Summary of previous lectures

- Probability and statistics in general terms
- Subjective, classical, and relative frequency approaches to probability
- Probability is a numerical quantification of our beliefs about future occurrences
- An event of probability p is expected to occur roughly pN times on N trials
- Axiomatic approach to probability

Show me where it is written …

- Where in the axiomatic theory does it say that an event of probability p is expected to occur pN times on N future trials?
- To understand the where, we need to develop two important concepts:
  - random variables
  - independent trials
- Both concepts have been known, and even well-understood, since ancient times

It's not whether you won or lost, …

- Consider an experiment of rolling a fair die
- Suppose that you bet one dollar on the outcome of the trial
- If the outcome is a perfect square {1, 4} you win one dollar — of course, you do get your original dollar back too!
- Otherwise, if the outcome is {2, 3, 5, 6}, you lose the dollar that you bet
- Your wealth increases/decreases by $1

… it's how you played the game

- This game can be played in different ways
- Scenario #1: You give your bookie $1 as a bet. If your winning number comes up, he returns $2 to you — else he waves to you while he laughs all the way to the bank
- Scenario #2: You tell your bookie you want to bet $1. If your winning number comes up, he sends you $1 — else he sends two large men with baseball bats to collect $1

In the end, it is all the same

- Regardless of which scenario is your bookie's ***modus operandi,*** it is always true that at the end of the game, your net worth has either increased **or** decreased by $1
- An alternative viewpoint of Scenario #1 is that your $1 bet is a nonreturnable entry fee to play the game, and you win $2 or $0
- Scenario #1 is the one commonly used in lotteries, bets on horse races, and the like

You win some, you lose some …

- Let **X** denote your "winnings" in this game
- **X** takes on values +1 or –1 depending on the outcome of the trial
- Positive values of **X** denote an increase in your net worth; negative values of **X** denote a decrease in your net worth
- Long before the "invention" of negative numbers in ancient times, many people (even illiterates) understood these ideas

I hate algebra: I never know what X is

- **X** denotes your "winnings" in this game
- **X** takes on values +1 or –1 depending on the outcome of the trial
- So, what is the value of **X** going to be on the next trial?
- We cannot say for sure: the value that **X** will have on the next trial depends on the outcome of the die roll — and that is something that we cannot predict exactly

X is called a random variable

- **X** denotes your "winnings" in this game
- **X** takes on values +1 or –1 depending on the outcome of the trial
- The observed value of **X** sometimes is +1 and at other times it is –1, and we cannot tell ahead of time which it is going to be
- The observed value of **X** varies at random
- **X** is called a random variable or chance variable

So, what's so random about it?

- There is nothing random about the set of all possible values that might be taken on by **X**
- In our example, the values are +1 and –1
- The randomness lies in the unpredictability of the specific value that **X** is going to take on next: we don't know <u>which</u> of the values +1 and –1 will occur on the next trial
- We can never say "Let **X** = 1"

I hated geography, and I hate maps

- Formally, a random variable is a function whose domain is the sample space Ω and whose co-domain is the real line
- In everyday language, a random variable assigns a real number to each outcome in the sample space Ω
- The random variable is said to map an outcome to its assigned real number
- **X** maps ω Ω to the number **X**(ω)

A function as a table of values

- **X** maps ω Ω to the number **X**(ω)
- The table below illustrates the mapping

ω	**X**(ω)
1	+1
2	–1
3	–1
4	+1
5	–1
6	–1

A'int nutt'n random about that!

- **X** maps ω Ω to the number **X**(ω)
- The table below illustrates the mapping

ω	**X**(ω)
1	+1
2	–1
3	–1
4	+1
5	–1
6	–1

There is nothing random about the mapping; it is a fixed mapping — e.g. the outcome 3 is always mapped to –1, and 4 is always mapped to +1, etc

Brevity is the soul of wit

- **X** maps ω Ω to the number **X**(ω)
- The random variable is always denoted as **X**, never as **X**(ω)
- **X**(ω) means the number assigned to the outcome ω, e.g. **X**(4) is the number +1 for our random variable. There is nothing random about this
- **X** is the random variable — one of its possible values is +1

A feeling of déjà vu all over again…

- It is often convenient to not display the arguments of the functions when it is the functional relationship that is of interest
- d(uv) = u•dv + v•du
- y = h∗x
- x(t) = value of the input at the fixed time t
- Similarly, we always use **X** to denote a random variable (remember it is a function) and do not include the argument

Teaching by repetition

- A random variable **X** is an assignment of real numbers to the outcomes ω Ω
- Every outcome ω Ω is mapped to some number — **X**(ω) is the image of ω Ω
- Two or more outcomes could have the same image — but each outcome has exactly one image
- **X** denotes the random variable, **X**(ω) the value that **X** takes on if ω is the outcome

Examples of random variables I

- Example: The experiment consists of rolling a die. The random variable **Y** is just the number that shows on the top face; the random variable **Z** is the number on the opposite (bottom) face
- **Y** can be called the identity map
- Both **Y** and **Z** take on values {1,2,3,4,5,6}
- For a standard die, **Z** = 7 – **Y**, that is, for each outcome ω, **Z**(ω) = 7 – **Y**(ω)

Examples of random variables II

- Example: The experiment consists of tossing two coins. **X** is the number of Heads; **Y** is the number of Tails
- Both **X** and **Y** take on values {0, 1, 2}
- **Y** = 2 – **X**, that is, **Y**(ω) = 2 – **X**(ω) for all ω
- **Z** is 3.1415926 if the first coin results in a Head and 0.110210313 otherwise
- **Z**(HH)= **Z**(HT) = 3.1415926
  Z(TH) = **Z**(TT) = 0.110210313

Examples of random variables III

- Example: The experiment consists of tossing a coin till a Tail appears for the first time. **X** is the number of tosses on the trial; **Y** is the number of Heads observed
- **X** takes on values 1, 2, 3, 4, …
  Y takes on values 0, 1, 2, 3, …
- The number of possible values taken on by each of **X** and **Y** is countably infinite
- Exercise: What (if any) is the functional relationship between **X** and **Y**?

Examples of random variables IV

- Example: The experiment consists of choosing a real number at random in the interval (0, 1). **X** is the number chosen
- Once again, **X** is the identity map
- **X** is called a continuous random variable: its possible values form a continuum (that is, an interval of the real line)
- We will defer discussion of continuous random variables till later in the course

Discrete random variables

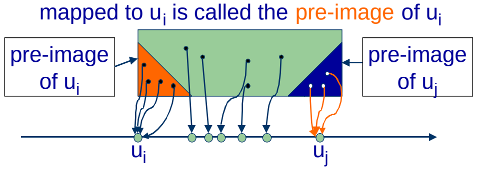
- A discrete random variable is one that takes on only a finite set of values in any interval of finite length
- The random variables in Examples I-III are all discrete random variables
- A discrete random variable takes on a finite, or a countably infinite, set of values
- The values taken on by a discrete random variable are discretely spaced

Indiscreet random variables

- A discrete random variable takes on finitely many or countably infinitely many values that are discretely spaced: $u_i < u_{i+1}$
- The rational numbers in the interval (0, 1) are a countably infinite set. However, a random variable that takes on all rational values in (0, 1) is not called discrete
- The rational number (x+y)/2 lies between any two given rational numbers x and y

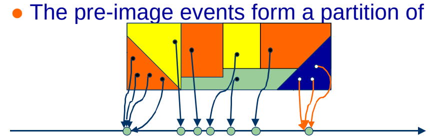
Pre-images

- Let $\{u_1, u_2, u_3, \ldots, u_n, \ldots\}$ denote the set of values taken on by a discrete random variable
- The set of all outcomes in Ω that are mapped to u_i is called the pre-image of u_i



Partitions, again

- Technical quibble: The pre-image of each u_i must be an event in the σ-field
- Quibble is irrelevant when is all the subsets of a finite or countably infinite Ω
- The pre-image of each u_i is an event
- The pre-image events form a partition of Ω



Special discrete random variables

- A random variable that takes on only one value (regardless of the outcome of the experiment) is called a degenerate or trivial random variable
- Non-experts in probability theory usually call it a constant!
- A random variable that takes on only the two values 0 and 1 is called a 0-1 random variable or a Bernoulli random variable

Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign

Indicator functions

- Let A denote an event defined on the sample space
- A random variable that takes on value 1 for all ω A, and value 0 for all ω A^c is called the indicator function of the event A
- The indicator function of the event A is denoted by I_A
- I_A has value 1 whenever A occurs and it has value 0 if A does not occur

Understanding indicator functions

- Indicator functions are 0-1 or Bernoulli random variables
- Exercises:
  What is the indicator function of A^c?
  If A and B are events with indicator functions I_A and I_B respectively
  - what is the indicator function of AB?
  - what is the indicator function of A B?
  - what is the indicator function of A B?

I left my heart in Decatur …

- Every discrete random variable is built up out of indicator functions
- Let A_j denote the pre-image of u_j and let I_j denote the indicator function of A_j
- Then, $\mathbf{X} = u_1 I_1 + u_2 I_2 + \ldots$
- A_1, A_2 … form a partition of Ω
- Let ω Ω belong to A_j
- When ω occurs, all the I_k except I_j have value 0, and **X** takes on value $u_j I_j = u_j$

Random is as random does …

- What about probabilities associated with random variables?
- For our die rolling experiment, what is the probability that **X** = +1? that **X** = –1?
- More generally, $P\{\mathbf{X} = u_j\} = P(A_j)$
- Since the A_j's are a partition of Ω, their probabilities sum to 1

$$\sum_j P\{X = u_j\} = 1$$

The probability mass function (pmf)

- $P\{\mathbf{X} = u_j\} = P(A_j)$
- Since the A_j's are a partition of Ω, their probabilities sum to 1

$$\sum_j P\{X = u_j\} = 1$$

- The function $p_X(\bullet)$ (or more simply $p(\bullet)$) is called the probability mass function of the discrete random variable **X**
- $p_X(u_j) = p(u_j) = P\{\mathbf{X} = u_j\}$

Pmf of a Bernoulli random variable

- **X** is called a Bernoulli random variable with parameter p if its pmf is given by
  $p_X(1) = p(1) = P\{\mathbf{X} = 1\} = p$
  $p_X(0) = p(0) = P\{\mathbf{X} = 0\} = 1–p$
- A Bernoulli random variable with parameter p is the indicator function of an event of probability p

Properties of the pmf

- We can think of the function $p_X(u)$ or $p(u)$ as being defined for all real numbers u
- The function $p_X(u)$ or $p(u)$ has value 0 for almost all choices of the argument u
- But if $u = u_j$ where u_j is one of the values taken on by the random variable **X**, then $p_X(u) = p(u) = P\{\mathbf{X} = u_j\}$
- $p(u) \geq 0$ for all values of u

$$\sum_j p(u_j) = 1$$

A physical interpretation

- The probability mass function is thought of as defining a set of point masses on the real line
- The masses denote probabilities
- For each j, we imagine that there is a mass of $p(u_j)$ at location u_j
- The total probability mass is 1
- There are no negative masses

Graphical representations

Is the mass proportional to the radius or the area of the blob?

The mass is proportional to the height of the stick. Total length of all sticks represents the probability of the sample space, i.e. 1

More physical interpretations

- The probability mass function defines a set of point masses on the real line
- The center of mass of this set of point masses has an important probabilistic interpretation
- The moment of inertia and the radius of gyration (about the origin, and also about the center of mass) have important probabilistic interpretations as well

So, what's the pmf of X anyway?

- Consider an experiment of rolling a fair die
- If the outcome is a perfect square {1, 4} you win $1
- If the outcome is {2, 3, 5, 6}, you lose $1
- The random variable **X** that denotes your winnings takes on values +1 and –1 with probabilities 1/3 and 2/3
- $p(u) = 1/3$ if $u = +1$; $p(u) = 2/3$ if $u = -1$; $p(u) = 0$ for all other values of u

Summary

- We studied the concept of a random variable, viz. numerical values associated with the random outcomes of trials of an experiment
- We defined discrete random variables and their probability mass functions
- The concepts of center of mass and moment of inertia also have important probabilistic interpretations

