

Homework Solutions for HW #1

1. A fair - sided die (each side is equally likely to come up) with sides numbered from 1 to n is rolled until the value 5 appears on top. How many rolls are expected?

For the n -sided fair die the probability of rolling one of the n faces (e.g. 5) is:

$P(\text{roll one of } n \text{ faces}) = \frac{1}{n}$. So in order to calculate the expected number of tosses we sum up the

probabilities that we roll 5 in one toss, two tosses and so on. More formally:

We define independent identically distributed random variables (r.vs) $N_i = 1$ if a 5 is rolled in i th toss, 0 otherwise. Let S be the sum of all $N_i \forall i > 0$. Then

$$E(S) = 1 \cdot P(N_1 = 1) + 2 \cdot P(N_1 = 0) \cdot P(N_2 = 1) + \dots + k \cdot P(N_{k-1} = 0)^{k-1} \cdot P(N_k = 1) + \dots$$

Therefore:

$$\begin{aligned} E(S) &= \sum_{k=1}^{\infty} k \cdot P(N_{k-1} = 0)^{k-1} \cdot P(N_k = 1) = P(N_k = 1) \cdot \sum_{k=1}^{\infty} k \cdot P(N_{k-1} = 0)^{k-1} \Rightarrow \\ E(S) &= \frac{P(N_1 = 1)}{(1 - P(N_1 = 0))^2} = \frac{P(N_1 = 1)}{(P(N_1 = 1))^2} = \frac{1}{P(N_1 = 1)} = \frac{1}{\frac{1}{n}} = n \Rightarrow \end{aligned}$$

$$E(S) = n \text{ (the number of faces in the die).}$$

You can also think of the result as the expected value of a Bernoulli process with probability of success $1/n$.

Second solution:

$$E[N] = \sum_{i=0}^{\infty} P(N > i) \text{ where } P(N > i) = \left(\frac{n-1}{n}\right)^i, E[N] = \sum_{i=0}^{\infty} \left(\frac{n-1}{n}\right)^i, \text{ (geometric regression)}$$

$$\text{So } E[N] = \frac{1}{1 - \left(\frac{n-1}{n}\right)} = n$$

2. A fair coin is tossed 10 times. What is the probability that either the first five tosses are heads or the last 5 tosses are tails?

The tosses are completely independent so the probability that the first five tosses are heads is the same as the probability that the last five tosses are heads. The probability that the first five tosses are heads is:

$$P(X_1 = H, X_2 = H, X_3 = H, X_4 = H, X_5 = H) = P(X_1 = H) \cdot P(X_2 = H) \cdot P(X_3 = H) \cdot P(X_4 = H) \cdot P(X_5 = H)$$

$$P(X_1 = H, X_2 = H, X_3 = H, X_4 = H, X_5 = H) = P(X_1 = H)^5 = \left(\frac{1}{2}\right)^5$$

So the total probability:

$$\begin{aligned} P(X_1 = H, X_2 = H, X_3 = H, X_4 = H, X_5 = H \cup X_6 = T, X_7 = T, X_8 = T, X_9 = T, X_{10} = T) &= \\ &= P(X_1 = H, X_2 = H, X_3 = H, X_4 = H, X_5 = H) + P(X_6 = T, X_7 = T, X_8 = T, X_9 = T, X_{10} = T) - \\ &\quad - P(X_1 = H, X_2 = H, X_3 = H, X_4 = H, X_5 = H, X_6 = T, X_7 = T, X_8 = T, X_9 = T, X_{10} = T) = \\ &= \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 - \left(\frac{1}{2}\right)^{10} = 2 \cdot \left(\frac{1}{2}\right)^5 - \left(\frac{1}{2}\right)^{10} \quad (\text{using inclusion-exclusion}) \end{aligned}$$

(The tosses are all independent to each other).

3. You lose at craps if you roll 2,3 or 12 on your first roll. You win if on your first roll, you roll 7 or 11. You win if you roll anything else (4-6, 8-10) on your first roll and roll this value again before rolling a 7. Otherwise you lose.

(a) What is the probability that you will win? If you do this right, the odds should be in the house's favor (i.e., less than .5), but by very little.

(b) If you gain a dollar for each win and lose a dollar for each loss, how much money will you have left on average after playing 100 games?

a) In order to compute the probability to win we first have to compute the probabilities that we will win on the first roll, on the second roll and so on.

$$P(\text{win in any roll}) = P(\text{win on first roll}) + P(\text{win on all other rolls}) \quad (1)$$

Let R_i = the value of the i^{th} roll.

$$\text{Then } P(R_i = k) = P(R_1 = k) \text{ for every } i \in \mathbb{N}^* \quad (1b)$$

The probability that we are going to win on the first roll is:

$$P(\text{win on first roll}) = P(R_1 = 7) + P(R_1 = 11).$$

But we have that :

$$P(R_1 = 7) = P\left(\begin{array}{cc} \text{1st die} & \text{2nd die} \\ 1 & 6 \\ 2 & 5 \\ 3 & 4 \\ 4 & 3 \\ 5 & 2 \\ 6 & 1 \end{array}\right) = 6 \cdot \frac{1}{36} = \frac{1}{6} \quad (\text{all rolls are independent with probability } \frac{1}{36})$$

$$P(R_1 = 11) = P\left(\begin{array}{cc} \text{1st die} & \text{2nd die} \\ 5 & 6 \\ 6 & 5 \end{array}\right) = 2 \cdot \frac{1}{36} = \frac{1}{18} \quad (\text{all rolls are independent with probability } \frac{1}{36})$$

So (1) now becomes $P(\text{win of first roll}) = P(R_1=7) + P(R_1=11) = \frac{3}{18}$ (2)

Now we continue to compute the probability that we are going to win in two rolls, three rolls, ... etc. We have to compute six groups of probabilities because we can win if we roll a 4,5,6,8,9,10 and then we repeat the same number without bringing a 7. So:

$$P(\text{two rolls of 4}) = \underbrace{P(R_1=4) \cdot P(R_2=4)}_{\text{two rolls}} + \underbrace{P(R_1=4) \cdot P(R_2 \neq 4, R_2 \neq 7) \cdot P(R_3=4)}_{\text{three rolls}} +$$

$$+ \dots + \underbrace{P(R_1=4) \cdot P(R_2 \neq 4, R_2 \neq 7)^{k-2} \cdot P(R_k=4)}_{k \text{ rolls}} + \dots$$

$$P(\text{two rolls of 4}) = \sum_{k=0}^{\infty} P(R_1=4) \cdot P(R_k \neq 4, R_k \neq 7)^k \cdot P(R_{k+2}=4) \Rightarrow \text{using (1b)}$$

$$P(\text{two rolls of 4}) = P(R_1=4)^2 \cdot \sum_{k=0}^{\infty} P(R_2 \neq 4, R_2 \neq 7)^k = \frac{P(R_1=4)^2}{1 - P(R_2 \neq 4, R_2 \neq 7)}$$

$$P(\text{two rolls of 4}) = \frac{P(R_1=4)^2}{1 - P(R_2 \neq 4, R_2 \neq 7)} = \frac{P(R_1=4)^2}{P(R_2=4 \cup R_2=7)} = \frac{P(R_1=4)^2}{P(R_2=4) + P(R_2=7)} \quad (3)$$

With the same method we find that:

$$P(\text{two rolls of 5}) = \frac{P(R_1=5)^2}{P(R_2=5) + P(R_2=7)}$$

$$P(\text{two rolls of 6}) = \frac{P(R_1=6)^2}{P(R_2=6) + P(R_2=7)}$$

$$P(\text{two rolls of 8}) = \frac{P(R_1=8)^2}{P(R_2=8) + P(R_2=7)}$$

$$P(\text{two rolls of 9}) = \frac{P(R_1=9)^2}{P(R_2=9) + P(R_2=7)}$$

$$P(\text{two rolls of 10}) = \frac{P(R_1=10)^2}{P(R_2=10) + P(R_2=7)}$$

We now have to compute $P(R_1=4)$, $P(R_1=5)$, $P(R_1=6)$, $P(R_1=8)$, $P(R_1=9)$, $P(R_1=10)$.

$$\begin{aligned}
 P(R_1=4) &= \left\{ \begin{array}{cc} \text{1st die} & \text{2nd die} \\ 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{array} \right\} = \frac{3}{36}, \quad P(R_1=5) = \left\{ \begin{array}{cc} \text{1st die} & \text{2nd die} \\ 1 & 4 \\ 2 & 3 \\ 3 & 2 \\ 4 & 1 \end{array} \right\} = \frac{4}{36} \\
 P(R_1=6) &= \left\{ \begin{array}{cc} \text{1st die} & \text{2nd die} \\ 1 & 5 \\ 2 & 4 \\ 3 & 3 \\ 4 & 2 \\ 5 & 1 \end{array} \right\} = \frac{5}{36}, \quad P(R_1=8) = \left\{ \begin{array}{cc} \text{1st die} & \text{2nd die} \\ 2 & 6 \\ 3 & 5 \\ 4 & 4 \\ 5 & 3 \\ 6 & 2 \end{array} \right\} = \frac{5}{36} \\
 P(R_1=9) &= \left\{ \begin{array}{cc} \text{1st die} & \text{2nd die} \\ 3 & 6 \\ 4 & 5 \\ 5 & 4 \\ 6 & 3 \end{array} \right\} = \frac{4}{36}, \quad P(R_1=10) = \left\{ \begin{array}{cc} \text{1st die} & \text{2nd die} \\ 4 & 6 \\ 5 & 5 \\ 6 & 4 \end{array} \right\} = \frac{3}{36}
 \end{aligned}$$

Now from (1),(2) using all the previous we have that:

$$\begin{aligned}
 P(\text{win in any roll}) &= P(\text{win in first roll}) + P(\text{two rolls of 4}) + P(\text{two rolls of 5}) + P(\text{two rolls of 6}) \\
 &\quad + P(\text{two rolls of 8}) + P(\text{two rolls of 9}) + P(\text{two rolls of 10}).
 \end{aligned}$$

$$P(\text{win in any roll}) = \frac{3}{36} + \frac{\left(\frac{3}{36}\right)^2}{\frac{3}{36} + \frac{6}{36}} + \frac{\left(\frac{4}{36}\right)^2}{\frac{4}{36} + \frac{6}{36}} + \frac{\left(\frac{5}{36}\right)^2}{\frac{5}{36} + \frac{6}{36}} + \frac{\left(\frac{5}{36}\right)^2}{\frac{5}{36} + \frac{6}{36}} + \frac{\left(\frac{4}{36}\right)^2}{\frac{4}{36} + \frac{6}{36}} + \frac{\left(\frac{3}{36}\right)^2}{\frac{3}{36} + \frac{6}{36}} =$$

$$P(\text{win in any roll}) = \frac{3}{36} + 2 \cdot \frac{1}{36} + 2 \cdot \frac{16}{360} + 2 \cdot \frac{25}{396} = 0.4929$$

b) For 100 games we have 100 independent games with probability p of winning.

Let X be a random variable that equals the earnings, let W be the number of games won, L be the number of games lost. Then $E[X] = E[W-L] = E[W] - E[L]$. Moreover we know that the total number of games is 100 so $E[L] = 100 - E[W]$. Now $E[X] = 2E[W] - 100$. But this is a Bernoulli distribution with probability of success p . So we know that $E[W] = n \cdot p = 100 \cdot 0.4929 = 49.29$ and the expected number of dollars that we will win is $E[X] = 2 \cdot 49.29 - 100 = -1.42\$$ (We lose 1.42\$).

In other words on average the expected value of the games that we will win is $100 \cdot p$ (because it is the sum of 100 independent identically distributed random variables). So on average we will win 49.29 games (and lose $100 - 49.29 = 50.71$) and thus we will lose $49.29 - 50.71 = 1.42\$$.

4. Consider a room containing n people in which each person is equally likely to be born on any day of the year (i.e., assume no leap years and $P(G_i = k) = \frac{1}{365}$, where G_i is the birth date of the i th person

and k is a number between 1 and 365 indicating that person's birth date).

(a) What is the probability that no two people have the same birthday?

(b) What is the probability that k people were born on January 24 where $k < n$?

(c) How many pairs of people can we expect to have the same birthday? Give a closed-form solution without summations.

a) To solve this part we have to think that each time we fix a birthday in a day this day cannot be the birthday of another person. Therefore the probability that no two people have the same birthday is:

$P(\text{no two people have same birthday}) =$

$$P(X_2 = b_2 | X_1 = b_1) \cdot P(X_3 = b_3 | X_1 = b_1, X_2 = b_2) \cdot \dots \cdot P(X_n = b_n | X_i = b_i, \forall i < n) =$$

$$P(\text{no two people same birth.}) = \frac{364}{365} \cdot \frac{363}{365} \cdot \dots \cdot \frac{364 - n + 1}{365} = \prod_{i=1}^{n-1} \frac{365 - i}{365}, \text{ where } n \text{ is the number of people}$$

If $n > 365$ we will have $P(\text{no two people same birth.}) = 0$.

Second solution:

There are $\frac{365!}{(365 - n)!}$ ways to choose distinct days for n people (allowing same n -tuples with different

ordering), and there are 365^n ways to pick n birthdays for n people (allowing repetition). Thus we have

$$\text{that } P(\text{no two people same birth.}) = \frac{365!}{(365 - n)!} = \frac{365!}{(365 - n)! \cdot 365^n}.$$

b) For this part we have to remember the Binomial distribution because that is what we have.

There are $\binom{n}{k}$, (n choose k) ways to select k people from n (without allowing reordering of the k -

tuples) whose birthday falls on January 24th. Then, for each possible selection, there is a probability

of $\left(\frac{1}{365}\right)^k \cdot \left(\frac{364}{365}\right)^{n-k}$ that exactly those k people are the only ones to have their birthday on Jan 24.

Since it is not possible for two different sets of k people to be the only people to have their birthday on

Jan 24, each of these $\binom{n}{k}$ probabilities covers a disjoint set of events, and hence they can be added

together.

$$\text{Formally: } P(k \text{ persons born in 24th of Jan.}) = \binom{n}{k} \cdot \left(\frac{1}{365}\right)^k \cdot \left(\frac{364}{365}\right)^{n-k}.$$

In other words the probability of two people having the same birthday is the probability of success in terms of binomial distribution. So the formula is the probability of k successes when we have a binomial distribution with n trials.

c) We define the indicator random variable: $I_{ij} = \begin{cases} 1 & \text{if birthday of } i \text{ is the same as } j\text{'s} \\ 0 & \text{otherwise} \end{cases}$

The expected value of the pairs is:

$E\left(\sum_{i=1}^{n-1} \sum_{k=i+1}^n I_{ij}\right) = \sum_{i=1}^{n-1} \sum_{k=i+1}^n E(I_{ij}) = \sum_{i=1}^{n-1} \sum_{k=i+1}^n p$, since the expected value of a sum of random variables equals the sum of the expected values of the random variables. Also $E(I_{ij}) = p = 1/365$.

Now the double summation is nothing more than an arithmetic regression $(n-1) + (n-2) + \dots + 2 + 1$. Eventually we have:

$$E\left(\sum_{i=1}^{n-1} \sum_{k=i+1}^n I_{ij}\right) = p \cdot (1 + 2 + \dots + n - 1) = p \cdot \frac{n(n-1)}{2} = \frac{1}{365} \cdot \frac{n(n-1)}{2}$$

5. Let X , Y , and Z be random variables such that X and Y are independent, X and Z are independent, and Y and Z are independent (i.e., $P(X, Y) = P(X) \cdot P(Y)$, $P(X, Z) = P(X) \cdot P(Z)$, $P(Y, Z) = P(Y) \cdot P(Z)$). Show that this does not imply that X , Y and Z are independent (i.e., that it is not always true that $P(X, Y, Z) = P(X) \cdot P(Y) \cdot P(Z)$). Think about rolling 2 dice...

For this exercise we have to find a counter example. We present the following an example (others exist):

Define X, Y to be $X = \begin{cases} 1 & \text{if the first die rolls even} \\ 0 & \text{otherwise (rolls odd)} \end{cases}$ and $Y = \begin{cases} 1 & \text{if the second die rolls even} \\ 0 & \text{otherwise (rolls odd)} \end{cases}$

Define also Z to be $Z = \begin{cases} 1 & \text{if } X + Y \text{ is even} \\ 0 & \text{otherwise } (X + Y \text{ odd}) \end{cases}$

Now $P(X=1, Y=1) = P(X=1) P(Y=1)$ since these are given to be independent random rolls $P(X=1, Z=1) = P(X=1, Y=1) = P(X=1) P(Y=1)$, and $P(Y=1) = 1/2 = P(Z=1)$

A similar proof holds for $P(Y=1, Z=1)$. We have showed that two of them are independent but all of them are not because the value of one (e.g. Z) depends on the values of the others (e.g. X, Y). A numerical example is the following:

$$P(X=1, Y=1, Z=0) = 0 \text{ (} X+Y=2 \text{ even)} \text{ but } P(X=1) \cdot P(Y=1) \cdot P(Z=0) = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) = \left(\frac{1}{8}\right).$$

6. With probability, p_1 , TA Stavrou has a message to give to the class. With probability p_2 , Professor Rubenstein will remember to deliver the message. With probability, p_3 , Mary, a student in the class arrives on time to hear any announcements Professor Rubenstein makes. Assuming these three events are independent (such that the probability that Mary hears an announcement is, $p_1 p_2 p_3$), what is the

probability that TA Stavrou gave an announcement to Professor Rubenstein, given that Mary did not hear an announcement?

We define the following random variables to facilitate our solution to the problem:

A: TA Stavrou made an announcement (takes value 1 with probability p_1 , 0 otherwise)

B: Professor Rubenstein remember (takes value 1 with probability p_2 , 0 otherwise)

C: Mary arrives on time (takes value 1 with probability p_3 , 0 otherwise)

H: Mary hears the announcement is $H=A \cdot B \cdot C$

Now we have to compute the conditional probability:

$$P(A=1 | A \cdot B \cdot C = 0) = \frac{P(A=1, A \cdot B \cdot C = 0)}{P(A \cdot B \cdot C = 0)} = \frac{P(A=1) \cdot P(B \cdot C = 0)}{P(A \cdot B \cdot C = 0)} = \frac{P(A=1) \cdot (1 - P(B \cdot C = 1))}{1 - P(A \cdot B \cdot C = 1)}$$

$$P(A=1 | A \cdot B \cdot C = 0) = \frac{p_1 \cdot (1 - p_2 \cdot p_3)}{1 - p_1 \cdot p_2 \cdot p_3}$$

Another equivalent solution is:

$$P(A=1 | A \cdot B \cdot C = 0) = \frac{P(A=1, A \cdot B \cdot C = 0)}{P(A \cdot B \cdot C = 0)} = \frac{P(A=1) \cdot P(B \cdot C = 0)}{P(A \cdot B \cdot C = 0)} = \frac{P(A=1) \cdot (P(B=1) \cdot P(C=0) + P(B=0))}{1 - P(A \cdot B \cdot C = 1)}$$

$$P(A=1 | A \cdot B \cdot C = 0) = \frac{p_1 \cdot (p_2 \cdot (1 - p_3) + (1 - p_2))}{1 - p_1 \cdot p_2 \cdot p_3} = \frac{p_1 \cdot (1 - p_2 \cdot p_3)}{1 - p_1 \cdot p_2 \cdot p_3}$$