ECS509U - Probability & Matrices

Tassos Tombros

Week 4

Week 4: Learning Objectives

- Basics of **Probability Distributions**
- At the end of Week 4 you should be able to:
  - Understand the concept of a **random variable** and how it relates to probability problems
  - Work with **probability mass functions**
  - Calculate the **expected value, variance** and **standard deviation** of random variables
  - Solve probability problems involving the **binomial probability model**

Random Variables

- A **random variable** is a function from the sample space S to a set of numbers that makes sense for the sample space
  - we will use capital letters to denote a r.v., e.g. X
  - you can define many r.v. for the same sample space
- Suppose you roll two dice
  - sample space consists of 36 outcomes S={(1,1), (1,2), (2,1), . . ., (6,6)}
  - one possible r.v. will represent **the sum of the potential outcomes on the two dice**, let's call the r.v. **X**
  - X will have a different value for each point of the sample space (each pair of dice), but is constrained to range from 2 to 12

The two-dice example

Out-come	X value	Out-come	X value	Out-come	X value	Out-come	X value	Out-come	X value	Out-come	X value
(1,1)	2	(2,1)	3	(3,1)	4	(4,1)	5	(5,1)	6	(6,1)	7
(1,2)	3	(2,2)	4	(3,2)	5	(4,2)	6	(5,2)	7	(6,2)	8
(1,3)	4	(2,3)	5	(3,3)	6	(4,3)	7	(5,3)	8	(6,3)	9
(1,4)	5	(2,4)	6	(3,4)	7	(4,4)	8	(5,4)	9	(6,4)	10
(1,5)	6	(2,5)	7	(3,5)	8	(4,5)	9	(5,5)	10	(6,5)	11
(1,6)	7	(2,6)	8	(3,6)	9	(4,6)	10	(5,6)	11	(6,6)	12

X is the random variable

It is a function from S to the set of numbers that correspond to the sum of two dice = {2, 3, …, 12}

Some more examples

- Select a random student amongst the class, and measure his/her height in cm.
  - S=set of students
  - r.v. X is ‘height’ which is a function from the set of students to the set of real numbers that correspond to heights of students
- I toss a coin 3 times, and measure the number of heads
  - S= {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
  - Random variable Y is a function from S to {0, 1, 2, 3}

Types of random variables

- **Discrete**
  - A r.v. is discrete if the values it can take are **finite** or **countably infinite** (link to week 1 slides for these)
  - The examples with the coin and the dice correspond to finite cases
  - A countably infinite case would be e.g. a r.v. measuring the number of accidents in a year in the M1
- **Continuous**
  - A r.v. is continuous if the values it can take are **uncountably infinite**
  - e.g. the r.v. measuring the height of students could take any real number between certain extreme limits
  - e.g. r.v. measuring the length of phone calls

Probability mass function (pmf)

- Let X be a **discrete random variable**
  - Given any value *a* in the set of possible values for X, **what is the probability that X will have that value *a*?**
- One way to write this:
  - $\mathbf{A=\{x \in S : X(x)=a\}}$, and we would then be interested in P(A)
- Another way to write this:
  - **P(X=a)** : the probability that X will take the value *a*, **or** even simpler **P(*a*)**

Definitions and properties of pmf

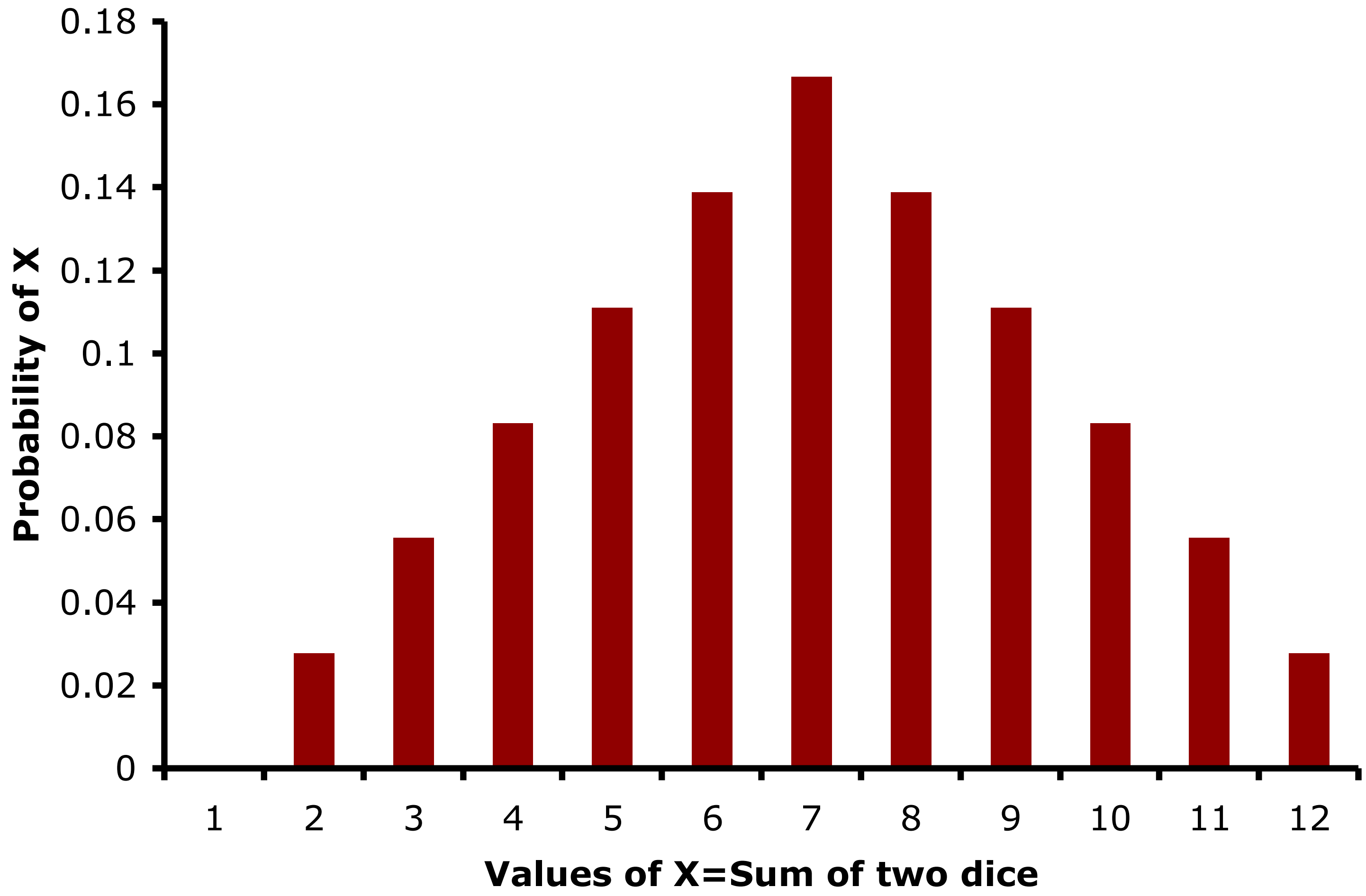
- The **probability mass function (pmf)** of a discrete r.v. will give us the value of P(X=a) for each element *a* in the set of possible values of the random variable X
- For a function to be a probability mass function of a discrete r.v., 2 conditions must hold:
  - **$P(X=a) \geq 0$**, for all possible *a*
  - **$\sum (P(X=a)) = 1$**, for all possible *a*
  - Both conditions should feel familiar!!!
  - Also, to find the probability that the r.v takes the values a OR b: **$P(X=a \text{ OR } X=b) = P(X=a) + P(X=b)$**

Example

X (the r.v.)	p.m.f of X=Sum of two dice P(X=a) or P(a)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

- What is the p.m.f. for the dice-rolling example, where the random variable X is the sum of the two dice?
- How did we find these probability values?
- Do they sum up to 1?

A graphical illustration



Another example

- Consider the case of tossing a coin 3 times
  - S= {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- Let X be a r.v. that records the number of heads {0, 1, 2, 3}. Find the p.m.f. of X

	p.m.f of X=number of heads in 3 coin tosses
X (the r.v.)	P(X=a) or P(*a*)
0	1/8
1	3/8
2	3/8
3	1/8

- To find the probability that X takes the values 1 **OR** 2:

P(X=1) + P(X=2) = P(1)+P(2) = 3/8 + 3/8 = 6/8

Expectation & distribution parameters

- Using the probability mass function you can figure three important parameters:
- **Expected Value**
  - The long-term average value that you would expect to see after an experiment is repeated a theoretically infinite number of times
- **Variance**
  - The amount of variability you would need to expect from one set of results of the experiment to another
- **Standard Deviation**
  - Helps us to interpret the variance of the results

Expected value of a random variable

- **The expected value** of a random variable X, E(X), is the **long-run theoretical average** value **of X**
- You can also see it as the weighted average of all possible values of X, weighted by how often we expect each value to occur over the long-term

$$E(X) = \sum_{\text{all x}} xp(x)$$, where x is a value of the r.v. X, and p(x) the probability of observing that value

- Another way to describe E(X) is as the mean of X, denoted with the letter μ

Methodology & Example

- To find the expected value of a random variable:
  - Multiply the value of X by its probability
  - Repeat the step for all values of X
  - Sum the results
- In the example of rolling the two dice and the r.v. X recording the sum of the dice:

$$E(X) = 2(1/36) + 3(2/36) + 4(3/36) + 5(4/36) + 6(5/36) + 7(6/36) +$$

$$8(5/36) + 9(4/36) + 10(3/36) + 11/2/36 + 12(1/36) = 7$$

- so, if we were to throw two dice a theoretically infinite number of times and every time record the sum, the average sum of all the rolls would be 7

The variance of a random variable

- **The variance V(X) of a r.v.** is the amount of variability you would expect in the results after repeating the experiment a theoretically infinite number of times

  $$\mathbf{V(X) = E(X^2)-\mu^2 = E(X^2)-[E(X)]^2}$$

- In other words, this is the difference between the expected value of X^2 and the square of the expected value of X
- To find $\mathbf{E(X^2)}$ you just need to calculate $\Sigma \mathbf{x^2 p(x)}$ for all x that are values of the r.v. X
- **Always, V(X)$\geq$0**

An example

- Consider the case where we toss a coin 3 times and the r.v. X counts the number of heads. Calculate E(X) and V(X)
- $E(X) = \mu = \sum_{\text{all x}} xp(x)$ = 0x(1/8) + 1x(3/8) + 2x(3/8) + 3x(1/8) = 3/2

$$V(X) = E(X^2) - (E(X))^2 = (\sum_{\text{all x}} x^2 p(x)) - (E(X))^2$$

$$= [(0^2 \times (1/8)) + (1^2 \times (3/8)) + (2^2 \times (3/8)) + (3^2 \times (1/8))] - (3/2)^2$$

$$= 3/4$$

Standard deviation

- **The standard deviation of a random variable** is simply the **square root of its variance**
  - we use the letter σ to denote the standard deviation

$$\sigma = \sqrt{V(X)}$$

- Standard deviation allows us to easier interpret the variation within the outcomes of the random variable.
  - It shows us the variability of X in the original units of X and not in their square, as does V(X)

“Celebrity” Probability Distributions

- A number of probability distributions have their own names and their own characteristics
  - they are distributions that occur in many probability problems, so people have studied them and their properties well
- Discrete Uniform, Binomial, Normal, Bernoulli, Poisson, Gaussian, etc.
- For each distribution there are:
  - **Conditions** that must be met
  - **Formulas** for pmf, E(X), V(X)

The Binomial probability model

- Binomial means ‘two names’ and is associated with situations involving two outcomes, e.g. success/failure
- The conditions to have a **Binomial** model are:
  - A **fixed number of trials**, ***n***
  - **The outcome of each trial** is either in one of two groups, e.g. success or failure
  - **The probability of success is the same in each trial**, let it be ***p***, and therefore ***1-p*** the probability of failure
  - **The trials are independent** of each other

Checking the conditions

- Do the following examples satisfy the conditions of the Binomial model?
  - You toss a coin 10 times and count the number of heads
  - You toss a coin until you get 4 heads
  - You have a drawer with 10 red pens, 10 blue and 10 black. You take a pen out and record its colour (then you do not put it back in). You repeat the process 5 times, and you are interested in the total number of red pens that you will pick from the jar

The pmf of the Binomial

- **The probability mass function of the Binomial** model for a r.v. X is given by:
  - **P(X=x)** is the probability of having exactly x successes, **$P(X=x) = C(n,x)p^x(1-p)^{n-x}$**, where:
  - **n** is the fixed number of trials
  - **x** is the specified number of successes, so **n-x** is the number of failures
  - **p** is the probability of success in any given trial, so **1-p** is the probability of failure in any given trial
  - **C(n,x)** is our known formula for combinations of x items chosen from n items (**here:** x successes chosen from the total of n trials)

Examples

- You flip a coin 3 times, the r.v. X records the number of heads. What is the probability that we get:
- **All heads?**
  - n=3, x=3, n-x=0, p=1/2, 1-p=1/2
  - The chance of getting all heads is P(X=3) or P(3) since X is recording the number of heads
  - $P(3) = C(3,3)(1/2)^3(1/2)^0 = (1)(1/8)(1)=1/8$
- **More than one tail?**
  - **Remember** that X counts the number of heads. More than one tail means 2 or 3 tails, which means 0 or 1 heads!!! So the answer will be given by P(0) + P(1)
  - Apply the formula for n=3, p=1/2, 1-p=1/2 and x=0 or 1 (and n-x=3 or 2 respectively)
  - P(more than one tail) = P(1)+P(0)=3/8+1/8=4/8=1/2

Expected value and variance

- In the Binomial model, for a r.v. X:
  - **E(X)=np**, where:
    - **n** is the number of trials, and **p** the probability of success in any given trial
  - **V(X)=np(1-p)**, where:
    - **n** is the number of trials, and **p** the probability of success in any given trial
  - The standard deviation $\sigma = \sqrt{np(1-p)}$

Summary of lecture

- In Week 4 we covered:
  - Random variables
  - Probability mass function
  - Expected value, variance and standard deviation
  - The Binomial probability distribution
- Don't forget:
  - Work on the exercises for Friday **BEFORE** you come to the class

Exercises on expected values

1. Two fair dice are rolled, and the r.v. X records the max of the two numbers that show on the dice. Find E(X)
2. A coin is tossed until a head, or 5 tails occur. Find the expected number of tosses of the coin.
   - **Hint:** what is the sample space? What does the random variable represent in this problem?
3. A player tosses two coins. The player wins £2 if 2 heads occur, and £1 if 1 head occurs. The player loses £3 if no head occurs. Is the game fair?
   - **Hint:** the game is favorable, fair or unfavorable to the player respectively, if $E>0$, $E=0$, or $E<0$

Sample answers

- First we would need to find the distribution of the r.v. X:
  $P(X=1)=1/36$ since only one toss (1,1) has the max value of 1, $P(X=2)=3/36$ for (1,2), (2,1) and (2,2) and similarly $P(X=3)=5/36$, $P(X=4)=7/36$, $P(X=5)=9/36$, $P(X=6)=11/36$ (**hint:** check if they sum up to 1)
  Now we can easily calculate
  $E(X)=1(1/36)+2(3/36)+3(5/36)+4(7/36)+5(9/36)+6(11/36)\approx 4.47$

- S={H,TH,TTH,TTTH,TTTTH,TTTTT} with respective probabilities (we multiply because of independent trials): $1/2$, $(1/2)^2=1/4$, $(1/2)^3=1/8$, $(1/2)^4=1/16$, $(1/2)^5=1/32$, $(1/2)^5=1/32$
  $P(1)=1/2$, $P(2)=1/4$, $P(3)=1/8$, $P(4)=1/16$,
  $P(5)=P(TTTTH)+P(TTTTT)= 1/32 + 1/32 = 1/16$
  It follows that $E(X)=1(1/2)+2(1/4)+3(1/8)+4(1/16)+5(1/16) =31/16 = 1.9375$

Sample answers cntd.

- S={HH, HT, TH, TT}. Let the r.v. X denote the player's gain: X(HH)=£2, X(HT)=X(TH)=£1, X(TT)=-£3

  P(£2)=1/4, P(£1)=2/4, P(-£3)=1/4

  It can then be calculated that E(X)= £2(1/4)+ £1(2/4)+ (-£3) (1/4) = £0.25

  So the game is favorable to the player because E(X)>0

Exercises on the binomial distribution

1. West Ham have a probability 2/3 of winning every time they play!!! Suppose they play 6 games. Find the probability that they win more than half of the games.
2. A student takes an 18-question multiple choice exam, with 4 choices per question. Suppose 1 of the 4 choices is obviously wrong, and the student decides to choose randomly amongst the other 3 choices for each question.
   - What is the probability that the student gets 7 questions correctly and thus scrapes a pass?
   - What is the expected number of correct answers and the standard deviation σ?

Sample answers

- This is a binomial model, with n=6, p=2/3, 1-p=1/3, number of successes (wins) x= 4, 5, 6

  $P(X>3)=P(4)+P(5)+P(6)= C(6,4)(2/3)^4(1/3)^2 + C(6,5)(2/3)^5(1/3)^1 + C(6,6)(2/3)^6(1/3)^0$ (you are expected to continue the calculations)
- This is a binomial experiment with n=18, p=1/3, 1-p=2/3, x=7, n-x=11

  $P(7)=C(18,7)(1/3)^7(2/3)^{11}$ = **0.168**

  The expected number of correct answers is given by $E(X)= np =18(1/3)= 6$

  The standard deviation $\sigma = \sqrt{np(1-p)} = \sqrt{18(1/3)(2/3)} = 2$