

# HW1-3

## Problem 2

Prove that

$$\sum_{j=0}^k \binom{n}{j} \binom{m}{k-j} = \binom{m+n}{k}$$

## Problem 3

A fair die is rolled  $n$  times. What is the probability that at least 1 of the 6 values never appears? Give a formula answer with derivation. Hint: use inclusion-exclusion formula.

## Solution 3

Let  $A_i$  be the event that  $i$ -th value does not appear. Then,  $\bigcup_{i=1}^6 A_i$  is the event that at least one values does not appear. Let  $B_{ij}$  be the event that  $i$ -th value does not appear on  $j$ -th roll. Then,

$$P(A_i) = \prod_{j=1}^n P(B_{ij}) = (P(B_{i1}))^n = \left(\frac{5}{6}\right)^n$$

Now, in order to use the inclusion-exclusion formula, we will need the probabilities of intersections. To get those, consider the probability of the event  $A_i \cap A_j$ , i.e. that values  $i$  and  $j$  did not appear:

$$P(A_i \cap A_j) = \left(\frac{4}{6}\right)^n$$

Finally, consider the fact that the values are free to be any of the six, so we need to choose them. Then,

$$P\left(\bigcup_{i=1}^6 A_i\right) = \binom{6}{1} \left(\frac{5}{6}\right)^n - \binom{6}{2} \left(\frac{4}{6}\right)^n + \binom{6}{3} \left(\frac{3}{6}\right)^n - \binom{6}{4} \left(\frac{2}{6}\right)^n + \binom{6}{5} \left(\frac{1}{6}\right)^n$$

## Problem 4

There are two baskets. The first basket contains one white ball, the second basket contains one black ball. One basket is chosen randomly and a white ball is put into the chosen basket. The balls in this basket are shuffled. Then one ball is extracted from this basket. This ball turns out to be white. What is the posterior probability that the second ball drawn from this basket is also white?

## Solution 4

Probability that the second ball drawn from this basket is also white is the same as probability that the basket containing a white ball was chosen initially. Denote the event that the basket with white ball was chosen as  $B$ , and the event that the first drawn ball was white as  $A$ , then using the Bayes rule

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\bar{B})\mathbb{P}(\bar{B})} = \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}} = \frac{2}{3}$$

## Problem 5

If you get a positive result on a COVID test that only gives a false positive with probability 0.001 (true positive with probability 0.999), what's the chance that you've actually got COVID, if

- (1 point) The prior probability that a person has COVID is 0.01
- (1 point) The prior probability that a person has COVID is 0.0001

## Problem 8

(2 bonus points) The cloakroom of a theater has randomly permuted all  $n$  visitors' hats. Find the probability that at least one visitor gets his hat. Give a formula answer with derivation. Given  $n = 4$ , give a number answer. Hint: use inclusion-exclusion formula.

## Solution 8

Denote  $A_i$  the event that visitor  $i$  gets his hat. Probability of this event is easily

obtained: if we fix this visitor as the first, he has the choice of all the hats

$$\mathbb{P}(A_i) = \frac{1}{n}$$

Probability of intersection of two of these events is also easy, the second visitor has the choice of all but one hat (picked by first visitor)

$$\mathbb{P}(A_i) = \frac{1}{n} \frac{1}{n-1}$$

Intersections of more events follow similar argumentation. Then, using these events we write

$$\mathbb{P}(\text{at least one visitor get their hats}) = \mathbb{P}\left(\bigcup_{i=1}^n A_i\right)$$

Using inclusion-exclusion formula:

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j) + \dots + (-1)^{n+1} \mathbb{P}(A_1 \cap \dots \cap A_n) = \\ &= \sum_{i=1}^n \frac{1}{n} - \sum_{1 \leq i < j \leq n} \frac{1}{n} \frac{1}{n-1} + \dots + (-1)^{n+1} \frac{1}{n!} = \\ &= n \cdot \frac{1}{n} - \frac{n(n-1)}{2!} \frac{1}{n} \frac{1}{n-1} + \dots + (-1)^{n+1} \frac{1}{n!} = \\ &= 1 - \frac{1}{2!} + \dots + (-1)^{n+1} \frac{1}{n!} \end{aligned}$$

## Problem 9

(2 bonus points) You are the contestant on the Monty Hall show. Monty is trying out a new version of his game, with rules as follows. You get to choose one of three doors. One door has a car behind it, another has a computer, and the other door has a goat (with all permutations equally likely). Monty, who knows which prize is behind each door, will open a door (but not the one you chose) and then let you choose whether to switch from your current choice to the other unopened door.

Suppose that Monty always opens the door that reveals your less preferred prize out of the two alternatives, e.g., if he is faced with the choice between revealing the goat or the computer, he will reveal the goat. Monty opens a door, revealing a goat. Given this information, should you switch? If you do switch, what is your probability of success in getting the car?

## Solution 9

Denote

- Door of original choice as door 1, door that Monty opens as door 2, door that is left as door 3
- $C_i$  the event that car is behind door  $i$
- $M_i$  the event that Monty opens door  $i$
- $G_i$  the event that goat is behind door  $i$

Using Bayes rule:

$$\mathbb{P}(C_1|M_2, G_2) = \frac{\mathbb{P}(M_2, G_2|C_1)\mathbb{P}(C_1)}{(\mathbb{P}(M_2, G_2|C_1)\mathbb{P}(C_1) + \mathbb{P}(M_2, G_2|C_2)\mathbb{P}(C_2) + \mathbb{P}(M_2, G_2|C_3)\mathbb{P}(C_3)}$$

$$\mathbb{P}(M_2, G_2|C_1) = \mathbb{P}(M_2|G_2, C_1)\mathbb{P}(G_2|C_1) = 1 * 1/2$$

$$\mathbb{P}(M_2, G_2|C_2) = \mathbb{P}(M_2|G_2, C_2)\mathbb{P}(G_2|C_2) = 0$$

$$\mathbb{P}(M_2, G_2|C_3) = \mathbb{P}(M_2|G_2, C_3)\mathbb{P}(G_2|C_3) = 1 * 1/2$$

$$\mathbb{P}(C_1|M_2, G_2) = 1/2 * 1/3 / (2 * 1/2 * 1/3) = 1/2$$

## HW4-5

### Problem 4

(2 points) Consider two independent random variables  $X \sim F_X$  and  $Y \sim F_Y$ . Find the CDF of random variables  $Z_1 = \max(X, Y)$  (it means that for every outcome  $w$  we have  $Z_1(w) = \max(X(w), Y(w))$ , so  $Z_1$  jumps between values of  $X$  and  $Y$ ) and  $Z_2 = \min(X, Y)$  (same reasoning applies).

### Solution 4

$$1. Z_1 = \max(X, Y)$$

$$\begin{aligned} F_{Z_1}(z) &= \mathbb{P}(Z_1 \leq z) = \mathbb{P}(\max(X, Y) \leq z) = \\ &= \mathbb{P}(X \leq z \text{ and } Y \leq z) = \mathbb{P}(X \leq z)\mathbb{P}(Y \leq z) = F_X(z)F_Y(z) \end{aligned}$$

$$2. Z_2 = \min(X, Y)$$

$$\begin{aligned} F_{Z_2}(z) &= \mathbb{P}(Z_2 \leq z) = \mathbb{P}(\min(X, Y) \leq z) = \\ &= \mathbb{P}(X \leq z \text{ or } Y \leq z) = \mathbb{P}(X \leq z) + \mathbb{P}(Y \leq z) - \mathbb{P}(X \leq z \text{ and } Y \leq z) = \\ &= F_X(z) + F_Y(z) - F_X(z)F_Y(z) \end{aligned}$$

## Problem 6

(2 bonus points) A good student solves a problem correctly with probability 0.95, while a bad student — with probability 0.15. What is the minimal number of problems that the test should include so that the probability that good student does not pass the test does not exceed 0.01, and the probability that the bad student passes the test does not exceed 0.1? Passing the test means solving strictly more than half of the problems. You can use Python to solve this problem.

## Solution 6

Let's have  $Bi(n, 0.95)$  describing how many problems a good student solves and  $Bi(n, 0.15)$  describing how many problems a bad student solves. We then need  $n$  such that:

$$\begin{cases} \text{CDF}_{Bi(n, 0.95)}(\lfloor \frac{n}{2} \rfloor) \leq 0.01, \\ \text{SF}_{Bi(n, 0.15)}(\lceil \frac{n}{2} \rceil) \leq 0.1. \end{cases}$$

## Problem 7

(3 bonus points) Two players are playing a game. The first player says a number  $p_1$  between 0 and 1. The second player, knowing the number of the first player, says a number  $p_2$  between 0 and 1. Then, with probability  $p_1$  the first number becomes zero, and with probability  $p_2$  the second number becomes zero. The player whose number is greater, wins.

- (2 bonus points) What is the optimal winning strategy of player two?
- (1 bonus point) What is the optimal winning strategy of player one, if player two follows strategy from the previous point?

1. Denote  $n_i$  the numbers announced by player  $i$ . Let's consider four cases.

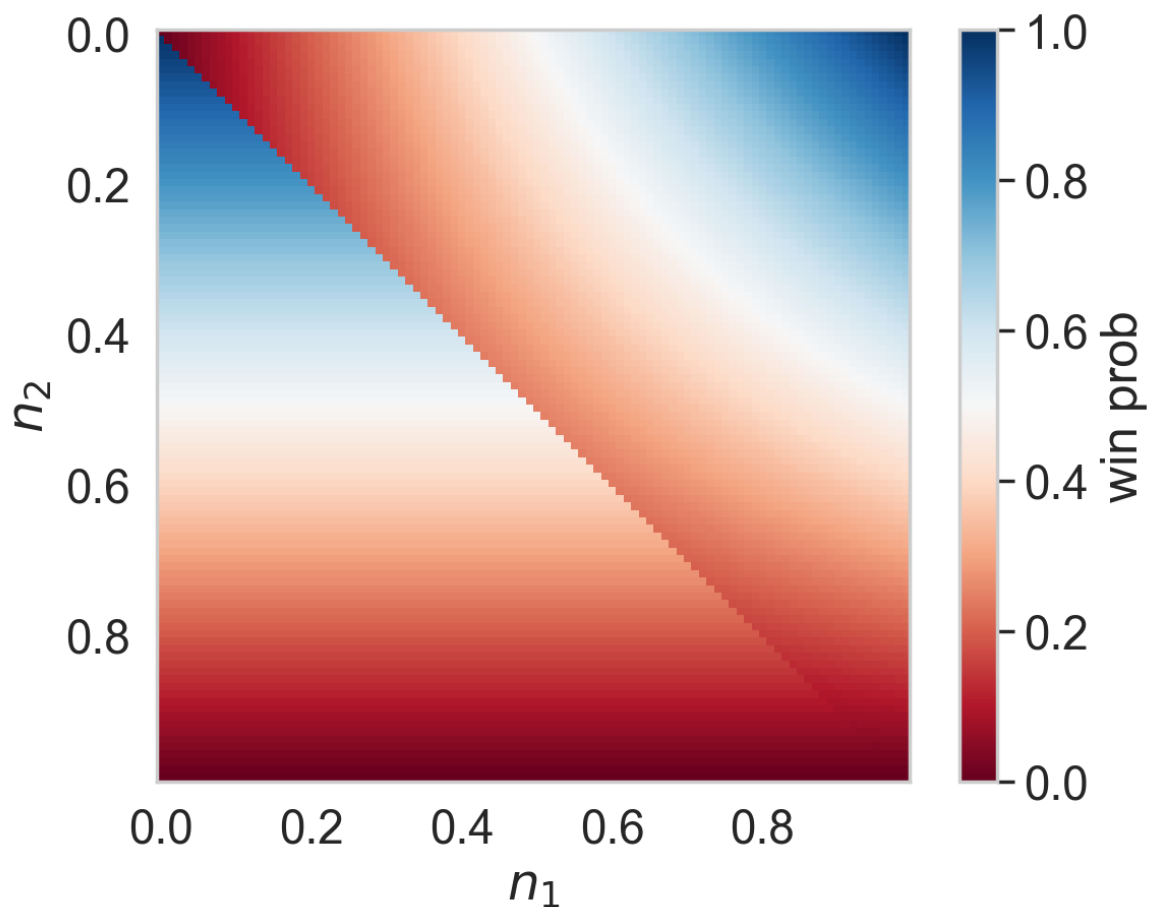
Consider that the first player has announced a number close to 1. If the second player announces a number \textbf{smaller} than the first player, he wins only

when the number of the first player is nullified:

$\mathbb{P}(\text{player 2 wins} | n_2 < n_1) = n_1 \cdot (1 - n_2)$ . If, however, the second player announces a number **larger** than the first player, he wins in both cases (if his number is not nullified):

$\mathbb{P}(\text{player 2 wins} | n_2 > n_1) = n_1 \cdot (1 - n_2) + (1 - n_1) \cdot (1 - n_2) = (1 - n_2)$ , which is **higher**, but gets smaller with  $n_2$ . So, for large  $n_1$ , the best  $n_2$  is just above  $n_1$  or  $n_2 = n_1 + \varepsilon$ .

If the  $n_1$  is small, then the inequality changes sign, so the best  $n_2$  is just above zero or  $n_2 = \varepsilon$ .



2. By looking at the image above, it is obvious that the smallest probability of winning is when  $n_1 = 0.5$ .

## HW6-7

### Problem 1

(1 points) You have  $n$  enumerated letters and  $n$  enumerated envelopes. You randomly put letters into envelopes. What is the expected value of the number of coinciding numbers of the letter and its envelope?

## Solution 1

Denote  $A_i$  the event that one letter coincides with envelope. The probability of this event is

$$\mathbb{P}(A_i) = \frac{1}{n}$$

Denote  $X_i$  the indicator r.v. of event  $A_i$ . Denote  $N$  the number of coincidences, then

$$\mathbb{E}[N] = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}[X_i] = n \frac{1}{n} = 1$$

## Problem 5

(2 points) Consider independent random variables  $X \sim \text{Exp}(\lambda_X)$  and  $Y \sim \text{Exp}(\lambda_Y)$ .

1. Find the CDF of  $Z = \min\{X, Y\}$ .
2. Let  $\lambda_X = \lambda_Y$ . Find the CDF of  $Z = X + Y$ . This will be a (very simple case of) **Gamma distribution**. Hint: use the convolution formula.

## Solution 5

1. Based on HW4-5 Problem 4:

$$\begin{aligned} F_Z(z) &= F_X(z) + F_Y(z) - F_X(z)F_Y(z) = \\ &= 1 - e^{-\lambda_X z} + 1 - e^{-\lambda_Y z} - (1 - e^{-\lambda_X z})(1 - e^{-\lambda_Y z}) = \\ &= 1 - e^{-(\lambda_X + \lambda_Y)z} \end{aligned}$$

2. Let  $\lambda_X = \lambda_Y = \lambda$ . Convolution formula for PDF:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \lambda e^{-\lambda t} \lambda e^{-\lambda(z-t)} dt = \\ &= \lambda^2 e^{-\lambda z} \int_0^z dt = \lambda^2 z e^{-\lambda z} \end{aligned}$$

CDF:

$$F_Z(z) = \int_0^z \lambda^2 x e^{-\lambda x} dx = \lambda^2 \frac{1 - e^{-\lambda z}(\lambda z + 1)}{\lambda^2} = 1 - e^{-\lambda z}(\lambda z + 1)$$

## Problem 6

(4 bonus points) Compute expectation (2 bonus points) and variance (2 bonus points) of Geometric distribution. Hints:

1. Obtain the series for expectation in terms of  $p$  and  $q$ .
2. Compare the series with geometric series in  $q$ , for which we know that it converges, and we know the limit.
3. Differentiate the geometric series. Because the geometric series converges, we can interchange sum and differential operators. Compare the series for expectation with differentiated geometric series.
4. Make the necessary changes and obtain the expectation.
5. Use LOTUS to get series for  $\mathbb{E}[X^2]$ .
6. Compare with differentiated geometric series. Make adjustments, differentiate again.
7. Obtain the variance.

## Solution 6

1. Series for expectation:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k p q^k$$

2. Geometric series (since  $0 < q < 1$ ):

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

3. Differentiate w.r.t.  $q$  (we can do that because the series converges):

$$\sum_{k=0}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$



4. Multiply by  $pq$ :

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k p q^k = \frac{pq}{p^2} = \frac{q}{p}$$

5. Using LOTUS:

$$\mathbb{E}[X^2] = \sum_{k=0}^{\infty} k^2 \mathbb{P}(X = k) = \sum_{k=0}^{\infty} k^2 p q^k$$

6. Differentiated geometric series:

$$\sum_{k=0}^{\infty} k q^{k-1} = \frac{1}{(1-q)^2}$$

Multiply by  $q$ :

$$\sum_{k=1}^{\infty} k q^k = \frac{q}{(1-q)^2}$$

Differentiate again:

$$\sum_{k=1}^{\infty} k^2 q^{k-1} = \frac{1+q}{(1-q)^3}$$

Multiply by  $pq$ :

$$\mathbb{E}[X^2] = \sum_{k=1}^{\infty} k^2 p q^k = pq \frac{1+q}{(1-q)^3} = \frac{q(1+q)}{p^2}$$

7. Obtain variance:

$$\mathbb{V}\text{ar}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{q(1+q)}{p^2} - \frac{q^2}{p^2} = \frac{q}{p^2}$$

## Problem 7

(1 bonus point) Compute expectation and variance of Negative binomial distribution.

Hint: there is a relation between Negative binomial and Geometric distributions.

## Solution 7

The relation is

$$NBi(r, p) = \sum_{k=1}^r Geom(p)$$

Denote  $X \sim NBi(r, p)$  and  $X_k \sim Geom(p), k = 1, \dots, r$ . Then,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{k=1}^r X_k\right] = \sum_{k=1}^r \mathbb{E}[X_k] = \frac{rq}{p}$$

Also because  $X_k$  are independent, they are uncorrelated and the variance is also linear

$$\mathbb{V}\text{ar}(X) = \mathbb{V}\text{ar}\left(\sum_{k=1}^r X_k\right) = \sum_{k=1}^r \mathbb{V}\text{ar}(X_k) = \frac{rq}{p^2}$$