MATH 323 Class Notes

Owen Lewis

Summer 2018

Contents

1 May 1, 2018

1.1 Definitions

Let Ω be the set of all possible outcomes. We call Ω the *Sample Space*.

Ex 1: Flipping a coin. The possible outcomes are H and T
$\Rightarrow \Omega = \{H, T\}$.
Ex 2: Tossing a die. We list all the outcomes as ω_i where i is the face of the die that we land on. We'll assume a normal 6-sided die
$\Rightarrow \Omega = \{\omega_1, \omega_2, \ldots, \omega_6\}$.
Ex 3: Flipping a coin until an H appears. The possible outcomes are H, TH, $TTH, \ldots, TT\ldots TH$ (with $n-1$ Ts), $\ldots$ ad infinitum.

$\Rightarrow \Omega = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$.
Note that in this case, Ω is a (countably) infinite set!

Let Ω be a sample space. Any subset A of Ω is called an *Event*.

- If $A = \emptyset$, then we call A the *Null Event*
- If $A = \Omega$, then we call A the *Certain Event*
- If $|A| = 1$, then we call A an *Elementary Event*

Ex: Tossing a die. Let $\Omega := \{\omega_1, \omega_2, \ldots, \omega_6\}$, $A := \{\omega_1, \omega_2\}$. Then A is an event but is not an elementary event.

If A is an event, then A^c is also an event called the *complement event* of A.
If A, B are two disjoint events then we call A, and B *mutually exclusive*, or *disjoint*.

Let Ω be a sample space, $\mathcal{P}$ be the power set of Ω. A *Probability* $\mathbb{P}$ on Ω is a function $\mathbb{P} : \mathcal{P}(\Omega) \to [0, 1]$, such that:

1. $\forall\ A \subseteq \Omega$, $0 \leq \mathbb{P}(A) \leq 1$
2. $\mathbb{P}(\Omega) = 1$
3. If $A_1, A_2, \ldots, A_n, \ldots$ is a sequence of pairwise disjoint events then

$$\mathbb{P}(\bigcup_{i=1}^{\infty} p_i) = \sum_{i=0}^{\infty} \mathbb{P}(A_i)$$

1.2 How do we Apply this?

Let Ω be a discrete set, $E_i = \{\omega_i\}$ be an elemental event, with $E_i \subseteq \Omega$. A probability on Ω is given by a sequence $\mathbb{P}_1$, $\mathbb{P}_2, \ldots$, $\mathbb{P}_n, \ldots$ of positive numbers such that

$$\mathbb{P}(E_i) = p_i, \ and \ \sum_i \mathbb{P}(p_i) = 1$$

If $A \subseteq \Omega$, then

$$\mathbb{P}(A) = \sum_{\omega_i \in A} \mathbb{P}(p_i)$$

Ex 1: Toss a die.
a) Given that $\mathbb{P}(\omega_2) = \mathbb{P}(\omega_4) = \mathbb{P}(\omega_5) = \mathbb{P}(\omega_6) = \frac{1}{6}$, $\mathbb{P}(\omega_1) = \frac{1}{4}$, find $\mathbb{P}(\omega_3)$.
b) Find the probability that the die will land on an odd face.
Solution:
a) $\Omega = \{\omega_1, \ \omega_2, \ldots, \ \omega_6\}$. $\mathbb{P}(\omega_3)$ is a singleton, and as we know $\sum_i \mathbb{P}(p_i) = 1$ then $\sum_{i=1}^{6} \mathbb{P}(\omega_i) = 1 \implies \mathbb{P}(\omega_2) + \mathbb{P}(\omega_4) + \mathbb{P}(\omega_5) + \mathbb{P}(\omega_6) + \mathbb{P}(\omega_1) + \mathbb{P}(\omega_3)$ $= 1 \implies \frac{4}{6} + \frac{1}{4} + \mathbb{P}(\omega_3) = 1 \implies \mathbb{P}(\omega_3) = 1$ - $\frac{11}{12}$
$\implies \mathbb{P}(\omega_3) = \frac{1}{12}$.
b) Let $A \subseteq \Omega$ be the subset of Ω containing all the odd faces. The total probablilty of A is therefore the sum of all the probabilities of the elements of $A \Rightarrow$ $\mathbb{P}(A_i) = \frac{1}{4} + \frac{1}{12} + \frac{1}{6} = \frac{1}{2}$.

Ex 2: Given a countably infinite sample space, find a constant c such that $\mathbb{P}(\{\omega_n\}) = c(\frac{1}{5})^n$ for some n.
Solution: $\sum_{n=1}^{\infty} \mathbb{P}(\{\omega_n\}) = 1 \Rightarrow \sum_{n=1}^{\infty} c(\frac{1}{5})^n = 1 \Rightarrow \sum_{n=1}^{\infty}(\frac{1}{5})^n = \frac{1}{c}$. Notice how $\sum_{n=1}^{\infty}(\frac{1}{5})^n$ is a geometric series that converges to $\frac{1}{4} \Rightarrow$ therefore $\frac{1}{c} = \frac{1}{4}$
$\Rightarrow c = 4$.

2 May 2, 2018

2.1 Properties of $\mathbb{P}$

Let Ω be a sample space and let $\mathbb{P}$ be a probability on Ω. Then:

1. $\mathbb{P}(\emptyset) = 0$
2. $\mathbb{P}(A^c) = 1$ - $\mathbb{P}(A)$
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ - $\mathbb{P}(A \cap B)$

Proof:

1. Set $A_1 = \Omega$, $A_2 = \emptyset$. Then $A_1 \cup A_2 = \Omega$, and $A_1 \cap A_2 = \emptyset$. Therefore $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) \Rightarrow \mathbb{P}(\Omega) = \mathbb{P}(\emptyset) + \mathbb{P}(\Omega) \Rightarrow 1 = \mathbb{P}(\emptyset) + 1$
$\Rightarrow 0 = \mathbb{P}(\emptyset)$. $\square$
2. $A^c \cup A = \Omega$, and $A^c \cap A = \emptyset$. Then $= \mathbb{P}(\Omega) \Rightarrow 1 = \mathbb{P}(A^c \cup A)$
$\Rightarrow 1 = \mathbb{P}(A) + \mathbb{P}(A^c) \Rightarrow \mathbb{P}(A^c) = 1$ - $\mathbb{P}(A)$. $\square$
3. It's easy to show that $A = (A \setminus B) \cup (A \cap B)$, and $\emptyset = (A \setminus B) \cap (A \cap B)$. Similarily, $\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)$. From these, it follows that
$A \cup B = A \cup (B \setminus A) \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A)$
$\Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ - $\mathbb{P}(A \cap B)$. $\square$

2.2 Equiprobability

Let Ω be a finite sample space. Set $N := |\Omega|$. Equiprobability means that all outcomes have the same probability $\mathbb{P} = \frac{1}{N}$. Let $A \subseteq \Omega$ be an event. Then we have $\mathbb{P}(A) = |A| \cdot \frac{1}{N}$, or

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

This is great because it means that in **equiprobability** problems we just need to count the cardinality of A, count the cardinality of Ω, and divide them, and we're done. Too bad counting isn't really all that easy.

2.3 Counting Tools

We have here three tools to help us calculate the cardinalities of huge (finite) subsets of huge sample spaces, each with their own specific situations that require its use

1. The Cartesian Product
2. Permutations
3. Combinations

2.3.1 The Cartesian Product

We all know what the cartesian product *is*. In probability we use it in our experiment when we have more than one """input""" each with its own possible outcome, for example, we roll three dice, or flip two coins.

Let A, B, be sets such that $|A| = a$, and $|B| = b$. Then the cardinality of the cartesian product $|(A \times B)|$ is $|(A \times B)| = a \cdot b$.

Ex 1: Suppose we roll a die twice. What is the cardinality of the sample space Ω?
Solution: If we roll a die once we have $\Omega = \{\omega_1, \ldots, \omega_6\}$. Therefore, the sample space for rolling a die twice is $\Omega \times \Omega$. The cardinality of our sample space $\Omega \times \Omega$ is $|\Omega \times \Omega| = 6 \cdot 6 = 36$.

Ex 2: Suppose we roll a fair die twice. What is the probability that the outcome is even?
Solution: If the sum of the two outcomes is even then both must either be even or both must be odd. Let A be the event where the sum of the rolls is even, and let A_1 be event where both individual rolls are even, A_2 be the event where both individual rolls are odd. For example,

$$A_1 = \{(2,2), (2,4), (2,6), (4,2), (4,4), (4,6), (6,2), (6,4), (6,6)\}$$

Where the 1^{st} element in each ordered pair is the outcome of the 1^{st} roll and the 2^{nd} element in each ordered pair is the outcome of the 2^{nd} roll. Therefore, we have that $|A_1| = |A_2| = 9$.
Since naturally $A = A_1 \cup A_2$ then $\mathbb{P}(A) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$. We also know that the die was fair, so we can use our equiprobability formula here.

$$\mathbb{P}(A) = \mathbb{P}(A_1) + \mathbb{P}(A_2) = \frac{|A_1|}{|\Omega|} + \frac{|A_2|}{|\Omega|} = \frac{9}{36} + \frac{9}{36} = \frac{18}{36} = \frac{1}{2}.$$

2.3.2 Permutations

A permutation of r integer elements chosen from n (possible elements) is equivalent to a successive draw, without replacement, of r elements from a list of n elements. We denote the number of possibilities by P^n_r. The general formula for P^n_r is

$$P^n_r = n \cdot (n-1) \cdot \dots \cdot (n-r+1) = \frac{n!}{(n-r)!}$$

Ex:
a) A thick black bag contains 4 balls: 1 green, 1 blue, 1 red, 1 yellow. The bag is made of lead, or something, and also light cannot exist in this bag. You couldn't see into this bag if your life depended on it. Draw successively two balls from the bag without putting them back in. What is the probability that

the second ball drawn is green?
b) What is the probability that one of the two balls drawn will be green?
Solution:
a) Let Ω be the set of permutations of 2 balls chosen from the bag containing 4 balls. Then, $|P_2^4| = \frac{4!}{2!} = \frac{24}{2} = 12$. Then let A be the event where the second ball is green. The cardinality of A is 3, as if we take for granted that the second ball is green, then there are 3 other different-coloured balls in the bag to accompany it. Since the bag is so dark that it's physically impossible to see inside it, all the balls in the bag have an equal probability of being drawn. Therefore,

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{12} = \frac{1}{4}$$

b) Let B be the event where one of the balls drawn is green. Suppose B_1 is the event where the first ball drawn is green, and B_2 is the event where the second ball drawn is green. In part a) we found that $|B_2| = 3$, and a parallel argument shows that $|B_2| = 3$ as well. Thus,

$$\mathbb{P}(B) = \frac{|B_1|}{|\Omega|} + \frac{|B_2|}{|\Omega|} = \frac{3}{12} + \frac{3}{12} = \frac{6}{12} = \frac{1}{2}$$

2.3.3 Combinations

Consider a set Ω with n elements. Let r be an integer such that $0 \leq r \leq n$. A combination C_r^n, also denoted $\binom{n}{r}$ (pronounced n choose r), is the number of subsets of Ω containing r elements. The general formula is

$$C_r^n = \frac{P_n^r}{r!} = \frac{n!}{r! \cdot (n-r)!}$$

Ex:
a) Out of a deck of 52 cards how many distinct 5-card hands are possible?
b) What is the probability that a given hand contains at least one ace?
Solution:
a) $\binom{52}{5} = \frac{52!}{5! \cdot (47!)} = 2{,}598{,}960$
b) In this case it is easier to calculate the probability where the hand contains no aces and then subtract that from 1 to find the probability that we have an ace. If A is the event where the hand contains an ace, then A^c is the event where a hand contains no ace. Then

$$\mathbb{P}(A^c) = \frac{|A^c|}{|\Omega|} = \frac{C_5^{48}}{C_5^{52}} = \frac{1,712,304}{2,598,960} = \frac{35,673}{54,145}$$

Then we need to subtract this from 1 and we're done

$$\mathbb{P}(A) = 1 - \frac{35,673}{54,145} \approx 0.3412$$

2.4 Properties of Combinations

Here are some properties of combinations

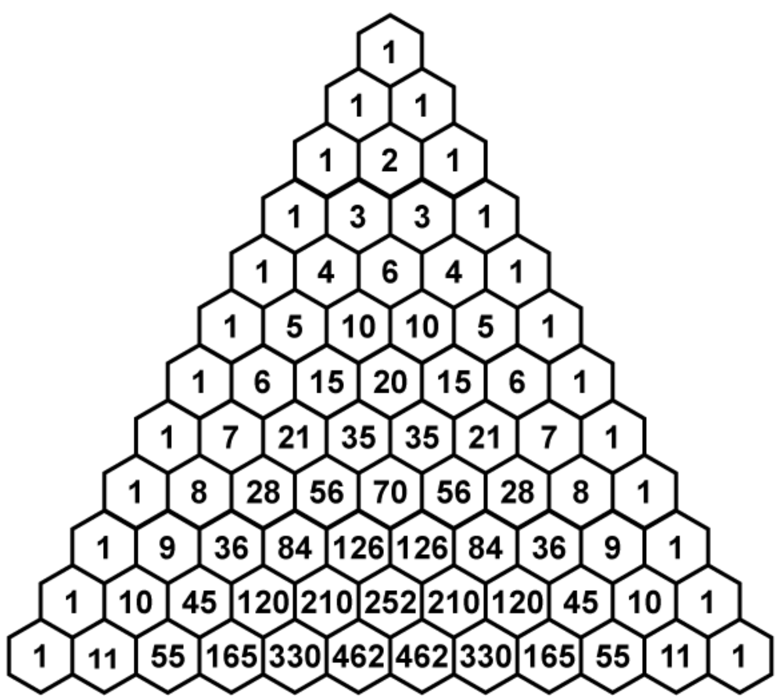
1. $C_0^n = C_n^n = 1$
2. $C_r^n = C_{n-r}^n \; 0 \leq r \leq n$
3. $C_r^n = C_r^{n-1} + C_{r-1}^{n-1} \; 0 \leq r \leq n$

The proofs are really easy and I dont wan't to bother typing them out but here is the idea behind each of them:

1. Trivial
2. Induction on n
3. Also induction on n

2.4.1 Pascal's Triangle

Pascal's triangle is a table of combinations C_r^n. The rows of the triangle represent n, starting at 0 at the tip and working down, and the r^{th} element from the left (starting at 0) represents r. Each number in the triangle is determined by summing the two numbers directly above it.



Pascal's Triangle

3 May 3, 2018

3.1 Binomial Theorem

$(a+b)^n = \sum_{k=0}^{n} C_k^n \cdot a^k \cdot b^{n-k}$

Ex 1: $(a+b)^3 = C_0^3 \cdot a^6 b^3 + C_1^3 \cdot ab^2 + C_2^3 \cdot a^2 b + C_3^3 \cdot a^3 b = b^3 + 3ab^2 + 3ba^3 + a^3$

Ex 2: Find the coefficient of x^6 in the expansion of $(x^2+2)^7$.
Solution: If we let $x^2 := a$ and $2 := b$, then from the binomial theorem:

$$(a+b)^n = \sum_{k=0}^{n} C_k^n \cdot a^k \cdot b^{n-k} = \sum_{k=0}^{n} C_k^7 \cdot a^k \cdot b^{7-k} = C_3^7 \cdot 2^4 = 560$$

3.2 Conditional Probability

Let Ω be a sample space and let $\mathbb{P}$ be a probability on Ω. Let $A \subseteq \Omega$ be an event such that $\mathbb{P}(A) > 0$, and let $B \subseteq \Omega$ be another event. The *conditional probability of B given A*, is defined as

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

Ex 1: We have 2 urns. Urn 1 contains 7 red balls and 4 blue balls, and urn 2 contains 5 red balls and 6 blue balls. First, we choose a ball from urn 1 and place it into urn 2, then we remove a ball from urn 2. What is the probability that the ball that we remove from urn 2 will be blue?
Solution: Let B be the event where the ball drawn from urn 2 is blue, and let A_r, A_b be the events where we draw a red or a blue ball from urn 1 respectively. Then $\mathbb{P}(A_r) = \frac{7}{11}$ and $\mathbb{P}(A_b) = \frac{4}{11}$. Also,

$$B = B \cap \Omega \iff B = B \cap (A_r \cup A_b) \iff B = (B \cap A_r) \cup (B \cap A_b)$$

Which then implies

$$\mathbb{P}(B) = \mathbb{P}(B \cap A_r) + \mathbb{P}(B \cap A_b) \iff \mathbb{P}(B) = \mathbb{P}(B|A_r) \cdot \mathbb{P}(A_r) + \mathbb{P}(B|A_b) \cdot \mathbb{P}(A_b)$$

So all that is left is to compute $\mathbb{P}(B|A_r)$, and $\mathbb{P}(B|A_b)$.
From the definition, we can see that $\mathbb{P}(B|A_r) = \frac{1}{2}$, and $\mathbb{P}(B|A_b) = \frac{7}{12}$, so finally

$$\mathbb{P}(B) = \frac{6}{12} \cdot \frac{7}{11} + \frac{7}{12} \cdot \frac{4}{11} = \frac{70}{132}$$

Ex: 2 Roll a fair die. Let $A \subseteq \Omega$ be the event where the outcome is even, $B \subseteq \Omega$ be the event where the outcome is odd, $C \subseteq \Omega$ be the event where the outcome is either a 1 or a 2. Compute all the conditional probabilities.
Solution First $\mathbb{P}(A) = \frac{1}{2}$, $\mathbb{P}(B) = \frac{1}{2}$, and $\mathbb{P}(C) = \frac{1}{3}$ Then:

$\mathbb{P}(A|B) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$, and $\mathbb{P}(B|A) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$
$\mathbb{P}(A|C) = \frac{1}{2}$, and $\mathbb{P}(C|A) = \frac{1}{3}$
$\mathbb{P}(B|C) = 1$, and $\mathbb{P}(C|B) = \frac{2}{3}$

3.3 Independent Events

Let A, and B be two events over some sample space Ω. A, and B are said to be *independent* if and only if:

1. $\mathbb{P}(A|B) = \mathbb{P}(A)$
2. $\mathbb{P}(B|A) = \mathbb{P}(B)$
3. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

Ex:
a) From the previous example, are A and B independent?
b) What about A and C?
c) B and C?
Answers:
a) No.
b) Yes.
c) No.

3.4 Baye's Rule

Let $A, B \subseteq \Omega$ be events on a sample space Ω. Then the equation

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

is calles Baye's Rule, or Baye's Theorem.

Theorem: Total Probability Rule
Let Ω be a sample space, and let $A \subseteq \Omega$ be an event. Suppose we partition Ω like $\Omega = \{B_1, B_2, \ldots, B_n\}$, where the B_is are pairwise disjoint events such that

$$\bigcup_{i=1}^{n} B_i = \Omega$$

Then:

1. $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$, and
2. If $k \in \mathbb{N}$ is fixed, then $\mathbb{P}(B_k|A) = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}$

Proof:

1. We know that $A = A \cap \Omega \iff A = A \cap (\bigcup_{i=1}^{n} B_i) \iff A = \bigcup_{i=1}^{n} A \cap B_i$. Therefore, $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) \iff \mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$ $\square$

2. $\mathbb{P}(B_k|A) = \frac{\mathbb{P}(B_k \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}{\sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}$ $\square$

Ex: A (very simple) forecast model
Suppose that on a given day the weather is one of two states: sunny or rainy. Let $R \subseteq \Omega$ be the event where it's rainy and $S \subseteq \Omega$ be the event where it's sunny. If today is rainy then the probability that tomorrow will also be rainy is 60%. On the other hand, if today is sunny then the probability that tomorrow will be sunny is 70%.
a) If Monday is sunny then what is the probability that Wednesday will also be sunny?
b) If Wednesday is sunny then what is the probability that Tuesday was rainy?
Solutions:
a) Let A be the event where it is sunny on Wednesday. Then

$$\mathbb{P}(A) = \mathbb{P}(A|A_1) \cdot \mathbb{P}(A_1) + \mathbb{P}(A|A_2) \cdot \mathbb{P}(A_2) = 0.7 \cdot 0.7 + 0.4 \cdot 0.3 = 0.61$$

where A_1 is the event where it's sunny on Tuesday, and A_2 is the event where it's rainy on Tuesday.
b) $\mathbb{P}(A_2|A) = \frac{\mathbb{P}(A|A_2) \cdot \mathbb{P}(A_2)}{\mathbb{P}(A)} = \frac{0.12}{0.61} = \frac{12}{61}$

4 May 7, 2018

4.1 When to use Permutations vs Combinations

To determine whether we need to use a permutation or a combination we first need to consider two things. First, the sampling method:

1. Without Replacement

2. With Replacement

and whether we care about the order of the sample points:

3. Order doesn't matter

4. Order matters

We use permutations when we have 1&4, and
we use combinations when we have 1&3 *or* 2&4.

4.2 Multinomial Coefficients

The number of ways we can partition n objects into k distinct groups containing n_1, n_n, ..., n_k objects, respectively, where each object appears in exactly one group and $\sum_{i=1}^{k} n_i = n$, is

$$N := \binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!n_2!\ldots n_k!}$$

where we call

$$\binom{n}{n_1, n_2, \ldots, n_k}$$

the *multinomial coefficient.*

Ex 1: We wish to expand $(x+y+z)^{17}$. What is the coefficient on the $x^2y^5z^{10}$ term?
Solution: The coefficient is $\binom{17}{2,5,10} = \frac{17!}{2!5!10!}$

Ex 2: #62 from the first assignment.
Solution: Total number $= \binom{9}{3,3,3}$, Desired number $= \binom{7}{1,3,3}$.
Therefore, $\mathbb{P}(A) = \frac{\binom{7}{1,3,3}}{\binom{9}{3,3,3}}$

4.3 Discrete Random Variables

Random Variables are variables that take on random values based on the outcome of the experiment. Each random variable is associated with a *probability distribution* that specifies the possible random variable values and the probability each value will occur. A random variable is said to be *discrete* if it can get only a finite or countably infinite number of possible distinct values.

The probability that the random variable Y will take on the value y, denoted by $\mathbb{P}(Y = y)$, is defined as the sum of the probabilities of all the sample points $\omega \in \Omega$ that are assigned the value of y.

4.4 Probability Distribution Representations

The probability distribution for a discrete random variable Y can be represented as a rule (formula), a table, or a graph.
Ex:

y	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(Y = y)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

is a probability distribution of y.

Theorem: For any discrete probability distribution, the following hold:

1. $\mathbb{P}(Y = y) \geq 0$, $\forall y$
2. $\sum_y \mathbb{P}(Y = y) = 1$

4.5 Expected Value

Let Y be a discrete random variable with a probability function $\mathbb{P}(Y = y)$. Then the expected value of Y, $E(Y)$, is defined to be

$$E(Y) = \sum_y y \cdot \mathbb{P}(Y = y).$$

The expected value exists only if the above summation is *absolutely convergent*. We often denote $E(Y)$ by μ.

Ex: Consider a random variable Y with the following probability distribution.

y	0	1	2
$\mathbb{P}(Y = y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Then

$$\mu = \sum_y y \cdot \mathbb{P}(Y = y) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = \frac{1}{2} + \frac{1}{2} = 1$$

4.6 Expected Value of a Function of a Random Variable

Theorem: Let Y be the discrete random variable with probability function $\mathbb{P}(Y = y)$, and let $g(Y)$ be a real-valued function of Y. Then, the expected value of $g(Y)$ is

$$E(g(Y)) = \sum_y g(Y) \cdot \mathbb{P}(Y = y).$$

4.7 Varience of a Random Variable

For a random variable Y with mean $E(Y) \equiv \mu$, the *varience* of Y, $V(Y)$, is defined as the expected value of $(Y-\mu)^2$. That is

$$V(Y) = E((y-\mu)^2).$$

The varience represents the "average squared deviation of Y from its mean". Intuitively, it's a measure of how variable a random variable is. The bigger the value of the varience is, the more spread out the values that the random variable can take on are.

4.8 Standard Deviation of a Random Variable

The *standard deviation*, σ, of a random variable Y, is defined as the principle square root of $V(Y)$. That is

$$\sigma^2 = V(Y) \iff \sigma = |\sqrt{V(Y)}|$$

This is a little easier to visualize than varience.

4.9 Some Theorems (without proof)

<u>Theorem:</u> Let Y be a discrete random variable with the probability function $\mathbb{P}(Y=y)$, and let $c \in \mathbb{R}$ be a constant. Then

$$E(c) = c$$

<u>Theorem:</u> Let Y be a discrete random variable with the probability function $\mathbb{P}(Y=y)$, let $g(Y)$ be a function of Y, and let c be a constant. Then

$$E(c \cdot g(Y)) = c \cdot E(g(Y))$$

<u>Theorem:</u> Let Y be a discrete random variable with the probability function $\mathbb{P}(Y=y)$, and let $g_1(Y), g_2(Y), \ldots, g_k(Y)$ be k functions of Y. Then

$$E(g_1(Y) + g_2(Y) + \cdots + g_k(Y)) = E(g_1(Y)) + E(g_2(Y)) + \cdots + E(g_k(Y))$$

<u>Theorem:</u> Let Y be a discrete random variable with the probability function $\mathbb{P}(Y=y)$ and mean μ. Then

$$V(Y) \equiv \sigma^2 = E((Y-\mu)^2) = E(Y^2) - \mu^2$$

4.10 Binomial Probability Distribution

The *binomial probability distribution* is a probability distribution that comes up pretty commonly. It applies where

- There's a sequence of independent or identical trials

- Each trial can result in one of two outcomes (flipping a coin, for example).

More precisely, a binomial experiment is such that

- Consists of a fixed number of n trials
- Each trial results in one of two possible outcomes: S(uccess), or F(ailure).
- The probability of S on any trial is p, and the probability of F on any trial is $q = 1 - p$.
- All trials are independent
- The random variable Y is the number of successes out of n trials.

A random variable Y is said to have binomial distribution based on n trials iff

$$\mathbb{P}(Y = y) = \binom{n}{y} p^n q^{n-y}$$

5 May 8, 2018

5.1 Bernoulli Distribution

Let X be a random variable. Toss a coin. The outcomes are H, and T. Set $X(H) = 1$, and $X(T) = 0$. We have $X(\Omega) = \{0, 1\}$.
Assume $\mathbb{P}(H) = p$, for $0 < p < 1$. Then the probability function $\mathbb{P}(X = x)$ of X is

x	0	1
$\mathbb{P}(X = x)$	1-p	p

A random variable X that has the above probability function is called a *Bernoulli random variable* on X. Also, X is said to have the *Bernoulli Distribution* with parameter p. We write $X \sim \text{Ber}(p)$.

Proposition: If $X \sim \text{Ber}(p)$, then

1. $E(X) = p$
2. $V(X) = p(1 - p)$

Proof of (1):
$E(X) = \sum_x x \cdot \mathbb{P}(X = x) = 0 \cdot \mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1) = p.$ □

Proof of (2):
$V(X) = E(X - E(X)^2) = E(X^2) - (E(X))^2.$
Also, $E(X^2) = \sum_x x^2 \cdot \mathbb{P}(X = x) = p.$
Therefore, $V(X) = p - p^2 = p(1 - p).$ □

5.2 Some Properties of E, and V

Let X be a random variable, and let $c, \alpha \in \mathbb{R}$ be constants.

1. $E(X + c) = E(X) + c$
2. $V(X + c) = V(X)$
3. $E(\alpha X) = \alpha E(X)$
4. $V(\alpha X) = \alpha^2 V(X)$

5.3 A Proof that $V(X) = E(X^2) - (E(X))^2$

Property: If f, g are functions, then $E(f(x) + g(x)) = E(f(x)) + E(g(x))$.

Set $\mu := E(X)$. Then, by definition,

$$\begin{aligned}
V(X) &= E((X - E(X))^2) \\
&= E(x^2 - 2\mu x + \mu^2) \\
&= E(X^2) - 2\mu E(X) + \mu^2 \\
&= E(X^2) - 2\mu^2 + \mu^2 \\
&= E(X^2) - \mu^2 \\
&= E(X^2) - (E(X))^2
\end{aligned}$$

□

5.4 More on Binomial Distribution

Proposition: If $X \sim \mathrm{Bin}(n, p)$, then

1. $E(X) = np$
2. $V(X) = np(1-p)$

To prove there we first have to note that

$$k \cdot \binom{n}{k} = k \cdot \frac{n!}{k!(n-k)!} = n \cdot \frac{(n-1)!}{(k-1)!(n-k)!} = n \cdot \binom{n-1}{k-1}$$

Proof of (1): By definition,

$$\begin{aligned}
E(X) &= \sum_{k=0}^{n} k \cdot \mathbb{P}(X = k) \\
&= \sum_{k=0}^{n} k \cdot \binom{n}{k} \cdot p^k (1-p)^{n-k} \\
&= \sum_{k=0}^{n} n \cdot \binom{n-1}{k-1} \cdot p^k (1-p)^{n-k}
\end{aligned}$$

Let $\ell = k - 1$

$$\begin{aligned}
&= n \cdot \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} \cdot p^{\ell+1} (1-p)^{(n-1)-\ell} \\
&= np \cdot \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} \cdot p^{\ell} (1-p)^{(n-1)-\ell} \\
&= np(p + (1-p))^{n-1} \\
&= np
\end{aligned}$$

$\square$

<u>Proof</u> of (2): By definition,

$$\begin{aligned}
V(X) &= E(X^2) - n^2p^2 \\
&= E(X(X-1) + X) - n^2p^2 \\
&= E(X(X-1)) + np - n^2p^2
\end{aligned}$$

Now we need to reduce $E(X(X-1))$ down.

$$\begin{aligned}
E(X(X-1)) &= \sum_{k=0}^{n} k \cdot (k-1) \cdot \binom{n}{k} \cdot p^k(1-p)^{n-k} \\
&= \sum_{k=2}^{n} k \cdot (k-1) \cdot \binom{n}{k} \cdot p^k(1-p)^{n-k}
\end{aligned}$$

Now we need to change the variables in $k(k-1) \cdot \binom{n}{k}$.

$$k \cdot (k-1) \cdot \binom{n}{k} = n \cdot (n-1) \cdot \binom{n-2}{k-2}$$

so now plugging that back into $E(X)$ we get

$$\begin{aligned}
E(X(X-1)) &= \sum_{k=2}^{n} k \cdot (k-1) \cdot \binom{n}{k} \cdot p^k(1-p)^{n-k} \\
&= \sum_{k=2}^{n} n \cdot (n-1) \cdot \binom{n-2}{k-2} \cdot p^k(1-p)^{n-k}
\end{aligned}$$

Let $\ell = k - 2$

$$\begin{aligned}
&= n \cdot (n-1) \cdot \sum_{\ell=0}^{n-2} \binom{n-2}{\ell} \cdot p^\ell(1-p)^{(n-2)-\ell} \\
&= n \cdot (n-1) \cdot p^2 \cdot (p + (1-p))^{n-2} \\
&= n \cdot (n-1) \cdot p^2
\end{aligned}$$

Plugging this back into $V(X)$, we get

$$\begin{aligned}
V(X) &= E(X(X-1)) + np - n^2p^2 \\
&= n \cdot (n-1) \cdot p^2 + np - n^2p^2 \\
&= np - np^2 \\
&= np(1-p)
\end{aligned}$$

$\square$

Ex 1: 3.60 from the textbook

Solution Fish die with $\mathbb{P} = 0.2$. Therefore the prob of a success (they survive), $\mathbb{P}(S) = 0.8$. Let the random variable X be the number of fish that survive. There are 20 fish. Thus $X \sim \text{Bin}(n = 20,\, p = 0.8)$. The probability that 14 fish survive is

$$\mathbb{P}(X = 14) = \binom{20}{14} \cdot (0.8)^{14} \cdot (0.2)^{6} \approx 0.1091 = 10.91\%$$

5.5 Geometric Distribution

FIrst recall that

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x},\ |x| < 1.$$

Suppose an experiment leads to either an $S(uccess)$ or an $F(ailure)$. Suppose we want to repeat the experiment until an S occurs

Let $\mathbb{P}(S) = p$, when $0 < p < 1$, and of course, $\mathbb{P}(F) = 1 - p$. Then

$$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$$

where

$$\omega_n = FFF \ldots FS$$

with $n - 1$ Fs.

Let the random variable X reperesent the number of trials needed for as S to occur. Then

$$\Omega(X) = \mathbb{N}$$

The probability function of X is then

$$\mathbb{P}(X = k) = p \cdot (1-p)^{k-1}$$

If X satisfies this then we write $X \sim \text{Geometric}(p)$.

<u>Proposition:</u>

1. $E(X) = \frac{1}{p}$
2. $V(X) = \frac{1}{p} \cdot (\frac{1}{p} - 1)$

Proof of (1): By definition

$$
\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X = k) \\
&= \sum_{k=1}^{\infty} k \cdot p \cdot (1-p)^{k-1} \\
&= p \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \\
&= p \cdot \frac{1}{(1-(1-p))^2} \\
&= \frac{p}{p^2} \\
&= \frac{1}{p}
\end{aligned}
$$

□

Proof of (2): By definition

$$
\begin{aligned}
V(X) &= E(X^2) - (E(X))^2 \\
&= E(X^2) - \frac{1}{p^2} \\
&= E(X(X-1)) + E(X) - \frac{1}{p^2} \\
&= E(X(X-1)) + \frac{1}{p} - \frac{1}{p^2}
\end{aligned}
$$

Now we need to reduce down $E(X(X-1))$

$$
\begin{aligned}
E(X(X-1)) &= \sum_{k=1}^{\infty} k \cdot p \cdot (k-1) \cdot (1-p)^{k-1} \\
&= p \cdot (1-p) \cdot \sum_{k=2}^{\infty} k \cdot (k-1)(1-p)^{k-2} \\
&= p \cdot (1-p) \cdot \frac{2}{p^3} \\
&= \frac{2(1-p)}{p^2}
\end{aligned}
$$

Plugging this back into $V(X)$, we get

$$\begin{aligned} V(X) &= E(X(X-1)) + \frac{1}{p} - \frac{1}{p^2} \\ &= \frac{2(1-p)}{p^2} + \frac{1}{p} - \frac{1}{p^2} \\ &= \frac{2-2p+1-1}{p^2} \\ &= \frac{1-p}{p^2} \\ &= \frac{1}{p} \cdot \left(\frac{1}{p} - 1\right) \end{aligned}$$

$\square$

6 May 9, 2018

7 May 10, 2018

8 May 14, 2018

9 May 16, 2018

10 May 17, 2018

11 May 21, 2018

12 May 22, 2018

13 May 23, 2018

14 May 24, 2018

15 May 28, 2018

16 May 29, 2018

17 May 30, 2018

18 May 31, 2018