successes

Expectation

- The <u>expected value</u> or <u>mean</u> of a discrete r.v. X, denoted by $\mathrm{E}(X)$, is defined as:

$$\boxed{\mathrm{E}[X] = \sum_x x \cdot p_X(x)}$$

- Interpretations:
  - Center of gravity of pmf.
  - Average in large number of repetitions of the experiment. (to be substantiated later in this course)
- This is essentially a weighted average of the possible values the r.v. can assume, weights = $f(x)$

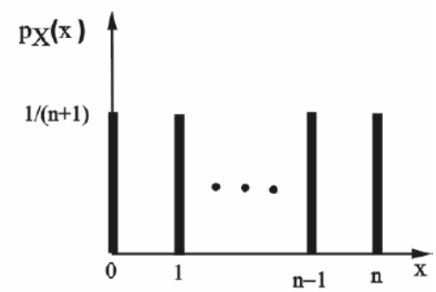
Expectation

- Example: Uniform on 0, 1, …, n

$$\mathrm{E}[X] = \sum_x x \cdot p_X(x)$$

- E[X] =

$$= 0 \times \frac{1}{n+1} + 1 \times \frac{1}{n+1} + \cdots + n \times \frac{1}{n+1}$$



Properties of Expectations -1

- Let X be a r.v. and let $Y = g(X)$
  - $E(Y) = \sum_y y \cdot p_Y(y)$
  - $E(Y) = \sum_x g(x) \cdot p_X(x)$
- "Second Moment": $E(X^2) = \sum_x x^2 p_X(x)$
- Caution: In general,

$$E[g(X)] \neq g(E[x])$$

Expected Values

- $E(x)$ given a p.m.f. provides some sense of the center of mass of the pmf.
- Variance is another measure that provides some measure of the distribution of a pmf/pdf around its expected value.
- Variance: $\mathrm{var}(X) = E[(X\text{-}E[X])^2]$

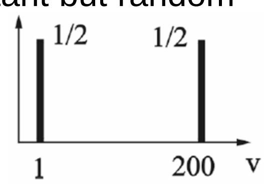
$$= \sum_x (x - E[X])^2 p_X(x)$$

Properties of Expectations -2

- If α, β are constants, then:
  - $\mathrm{E}[\alpha] =$
  - $\mathrm{E}[\alpha X] =$
  - $\mathrm{E}[\alpha X + \beta] =$
  - $\mathrm{E}[X - E(X)] =$

Average Speed vs. Average Time

- Traverse a 200 mile distance at constant but random speed V:



- $d = 200$, $T = t(V) = 200/V$
  - $\mathrm{E}[V] = 1.\frac{1}{2} + 200.\frac{1}{2} = 100.5$
  - $\mathrm{E}[T] = \sum_v t(v) p_V(v)$

$$= \frac{200}{1}.\frac{1}{2} + \frac{200}{2}.\frac{1}{2} = 100.5$$

Average Speed vs. Average Time vs. Expected Distance

- E[d] =
- E[T] $\neq$ 200/E[V]
- Var (V) = $\sum_v (v - E[V])^2 \, p_V(v)$

  $= (1-100.5)^2 . \frac{1}{2} + (200-100.5)^2 . \frac{1}{2}$

  $\approx 10\ 000$
- Standard Deviation $\sigma_V = \sqrt{\text{var}(V)} \approx 100$

Expectations:

- 4 buses carrying 148 job-seeking students arrive at a job fair. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying this randomly selected student. Also, one of the 4 bus drivers is also randomly selected. Let Y denote the number of students on his bus.

Do you think $E[X]$ and $E[Y]$ are equal?

$$E[X] = \frac{40}{148} \cdot 40 + \frac{33}{148} \cdot 33 + \frac{25}{148} \cdot 25 + \frac{50}{148} \cdot 50 \approx 39.3$$

$$E[Y] = \frac{1}{4} \cdot 40 + \frac{1}{4} \cdot 33 + \frac{1}{4} \cdot 25 + \frac{1}{4} \cdot 50 = 37$$

Variance and Standard Deviation

- The variance of an r.v. X, denoted by Var(X), or simply σ^2, is defined as:
  - $\mathrm{Var}(X) = \sigma^2 = E[(X - E[X])^2]$

    $\mathrm{Var}(X) = E(X^2 - 2E[X]X + (E[X])^2)$

    $= E(X^2) - 2E[X].E(X) + E((E[X])^2)$

    $= E(X^2) - 2(E[X])^2 + (E[X])^2$

    $= E(X^2) - (E[X])^2$
- The standard deviation (SD) is the square root of the variance.
- Note that the variance is in the square of the original units, while the SD is in the original units.

Next Time…

- Conditional Expectations
- Bernoulli & Binomial Random Variables