Notes - Math 135A

May 6, 2014

1/5

Question

I have a deck of cards and I want to distribute all 52 cards to 4 players equally. What is the probability that each player gets an Ace.

Answer

$\binom{52}{4}$ for positions for 4 aces.
13^4 for positions to pick for first slot, second slot, third slot, and fourth slot.
Therefore, it's $\frac{13^4}{\binom{52}{4}}$.

History of Probability

1654: Chevabu de Mere, Blasse Pascal.
Fair coin: H or T, with equal probability
Die: 1, 2, 3, 4, 5, 6 with equal probability
Shuffled deck of cards: Any ordering fo cards equally likely
In the letter between Chevabu de Mere and Balsse Pascal, de Mere wants a 6 from 4 rolls of a die. According to de mere, of a fair die, the probability of getting at least one 6 from 4 rolls of a die is $4 * \frac{1}{6} = \frac{2}{3}$. De Mere won money on this. The probability of getting at least one double 6 in 24 rolls of a pair of dice is $24 * \frac{1}{36} = \frac{2}{3}$. De mere lost money on this.
Why? He was WRONG! The logic is bad. We'll come back to this later.

Question 2

In a family with 4 children, what is the probability of a 2:2 boy:girl split.
Wrong Answers: One guess: $\frac{1}{5}$, WRONG! 5 possibilities for number of boys are not equally likely.
Another guess: Close to 1. They're almost equally likely. WRONG! This idea shows the misunderstanding that probability = certainty. That is wrong.

Equally likely outcomes

Suppose an experiment is performed, with n possible outcomes. ASSUME also, that each outcome are equally likely. (Whether this concept is realistic, *shrug*. In Question 2, this was a bad assumption.) If an event E (= a set of outcomes) consists of m different outcomes ("good" outcomes for E), then the probability of E, $P(E) = \frac{m}{n}$

Example

Fair die has 6 outcomes, $E = \{2,4,6\}$, $P(E) = \frac{1}{2}$.
What does this mean? For arbitrary large number, N, of rolls, about half of the outcomes will be even.

1/7

Example: Roll two dice. What's the most likely sum? Outcomes: (i, j), $1 \le i \le 6$ and $1 \le j \le 6$

sum	Frequency
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1

Our answer is 7.
$P(sum = 7) = \frac{6}{36} = \frac{1}{6}$.
Example: 4 kids in a family. Probability of two sons and two daughters.
We have 16 outcomes. We can list them, but... there's got to be a better way.

BBBB	BBBG	BBGB	BBGG
BGBB	BGBG	BGGB	BGGG
GBBB	GBBG	GBGB	GBGG
GGBB	GGBG	GGGB	GGGG

6 of them are two boys and two girls.
$P(2-2 \text{ split}) = \frac{6}{16} = \frac{3}{8}$
$P(1:3 \text{ split or } 3:1 \text{ split}) = \frac{8}{16} = \frac{1}{2}$
$P(4:0 \text{ split or } 0:4 \text{ split}) = \frac{2}{16} = \frac{1}{8}$

How to count?

Listing them all is very inefficient especially as the sample gets bigger. There's a better way.

> Basic principle: If an experiment consists of 2 parts, first part has m outcomes, and the second part has n outcomes regardless of the outcome in the first part, then the experiment as a whole has mn outcomes.

Example: roll a dice 4 times.
P(get different numbers)?
Number of outcomes: 6^4 tuples
Number of good outcomes: $6 * 5 * 4 * 3$, 6 options for the first roll, 5 options for the second roll since

it must be not be the number in the first roll. 4 options for the third roll since it must not be the first two rolls. etc...
Example: Last time about the dice rolling game, Game 1: 4 rolls and get at least one 6.
Number of good events $= 6^4 - 5^4$ (Total - bad)
$P(mn) = 1 - (\frac{5}{6})^4 \approx .5177$
Game 2: 24 rolls of two die and get at least one pair of 6 rolled.
Number of outcomes $= 36^24$
Number of good outcomes $= 36^24 - 35^24$
$P(\text{win}) = 1 - \left(\frac{35}{36}\right)^{21} \approx .4914$
de Mere overcounted. $4 * 6^3$ means that I pick a 6 in one spot and any of the 6 in the next one.
Convergence can be slow sometimes.

1/9

n objects. The number of ways to fill n ordered slots with them is $n(n-1)\ldots 2 * 1$
the number of ways to fill $k \leq n$ slots is $n(n-1)\ldots(n-1+1) = \frac{n!}{(n-k)!}$
Example: Shuffle a deck of cards

1. $P(\text{top card is an Ace}) = \frac{1}{13} = \frac{4*51!}{52!}$
2. $P(\text{all card of the same suit end up stuck together}) = \frac{4!*(13!)^4}{52!}$
3. $P(\text{hearts are together}) = \frac{40!13!}{52!}$

Example: A bag has 6 pieces of paper with letters, E, E, P, P, P, and R, on them. Pull 6 papers blindly out of the bag (1) without, (2) with replacement. What is the probability that these pieces in order, spell PEPPER?

1. Outcome = ordering of pieces of paper. $E_1E_2P_1P_2P_3R$. The number of outcomes is 6!. The number of good outcomes is 3!2!. The probaility is $\frac{3!2!}{6!}$.
2. $\frac{3^3 2^2}{6^6}$

Example: 6 men and 6 women sit (1) on a row (2) around a table, at random. $P(\text{all women sit together})$

1. $\frac{4!6!}{12!}$
2. $\frac{6!6!}{11!}$

Example: We have 2 Norwegians, 3 Swedes, and 4 Finns. What is the probability of all groups end up sitting together with their own group?
Probability is $\frac{2*2*4!*3!}{8!}$. Fix a Norwegian. 2 choices for picking a place where the 2nd Norwegian sits. 4! choices for arranging the Finns. 3! choices for arranging the Swedes. 2 choices to pick which blocks are to the left of Norwegian.

1/12

Let $\binom{n}{r}$ be the number of choices of a subset with r elements from a set with n elements. Then,

$$\begin{aligned}\binom{n}{r}r! &= n(n-1)\dots(n-r+1)\\ \binom{n}{r} &= \frac{n(n-1)\dots(n-r+1)}{r!}\\ &= \frac{n!}{r!(n-r)!}\end{aligned}$$

Note: $\binom{n}{0} = 1$ and $\binom{n}{r} = \binom{n}{n-r}$
$\binom{n}{r} =$ "n choose r"

More generally, the number of ways to choose a set of n elements into r groups of $n_1, n_2, \dots, n_r$ elements where $n_1 + \dots + n_r = n$ is

$$\begin{aligned}\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\dots\binom{n-n_1-\dots-n_{r-1}}{n_r} &= \frac{n!}{n_1!n_2!\dots n_r!}\\ &= \binom{n}{n_1\dots n_r}\end{aligned}$$

Example: A fair coin is tossed 10 times.
$P(\text{exactly 5 heads}) = \frac{\binom{10}{5}}{2^{10}} \approx 0.246$

Example: We have a bag that contains 100 balls, with 50 red and 50 blue. Select 5 balls at random. What is the probability that 3 are blue and 2 are red?
$P(\text{3 are blue and 2 are red}) = \frac{\binom{50}{3}\binom{50}{2}}{\binom{100}{5}} \approx .319$
Why should it be less than half? The probability that 3 are blue and 2 are red is equal to the probability of 3 are red and 2 are blue, so if it's greater than half, it's nonsense since both add to over 1. It cannot be half either because there is a probability greater than 0 for other choices.

Example: We have 52 cards from a standard deck. We shuffle them and dealt them to 4 players.

$P(\text{each has an ace})$?
Way 1:

1. 52! total choices
2. Let first 13 cards go to the first player, second 13 cards to the second player, etc... Pick a spot with each of the 13 slots for the ace to be in. There are 13^4 possibility to where each ace resides.
3. Putting the aces into those positions are 4! since there are four different aces
4. Ordering the rest of the cards, 48! ways.

The probability is then $\frac{13^4 4! 48!}{52!}$

Way 2:

1. Outcomes = positions of 4 A's.
2. Total outcomes = $\binom{52}{4}$. Pick four aces from 52.
3. The number of ways to order the A's, 13^4

The probability is then $\frac{13^4}{\binom{52}{4}}$

P(one person has all four aces)?

1. total outcomes = $\binom{52}{4}$
2. Pick one player $\binom{4}{1}$
3. Pick four slots for aces for that player $\binom{13}{4}$

The probability is then $\frac{\binom{4}{1}\binom{13}{4}}{\binom{52}{4}}$

Example: Roll a die 12 times.
P(each number appears exactly two times)?

1. Total outcomes 6^{12}
2. Pick two for 1, pick two for 2, etc... $\binom{12}{2}\binom{10}{2}\ldots\binom{2}{2}$

The probability is then $\frac{\binom{12}{2}\binom{10}{2}\ldots\binom{2}{2}}{6^{12}}$

P(1 appears exactly 3 times, 2 appears exactly 2 times)?

1. total outcomes, 6^{12}
2. Pick three buckets for 1 and 2 for 2 and 7 for any number other than 1. $\binom{12}{3}\binom{9}{2}4^7$

The probability is then $\frac{\binom{12}{3}\binom{9}{2}4^7}{6^{12}}$

Example: We have 14 rooms with 4 colors, w, b, g, y. Let's paint each room at random.
$P($ 5w, 4b, 3g, 2y$) = \frac{\binom{14}{5}\binom{9}{4}\binom{5}{3}\binom{2}{2}}{4^14}$

1/16

Example: 3 married couples. Take seats around a table at random. we want to figure out P(no wife sits next to the
First compute the complement, but unions are hard compute, so we can compute this instead: $1 - (P(A_1) + P(A_2) + P(A_3) - P(A_1 \bigcap A_2) - P(A_1 \bigcap A_3) - P(A_2 \bigcap A_3) + P(A_1 \bigcap A_2 \bigcap A_3)$ where $A_i = \{\text{wife}_i \text{ and husband}_i \text{ sit together}\}$
$P(A_1) = \frac{2}{5} = P(A_2) = P(A_3)$
$P(A_1 \bigcap A_2) = P(A_1 \bigcap A_3) = P(A_2 \bigcap A_3) = \frac{2*2*3!}{5!} = \frac{1}{5}$.
Fix husband 1. 2 places for wife to be. 3! ways of picking where second couple sits. 2 ways to vary the

couples. $P(A_1 \bigcap A_2 \bigcap A_3) = \frac{2*2!*2^2}{5!}$
Our answer is $\frac{4}{15}$

Example: Birthday Problem
n possible birthdays, sample k times with replacement.
$P(\text{at least 2 people share a birthday}) = 1 - P(\text{no shared birthdays}) = 1 - \frac{n*(n-1)...(n-k+1)}{n^k}$
Example(Massachusetts lottery): $n = 10,000$, $k = 660$
Pick a 4 digit number a day. 660 days have passed with no numbers being the same. What is this probability? $P(\text{no shared birthdays or lottery numbers}) = 2.191 * 10^{-10}$

Example: $P(\text{all } n \text{ birthday are represented})$?
Suppose all birthdays are represented. That probability is $\frac{n!}{n^n}$.
What about our problem? The probability is then $1 - P(\bigcup_{i=1}^{n} A_i)$ where $A_i = \{i\text{th birthday is missing}\}$
$P(A_1) = \frac{(n-1)^k}{n^k} = P(A_i)\ \forall i$
$P(A_1 \bigcap A_2) = \frac{(n-2)^k}{n^k} = P(A_i \bigcap A_j)$ where $i < j$
$\vdots$
We then get $1 - n\left(\frac{n-1}{n}\right)^k + \binom{n}{2}\left(\frac{n-2}{n}\right)^k - \ldots = \sum_{i=0}^{n-1}\binom{n}{i}(-1)^i(1-\frac{i}{n})^k$
As it turns out for the above formula, the probability is $\frac{n!}{n^n}$ for $k = n$, but about 0 when $k < n$. This is similar to the sterling numbers of the second kind.

Example: Matching Problems
n employees, each buys a gift. These are then assigned at random.

$$\begin{aligned}
P(\text{somebody gets own gift}) &= P(\bigcup_{i=1}^{n} A_i) \text{ where } A_i = \{i\text{th person gets own gift}\} \\
P(A_1) &= \frac{1}{n} \\
&= P(A_i) \\
P(A_1 \bigcap A_2) &= \frac{(n-2)!}{n!} \\
&= \frac{1}{n(n-1)} \\
P(\ldots) &= n * \frac{1}{n} - \binom{n}{2}\frac{1}{n(n-1)} + \ldots \\
&= (\text{as } n \to \infty)\ 1 - \frac{1}{e}
\end{aligned}$$

1/20 (Discussion)

Sample Spaces

In Combinatorial sense,

1. Choosing w/o replacement. Given S, sample space is the subsets of S of given size.
   i.e. 10 keys, 1 key open your door. What's the probability that you can open door on 5th try?
   Total number of subsets of size 5: $\binom{10}{5}$ Number of ways to open on 5th: $\binom{9}{4}\binom{1}{1}$

So, $P = \frac{\binom{9}{4}\binom{1}{1}}{\binom{10}{5}}$

2. Choosing with replacement. k chairs in S, Sample space is $S \times \ldots \times S$.
   i.e. Again, 10 keys, try without replace, k to open door, What is the probability of opening on the 5th try?
   Total number of choices: 10^5
   Number of ways to suceed: $9^4 * 1$
   So, $P = \frac{9^4}{10^5}$

3. Inclusion-Exclusion Principle
   Given events, $A_1, \ldots, A_n$
   $P(A_1 \bigcup \ldots \bigcup A_n) = \sum P(A_i) - \sum P(A_j \bigcap A_k) + \sum P(A_l \bigcap A_m \bigcap A_p) \ldots$

4. Mutually Exclusive Events
   Two events are mutually exclusive if $A_i \bigcap A_j = \emptyset$

5. Roll a fair die 10 times
   P(at least one number appears exactly once) =?
   Total number of outcomes is 6^{10}
   Consider $B_1, \ldots, B_6$
   $P(B_1 \bigcup \ldots \bigcup B_6) 6P(B_1) - \binom{6}{2}P(B_1 \bigcap B_2) + \binom{6}{3}P(B_1 \bigcap B_2 \bigcap B_3) - \binom{6}{4}P(B_1 \bigcap \ldots \bigcap B_4) + \binom{6}{5}P(B_1 \bigcap \ldots \bigcap B_5) - \binom{6}{6}P(B_1 \bigcap \ldots \bigcap B_6)$

<u>Problem 2.40</u>: $\exists 4$ TV repairmen
4 TV sets break down
P(exactly i are called)?
Total outcomes $= 4^4$
$P_1 = \frac{\binom{4}{1}}{4^4}$
$P_2 = \frac{\binom{4}{2}(\binom{4}{1}+\binom{4}{2}+\binom{4}{3})}{4^4}$
The $\binom{4}{2}$ ways is to choose repairman and the rest is to choose customers.

1/21

<u>Example</u>: Ten keys with only one that opens the door. Try one by one and discard after each unsuccessful try.
P(succeed on 5th try) $= \frac{1}{10}$
P(suceed on or before 5th try) $= \frac{1}{2}$

<u>Example</u>: Roll a die 12 times.
P(a number occurs 6 times and two other number occur three times each)

1. Sample Size: 6^{12}

2. Choices to pick the first number that occurs 6 times: $\binom{6}{1} = 6$

3. Choices to pick the next two that occurs 3 times each: $\binom{5}{2}$

4. Pick spots for the number that occurs 6 times: $\binom{12}{6}$

5. Pick spots for one of the numbers that occur 3 times: $\binom{6}{3}$

Therefore, our probability is $\frac{6\binom{5}{2}\binom{12}{6}\binom{6}{3}}{6^{12}}$

Example: We have 10 pairs of socks in the closet. I will pick 8 socks at random.
$P(i$ complete pair of socks)?

1. Samples size: $\binom{20}{8}$ Pick 8 socks out of 20 socks (10 pairs)

2. Pick i pairs of socks of the 10: $\binom{10}{i}$

3. Pick other socks. Make sure that you don't pick pairs again: $\binom{10-i}{8-2i}$

4. Vary the complete pairs: 2^{8-2i}

Therefore, our probability is $\frac{2^{8-2i}\binom{10-i}{8-2i}\binom{10}{i}}{\binom{20}{8}}$

Example: **Poker Hands**
In the definitions, the word *value* refers to A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2. This sequence also describes the relative rankings of cards, with one exception: an Ace may be regarded as 1 for the purposes of making a straight:
(a) *one pair*: two cards of same value plus 3 cards with different values

$$J\spadesuit\ J\clubsuit\ 9\heartsuit\ Q\clubsuit\ 3\spadesuit$$

(b) *two pairs*: two pairs plus another card of different value

$$J\spadesuit\ J\clubsuit\ 9\heartsuit\ 9\clubsuit\ 3\spadesuit$$

(c) *three of a kind*: three cards of the same value plus two with different values

$$J\spadesuit\ J\clubsuit\ J\heartsuit\ 9\clubsuit\ 3\spadesuit$$

(d) *straight*: five cards with consecutive values

$$5\heartsuit\ 4\clubsuit\ 3\clubsuit\ 2\heartsuit\ A\spadesuit$$

(e) *flush*: five cards of the same suit

$$K\clubsuit\ 9\clubsuit\ 7\clubsuit\ 6\clubsuit\ 3\clubsuit$$

(f) *full house*: a three of a kind and a pair

$$J\clubsuit\ J\diamondsuit\ J\heartsuit\ 9\clubsuit\ 9\spadesuit$$

(g) *four of a kind*: four cards of the same value

$$J\clubsuit\ J\diamondsuit\ J\heartsuit\ J\clubsuit\ 9\spadesuit$$

(e) *straight flush*: five cards of the same suit with consecutive values

$$A\clubsuit\ K\clubsuit\ Q\clubsuit\ J\clubsuit\ 10\clubsuit$$

hand	no. combinations	approx. prob.
one pair	$13 \cdot \binom{12}{3} \cdot \binom{4}{2} \cdot 4^3$	0.422569
two pairs	$\binom{13}{2} \cdot 11 \cdot \binom{4}{2} \cdot \binom{4}{2} \cdot 4$	0.047539
three of a kind	$13 \cdot \binom{12}{2} \cdot \binom{4}{3} \cdot 4^2$	0.021128
straight	$10 \cdot 4^5$	0.003940
flush	$4 \cdot \binom{13}{5}$	0.001981
full house	$13 \cdot 12 \cdot \binom{4}{3} \cdot \binom{4}{2}$	0.001441
four of a kind	$13 \cdot 12 \cdot 4$	0.000240
straight flush	$10 \cdot 4$	0.000015

Note. The probabilities of straight and flush above include the possibility of a straight flush. The number of all outcomes is $\binom{52}{5} = 2,598,960$. Then, for example, for the *three of a kind*, the number of good outcomes is obtained by choosing the value for the three cards with the same value, then the values of other two cards, then three cards from the four of the same chosen value, then a card from each of the two remaining chosen values.

Assumption: Every hand is dealt at random
$P(\text{one pair})$?

1. Sample Size: $\binom{52}{5}$ Pick five cards out of 52
2. Pick a number: 13 choices
3. Pick the other three numbers that don't make a pair: $\binom{12}{3}$
4. Pick which pairs of the chosen number: $\binom{4}{2}$
5. Vary the suits of the other three number: 4^3

Therefore, our probability is $\frac{13\binom{4}{2}\binom{12}{3}}{\binom{52}{5}}$

$P(\text{flush})$?

1. Sample Size: $\binom{52}{5}$ Pick five cards out of 52
2. Pick a suit: 4 choices
3. Pick five numbers: $\binom{13}{5}$

Our probability is $\frac{4\binom{52}{5}}{\binom{13}{5}}$.
This includes straight flushes.

$P(\text{straight flush})$?

1. Sample Size: $\binom{52}{5}$ Pick five cards out of 52
2. Pick a suit: 4
3. Pick the beginning number: 10. Not cyclic

Therefore, our probability is $\frac{4*10}{\binom{52}{5}}$

$P(\text{nothing}) = P(\text{all cards with different values}) - P(\text{straight or flush})$?

1. Sample Size: $\binom{52}{5}$ Pick five cards out of 52
2. Pick 5 numbers of 13: $\binom{13}{5}$
3. Vary the suits: 4^5

Therefore, our probability is $\frac{\binom{13}{5}4^5}{\binom{52}{5}} - (P(\text{straight}) + P(\text{flush}) - P(\text{straight flush}))$
It's a little over half the time.

Example:**Football players example.**

Assume that 20 players, 10 offensive and 10 defensive, are to be distributed at random into 10 rooms, 2 per room. What is the probability that exactly $2i$ rooms are mixed, $i = 0, \ldots 5$?

This is an example when careful thinking about what outcomes should be really pays off. Consider the following model for distributing players into rooms. First arrange them at random onto a row of 20 slots $S1, S2, \ldots, S20$. Room 1 then takes players in slots $S1, S2$, so let's call these two slots $R1$. Similarly, room 2 takes players in slots $S3, S4$, so let's call these two slots $R2$, etc.

Now it is clear that we only need to keep track of distribution of 10 d's into the 20 slots, corresponding to the positions of 10 defensive players. Any such distribution constitutes an outcome, and they are equally likely. Their number is $\binom{20}{10}$.

To get $2i$ (for example, 4) mixed rooms, start by choosing $2i$ (ex., 4) out of 10 rooms which are going to be mixed. There are $\binom{10}{2i}$ of these choices. You also need to make a choice into which slot in each of the $2i$ chosen mixed rooms the d goes, for 2^{2i} choices.

Once these two choices are made, you still have $10 - 2i$ (ex., 6) d's to distribute into $5 - i$ (ex., 3) rooms, as there are two d's in each of those rooms. For this, you need to choose $5 - i$ (ex., 3) rooms out of the remaining $10 - 2i$ (ex., 6), for $\binom{10-2i}{5-i}$ choices, and this choice fixes a good outcome.

The final answer therefore is

$$\frac{\binom{10}{2i}2^{2i}\binom{10-2i}{5-i}}{\binom{20}{10}}$$

1/23

In general, for events, E and F, $P(F) > 0$, the conditional probability of E given F is $P(E|F) = \frac{P(E \bigcap F)}{P(F)}$
Conditional Probability are somtimes given especially in sequential random experiments.

Then, you can use $P(A_1 \bigcap A_2) = P(A_1)P(A_2|A_1)$ and $P(A_1 \bigcap A_2 \bigcap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \bigcap A_2)$ etc...

Example: An urn contains 10 black and 10 white balls. Draw 3 without replacement. $P($ all 3 are white$)$?

1. Sample size: $\binom{20}{3}$
2. pick 3 balls out of 10 whites: $\binom{10}{3}$

Our probability is then $\frac{\binom{10}{3}}{\binom{20}{3}}$

Let's try it another way.
If you do the drawing sequentially, $P($ 1st is white, 2nd is white, and 3rd is white$)$

1. 1st ball is white: $\frac{1}{2}$
2. 2nd ball is white after the first one picked is white: $\frac{9}{19}$
3. 3rd ball is white after the first two picked are white: $\frac{8}{18}$

Our probability is also $\frac{1}{2} \times \frac{9}{19} \times \frac{8}{18}$. Turns out to be the same.

Let's try the example with replacement: $\left(\frac{1}{2}\right)^3$

Example: Flip a fair coin: If you get a heads, roll 1 dice. If you get a tails, roll 2 die.
$P($ get a T and roll one 6$) = P(\text{get T})P(\text{exactly one 6} \mid \text{get T})$?

1. Get tails: $\frac{1}{2}$
2. Exactly one 6 given that I get T $\frac{2\times 5}{36}$

Our probability is then $\frac{1}{2} \times \frac{2\times 5}{36}$

Theorem 1. *Bayes' formula Assume that $F_1, \ldots, F_n$ are distinct and $F_1 \bigcup \ldots \bigcup F_n =$ $P($one of them always happens$)$. Then, for any event, E,*

$$\begin{aligned} P(E) &= P(F_1)P(E|F_1) + P(F_2)P(E|F_2) + \ldots + P(F_n)P(E|F_n) \\ &= P(E \bigcap F_1) + \ldots + P(E \bigcap F_n) \\ &= P((E \bigcap F_1) \bigcup \ldots \bigcup (E \bigcap F_n) \\ &= P(E \bigcap (F_1 \bigcup \ldots \bigcup F_n)) \end{aligned}$$

Example: Roll a die, pull as many cards from the deck as the die shows.
$F_i = \{\text{number on the die} = i\}$ where $i = 1, \ldots, 6$, which is $\frac{1}{6}$.

1. 1 card: $\frac{1}{13}$
2. 2 cards: $\frac{\binom{48}{2}}{\binom{52}{2}}$
3. etc...

$P(\text{get at least one ace}) = \frac{1}{6}\frac{1}{13} + \ldots$

1/26

<u>Example</u>: You roll a die, your friend tosses a coin.

- Roll six $\rightarrow$ win outright
- If you do not roll a six, your friend tosses a head $\rightarrow$ lose outright
- If neither, the game is repated until decided

$P(\text{you win}) = ?$
One way: $P(\text{you win}) = \frac{1}{6} + \frac{5}{6}\frac{1}{2}\frac{1}{6} + (\frac{5}{6}\frac{1}{2})^2\frac{1}{6} + \ldots$
$\frac{1}{6}$: Probability of getting a 6 $\frac{5}{6}\frac{1}{2}\frac{1}{6}$: Do not roll a 6, get tails, and roll a 6.

$D = \{\text{game is decided on 1st step}\}$
$W = \{\text{you win}\}$

$$\begin{aligned} P(W) &= P(W|D)P(D) + P(W|D^c)P(D^c) \\ &= P(W|D)P(D) + P(W)(1 - P(D)) \\ P(W) &= P(W|D) \text{ (with a little algebra)} \end{aligned}$$

$P(W|D^c) = P(W)$ because your probability of winning on the next round is the same.
$P(W|D) = \frac{P(W\bigcap D)}{P(D)} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{5}{6}\frac{1}{2}} = \frac{2}{7}$
$\frac{1}{6}$ means that you win on the first round by rolling a 6
$\frac{5}{6}\frac{1}{2}$? Probability of winning the first round by not rolling a 6, but getting heads.

<u>Example</u>: "Craps"
Two dice are rolled, the sum, S, is observed.

- $S \in \{7, 11\} \rightarrow$ win
- $S \in \{2, 3, 12\} \rightarrow$ lose
- $S \in \{4, 5, 6, 8, 9, 10\} \rightarrow$ dice are rolled until the following
  - S repeats $\rightarrow$ *win*
  - 7 appears $\rightarrow$ *win*

1. Win on the first roll: $\frac{8}{36}$
2. Lose on the first roll: $\frac{4}{36}$
3. Otherwise,
   - roll a 4: $\frac{3}{36} \rightarrow$ win: $\frac{\frac{3}{36}}{\frac{3}{36}+\frac{6}{36}} = \frac{3}{3+6} = \frac{1}{3}$
   - roll a 5: $\frac{4}{36}$
   - roll a 6: $\frac{5}{36}$
   - roll a 8: $\frac{5}{36}$

- roll a 9: $\frac{4}{36}$
- roll a 10: $\frac{3}{36}$

Using the 1st Bayes' formula: $P(\text{win}) = \frac{8}{36} + 2(\frac{3}{36}\frac{3}{3+6} + \frac{4}{36}\frac{4}{4+6} + \frac{5}{36}\frac{5}{5+6}) = .493$

Theorem 2. *2nd Bayes' formula:*
$S = F_1 \bigcup \ldots \bigcup F_n$, *where* F_i *is pairwise disjoint and* E *is an event.*
$P(F_j|E) = \frac{P(F_j \bigcap E)}{P(E)} = \frac{P(E|F_j)P(F_j)}{P(E|F_1)P(F_1)+\ldots+P(E|F_n)P(F_n)}$

Often times, F_j is called an hypothesis and $P(F_j)$ is called the prior probability and $P(F_j|E)$ is called the posterior probability.

Example:

Machines	percent items	item that comes from machine is defective with probability
M_1	20	.001
M_2	30	.002
M_3	50	.003

Pick an item, test it, and you find it's defective. What is the probability that it's made by M_2
$P($ item is defective $|$ made by $M_1) = .001$
$P($ item is defective $|$ made by $M_2) = .002$
$P($ item is defective $|$ made by $M_3) = .003$
$P(M_2|D) = \frac{.002 * .3}{.001 * .2 + .003 * .5 + .002 * .3} \approx .26$ where D is the event where a product is defective.

Example: Assume 10% of people have a certain disease. Test gives the correct diagnosis with probability of .8. A random person from the population has tested positive for the disease.
$P(\text{He's actually sick})$?
Let

1. $S = \{$ sick $\}$
2. $H = \{$ healthy $\}$
3. $T = \{$ tested positive $\}$

Now,

1. $P(H) = .9$
2. $P(S) = .1$
3. $P(T|H) = .2$
4. $P(T|S) = .8$

Now, $P(S|T) = \frac{P(T|S)P(S)}{P(T|S)P(S)+P(T|H)P(H)} = \frac{8}{31} \approx 31\%$

1/28

Example: We have a fair coin and an unfair coin which always tosses head. Choose one at random, toss it twice. It comes out heads.
P(the coin is fair)?
Let $F = \{\text{fair}\}$, $U = \{\text{unfair}\}$, and $E = \{$both tosses H.
Using the First Bayes' formula, you can compute E a-priori.
$P(F) = P(U) = \frac{1}{2}$. You have equal probability of picking fair and unfair coin a-priori.
$P(E|F) = \frac{1}{4}$. Given you picked fair coin, probability of both tosses heads
$P(E|U) = 1$. Given you picked unfair coin, probability of both tosses tails.
Our probability is $\frac{\frac{1}{2}\frac{1}{4}}{\frac{1}{2}\times\frac{1}{4}+\frac{1}{2}\times 1} = \frac{1}{5}$.

Example: OJ Simpson
A Dershowitz: not probable because $P(\text{an wife abuser kills wife}) = .001$
JF Merz and JP Caulkins cancels his logic because Dershowiz did not take into account the fact that someone actually murdered someone.
Let $S = \{\text{murdered women}\}$ with size 4936 (in 1992), out of which 1430 were killed by partners.
Let $A = \{\text{partner abused}\}$
Let $M = \{\text{partner murdered}\}$
$P(M) = .29$
$P(M^c) = .71$
You're interested in $P(M|A)$.
$P(A|M) = .5$
$P(A|M^c) = .05$
$P(M|A) = \frac{P(M)P(A|M)}{P(M)P(A|M)+P(M^c)P(A|M^c)} \approx 80\%$

Independence

Events, E and F are independent if

- $P(E|F) = P(E)$
- $P(E \bigcup F) = P(E)P(F)$

Note: If E and F are independent, so are E and F^c. Then,

- $P(E \bigcup F) = P(E)P(F^c)$
- Then, $P(E) - P(E \bigcap F) = P(E) - P(E)P(F)$

Note: Independence is often an assumption rather than something we find.

Example: Pull a random card from a full deck.
Let $E = \{\text{red}\}$
Let $F = \{\text{ace}\}$
$P(E) = \frac{1}{2}$, $P(E) = \frac{1}{13}$, and $P(E \bigcap F) = \frac{2}{52}$, so they're independent.

Pull two cards out of the deck sequentially.

$E = \{\text{first card is ace}\}$
$F = \{\text{2nd card is ace}\}$
$P(E) = P(F) = \frac{1}{13}$
$P(F|E) = \frac{3}{51}$, so not independent.

Example: Toss 2 fair coins
$E = \{$ head on 1st try$\}$
$F = \{$ head on 2nd try$\}$
This is independent.

Independence of $E_1, E_2, \ldots, E_n$
$P(E_{i_1} \bigcap \ldots \bigcap E_{i_k}) = P(E_{i_1})P(E_{i_2}) \ldots P(E_{i_k})$ where $1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq n$.

Example: Roll a four sided fair die.
$A = \{1, 2\}$, $B = \{2, 3\}$, $C = \{1, 4\}$
$P(A) = P(B) = P(C) = \frac{1}{2}$
$P(A \bigcap B) = P(A \bigcap C) = P(B \bigcap C)$.
These events are pairwise independent $= \frac{1}{4}$
Are all 3 independent?
$P(A \bigcap B \bigcap C) = \frac{1}{4} \neq \frac{1}{8}$
$A \bigcap B \subset C$
A and B cannot be independent if they are disjoint. If one happens, the other cannot happen. This is not independent because that would mean that it can still happen with the same probability.

1/30

Bernoulli trials

n indepdent experiments, success in each occurs with probability, p. (Failure is always with probability $1 - p$).
Then, $P(\text{exactly } k \text{ successes}) = \binom{n}{k} p^k (1-p)^{n-k}$

Example: A machine produces items which are independently defective with probability p.

1. Take 6 items. $P(\text{exactly two are defective}) = \binom{6}{2} p^2 (1-p)^4$
2. Take 6 items. $P(\text{at least one is defective}) = 1 - P(\text{ no defects}) = 1 - (1-p)^6$
3. $P(\text{at least 2 defective}) = 1 - (1-p)^6 - 6p(1-p)^5$
4. $P(\text{exactly 100 items are made before 6 defective are found})$
   $= P(\text{100th item defective, exactly 5 items among 1st 99}) = p \times \binom{99}{5} p^5 (1-p)^{99}$

Problem of points

We have independent trials with success with probability p.

1. $P(n \text{ successes occur before } m \text{ failures}) = P(\text{in first } m + m - 1 \text{ trials, the number of successes is} \geq n) = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}$

Example: "Best of 7".
Team A loses 1st game. (Assume Bernoulli trials with probability of $\frac{1}{2}$). Now,
$P(A \text{ wins the series}) = P(4successsbefore3failures) = \sum_{k=4}^{6} \binom{6}{k} p^k (1-p)^{6-k} = \frac{1}{2}$
Other notes:
$p_{n,m} = P(n \text{ successful before } m \text{ failures}) = P(\text{ first try and success}) P(n-1 \text{ success before } m \text{ fails}) + P(\text{ first try is failure}) P(n \text{ success before } m-1 \text{ fails}) = pp_{n-1,m} + (1-p)p_{n,m-1}$
Example: n birthdays and k people.
$p_{m,k} = P(\text{exactly } m \text{ days are not represented among the } n \text{ birthdays})$?
Notes: $P(k=0 \text{ and } m \text{ be anything except } n) = 0$
$P(m=n \text{ and } k=0) = 1$
$P(m=n \text{ and } k>0) = 0$.
$p_{m,k} = p_{m,k-1}\frac{n-m}{n} + p_{m+1,k-1}\frac{m+1}{n}$

Example: Each day, you decide with probability, p, independently to flip a fair coin. Otherwise, do nothing.
$P(\text{get exactly 10 heads in first 20 days}) = \binom{20}{10}(\frac{p}{2})^{10}(1-\frac{p}{2})^{10}$
$P(\text{10 heads before 5 tails}) = $ (just did disregard the days where you don't flip) $\sum_{k=10}^{14}(\binom{14}{k})\frac{1}{2^{14}}$

2/2

Example: Check pic circuit problem.png.
Each relay is independently conducting with probability p.
Then, $P($ current flows from A to B$)$?
Conditioned on 3 conditions, $\{1 \text{ or } 2 \text{ conduct}\} \bigcap \{4 \text{ or } 5 \text{ conducts }\}$
$(1-(1-p)^2)^2$
3 non-conducting: $\{ \text{ 1 and 4 conduct } \} \bigcup \{ \text{ 2 and 5 conducts}\}$
$1-(1-p^2)^2$
$P(\{\text{1 and 4 conducts}\}^c) = 1 - P(\text{1 and 4 conducts}) = 1-p^2$
Answer: $p(2p-p^2)^2 + (1-p)(2p^2-p^4)$

4 random variables

A "random variable" is a number whose value is the result of a random experiment.

Example:

1. Toss a coin 10 times, X = number of heads.
2. Choose a random point on $(0,1)$. X = distance from origin
3. Choose a random person, X = height of person

Mathematically, X is a real valued funciton on S, the space of outcomes.

Definition 3. *Discrete random variables are those with only countably many values, x_i with $i = 1, 2, \ldots$ and $p(x_i) = P(X = x_i)$ with $i = 1, 2, \ldots$ is known as the probability mass function*

Properties:

1. $p(x_i) > 0$
2. $P(X \in A) = \sum_{x_i \in A} p(x_i)$
3. $\sum_{i=1}^{\infty} p(x_i) = 1$

Example: X = number of heads in 2 pair coin tosses.
$P(X = 0) = \frac{1}{4} = P(X = 2)$
$P(X = 1) = \frac{1}{2}$
Example: An urn contains 20 balls, numbered $1, \dots, 20$. Pull out 5 balls at random.
X = largest number among selected balls.
$P(X = i)$ where $i = 5, \dots, 20$?

1. Sample space: $\binom{20}{5}$. Pick 5 balls of 20
2. $\binom{i-1}{4}$

$P(\text{at least one number at least 15}) = P(X \geq 15) = \sum_{i=15}^{20} P(X = i)$ Example: An urn contains 11 balls. We have 3 white balls, 3 red balls, and 5 blue balls.
Take out 3 balls at random.
You win \$1 for each red ball you get and lose a \$1 for each white ball you get.
$P(X = 0) =$

1. Sample Space: $\binom{11}{3}$ Pick three balls
2. Choose 2 red and 1 white (or vice versa for -1): $\binom{3}{2}\binom{3}{1}$
3. Choose 1 red/white and 2 blue: $\binom{3}{1}\binom{5}{2}$

$$P(X = -1) = P(X = 1) = \frac{\binom{3}{2}\binom{3}{1}\binom{3}{1}\binom{5}{2}}{\binom{11}{3}}$$

1. Sample Space: $\binom{11}{3}$ Pick three balls
2. Choose 2 red and 1 white (or vice versa for -1): $\binom{3}{2}\binom{3}{1}$
3. Choose 1 red/white and 2 blue: $\binom{3}{1}\binom{5}{2}$

$P(X = -2) = P(X = 2) = \frac{\binom{3}{2}\binom{5}{1}}{\binom{11}{3}}$

1. Sample Space: $\binom{11}{3}$ Pick three balls
2. Only way to get 2 or -2 is to get 2 white/red and a blue, so $\binom{3}{2}\binom{5}{1}$

$P(X = -3) = P(X = 3) = \frac{1}{\binom{11}{3}}$

1. Sample Space: $\binom{11}{3}$ Pick three balls
2. only one possibility of picking

2/3 (Discussion)

What we have so far

1. Inclusion-exclusion: $P(A_1 \bigcup \ldots \bigcup A_n) = P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n) - P(A_1 \bigcap A_2) - \ldots - P(A_1 \bigcap A_n) - P(A_2 \bigcap A_3) - \ldots P(A_{n-1} \bigcap A_n) + P(A_1 \bigcap A_2 \bigcap A_3) + \ldots$

2. $P(A|B) = \frac{P(A \bigcap B)}{P(B)}$

Problem 3.41: Let A be the event that an ace is drawn and I be the even when transferred ace is drawn. $P(A) = P(A \bigcap I) + P(A \bigcap I^c) = P(A|I)P(I) + P(A|I^c)P(I^c)$

2/6

Definition 4. *X is a discrete random variable with possible values, x_i. Then, the expected values (average, mean) of X is $EX = \sum_i x_i P(x = x_i) = \sum_i x_i p(x_i)$*
For any function, g, $Eg(X) = \sum_i g(x_i)P(X = x_i)$

Example: X, a random variable with $P(X = 1) = .2, P(X = 2) = .3, and P(X = 3) = .5$
What should be the average value of X?
In a population of n independent realization of X, about $.2n$ will have $X = 1$, $.3n$ will have $X = 2$, $.5n$ will have $X = 3$.

Average: $\frac{1 \times .2n + 2 \times .3n + 3 \times .5n}{n} = 2.3$
The expected value is 2.3.

X is a discrete random variable.
$\mu = EX$ where μ is the averag where μ is the average.

How about deviations? One way to find it, $E|X - \mu|$. Problem is that absolute values are annoying.
Therefore, we devise, $E(x - \mu)^2$. This is known as the Variance of X.
Let the standard devision, $SD(X) = \sigma(X) = \sqrt{Var(X)} = \sqrt{E(x - \mu)^2}$

$$\begin{aligned} Var(X) &= E[(X - \mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - \mu^2 = E(X^2) - (EX)^2 \end{aligned}$$

Example: Previous example, continued
$E(X^2) = 1^2 \times .2 + 2^2 \times .3 + 3^2 \times .5 = 6.7$
$Var(X) = 6.7 - (2.3)^2 = 1.49$
$\sigma(X) = \sqrt{Var(X)} = 1.19$

Example: X = number shown on a die
$E(X) = \frac{1}{6}(1 + 2 + \ldots + 6 = \frac{7}{2} = 3.5$

$EX^2 = \frac{1}{6}(1 + 2^2 + \ldots + 6^2) = \frac{91}{6} = 15.16$

Probability Mass functions

1. Uniform
   (a) $P(X = X_1) = \frac{1}{n}$, where $i = 1, \ldots, n$.
   (b) $EX = \frac{x_1 + \ldots x_n}{n}$
2. Bernoulli: An indicator random variable. Assume A is an event with probability p.
   (a) Indicator of A: $I_A - X_A = 1_A = \begin{cases} 1 & \text{if satisfied} \\ 0 & \text{otherwise} \end{cases}$.
   (b) $E(I_A) = p$
3. Binomial random variable: Number of successes in n independent trials, each of which is a success with probability p.
   (a) X is $Binomial(n, p)$, so $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ where $i = 0, \ldots, n$.
   (b) $EX = np$
   (c) $Var(X) = np(1-p)$
4. Poisson random variable. $X = Poisson(\lambda)$
   (a) $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ where $k = 0, 1, 2, \ldots$

Example: Let X = number of heads in 50 tosses of a fair coin.
X is $Binomial(50, \frac{1}{2})$
$P(X \leq 10) = \sum_{i=0}^{10} \binom{10}{1} \frac{1}{2^{50}}$

Example: Let d = dominant genes and r = recessive gene.
dd is called the pure dominant gene, dr is called hybrid, and rr puure recessive recessive". dd and dr produces the quality of dominant genes. rr procues the quality of recessive genes.

Assume that both parents are hybrid and have n children, what is P(at least two will be rr)?
Each child, independent, gets one of the genes at random from each parent.
$P(\text{pure recessive child}) = \frac{1}{4}$
X is the number of rr children.
$X = Binomial(n, \frac{1}{4})$
$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - (\frac{3}{4})^n - n\frac{1}{4}(\frac{3}{4})^{n-1}$

2/9

Poisson random variable

$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$ where $i = 0, 1, 2, \ldots$
$EX = \lambda$, $Var(X) = \lambda$

If X is $Binomial(n,p)$, n large, $p = \frac{\lambda}{n}$ is small, then $P(X = i) \to_{n\to\infty} e^{\lambda}\frac{\lambda^i}{i!}$ for each fixed i.
$EX = \sum_{i=1}^{\infty} ie^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\lambda\sum_{i=1}^{\infty}\frac{\lambda^{i-1}}{(i-1)!} = e^{-\lambda}\lambda e^{\lambda} = \lambda$
$pn = \lambda$, a constant.

Proof.

$$\begin{aligned} P(X = i) &= \binom{n}{i}\left(\frac{\lambda}{n}\right)^i\left(1-\frac{\lambda}{n}\right)^{n-i} \\ &= \frac{n(n-1)\dots(n-i+1)}{i!}\frac{\lambda^i}{n^i}\left(1-\frac{\lambda}{n}\right)^n \\ &\to_{\infty} \frac{\lambda^i}{i!}e^{-\lambda} \end{aligned}$$

So, basically, as $i \to \infty$, binomial becomes poisson. $\square$

Example: Suppose that probability that a person to be killed by lighting in a year is 1/500 million. Assume that the US population is 300 million. (Again, assume indepedence)

1. P(3 or more people will be killed by lightning in 2001)
   (a) X is the number of people killed by lightning.
   (b) $X = Binomial(n,p)$ where $n = 300$ million and $p = 1/$ 500 million
   (c) answer is $1-(1-p)^n - np(1-p)^{n-1} - \binom{n}{2}p^2(1-p)^{n-2}$
2. Approximate probability
   (a) $np = \frac{3}{5}$, so the approximately $Poisson(\frac{3}{5})$
   (b) Answer is $1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2}{2}e^{-\lambda} \approx .0231$
3. Interpretation of λ as "rate". P(two or more people are killed within the first 6 months of 2005)
   (a) p changes to $\frac{1}{100 \text{ million}}$ and $\lambda = \frac{3}{10}$.
   (b) Again, answer is $1 - e^{-\lambda} - \lambda e^{-\lambda}$
4. P(in exactly 3 of the 10 years exactly 3 people are killed)
   (a) Use Poisson approximation
   (b) Number of years with exactly people killed is $Binomial(10, \frac{\lambda^3}{3!}e^{-\lambda})$. Again, $\lambda = \frac{3}{5}$.
   (c) The answer is then $\binom{10}{3}(\frac{\lambda^3}{3!}e^{-\lambda})^3(1-\frac{\lambda^3}{3!}e^{-\lambda})^7$

Example: Expeted number of years among the next 10, in which 2 or more people get killed.
$10(1 - e^{-\lambda} - \lambda e^{-\lambda})$

Example: A crime is committed. The criminal is known to have a characterstics which occur with a small probability, p (very small). Assume we know nothing. A random person, among n, is arrested and charged with a crime. Defense?

2/10 (Discussion)

Random Variables

$X : S \to \mathbb{R}$
events $\mapsto$ subsets of $\mathbb{R}$

Homework 4.5

Let X be the difference between heads and tails when a coin is tossed n times. What are the possible values of X?

Answer

We'll define X as the number of heads - number of tails.
If $n = 1$, $X_i = \begin{cases} 1 & \text{if you get heads} \\ -1 & \text{if you get tails} \end{cases}$
$X = X_1 + \ldots + X_n$
The possible outcomes of X above is $-n, -n+2, -n+4, \ldots, n-2, n$
$P(X = -n) = P(\text{all tails}) = (\frac{1}{2})^n$
$P(X = -n + l)$? where $l = 0, 1, \ldots, 2n$
$P(X = -n + l) = P(l \text{ heads}) = \frac{\binom{n}{l}}{2^n}$

Homework 4.15

Let's say we have 66 balls. Pick the 11 worst teams
Team 1 gets 1 ball, Team 2 gets 2 balls, ..., Team 11 gets 11 balls.
Pick 3 distinct balls from the urn.
Let X denote the draft pick of Team 11.
Find the probability mass function of X.
$P(X = 1)$? where $X = \{1, 2, 3, 4\}$
$P(X = 1) = \frac{11}{66}$
$P(X = 2) = P(\text{ Team i} \neq 11 \text{ up to } i \text{ times, then Team 11 choose}) = \sum_{i=1}^{10} \sum_{l=0}^{i-1} \pi_{k=1}^{l} \frac{i-k}{66-k} \times \frac{11}{66-l-1}$
Another way, $P(X = 2) = \sum_{i=1}^{10} \frac{1}{66} \frac{11}{66-i}$
$P(X = 3) = \sum_i \sum_j \frac{i}{66} \frac{j}{66-j} \frac{11}{66-i-j}$
$P(X = 4) = 1 - P(X = 1) - P(X = 2) - P(X = 3)$

2/11

Example: **Poisson distribution and law.**

Assume a crime has been committed. It is known that the perpetrator has certain characteristics, which occur with small frequency p (say, 10^{-8}) in the population of size n (say, 10^8). A person who matches these characteristics has been found at random (e.g., at a routine traffic stop, or by airport security) and, since p is so small, charged with the crime. There is no other evidence. What should the defense be?

Let's start with a mathematical model of this situation. Assume that N is the number of people with given characteristics. This is a Binomial random variable, but given the assumption we can easily assume it's Poisson with $\lambda = np$. Choose a random person from among these N, call that person C, the criminal. Then, independently, choose at random another person, A, who is arrested. The question is whether $C = A$, that is, whether the arrested person is guilty. Mathematically, we can formulate the problem like this: We need to condition as the experiment cannot even be performed when $N = 0$. Now by the first Bayes formula,

$$\begin{aligned} P(C = A, N \geq 1) &= \sum_{k=1}^{\infty} P(C = A, N \geq 1 \,|\, N = k) \cdot P(N = k) \\ &= \sum_{k=1}^{\infty} P(C = A \,|\, N = k) \cdot P(N = k) \end{aligned}$$

and

$$P(C = A \,|\, N = k) = \frac{1}{k},$$

so

$$P(C = A, N \geq 1) = \sum_{k=1}^{\infty} \frac{1}{k} \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda}.$$

The answer to the question then is

$$P(C = A \,|\, N \geq 1) = \frac{e^{-\lambda}}{1 - e^{-\lambda}} \cdots \sum_{k=1}^{\infty} \frac{\lambda^k}{k \cdot k!}.$$

There is no closed-form expression for the sum, but it can be easily computed numerically. The defense may claim that the probability of innocence, $1-$(the above probability), is about 0.2330 when $\lambda = 1$, quite enough for a reasonable doubt.

This model was in fact tested in court, in the famous *People v. Collins* case, a 1968 jury trial in Los Angeles. In this instance, it was claimed by prosecution (on flimsy grounds) that $p = 1/12,000,000$ and n would be the adult couples in the LA area, say $n = 5,000,000$. The jury convicted the charged couple in a robbery on the basis of the prosecutor's claim that, due to low p, "the chances of there being another couple [with specified characteristics, in the LA area] must be one in a billion." The Supreme Court of California reversed the conviction, and gave two reasons. The first reason was insufficient foundation for estimate for p. The second reason was that

$$P(N \geq 2 \,|\, N \geq 1) = \frac{1 - e^{-\lambda} - \lambda e^{-\lambda}}{1 - e^{-\lambda}}$$

from the prosecutor's claim is in fact much larger than he claimed, namely about 0.1939. This is about twice as big as the probability of innocence, which by the above formula would be about 0.1015.

Geometric random variables

$X =$ number trials required for the first success in independent trials with success probability p.
$P(X = n) = p(1-p)^{n-1}$ with $n = 1, 2, \ldots$
$EX = \frac{1}{p}$

$Var(X) = \frac{1-p}{p^2}$
$P(X \geq n) = \sum_n^\infty p(1-p)^{n-1} = (1-p)^{n-1}$
$P(X \geq n+k | X \geq k) = \frac{(1-p)^{n+k-1}}{(1-p)^k}$

Example: X = number of tosses required for 1st heads (in a fair coin).
$EX = 2$.

Example: you roll a die, your opponent tosses a coin.
Roll 6 $\rightarrow$ you win.
Now roll 6, opponent tosses Heads $\rightarrow$ you lose
Otherwise, the game repeats.
On the average, how many steps does the game last?
$P(\text{game decided on step 1}) = \frac{1}{6} + \frac{5}{6}\frac{1}{2}$
$EN = \frac{1}{\frac{1}{6}+\frac{5}{6}\frac{1}{2}} = \frac{12}{7}$

Continuous random variables

Let X be a continuous random variable.
If $P(X \in B) = \int f(x)\,dx$ where f is a nonnegative function which $\int_{-\infty}^\infty f(x)\,dx = 1$
The function $f = f_X$ is called the density of X.
$P(X \in [a,b]) = P(a \leq X \leq b)$
$P(X = a) = 0$
$P(X \leq a) = \int_{-\infty}^a f(x)\,dx = P(X < a)$.

distribution functions

$F(x) = P(X \leq x) = \int_{-\infty}^x f(s)\,ds$ where F is the distribution function.
$\frac{dF}{dx} = f$
$EX = \int_{-\infty}^\infty f(x)\,dx$

Example: $f(x) = \begin{cases} cx & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases}$

1. Determine c
2. $P(1 \leq X \leq 2)$
3. EX, EX^2

Determine c.
$\int_0^4 cx\,dx = 1$, so $c = \frac{1}{8}$
$\int_1^2 \frac{x}{8}\,dx = \frac{3}{16}$
$EX = \int_0^4 \frac{x^2}{8}\,dx = \frac{1}{3}$

2/13

Example: X has density $f_X(x) = \begin{cases} 3x^2 & x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$

Compute density of $Y = 1 - x^4$.
To start off problems like this, compute the distribution function of Y.
$F_Y(y) = P(Y \leq y) = P(1 - x^4 \leq y) = P(1 - y \leq x^4) = *$
$y \in [0,1]$ because Y only has values on $[0,1]$.
$* = P(1-y)^{\frac{1}{4}} \leq X) = \int_{1-y)^{\frac{1}{4}}}^{1} 3x^2\,dx$
So, $f_Y(y) = \frac{d}{dy} F_Y(y) = -3((1-y)^{\frac{1}{4}})^2 \frac{1}{4}(1-y)^{\frac{-3}{4}}(-1) = \frac{3}{4}\frac{1}{(1-y)^{\frac{1}{4}}}$ where $y \in (0,1)$.

Uniform random variable

Choosing a random number on $[\alpha, \beta]$" $f(x) = \begin{cases} \frac{1}{\beta - \alpha} & x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$

This is ideally the output of a call of a random number generator.
If X is uniform on $[\alpha, \beta]$
$EX = \frac{\alpha + \beta}{2}$
$Var(X) = \frac{(\beta - \alpha)^2}{12}$
X is uniform on $[0,1]$.
$P(X \in \mathbb{Q}) = 0$
Proof: $P(X \in \mathbb{Q}) = P(\bigcup_i \{X = q_i\}) = \sum_i P(x = q_i) = 0$. This is only true for countable sets.

Binary expansion of uniform random variable on $[0,1]$. For example, 0.010 refers to a number between $\frac{1}{4}$ to $\frac{3}{8}$ since it is left of $\frac{1}{2}$, right of $\frac{1}{4}$ and left of $\frac{3}{8}$.
Any first three digits of a primary expansion of X.
The binary digits of X use the result of an infinite sequence of independent fair coin tosses.
$P(\text{periodic sequence}) = 0$ because eventually, the periodic sequence will break.
Because of this, the decimal does not repeat, meaning the $P(x \in \mathbb{Q}) = 0$.
Example: A uniform random number X divides $[0,1]$ into two subintervals. Let R be the ratio of smaller versus larger segment.
Compute the density of R.
The possible value of r is $(0,1)$ since they can be broken up equally to one person takes the entire stick.

$$\begin{aligned}
F_R(r) &= P(R \leq r) = P(X \leq \frac{1}{2}, \frac{x}{1-x} \leq r) + P(X > \frac{1}{2}, \frac{1-x}{x} \leq r) \\
&= P(X \leq \frac{1}{2}, X \leq \frac{r}{r-1}) + P(X > \frac{1}{2}, X \geq \frac{1}{r+1}) \\
&= P(X \leq \frac{r}{r+1}) + P(X \geq \frac{1}{r+1})[\text{We know that } \frac{r}{r+1} \leq \frac{1}{2} \text{ and } \frac{1}{r+1} \geq \frac{1}{2}] \\
&= \frac{r}{r+1} + 1 - \frac{1}{r+1} = \frac{2r}{r+1}
\end{aligned}$$

Exponential random variable

We have a pararmeter, $\lambda > 0$
$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$
$P(X \geq x) = \int_x^\infty \lambda e^{-\lambda s}\, dx = e^{-\lambda x}$
$P(X \geq x + y | X \geq y) = \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} = -e^{-\lambda x}$.("Memoryless property": How long you have waited does not matter! The probability is still the same no matter how long you waited.)

Example: Assume that a lightbulbs last on average 100 hours.
Assuming exponential distribution and $\lambda 1100$.
$P(\text{lasts } \geq 200 \text{ hours}) = e^{-2} \approx .135$
$P(\text{lasts } \leq 50 \text{ hours}) = 1 - e^{-\frac{1}{2}} \approx 0.3935$

2/17 (Discussion)

Quiz answer

Part 1

The rate of suicide is 1 in 10^5 per month.
A city has 400,000 residents.
$\lambda = np = 4$
$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$

$$\begin{aligned} P(\text{ 8 or more suicides / month }) &= 1 - \sum_{i=0}^{7} P(X = i) \\ &= 1 - \sum_{i=0}^{7} e^{-\lambda} \frac{\lambda^i}{i!} \end{aligned}$$

Part 2

Probability of at least 2 month in year will have 8 or more suicide.
Get p from part (a).
This is $Binomial(12, p)$

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - (1 - p)^{12} - 12p(1 - p)^{11} \end{aligned}$$

Continuous Random Variable

$X : S \to \mathbb{R}$
$P(X \in B) = \int_B f(x)\, dx$
This allows you to forget about sample space.

Homework 5.5

You have a gas station. Its weekly volume of sales in thousands of gallons is a random variable given with pdf:

$f(x) = \begin{cases} 5(1-x)^4 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

What is the volume of the tank such that the probability of running out is 0.01?

$P(X \geq v) = .01$

In this case, $B = [v, 1]$. Question is what is v?

$$\begin{aligned} P(X \geq v) &= \int_v^1 f(x)\,dx \\ &= \int_v^1 5(1-x)^4\,dx \\ &= -(1-x)^5|_v^1 \\ v &= 1 - \left(\frac{1}{100}\right)^{\frac{1}{5}} \end{aligned}$$

2/18

X is Normal, $N(\mu, \sigma^2)$ means

$$f(x) = f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where $x \in (-\infty, \infty)$.

Such a random variable occurs frequently, i.e. a measurement with a random mistake or height of a random heron in a population.

Some properties:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= 1 \\ EX &= \mu \end{aligned}$$

Proof of EX:

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} (x-\mu+\mu) \frac{1}{\sigma - \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx \\ \int_{-\infty}^{\infty} t \frac{1}{\sigma - \sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}\,dx &= 0 \\ EX &= 0 \end{aligned}$$

<u>Example</u>: Let $Y = \alpha X + \beta$.

How is Y distributed? $(\alpha > 0)$

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) = P(\alpha X + \beta \leq y) \\
&= P(X \leq \frac{y-\beta}{\alpha}) \\
&= \int_{-\infty}^{\frac{y-\beta}{\alpha}} f(x)\,dx
\end{aligned}$$

$$\begin{aligned}
f_Y(y) &= f(\frac{y-\beta}{\alpha})\frac{1}{\alpha} \\
&= \frac{1}{\sqrt{2\pi}\sigma\alpha} e^{-\frac{(y-\beta-\alpha\mu)^2}{2\alpha^2\sigma^2}}
\end{aligned}$$

Y is normal with $EY = \alpha\mu + \beta$, $Var(Y) = (\alpha\sigma)^2$
In particular, $Z = \frac{x-\mu}{\sigma}$ has $EZ = 0$ and $Var(Z) = 1$.
Z is called the standard normal.

$$\begin{aligned}
f_Z(z) &= \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \\
F_Z(z) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\,dx = \Phi(Z)
\end{aligned}$$

You have to look at the table in the book for the values of $\Phi(Z)$

$z = \frac{X-\mu}{\sigma}$ Example: Let $z > 0$. What is $\Phi(Z)$?
$\Phi(-Z) = 1 - \Phi(Z)$

Example: If X is $N(\mu, s\sigma^2)$, then $P(|X-\mu| \geq \sigma)$?
$P(|X-\mu| \geq \sigma) = P(|\frac{x-\mu}{\sigma} \geq 1) = P(|Z| \geq 1) = 2P(z \geq 1) = 2(1 - \Phi(1))$

$P(|X-\mu| \geq 2\sigma) \approx 7\%$
$P(|X-\mu| \geq 3\sigma) \approx .26\%$

Example: X with $N(\mu = 2, \sigma^2 = 5^2)$.

$$\begin{aligned}
P(1 \geq X \geq 4) &= P(1 \geq X \geq 4) \\
&= P(\frac{1-2}{5} \leq \frac{X-2}{5} \leq \frac{4-2}{5}) \\
&= P(-.2 \leq z \leq .4) \\
&= P(z \leq .4) - P(z \leq -.2) \\
&= \Phi(.4) - (1 - \Phi(.2)) \\
&= .2347
\end{aligned}$$

De Moirre-Laplace Central Limit Theorem

Let S_n be a $Binomail(n, p)$ random variable.
Let p be fixed. (e.g. 0.5) and n be large (i.e. $n = 100$)
$\frac{S_n - np}{\sqrt{np(1-p)}}$ is distributed approximiately as $N(0, 1)$.

2/20

De Moire-Laplace Central Limit Theorem

Let S_n br $Binomial(n,p)$ where p is fixed and n is large.
Then, $\frac{S_n - np}{\sqrt{np(1-p)}} \approx N(0,1)$
The condition for using this is $np(1-p) \geq 10$.
This is analytical statement: $\sum_k \binom{n}{k} p^k (1-p)^{n-k} \frac{k-np}{\sqrt{np(1-p)}} \leq x$.

As $x \to \infty$, $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{s^2}{2}}\, ds$ for every $x \in \mathbb{R}$.
In connection with normal curves: the generalized function is equivalent is $P(\frac{S_n - np}{\sqrt{np(1-p)}} \leq x)$ and as $x \to \infty$, it's equivalent to $P(z \leq x)$.
Note: Interesting enough, The maximum of a Binomial probability mass function is np.

Note: When, say $n = 100$ and $p = \frac{1}{100}$ is a better approximation for S_n is Poisson with $\lambda = 1$.

Example: Roulette wheel has 38 slots, 18 red, 18 black, and 2 green. A ball ends at one of these at random.
Let's say you're a player who bet \$1 on every game on red.
If you win, you get a \$1. If you lose, you lose a \$1. After n games, what is the probability that you are ahead for $n = 100$ and $n = 1000$?
Let S_n be the number of times you win is a $Binomial(n, \frac{18}{38} = \frac{9}{19})$.

$$
\begin{aligned}
P(\text{ahead}) &= P(\text{win more than half of the games}) \\
&= P\left(S_n > \frac{n}{2}\right) \\
&= P\left(\frac{s\_np}{\sqrt{np(1-p)}} > \frac{\frac{1}{2}n - np}{\sqrt{np(1-p)}}\right) \\
&\approx P\left(Z > \frac{(\frac{1}{2} - p)\sqrt{n}}{\sqrt{p(1-p)}}\right)
\end{aligned}
$$

For $n = 100$, $P(Z > \frac{t}{\sqrt{90}}) \approx 30\%$
For $n = 1000$, $P(Z > \frac{5}{3}) \approx 4.8\%$
What if $p = \frac{1}{2}$?
$P(\text{ahead}) \to_{n \to \infty} P(z > 0) = .5$

Example: How many times do you need to toss a fair coin to get 100 heads with probability 90%?
Let n be the number of tosses that we're looking for.
Let S_n be $Binomial(n, \frac{1}{2})$.

$$\begin{aligned}
P(S_n \geq 100) &= .9 \\
P\left(\frac{S_n - \frac{1}{2}n}{\sqrt{n\frac{1}{4}}} \geq \frac{100 - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}\right) &= \\
P\left(z \geq \frac{100 - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}\right) &\approx \\
P\left(z \geq \frac{200 - n}{\sqrt{n}}\right) &\approx \\
P\left(z \geq -\left(\frac{n - 200}{\sqrt{n}}\right)\right) &\approx \\
P\left(z \geq -\left(\frac{n - 200}{\sqrt{n}}\right)\right) &\approx \\
\Phi(\frac{n - 200}{\sqrt{n}}) &\approx .9
\end{aligned}$$

According to the tables, $\Phi(1.28) \approx .9$
Solve $\frac{n-200}{\sqrt{n}} = 1.28$.
We know that $n - 1.28\sqrt{n} - 200 = 0$
$\sqrt{n} = \frac{1.28+\sqrt{1.28^2+800}}{2}$
So, $n = 219$.

6 Joint distributions

Discrete Cases

Two discrete random variables, X and Y.
Their joint probability mass function.
$P(x, y) = P(X = x, Y = y)$
$P((X, Y) \in A) = \sum_{(X,Y)\in A} P(x, y)$

The Marginal probability mass function:
$P(X = x) = \sum_y P(X = x, Y = y) = \sum_y p(x, y)$
$P(Y = y) = \sum_x P(X = x, Y = y) = \sum_x p(x, y)$

Example: An urn contains 2 red, 5 white, and 3 green balls.
Select 3 balls at random.
X = number of red balls selected.
Y = number of white balls selected.

1. Compute joint probability mass function of (X, Y)
2. Marginal probability mass functions
3. $P(X \geq Y)$

2/23

Example: An urn has 2 red, 5 white, and 3 green balls.
Select 3 at random and let X be the number of red balls and Y be the number of white balls.

1. joint pmf of (X, Y)
2. marginal pmfs
3. $P(X \geq Y)$
4. $P(X = 2|X \geq Y)$

Joint pmf is $P(X = x, Y = y)$ for all possible x and y.
In our case, X can be $0, 1, 2$ and Y can be $0, 1, 2, 3$

Y/X	0	1	2	Y
0	$1/120$	$2 \cdot 3/120$	$3/120$	$10/120$
1	$5 \cdot 3/120$	$2 \cdot 5 \cdot 3/120$	$5/120$	$50/120$
2	$10 \cdot 3/120$	$10 \cdot 2/120$	0	$50/120$
3	$10/120$	0	0	$10/120$
X	$56/120$	$56/120$	$8/120$	1

Part c)

$$P(X \geq Y) = \frac{1 + 6 + 3 + 30 + 5}{120} = \frac{3}{8}$$

Part d)

$$\frac{P(X = 2, X \geq Y)}{P(X \geq Y)} = \frac{\frac{8}{120}}{\frac{3}{8}} = \frac{8}{45}$$

Independence

Two random variables, are *independent* if

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

for all sets A and B of real numbers.
In discrete case, this is eqiulvane to

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

(Joint pmf is the product of marginals)

Example: In the previous problem, is X and Y independent?
No. Those 0s are dead giveaways. The marginal probabilities of $X = 2$ and $Y = 2$ are not 0, so if $X =$ and Y are independent, then $P(X = 2, Y = 2)$ is not 0.

Example: Most often, independence is an assumption!
Roll a die twice. Let X be the number from the 1st roll and Y be the second roll. You cannot really explain if something is independent. It's always an assumption just like this one.

Continuous Case

Let (X, Y) be jointly continuous random variables if there exists $f(x, y) \geq 0$, so that

$$P((X,Y) \in C) = \int\int_C f(x,y)\,dxdy$$

where C is some subset of $\mathbb{R}^2$

Example: Let (X, Y) is a random point on S where S is a region in $\mathbb{R}^2$.
Then,

$$f(x,y) = \begin{cases} \frac{1}{area(S)} & \text{if } (x,y) \in S \\ 0 & \text{otherwise} \end{cases}$$

The simplest example is a square of length of 1

$$f(x,y) = \begin{cases} 1 & \text{if } 0 \leq x \leq x \text{ and } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Example: Let

$$f(x,y) = \begin{cases} cx^2y & \text{if } x^2 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

1. What is c?
2. $P(X \geq Y)$
3. $P(X = Y)$
4. $P(X = 2Y)$

Part a)

$$\begin{aligned} \int_{-1}^{1} dx \int_{x^2}^{1} cx^2y\,dy &= 1 \\ c\frac{4}{21} &= 1 \\ c &= \frac{21}{4} \end{aligned}$$

Part b) Let S be the area between $x = y$ and $x^2 = y$ for $x \in (0, 1)$.

$$\begin{aligned} P(X \geq Y) &= P((X,Y) \in S) \\ &= \int_0^1 (\int_{x^2}^{X} \frac{21}{4}x^2y\,dy)\,dx \\ &= \frac{3}{20} \end{aligned}$$

Part c and d) 0 because it's a line.

2/24 (Discussion)

answer to Quiz 1

$$f(x) = \begin{cases} a + bx^2 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

By definition,

$$\begin{aligned} \int_0^1 ax + bx^3 \, dx &= 1 \\ E(X) &= \int_0^1 x f(x) \, dx \\ &= \int_0^1 ax + bx^3 \, dx \\ &= \frac{3}{5} \end{aligned}$$

answer to Quiz 2

Find pdf of Y.

$$P(X \le x) = \int_0^x e^{-x} \, dx = 1 - e^{-x}$$

Distribution function of Y

$$\begin{aligned} P(Y \le y) &= P(logX \le y) \\ &= P(X \le e^y) \\ &= \int_0^{e^y} e^{-x} \, dx \\ &= \int_0^y f_y(x) \, dx \end{aligned}$$

By the fundamental theorem of calculus,

$$f_Y(x) = \frac{d}{dy} \int_0^{e^y} e^{-x} \, dx = \frac{d}{dy}(1 - e^{(e^y)^{-1}}) = -e^{-e^y} - (-e^y)$$

Homework 5.18

Let X be a normal random variable with $\mu = 5$.
If $P(X > 9) = .2$, approximately what is $Var(X)$?
For normal random variable,

$$P(X \le a) = P(\frac{X - \mu}{\sigma}) \le \frac{a - \mu}{\sigma}$$

$$P(X > 9) = 1 - P(X \le 9) = 1 - P(\frac{x - 5}{\sigma} \le \frac{4}{\sigma}) = 1 - \Phi(\frac{4}{\sigma})$$

Now, we need find $\Phi(\frac{4}{\sigma}) = .8$, so $\frac{4}{\sigma} \approx .84$. Therefore,

$$Var(X) = \sigma^2 = \left(\frac{4}{.84}\right)^2$$

Laplace-De-Moire's Central Limit Theorem

$$P(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b) \to \Phi(b) - \Phi(a)$$

S_n is the sample result of Bernoulli trials of size n. $\frac{S_n - np}{\sqrt{np(1-p)}}$ is rescaling it to normal just like usual normals.

2/25

From last time,
Example:

$$f(x, y) = \begin{cases} \frac{21}{4}x^2y & x^2 \leq y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute marginal densities. Independence?

$$f_X(x) = \int_{x^2}^{1} \frac{21}{4}x^2y\,dy = \frac{21}{8}x^2(1 - x^4)$$

where $x \in (-1, 1]$.
Note: The range is from x^2 to 1 because we are measuring the area from $y = x^2$ to $y = 1$.

$$f_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{21}{4}x^2y\,dx = \frac{7}{2}y^{\frac{5}{2}}$$

where $y \in [0, 1]$
Note: The range is from $-\sqrt{y}$ to $\sqrt{y}$ because we are measuring the area from the curve $x = -\sqrt{y}$ and $x = \sqrt{y}$.
They are not independent because if you multiply the two marginal densities, you don't get the $f(x, y)$.

Example: (X, Y) are random points in the square of length of 1 with bottom left corner at the origin. Independence?

$$f(x, y) = \begin{cases} 1 & (x, y) \in [0, 1] \times [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

Marginal densities:

$$f_X(x) = 1$$

if $x \in [0, 1]$.

$$f_Y(y) = 1$$

if $y \in [0,1]$.
This is independent.

Let $f(x,y)$ be distributed under the curve $y = x$.

$$f(x,y) = \begin{cases} 2 & 0 \le y \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Marginal densitites:

$$f_X(x) = 2x$$

if $x \in [0,1]$

$$f_Y(y) = 2(1-y)$$

if $y \in [0,1]$
They are not distributed uniformly and they're no longer independent in this case.

Note: The only way for two variables to be indepdent in a continuous setting is if it distributed uniformly to a rectangle.

Example: Mr. and Mrs. Smith agree to meet at a place "between 5 and 6 p.m". Assume that they both go there at a random time and are indepdent.

1. Find the density for the time one of them will have to wait for the other.
2. Mrs. Smith later tells you she had to wait. Probability that Mr. Smith arrived before 5:30 given that.

Let X be the time when Mr. Smith arrives and Y be the time when Mrs. Smith arrives.
We will assume that (X,Y) is uniform on $[0,1] \times [0,1]$.
Our unit is 1 hr
(Part 1): Let $T = |X - Y|$.
Let $t \in [0,1]$ since t is the difference between x and y.

$$\begin{aligned} P(T \le t) &= P(|X-Y| \le t) \\ &= P(-t \le X - Y \le t) \\ &= P(X - t \le Y \le X + t) \\ &= 1 - (1-t)^2 \text{ Take the area of the two triangles and subtract their area off the unit square} \\ &= 2t - t^2 \end{aligned}$$

$$f_T(t) = 2 - 2t$$

(Part 2):

$$\begin{aligned} P(X \le .5 | X > Y) &= \frac{P(X \le .5, X > Y)}{P(X > Y)} \\ &= \frac{\frac{1}{8}}{\frac{1}{2}} \\ &= \frac{1}{4} \end{aligned}$$

Example: Assume X and Y are independent.
X is uniform on $[0,1]$ and Y has density, $f_Y(y) = 2y$, if $y \in [0,1]$.
Compute $P(X+Y \leq 1)$

$$f(x,y) = \begin{cases} 2y & (x,y) \in [0,1] \times [0,1] \\ 0 & \text{otherwise} \end{cases}$$

We know that it is $2y$ because X and Y are independent and $P(X) = 1$ and $P(Y) = 2y$ where $(X,Y) \in [0,1] \times [0,1]$.

$$\int_0^1 dx \int_0^{1-x} 2y\, dy$$

or

$$\int_0^1 dy \int_0^{1-y} 2y\, dx$$

2/27

Example: Two independent exponential random variables, T_1 with expectation of 10 minutes and T_2 with expectation 40 minutes. What is $P(T_1 < T_2)$?
Our unit is 10 minutes.
We know that

$$f(t) = \lambda e^{-\lambda t}$$

where $t \geq 0$ and expectation is $\frac{1}{\lambda}$!

$$f_{T_1}(t_1) = e^{-t_1}$$

with $t_1 \geq 0$

$$f_{T_2}(t_2) = \frac{1}{4} e^{-t_2/4}$$

with $t_2 \geq 0$

$$\begin{aligned} P(T_1 < T_2) &= \int_0^\infty dt_1 \int_{t_1}^\infty e^{-t_1} e^{-t_2/4}\, dt_2 \\ &= \int_0^\infty \int_0^\infty e^{-t_1}\, dt_1 e^{-t_1/4} \\ &= \int_0^\infty e^{\frac{-5t_1}{4}}\, dt_1 \\ &= \frac{4}{5} \end{aligned}$$

Example: **Buffon needle problem**

2-27_needle.png

Horizontal lines at distance 1. Drop a needle of length l. Probability that it intersects one of the lines.

Let D be the distance from the center of hte needle to the nearest line, and θ be the acute angle relative to the lines.

We will assume that D and θ are independent and uniform: $0 \leq D \leq \frac{1}{2}$ and $0 \leq \theta \leq \frac{\pi}{2}$.
The hypotenuse is $\frac{D}{\sin(\theta)}$

$$\begin{aligned} P(\text{intersects}) &= P\left(\frac{D}{\sin(\theta)} < \frac{l}{2}\right) \\ &= P(D < \frac{l}{2}\sin(\theta)) \end{aligned}$$

<u>Case 1</u>: Let $l \leq 1$.

$$\begin{aligned} P(\text{intersects}) &= \frac{\int_0^{\pi/2} \frac{l}{2}\sin(\theta)d\theta}{\pi/4} \\ &= \frac{l/2}{\pi/4} \\ &= \frac{2l}{\pi} \end{aligned}$$

When $l = 1$, you get $\frac{2}{\pi}$

<u>Case 2</u>: The curve, $D = \frac{l}{2}\sin(\theta)$ intersects $D = \frac{1}{2}$ at $\arctan(\frac{1}{l})$.

$$\begin{aligned} P(\text{intersects}) &= \frac{4}{\pi}\left[\frac{l}{2}\int_0^{\arcsin(\frac{1}{l})} \sin(\theta)\,d\theta + \left(\frac{\pi}{2} - \arcsin\frac{1}{l}\right)\frac{1}{2}\right] \\ &= \frac{4}{\pi}\left(\frac{l}{2} - \frac{1}{2}\sqrt{l^2-1} + \frac{\pi}{4} - \frac{1}{2}\arctan\frac{1}{l}\right) \end{aligned}$$

<u>Example</u>: X_1, X_2, X_3 are uniform on $[0,1]$ and independent.
$P(X_1 + X_2 + X_3 \leq 1)$?

$$f_{(X_1,X_2,X_3}(x_1,x_2,x_3) = \begin{cases} 1 & \text{if } (x_1,x_2,x_3) \in [0,1]^3 \\ 0 & \text{otherwise} \end{cases}$$

And that is equal to $f_{x_1}(x_1)f_{x_2}(x_2)f_{x_3}(x_3)$

$$P(X_1 + X_2 + X_3 \leq 1) = \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \int_0^{1-x_1-x_2} dx_3 = \frac{1}{6}$$

In general, if $X_1, \ldots, X_n$ are independently uniform on $[0,1]$.

$$P(X_1 + \ldots + X_n \leq 1) = \frac{1}{n!}$$

3/2

Conditional distributions

Discrete cases:

$$p_{x|y}(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

Continuous cases:

$$f_{x|y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

The latter is a trivial formula since numerator and denominator is always 0 if the denominator is 0. It should be:

$$\begin{aligned} f_{x|y}(x|y) &= \frac{P(X = x + dx, Y = y + dy)}{P(Y = y + dy)} \\ &= \frac{f(x, y)\,dx\,dy}{f_Y(y)dy} \\ &= \frac{f(x, y)}{f_Y(y)}dx \\ &\approx P(X = x + dx|Y = y + dy) \end{aligned}$$

Example: Let (x, y) be a random point on the isosceles triangle of sides, 1.
$f(x, y) = 2$ on the triangle.
Compute $f_{x|y}(x|y)$.
Let's say I give you y, then you know that x is between 0 and $1 - y$.

$$\begin{aligned} f_Y(y) &= \int_0^{1-y} 2\,dx \\ &= 2(1 - y) \end{aligned}$$

where $y \in [0, 1]$

$$f_{x|y}(x|y) = \begin{cases} \frac{2}{2(1-y)} & 0 \le x \le 1 - y \\ 0 & \text{otherwise} \end{cases}$$

Example: Suppose

$$f(x, y) = \begin{cases} \frac{21}{4}x^2y & x^2 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute $f_{x|y}(x, y)$

$$f_Y(y) = \frac{21}{4}y\int_{-\sqrt{y}}^{\sqrt{y}} x^2\,dx = \frac{7}{2}y^{\frac{5}{2}}$$

where $y \in [0, 1]$

$$f_{x|y}(x|y) = \frac{\frac{21}{4}x^2y}{\frac{7}{2}y^{5/2}} = \frac{3}{2}x^2y^{-3/2}$$

where $-\sqrt{y} \leq x \leq \sqrt{y}$

How about $P(X \geq Y|Y = y)$? Normally, this makes no sense because we're given something that has probability, 0. This is interpreted as

$$\int_y^{\sqrt{y}} \frac{3}{2} x^2 y^{-3/2}\, dx = \int_y^{\sqrt{y}} f_{X|Y}(x|y)\, dx$$

$$EX = \begin{cases} \sum_i x_i p(x_i) & \text{where } X \text{ is a discerete with probability, } p. \\ \int_{-\infty}^{\infty} x f(x)\, dx & \text{where } X \text{ is continuous with density, } f \end{cases}$$

$$Eg(X) = \begin{cases} \sum_i g(x_i) p(x_i) & \text{where } X \text{ is a discerete with probability, } p. \\ \int_{-\infty}^{\infty} g(x) f(x)\, dx & \text{where } X \text{ is continuous with density, } f \end{cases}$$

If we have (X, Y), then

$$Eg(X) = \begin{cases} \sum_{x,y} g(x,y) P(X = x, Y = y) & \text{where } X \text{ is a discerete with probability, } p. \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy & \text{where } X \text{ is continuous with density, } f \end{cases}$$

Example: Roll two dice, with X being the number on the 1st die and Y be the number on the 2nd die. Let Z be the difference between X and Y.
Compute $P(Z = i)$ where $i = 0, 1, 2, 3, 4, 5$
Then,

$$E|X - y| = \sum_{i=1}^{6} \sum_{i=1}^{6} |i - j| \frac{1}{36}$$

Alternatively,

$$EZ = \sum_{k=0}^{5} k P(Z = k)$$

Example: Choose a random point, (x, y) on the right half of the unit circle. Compute EX.

3/3(Discussion)

Discrete Random Variables

1. Bernoulli Random Variables

$$P(X = 0) = 1 - p$$
$$P(X = 1) = p$$
$$E(X) = p$$
$$Var(X) = E(X^2) - (E(X))^2 = p(1 - p)$$

2. Binomial Random Variable

$$X = X_1 + \ldots + X_n$$

Basically, the sum of Bernoulli variables

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

$$E(X) = np$$

$$Var(X) = np(1-p)$$

3. Poisson random variables

$$P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

It's good approximation to Binomial when $\lambda = np$ with n being large and p is small.

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

4. Geometric Random Variable

$$P(X = n) = (1-p)^{n-1} p$$

where $n \in \{1, 2, 3, \ldots\}$.
Models waiting time till first sucess.

$$E(X) = \frac{1}{p}$$

$$Var(X) = \frac{1-p}{p^2}$$

Continuous Random variables

The whole idea is to calculate $P(X \leq B) = \int_B f_X(x)\, dx$ where $f_X(x)$ is the probability density function.

1. Uniform Random Variable

$$f(X) = \begin{cases} \frac{1}{\beta - \alpha} & \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \frac{\alpha + \beta}{2}$$

$$Var(X) = \frac{(\beta - \alpha)^2}{12}$$

2. Normal Random Variable

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/\sigma^2}$$

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

3. Exponential Random Variable

$$f(X) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \mu$$
$$Var(X) = \sigma^2$$

Suppos we have two independent random variabes, X, Y both of which are continuous. Then,

$$\int \int_{C \in \mathbb{R}^2} f_{x,y}\, dx\, dy = P((X, Y) \in C)$$

Since X, Y are independent, $f_{XY}(x, y) = f_X(x) f_Y(y)$.

3/6

Expectation

$$Eg(X, Y) = \int \int g(x, y) f(x, y)\, dxdy (\text{Continuous case})$$
$$Eg(X, Y) = \sum_{x,y} g(x, y) p(x, y) (\text{discrete case})$$

Example: Hw8.6

Joint mass function with (product of i, j

$i \backslash j$	1	2	3	4
1	.1 (1)	.1 (2)	.1 (3)	.1 (4)
2	.1 (2)	.1 (4)	.1 (6)	.1 (8)
3	.1 (3)	.1 (6)	0	0
4	.1 (4)	0	0	0

Example: (X, Y) is a random point in the right triangle with both sides of length 1. Compute EX, EY, and EXY.

$$\begin{aligned} EX &= \int_0^1 dx \int_0^{1-x} 2\, dy \\ &= \int_0^1 2x(1-x)\, dx \\ &= 2(\frac{1}{2} - \frac{1}{3}) \\ &= \frac{1}{3} \end{aligned}$$

$EY = EX = \frac{1}{3}$

$$\begin{aligned} E(XY) &= \int_0^1 dx \int_0^{1-x} xy2\,dy \\ &= \int_0^1 2x\frac{(1-x)^2}{2}\,dx \\ &= \int_0^1 (1-x)x^2\,dx \\ &= \frac{1}{3}-\frac{1}{4} \\ &= \frac{1}{12} \end{aligned}$$

Properties of expectation

1. $E(aX+b) = aE(X)+b$
2. $E(X_1+\ldots+X_n) = E(X_1)+\ldots+E(X_n)$

Proof. Proof of property of two
Take $n=2$, $E(X+Y) = EX+EY$
In the continuous case,

$$\int\int (x+y)f(x,y)\,dx\,dy = \int\int xf(x,y)\,dx\,dy + \int\int yf(x,y)\,dx\,dy$$

Use induction. □

Example: Assume that an urn contains 10 black, 7 red, and 15 white balls. Take 5 balls out.

1. with replacement
2. without replacement

and X is the number of red balls pulled out. Compute EX.

"Indicator trick": Let $I_i = I\{i\text{th ball is red}\} = \begin{cases} 1 & \text{if } i\text{th ball is red} \\ 0 & \text{otherwise} \end{cases}$
Let $X = I_1+I_2+I_3+I_4+I_5$.

In part 1, this is $Binomial(5, \frac{7}{22})$.
You know that $EX = 5\cdot\frac{7}{22}$, but let's pretend we don't know.

$$EI_1 = 1\cdot P(\text{1st ball is red}) = \frac{7}{22} = EI_2 = \ldots = EI_5$$

Therefore, $EX = 5\cdot\frac{7}{22}$.

In part 2, one way you can do it:

$$P(X=i) = \frac{\binom{7}{i}\binom{15}{5-i}}{\binom{22}{5}}$$

where $i = 0, 1, \ldots, 5$.
Then, we know that

$$EX = \sum_{i=0}^{5} i \frac{\binom{7}{i}\binom{15}{5-i}}{\binom{22}{5}}$$

Now, the indicator trick, it turns out to be equal to $5 \cdot \frac{7}{22}$ with the same calculations as the replacement version.
Example: **"Matching problem"**
n people buys n gifts, which are then assigned at random.
Let X be the number of people who recive their own gift.
What is EX?
Let $I_i I\{\text{Person } i \text{ receives own gift}\}$
Let $X = I_1 + I_2 + \ldots + I_n$
$EI_i = \frac{1}{n}$
And so, $EX = 1$.
Let $X^2 = I_1 + \ldots + I_n + sum_{i \neq j} I_i I_j$

Example: 5 married couples are seated around a round table at random. Let X be the number of wives who sit next to their husbands. What is EX?
Let $I_i = I\{\text{wife } i \text{ sits next to her husband}\}$.
Then, $X = I_1 + \ldots + I_5$
$EI_i = \frac{2}{9}$
$EX = \frac{10}{9}$

Example: **Coupon collector problem**
We have n cards. Sample them with replacement.
Let N be the number of cards you need to sample for complete collection.
What is EN?
Let N_1 be the number of cards to receive of 1st card.
Well, this is trivial. The first card you buy is always the 1st card.
Let N_2 be the number of cards to get the second new card.
Let N_3 be the number of cards to get the third new card.
$\vdots$
Let N_n be the number of cards to get the nth new card.
We know that $N = N_1 + \ldots + N_n$.
N_i is geometric with success, $p = \frac{n-i+1}{n}$. $EN_i = \frac{n}{n-i+1}$

$EN = n(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n})$
Asymptotically, the harmonic sequence above $\approx \log(n)$.

3/9

Example: Assume that an urn contains 10 black, 7 red, and 15 white balls. Take 5 balls out.

1. with replacement

2. without replacement

and W is the number of white balls pulled out. Compute EW.

$$EW = 5 \cdot \frac{5}{22}$$

Let Y be the number of different colors.
Let I_b be the event where black is represented.
Let I_r be the event where red is represented.
Let I_w be the event where white is represented.
We know that $Y = I_b + I_r + I_w$.

$$EY = 1 - \frac{\binom{12}{5}}{\binom{22}{5}} + 1 - \frac{\binom{15}{5}}{\binom{22}{5}} + 1 - \frac{\binom{17}{5}}{\binom{22}{5}}$$

More Properties of expectation

If X and Y are independent,

$$E[g(X)h(Y)] = Eg(X) \cdot E[h(Y)]$$

Proof.

$$\begin{aligned}
E[g(X)h(Y)] &= \int\int g(x)h(y)f(x,y)\,dx\,dy \\
&= \int\int g(x)h(y)f_X(x)g_Y(y)\,dxdy \\
&= \int g(x)f_X(x)\,dx \int h(y)f_Y(y)\,dy \\
&= Eg(X) \cdot Eh(Y)
\end{aligned}$$

$\square$

Example: The right triangle from last time.
(X, Y) is a random point on the right triangle of sides of length 1.
We computed that $EXY = \frac{1}{12}$ and $EX = EY = \frac{1}{3}$.
Therefore, this problem is not independent $\frac{1}{3} \cdot \frac{1}{3} \neq \frac{1}{12}$.

Now, let's look at another problem. Let's say that (X, Y) is a random point on the square of length 1. This problem in fact has X and Y independent because $EXY = EX \cdot EY = \frac{1}{4}$.

Example: Roll a die and then toss as many coins as shown up on the die.
E(number of heads)?
Assume that (X, Y) is discrete. Say we know how to compute $E(Y|X = x)$. Now, I claim that

$$EY = \sum_x E(Y|X = x) \cdot P(X = x)$$

Now,

$$\begin{aligned}
\sum_y yP(Y=y|X=x) &= \sum_y y\frac{P(X=x|Y=y)}{P(X=x)} \\
\sum_x E(Y|X=x)P(X=x) &= \sum_x\sum_y y\frac{P(X=x,y=y)}{P(X=x)}P(X=x) \\
&= \sum_x\sum_y yP(X=x,y=y)
\end{aligned}$$

In the continuous case,

$$EY = \int E(Y|X=x)f_X(x)\,dx$$

$$P(Y|X=x) = x\frac{1}{2}$$

Let X be the number on the die when $x=1,2,\ldots,6$ and Y be the number of heads. Therefore,

$$\begin{aligned}
E(\text{number of heads}) &= E[Y] \\
&= \sum_{x=1}^{6} x\frac{1}{2}\cdot\frac{1}{6} \\
&= \frac{7\cdot 6}{2}\frac{1}{12} \\
&= \frac{7}{4}
\end{aligned}$$

Example: A person dies and presented with 3 doors. There's one that leads to heaven, one that leads to 1 day in purgatory, and one that leads to 2 days in purgatory. After staying in purgatory, you go back to the door and pick again. Choose one at random each time. The doors are reshuffled each time you come back.
What is your expected stay in purgatory?
Let N be the number of days in purgatory.
Conditional expectation of N given you picked door i where $i=1,2,3$.

$$E(N|\text{door}_1) = 0$$

$$E(N|\text{door}_2) = 1+EN$$

$$E(N|\text{door}_i) = 2+EN$$

Therefore,

$$\begin{aligned}
EN &= \frac{1}{3}(1+EN)+\frac{1}{3}(2+EN) \\
\frac{1}{3}EN &= 1 \\
EN &= 3
\end{aligned}$$

3/10 (discussion)

HW 6.42

$$f(x,y) = xe^{-x(y+1)}$$

for $x > 0$, $y > 0$.

a)

$$\begin{aligned}
P(X \le x | Y = y) &= \frac{P(X \le x, Y = y)}{P(Y = y)} \\
P(Y = y) &= \int_0^\infty f(x,y)\,dx \\
&= \int_0^\infty \frac{1}{(y+1)^2} \\
P(X \le x, Y = y) &= \int_0^\infty xe^{-x(y+1)}\,dx \\
&= -\frac{x}{y+1}e^{-x(y+1)} + \frac{1}{(y+1)^2} - \frac{e^{-x(y+1)}}{(y+1)^2} \\
P(X \le x | Y = y) &= 1 - ((y+1)x + 1)e^{-x(y+1)}
\end{aligned}$$

Therefore, $f(x) = (y+1)^2 xe^{-x(y+1)}$.
b)

$$\begin{aligned}
P(XY \le z) &= \int\int_{XY \le z} xe^{-x(y+1)}\,dxdy \\
&= \int\int xe^{-x(y+1)}\,dydx \\
&= \int_0^\infty \int_0^{\frac{z}{x}} xe^{-x(y+1)}\,dydx \\
&= \int_0^\infty e^{-x(y+1)}|_0^{\frac{z}{x}}\,dx \\
&= \int_0^\infty e^{-x} - e^{-z-x}\,dx \\
&= 1 - e^{-z} \\
f_Z(z) &= \frac{d}{dy}P(Z \le z) \\
&= e^{-z}
\end{aligned}$$

HW 6.45

X_1, X_2, X_3 are uniformly distributed over $[0,1]$.

$$\begin{aligned}
P(X_3 \geq X_1 + X_2) &= \int\int_{X_3 \geq X_1+X_2} f_{X_1,X_2,X_3}(x,y,z)\, dxdydz \\
&= \int_0^1 \int_0^{1-y} \int_{x+y}^1 dzdxdy \\
&= \int_0^1 \int_0^{1-y} 1-(x+y)\, dx\, dy \\
&= \int_0^1 (x - \frac{x^2}{2} - xy)|_0^{1-y}\, dy \\
&= \int_0^1 ((1-y) - \frac{1-2y+y^2}{2} - y + y^2)\, dy \\
&= \int_0^1 \frac{2-4y+2y^2-1+2y-y^2}{2}\, dy \\
&= \int_0^1 \frac{1-2y+y^2}{2}\, dy \\
&= \frac{1}{2}(y - y^2 + \frac{y^3}{3})|_0^1 \\
&= \frac{1}{2}(1-1+\frac{1}{3}) \\
&= \frac{1}{6}
\end{aligned}$$

HW 6.49

Let $X_1, \ldots, X_5$ be independently identically distributed on exponential with λ.

$$\begin{aligned}
P(min(X_1,\ldots,X_5) \leq a) &= 1 - P(min(X_1,\ldots,X_5) > a) \\
&= 1 - P(X_1 > a, X_2 > a, \ldots, X_5 > a) \\
&= 1 - \int_a^\infty \ldots \int_a^\infty -\lambda e^{-\lambda} dX_1 \ldots dX_5
\end{aligned}$$

3/11

Covariance

Let X, Y be random variables.

$$\begin{aligned}
Cov(X,Y) &= E((X-EX)(Y-EY)) \\
&= E(XY - (EX)\cdot Y - (EY)\cdot X + EX \cdot EY) \\
&= E(XY) - EX\cdot EY - EY \cdot EX + EX \cdot EY \\
&= E(XY) - EX \cdot EY
\end{aligned}$$

Note:

1. If X and Y are independent, then $Cor(X,Y)=0$.
2. The converse of the previous is false.

Example of why note 2 is true:
Suppose that we have a square(diamond) of length 1 with vertices on (0,1), (1,0), (0,-1), and (-1,0).
Let (X,Y) be a random point on the square (diamond).

$$EX = EY = 0$$

$$\begin{aligned}
EXY &= \frac{1}{2}\int_{-1}^{1} dx \int_{1-|x|}^{-1+|x|} xy\,dy \\
&= \frac{1}{2}\int_{-1}^{1} x\,dx \int_{-1+|x|}^{1-|x|} y\,dy \\
&= \frac{1}{2}\int_{-1}^{1} x\,dx(0) \\
&= 0
\end{aligned}$$

Are they independent? No

Let X and Y be indicator random variables, so $X = I_A$ and $Y = I_B$.
$EX = P(A)$, $EY = P(B)$, $E(XY) = E(I_{A\bigcap B}) = P(A\bigcap B)$.
We know then that the

$$Cov(X,Y) = P(A\bigcap B) - P(A)P(B) = P(A)[P(B|A) - P(B)]$$

If $P(B|A) > P(B)$, then covariance is positive (meaning that they are positively correlated)
Otherwise, they are negatively correlated.
Intuitively, $Cov(X,Y) > 0$ means that "on the average", increasing X will tend to make Y larger.

Variance-Covariance formula

$$E((\sum_{i=1}^{n} X_i)^2) = \sum_{i=1}^{n} EX_i^2 + \sum_{i\neq j} E(X_iX_j)$$

$$\begin{aligned}
Var(\sum_{i=1}^{n} X_i) &= E[\sum_{i=1}^{n}(X_i - EX_i)]^2 \\
&= \sum_{i=1}^{n} Var(X_i) + \sum_{i\neq j} Cov(X_i, X_j) \\
Var(S) &= E(S-ES)^2
\end{aligned}$$

Corollary 5. *If $X_1, X_2, \ldots, X_n$ are independent, then*

$$Var(X_1 + \ldots + X_n) = Var(X_1) + \ldots Var(X_n)$$

Example: Let S_n be $Bernoulli(n,p)$, $S_n = \sum_{i=1}^n I_i$.
Let I_i be $I\{i\text{th trial is a success}\}$ and I_i are independent!
$S_n = np$,

$$\begin{aligned} Var(S_n) &= \sum_{i=1}^n Var(I_i) \\ &= \sum_{i=1}^n (E(I_i) - E(I_i)^2) \\ &= n(p-p^2) \\ &= np(1-p) : \end{aligned}$$

Matching Problem: Let n people buy n gifts, which are distributed at random.
Let X be the number of people who gets their own gift.
Let $I_i = I\{i\text{th person gets own gift}\}$
We know that $X = \sum_{i=1}^n I_i$

$$EI_i = \frac{1}{n}$$

Therefore,

$$EX = 1$$

Now, what is $E(X^2)$?

$$\begin{aligned} E(X^2) &= n \cdot \frac{1}{n} + \sum_{i \neq j} E(I_i I_j) \\ &= 1 + \sum_{i \neq j} \frac{1}{n(n-1)} \\ &= 1 + 1 \\ &= 2 \end{aligned}$$

$E(I_i I_j)$ is the probabiliy that ith person gets his own gift and that jth person get his own gift.
Therefore, $Var(X) = 1$.
In fact, X is close to Poisson with $\lambda = 1$.

Example: Roll a die 10 times.
Let X be the number of 6 rolled and Y be the number of 5 rolled.
Let $X = \sum_{i=1}^1 0X_i$ where $X_i = I\{i\text{th roll is } 6\}$.
Let $Y = \sum_{i=1}^1 0Y_i$ where $Y_i = I\{i\text{th roll is } 5\}$.
Then, $EX = EY = \frac{10}{6}$.
Then,

$$\begin{aligned} EXY &= \sum_{i=1}^1 0 \sum_{j=1}^{10} E(X_i Y_j) \\ &= \sum_{i \neq j} \frac{1}{6^2} (\text{can't get both 5 and 6}) \\ &= \frac{10 \times 9}{36} \end{aligned}$$

3/13

Meaning of the statement, "probability of an event A is p". If you repeat the experiment, independently, a large number of times, the proportion of times, A happens converges to p.
More generally,

Theorem 6. *If $X_1, X_2, \dots$ are independent and identically distributed random variables, then $\frac{X_1+\ldots+X_n}{n}$ converges to EX in the sense that*

$$P\left(\left|\frac{X_1+\ldots+X_n}{n} - EX\right| \geq \epsilon\right) = \ldots$$

In particular, if S_n is $Binomial(n,p) = I_1 + \ldots + I_n$, where $I_i = I\{\text{success at trial}\}$
So,

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \to_{n\to\infty} 0$$

This is called the weak law of Large Numbers.

Theorem 7 (Markov Inequality)**.** *If $X \geq 0$ is a random variable,*

$$P(X \geq a) \leq \frac{1}{a} EX$$

Example: If $EX = 1$ and $X > 0$, $P(X \geq 10) \leq 0.1$.

Proof.

$$I\{X \geq a\} \leq \frac{1}{a} X$$

so

$$P(X \geq a) = E(I\{X \geq a\}) \leq \frac{1}{a} EX$$

□

Theorem 8 (Chebyshev's inequality)**.** *If $EX = \mu$, $VarX = \sigma^2$ are both finite,*

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Example: If $EX = 1$, $VarX = .1$,

$$P(|X - 1| \geq .5) \leq \frac{.1}{.5^2} = \frac{2}{5}$$

This is only useful if σ is really small and k is really big.

Proof.

$$P((x-\mu)^2 \geq k^2) \quad \leq \quad \frac{1}{k^2} E[(X-\mu)^2]$$

□

Proof. Proof of the Central Limit Theorem:

$$\begin{aligned}
\mu &= EX \\
\text{Let } Y_n &= X_1 + \ldots + X_n \\
EY_n &= EX_1 + \ldots + EX_n \\
&= n\mu \\
VarY_n &= n\sigma^2 \\
P(|X_1 + \ldots + X_n - n\mu| \geq n\epsilon) &\leq \frac{n\sigma^2}{n^2\epsilon^2} \\
&= \frac{\sigma^2}{n\epsilon^2} \to 0
\end{aligned}$$

□

Remark: $\frac{X_1+\ldots+X_n}{n}$ converges to EX at the rate of about $\frac{1}{\sqrt{n}}$.

Does $\frac{X_1+\ldots+X_n-\mu\cdot n}{n}$ converges? Yes

Theorem 9 (Central limit theorem)**.** *$X, X_1, X_2, \ldots$ are independent, identically distributed. Let $\mu = EX$, $\sigma^2 = Var(X)$.*

$$P\left(\frac{X_1 + \ldots + X_n - \mu \cdot n}{\sigma\sqrt{n}} \leq a\right) \to_{n\to\infty} P(Z \leq a)$$

where Z follows the normal distribution.

Example: X_n are independently uniform on $[0, 1]$
Let $S_n = X_1 + \ldots + X_n$.

1. Compute approximately $P(S_{200} \leq 90)$
2. Find n so that $P(S_n \geq 50) \geq .99$.

Part 1) $EX_n = \frac{1}{2}$
$VarX_n = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$.

3/16

Central Limit Theorem

Let $X, X_1, X_2, \ldots$ be independent and identical distribution.
Then, the distribution of

$$\frac{X_1 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}$$

if approximately $N(0, 1)$

Note: $E(X_1 + \ldots + X_n) = n\mu$, $Var(X_1 + \ldots + X_n) = \sigma^2 \cdot n$ Example: X_i is uniformly distributed on $[0, 1]$ independently.

1. Compute approximately $P(S_{200} \leq 90)$

2. Find n so that $P(S_n \geq 50)$ with probability of at least .99

$EX_i = .5$
$Var(X_i) = E(X_i^2) - E(X_i)^2 = \int_0^1 x^2\,dx - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$.
Part 1)

$$\begin{aligned} P(S_{200} \leq 90) &= P\left(\frac{S_{200} - 200 \cdot \frac{1}{2}}{\sqrt{200 \cdot \frac{1}{12}}} \leq \frac{90 - 200 \cdot \frac{1}{2}}{\sqrt{200 \cdot \frac{1}{12}}}\right) \\ &\approx P(Z \leq -\sqrt{6}) \text{ where } Z \text{ is the standard normal} \\ &= 1 - P(Z \leq \sqrt{6}) \\ &\approx 1 - .993 \\ &= .007 \end{aligned}$$

Part 2)

$$\begin{aligned} P\left(\frac{S_n - n \cdot \frac{1}{2}}{\sqrt{n \cdot \frac{1}{12}}} \geq \frac{50 - n \cdot \frac{1}{2}}{\sqrt{n\frac{1}{2}}}\right) &= .99 \\ P\left(z \geq -\frac{\frac{n}{2} - 50}{\sqrt{n \cdot \frac{1}{2}}}\right) &= .99 \\ P\left(z \leq \frac{\frac{n}{2} - 50}{\sqrt{n \cdot \frac{1}{12}}}\right) &= \Phi\left(\frac{\frac{n}{2} - 50}{\sqrt{n \cdot \frac{1}{12}}}\right) = .99 \\ &\vdots \quad \text{(Look up tables)} \\ n - 1.345\sqrt{n} - 100 &= 0 \\ n &= 115 \end{aligned}$$

Example: A casino charge \$1 for entrance. For promotion, they offer to first 70000 "guests" the following game:
Roll a die

1. If you get a 6, you get free entrance and \$2

2. If you get a 5, you get free entrance

3. Pay normal fee otherwise

Compute the revenue loss with probability .9.
In symbols, if L is lost revenue, find X so that

$$P(L \leq x) = .9$$

Let $L = X_1 + X_n$ where $n \geq 30000$ and $P(X_1 = 0) = \frac{4}{6}$, $P(X_1 = 1) = \frac{1}{6}$, and $P(X_2 = 3) = \frac{1}{6}$.

$EX_i = \frac{2}{3}$, $Var(X_i) = \frac{1}{6} + \frac{9}{6} - \left(\frac{2}{3}\right)^2 = \frac{11}{9}$

$$\begin{aligned}
P(L \le x) &= P\left(\frac{L - \frac{2}{3}\cdot n}{\sqrt{n \cdot \frac{10}{9}}} \le \frac{x - \frac{2}{3}\cdot n}{\sqrt{n \cdot \frac{11}{9}}}\right) = .9 \\
&\approx P\left(z \le \frac{x - \frac{2}{3}\cdot n}{\sqrt{n \cdot \frac{11}{9}}}\right) = .9 \\
\frac{x - \frac{2}{3}\cdot n}{\sqrt{n \cdot \frac{11}{9}}} &= .8159 \\
x &= 20000 + 235
\end{aligned}$$