problem 1. Harvard Law School courses often have assigned seating to facilitate the "Socratic method." Suppose that there are 100 first year Harvard Law students, and each takes two courses: Torts and Contracts. Both are held in the same lecture hall (which has 100 seats), and the seating is uniformly random and independent for the two courses ? (a) Find the probability that no one has the same seat for both courses (exactly; you should leave your answer as a sum). (b) Find a simple but accurate approximation to the probability that no one has the same seat for both courses. (c) Find a simple but accurate approximation to the probability that at least two students have the same seat for both courses.

solution 1. a=probability same seat for both course b=total same seat for both course

In [ ]:
```
a)
```

In [ ]:
```
p=1-p(a)/p(b)
```

In [ ]:
```
b)
```

In [4]:
```
import numpy as np

a=99
b=100
c=a/b
cc=1-c
cc
```

Out[4]:
```
0.010000000000000009
```

In [ ]:
```
c)
```

In [29]:
```
a=2
b=100
r=a/b
r
```

Out[29]:
```
0.02
```

problem 2. There are 100 passengers lined up to board an airplane with 100 seats (with each seat assigned to one of the passengers). The first passenger in line crazily decides to sit in a randomly chosen seat (with all seats equally likely). Each subsequent passenger takes his or her assigned seat if available, and otherwise sits in a random available seat. What is the probability that the last passenger in line gets to sit in his or her assigned seat?

solution 2:= last passenger get assigned seat or random seat

In [2]:
```
import numpy as np

def aer(n):
    for i in range(n):
        t=5
        p=1
        s=0
        s=np.sum(1/n)
        #r=p/t
```

```
        return s
        #return r
```

In [4]:
```
aer(2)
```

Out[4]:

```
0.5
```

Problem 3. Raindrops are falling at an average rate of 20 drops per square inch per minute. What would be a reasonable distribution to use for the number of raindrops hitting a particular region measuring 5 inches2 in t minutes? Why? Using your chosen distribution, compute the probability that the region has no rain drops in a given 3 second time interval. A reasonable choice of distribution is P ?

solution 3 A reasonable choice of distribution is Poisson where l=20*5=100 assuming this distribution.

$$P(noraindropsin1/20ofaminute) = e^-5$$

In [24]:
```
import numpy as np
def factorial(x):

    if x == 1 or x==0:
        return 1
    else:

        return (x * factorial(x-1))
num = 0

result = factorial(num)
#print("The factorial of", num, "is", result)

def poisson(l,x):


    p=(l**x)*np.exp(l)/result
    #p=np.exp(-100/20)
    #rp=1-p
    return p

poisson(-100/20,0)
```

Out[24]:

```
0.006737946999085467
```

problem 4. Let X be a random day of the week, coded so that Monday is 1, Tuesday is 2, etc. (so X takes values 1, 2,..., 7, with equal probabilities). Let Y be the next day after X (again represented as an integer between 1 and 7). Do X and Y have the same distribution? What is P(X) ?

solution 4:yes x and y have same distribution

```
p(x)=6/7 [p(x)!=7 => x<y]
```

problem 5 For a group of 7 people, find the probability that all 4 seasons (winter, spring, summer, fall) occur at least once each among their birthdays, assuming that all seasons are equally likely ?

solution 5 one or more season have no student their birthday

$$r = \binom{4}{1}3^7 - \binom{4}{2}2^7 + \binom{4}{3}1^7$$

Exclude 1 season - Exclude 2 season + exclude 3 season, also note that we can't exclude all the 4 seasons

$$where \binom{4}{1}3^7 = 4!/1!(4-1)!$$

In [21]:

```
a=np.math.factorial(4)/(np.math.factorial(1)*np.math.factorial(4-1))
b=3**7
a1=np.math.factorial(4)/(np.math.factorial(2)*np.math.factorial(4-2))
b1=2**7
a2=np.math.factorial(4)/(np.math.factorial(3)*np.math.factorial(4-3))
b2=1**7
c=a*b
d=a1*b1
e=a2*b2
f=4**7
```

In [22]:

```
r=(c-d+e)/f
fr=1-r
fr
```

Out[22]:

```
0.5126953125
```

problem 6 Alice attends a small college in which each class meets only once a week. She is deciding between 30 non-overlapping classes. There are 6 classes to choose from for each day of the week, Monday through Friday. Trusting in the benevolence of randomness, Alice decides to register for 7 randomly selected classes out of the 30, with all choices equally likely. What is the probability that she will have classes every day, Monday through Friday?

solution 6

a=possible outcome b=total outcome

$$a = \binom{5}{2} * \binom{6}{2} * \binom{6}{2} * 6^3 + \binom{5}{1} * \binom{6}{3} * 6^4$$

$$b = \binom{30}{7}$$

In [8]:

```
a1=np.math.factorial(5)/(np.math.factorial(2)*np.math.factorial(5-2))
a2=np.math.factorial(6)/(np.math.factorial(2)*np.math.factorial(6-2))
a3=np.math.factorial(5)/(np.math.factorial(1)*np.math.factorial(5-1))
a4=np.math.factorial(6)/(np.math.factorial(3)*np.math.factorial(6-3))
b=np.math.factorial(30)/(np.math.factorial(7)*np.math.factorial(30-7))

c=np.sum((a1)*(a2)*(a2)*(6**3)+(a3)*(a4)*(6**4))/(b)
c
```

Out[8]:

```
0.30238726790450926
```

problem 7 Is it possible that an event is independent of itself? If so, when?

solution 7 Let be an event. If A is independent of itself, then P(A) =

$$a \bigcup a = p(a)^2$$

```
        so P(A) is 0 or 1. So this is only possible in the extreme
cases that the event
        has probability 0 or 1
```

problem 8 Is it always true that if A and B are independent events, then A^c and B^c are independent events? Show that it is, or give a counterexample ?

solution 8 here 2 and 3 are not divided by each other so its independent so a power 2 and b power 2 are not divided by each other so it is independent

In [43]:
```
a = 2
b = 3
a= a**2
b= b**2
print(a)
print(b)
```

```
4
9
```

or

solution 8

Yes, because we have

$$P(A^c \cap B^c) = 1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B));$$

since A and B are independent, this becomes

$$1 - P(A) - P(B) + P(A)P(B) = (1 - P(A))(1 - P(B)) = P(A^c)P(B^c).$$

problem 9

Give an example of 3 events A, B, C which are pairwise independent but not independent. Hint: find an example where whether C occurs is completely determined if we know whether A occurred and whether B occurred, but completely undetermined if we know only one of these things ?

solution 9:

consider dice shuffled randomly A be the event that the first dice is 6 number.b be the event that the second dice is 6 number,and c be the event that the two dice have same result.then a,b,c are dependent but they are pairwise independent. a and b are independent by definition,b and c are independent and similarly b and c are independent.

or

Solution 9

Consider two fair, independent coin tosses, and let A be the event that the first toss is Heads, B be the event that the second toss is Heads, and C be the event that the two tosses have the same result. Then A, B, C are dependent since $P(A \cap B \cap C) = P(A \cap B) = P(A)P(B) = 1/4 \neq 1/8 = P(A)P(B)P(C)$, but they are pairwise independent: A and B are independent by definition; A and C are independent since $P(A \cap C) = P(A \cap B) = 1/4 = P(A)P(C)$, and similarly B and C are independent.

problem 10. A bag contains one marble which is either green or blue, with equal probabilities. A green marble is put in the bag (so there are 2 marbles now), and then a random marble is taken out. The marble taken out is green. What is the probability that the remaining marble is also green?

solution 10.
Here a bag contain marble which is green or blue and one marble is put in bag so total outcome=3 and most possible outcome is 2 for green marble a=possible outcome b=total outcome a=a/b

In [7]:
```
a=2/3
p=np.sum(2/3)
p
```

Out[7]:
```
0.6666666666666666
```

or

solution 10

Let E_1 and E_2 be the events that marble is green and blue respectively in the bag. Let A be the event of picking up a green marble

Then $P(E_1) = P(E_2) = 1/2, P(A/E_1) = 1, P(A/E_2) = 1/2$

Now, if the marble taken out is green, then probability that remaining marble is also green is $P(E_1/A)$

$$P(E_1/A) = \frac{P(E_1)P(A/E_1)}{P(E_1)P(A/E_1)+P(E_2)P(A/E_2)}$$

$$= \frac{\frac{1}{2}\times 1}{\frac{1}{2}\times 1+\frac{1}{2}\times\frac{1}{2}} = \frac{\frac{1}{2}}{\frac{2+1}{4}} = \frac{1}{2} \times \frac{4}{3} = \frac{2}{3}$$

problem 11. A group of n 2 people decide to play an exciting game of Rock-Paper Scissors. As you may recall, Rock smashes Scissors, Scissors cuts Paper, and Paper covers Rock (despite Bart Simpson saying "Good old rock, nothing beats that!"). Usually, this game is played with 2 players, but it can be extended to more players as follows. If exactly 2 of the 3 choices appear when everyone reveals their choice, say a, b 2 {Rock, Paper, Scissors} where a beats b, the game is decisive: the players who chose a win, and the players who chose b lose. Otherwise, the game is indecisive and the players play again. For example, with 5 players, if one player picks Rock, two pick Scissors, and two pick Paper, the round is indecisive and they play again. But if 3 pick Rock and 2 pick Scissors, then the Rock players win and the Scissors players lose the game. 1 Assume that the n players independently and randomly choose between Rock, Scissors, and Paper, with equal probabilities. Let X, Y, Z be the number of players who pick Rock, Scissors, Paper, respectively in one game. (a) Find the joint PMF of X, Y, Z. (b) Find the probability that the game is decisive. Simplify your answer (it should not involve a sum of many terms). (c) What is the probability that the game is decisive for n = 5? What is the limiting probability that a game is decisive as n ! 1? Explain briefly why your answer makes sense.

solution 11

Assume that the n players independently and randomly choose between Rock, Scissors, and Paper, with equal probabilities. Let X, Y, Z be the number of players who pick Rock, Scissors, Paper, respectively in one game.

(a) Find the joint PMF of X, Y, Z.

The joint distribution of X, Y, Z is

$$P(X = a, Y = b, Z = c) = \frac{n!}{a!b!c!}\left(\frac{1}{3}\right)^{a+b+c}$$

where a, b, c are any nonnegative integers with $a+b+c = n$, since $(1/3)^{a+b+c}$ is the probability of any specific configuration of choices for each player with the right numbers in each category, and the coefficient in front counts the number of distinct ways to permute such a configuration.

Alternatively, we can write the joint PMF as

$$P(X = a, Y = b, Z = c) = P(X = a)P(Y = b|X = a)P(Z = c|X = a, Y = b),$$

where for $a + b + c = n$, $P(X = a)$ can be found from the Bin$(n, 1/3)$ PMF, $P(Y = b|X = a)$ can be found from the Bin$(n - a, 1/2)$ PMF, and $P(Z = c|X = a, Y = b) = 1$.

This is a Multinomial$(n, (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}))$. distribution.

(b) Find the probability that the game is decisive. Simplify your answer (it should not involve a sum of many terms).

Hint: using symmetry, the probability can be written as 3 times a certain sum. To do the summation, use the binomial theorem or the fact that $\sum_{k=0}^{n}\binom{n}{k} = 2^n$.

The game is decisive if and only if exactly one of X, Y, Z is 0. These cases are disjoint so by symmetry, the probability is 3 times the probability that X is zero and Y and Z are nonzero. Note that if $X = 0$ and $Y = k$, then $Z = n-k$. This gives

$$\begin{aligned} P(\text{decisive}) &= 3\sum_{k=1}^{n-1} \frac{n!}{0!k!(n-k)!}\left(\frac{1}{3}\right)^n \\ &= 3\left(\frac{1}{3}\right)^n \sum_{k=1}^{n-1}\binom{n}{k} \\ &= \frac{2^n-2}{3^{n-1}} \end{aligned}$$

since $\sum_{k=1}^{n-1}\binom{n}{k} = -1-1+\sum_{k=0}^{n}\binom{n}{k} = 2^n - 2$ (by the binomial theorem or the fact that a set with n elements has 2^n subsets). As a check, when $n = 2$ this reduces to $2/3$, which makes sense since for 2 players, the game is decisive if and only if the two players do not pick the same choice.

(c) What is the probability that the game is decisive for $n = 5$? What is the limiting probability that a game is decisive as $n \to \infty$? Explain briefly why your answer makes sense.

For $n = 5$, the probability is $(2^5 - 2)/3^4 = 30/81 \approx 0.37$. As $n \to \infty$, $(2^n - 2)/3^{n-1} \to 0$, which make sense since if the number of players is very large, it is very likely that there will be at least one of each of Rock, Paper, and Scissors.

problem 12

A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase “free money” is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention “free money”. What is the probability that it is spam?

solution 12

According to question from 80% email 10% use phrase "free money" and 1% use non-spam email

```
a=possiblities to spm mail
b=possibilities for span and non span email
```

In [5]:

```
a=(10/100)*(80/100)
b=(80/100*10/100)+(1/100*20/100)

c=np.sum(a/b)
c
```

Out[5]: 0.9756097560975612

problem 13 A crime is committed by one of two suspects, A and B. Initially, there is equal evidence against both of them. In further investigation at the crime scene, it is found that the guilty party had a blood type found in 10% of the population. Suspect A does match this blood type, whereas the blood type of Suspect B is unknown. (a) Given this new information, what
is the probability that A is the guilty party? (b) Given this new information, what is the probability that B’s blood type matches that found at the crime scene?

solution 13

```
            here a) a is guilty party = possible outcome/total outcome
                            per=10%=1/10
                               ans = 10/11

                 b) b's blood type match at crime scene = possible
   outcome/total outcome
                            per=10%=1/10
                            a=2 b=11
                            ans=2/11
```

or

solution 13

(a) Given this new information, what is the probability that A is the guilty party?

Let M be the event that A's blood type matches the guilty party's and for brevity, write A for "A is guilty" and B for "B is guilty". By Bayes' Rule,

$$P(A|M) = \frac{P(M|A)P(A)}{P(M|A)P(A) + P(M|B)P(B)} = \frac{1/2}{1/2 + (1/10)(1/2)} = \frac{10}{11}.$$

(We have $P(M|B) = 1/10$ since, given that B is guilty, the probability that A's blood type matches the guilty party's is the same probability as for the general population.)

(b) Given this new information, what is the probability that B's blood type matches that found at the crime scene?

Let C be the event that B's blood type matches, and condition on whether B is guilty. This gives

$$P(C|M) = P(C|M, A)P(A|M) + P(C|M, B)P(B|M) = \frac{1}{10} \cdot \frac{10}{11} + \frac{1}{11} = \frac{2}{11}.$$

problem 14. You are going to play 2 games of chess with an opponent whom you have never played against before (for the sake of this problem). Your opponent is equally likely to be a beginner, intermediate, or a master. Depending on (a) What is your probability of winning the first game? (b) Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game (c) Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

solution 14. a) a=possible outcome b=total outcome b) a=possible outcome b=total outcome

a)

In [6]:
```
a=0.9+0.5+0.3
b=3
c=np.sum(a/b)
c
```

Out[6]:
```
0.5666666666666667
```

b)

In [11]:
```
a=((0.9*0.9)+(0.5*0.5)+(0.3*0.3))/3
d=c
cb=np.sum(a/d)
cb
```

Out[11]:
```
0.6764705882352942
```

```
  c)      Here outcome of the game are independent because one
game is not related to another game and the opponent skill level
is high game is challenging otherwise easy and player's skill
level is high then chance to game win.
```

or

problem 14

(a) What is your probability of winning the first game?

Let W_i be the event of winning the ith game. By the law of total probability,

$$P(W_1) = (0.9 + 0.5 + 0.3)/3 = 17/30.$$

(b) Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game (assume that, given the skill level of your opponent, the outcomes of the games are independent)?

We have $P(W_2|W_1) = P(W_2, W_1)/P(W_1)$. The denominator is known from (a), while the numerator can be found by conditioning on the skill level of the opponent:

$$P(W_1, W_2) = \frac{1}{3}P(W_1, W_2|\text{beginner}) + \frac{1}{3}P(W_1, W_2|\text{intermediate}) + \frac{1}{3}P(W_1, W_2|\text{expert}).$$

Since W_1 and W_2 are conditionally independent given the skill level of the opponent, this becomes

$$P(W_1, W_2) = (0.9^2 + 0.5^2 + 0.3^2)/3 = 23/60.$$

So

$$P(W_2|W_1) = \frac{23/60}{17/30} = 23/34.$$

c)Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

Independence here means that knowing one game's outcome gives no information about the other game's outcome, while conditional independence is the same statement where all probabilities are conditional on the opponent's skill level. Conditional independence given the opponent's skill level is a more reasonable assumption here. This is because winning the first game gives information about the opponent's skill level, which in turn gives information about the result of the second game. That is, if the opponent's skill level is treated as fixed and known, then it may be reasonable to assume independence of games given this information; with the opponent's skill level random, earlier games can be used to help infer the opponent's skill level, which a↵ects the probabilities for future games.

problem 15. A chicken lays n eggs. Each egg independently does or doesn't hatch, with probability p of hatching. For each egg that hatches, the chick does or doesn't survive (independently of the other eggs), with probability s of survival. Let N ⇠ Bin(n, p) be the number of eggs which hatch, X be the number of chicks which survive, and Y be the number of chicks which hatch but don't survive (so X + Y = N). Find the marginal PMF of X, and the joint PMF of X and Y . Are they independent ?

solution 15

Marginally we have $X \sim \text{Bin}(n, ps)$, as shown on a previous homework problem using a story proof (the eggs can be thought of as independent Bernoulli trials with probability ps of success for each). Here X and Y are *not* independent, unlike in the chicken-egg problem from class (where N was Poisson). This follows immediately from thinking about an *extreme case*: if $X = n$, then clearly $Y = 0$. So they are not independent: $P(Y = 0) < 1$, while $P(Y = 0|X = n) = 1$.

To find the joint distribution, condition on N and note that only the $N = i+j$ term is nonzero: for any nonnegative integers i, j with $i + j \leq n$,

$$
\begin{aligned}
P(X = i, Y = j) &= P(X = i, Y = j|N = i + j)P(N = i + j) \\
&= P(X = i|N = i + j)P(N = i + j) \\
&= \binom{i+j}{i} s^i (1-s)^j \binom{n}{i+j} p^{i+j}(1-p)^{n-i-j} \\
&= \frac{n!}{i!j!(n-i-j)!}(ps)^i (p(1-s))^j (1-p)^{n-i-j}.
\end{aligned}
$$

If we let Z be the number of eggs which don't hatch, then from the above we have that (X, Y, Z) has a Multinomial$(n, (ps, p(1-s), (1-p))$ distribution, which makes sense intuitively since each egg independently falls into 1 of 3 categories: hatch-and-survive, hatch-and-don't-survive, and don't-hatch, with probabilities $ps, p(1-s), 1-p$ respectively.