

Quiz 2: Give yourself 50 minutes to solve 5 of the following 6 problems. Each problem weights 20 point scores

1. Let X be a discrete random variable with the following p.m.f

$$P_X(x) = \begin{cases} 0.3, & \text{for } x = 3 \\ 0.2, & \text{for } x = 5 \\ 0.3, & \text{for } x = 8 \\ 0.2, & \text{for } x = 10 \\ 0, & \text{otherwise} \end{cases}$$

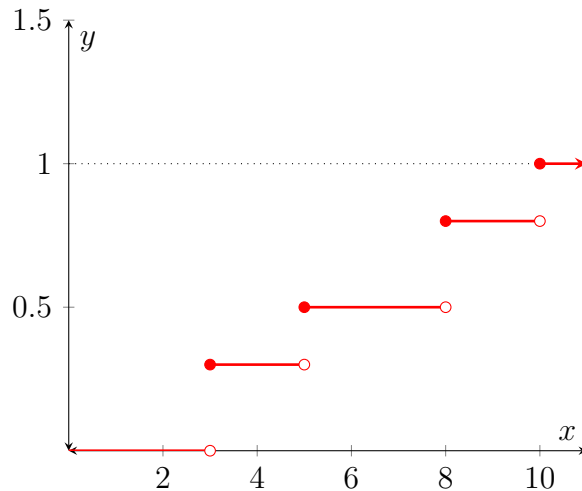
- (a) Find the function $F_X(x) = P(X \leq x)$, for $x \in \mathbb{N}$.

Solution.

$$F_X(x) = \begin{cases} 0, & \text{for } x = 1, 2 \\ 0.3, & \text{for } 3 \leq x < 5 \\ 0.5, & \text{for } 5 \leq x < 8 \\ 0.8, & \text{for } 8 \leq x < 10 \\ 1, & \text{for } x \geq 10 \end{cases}, x \in \mathbb{N}$$

Other equivalent form is permitted, too.

- (b) Plot the function $F_X(x)$.



2. Let $S = \{1, 2, \dots, n\}$ and suppose that A and B are, independently, equally likely to be any of the 2^n subsets (including the null set and S itself) of S .

- (a) Show that $P(A \subset B) = (\frac{3}{4})^n$ Hint: Let $N(B)$ denote the number of elements in B . Use $P(A \subset B) = \sum_{i=0}^n P(A \subset B \mid |B| = i)P(|B| = i)$

Solution.

As suggested in the hint, it is convenient to break $P(A \subset B, |B| = i)$ to the form of conditional probability $P(A \subset B \mid |B| = i)P(|B| = i)$.

The probability of $B \subset S$ having size i , $P(|B| = i)$, is $\binom{n}{i} / 2^n$.

The probability of $A \subset S$ being a subset of B given $|B| = i$, $P(A \subset B | |B| = i)$, is $\frac{2^i}{2^n} = 2^{i-n}$

Hence, the answer is

$$\sum_{i=0}^n P(A \subset B | |B| = i) P(|B| = i) = 4^{-n} \underbrace{\sum_{i=0}^n \binom{n}{i} 2^i}_{(1+2)^n} = \left(\frac{3}{4}\right)^n$$

Another Solution. You can also do it without using conditional probability. Observe that the elements of S either in A , (which is also $A \cap B$), $B - A$, or neither. Therefore, we can think of the number of desired event as placing n distinct balls to 3 distinct buckets, which is 3^n . Divide the number of possible A, B combination and we get the answer $(\frac{3}{4})^n$.

(b) Show that $P(A \cap B = \emptyset) = (\frac{3}{4})^n$

Solution.

$$\begin{aligned} P(A \cap B = \emptyset) &= \sum_{i=0}^n P(A \subset B^c | |B| = i) P(|B| = i) \\ &= \frac{1}{4^n} \underbrace{\sum_{i=0}^n 2^{n-i} \binom{n}{i}}_{(2+1)^n} \\ &= \left(\frac{3}{4}\right)^n \end{aligned}$$

□

3. When you bet on head in coin tossing, your chance of winning is $\frac{1}{2}$. Suppose also that bets over \$ 250 are not allowed. You decide to play the following strategy: you start with a \$1 bet and double the bet until either you win (the same amount as the bet) or the bet exceeds \$250; then you start again with a \$1 bet and repeat. We will call this sequence of bets in the strategy until restart with a \$1 bet ‘one round’. If you play 1000 rounds of this strategy

(a) How many rounds you have to win at least in 1000 rounds so that net revenue ≥ 0 .

Solution.

Suppose in the i -th round, $1 \leq i \leq 1000$, the money you’ll get if end up winning at the n -th bet is

$$-1 - 2 - \dots - 2^{n-2} + 2^{n-1} = 2^{n-1} - (2^{n-1} - 1) = 1,$$

which is independent to n . Otherwise, if you end up lose all the bets, then the money you'll earn is

$$-1 - 2 \cdots - 2^7 = -(2^8 - 1) = -255.$$

Note that the maximum number of bets is 8 times since the amount on the 9th bet is $2^{9-1} = 256 > 250$.

Therefore, suppose you want to earn some money in 1000 rounds, the number of rounds you win/lose, name W and L , satisfies

$$\begin{cases} L + W = 1000 \\ -255L + 1W \geq 0. \end{cases}$$

Solving the system of equation you will get $W \geq 996.09 \cdots \implies W \geq 997$
 \square

- (b) Write down the expression of the probability that net revenue ≥ 0 .

Solution. Clearly, you'll lose \$255 with probability $q = (1 - p)^8$. Therefore,

$$\mathbb{P}(\text{ends up getting 1 dollar in a round}) = 1 - q.$$

The number of loses is the sum of 1000 Bernoulli random variable with probability q . Given the answer in (a), we have win at least 997 times, i.e lose less than 4 times. Hence, the probability is

$$\sum_{i=0}^3 \binom{1000}{i} q^i (1 - q)^{1000-i}$$

\square

4. (a) For what value of C do the function $C2^x/x!$ define a probability mass functions on $1, 2, 3, \dots$?

Solution.

If a function f with countable range S meets the following conditions

- non-negative, $f(x) \geq 0, x \in S$
- sum to 1, $\sum_{x \in S} f(x) = 1$

we call it a p.m.f

Here, the first condition is trivial, and in order for $C2^x/x!$ to be a p.m.f,

$$C \underbrace{\sum_{x \in \mathbb{N}} 2^x/x!}_{e^2-1} = 1$$

Hence, $C = \frac{1}{e^2-1}$

□

- (b) If X is geometric show that $P(X = n + k | X > n) = P(X = k)$, for $k, n \geq 1$.

Solution

We call Geometric random variable **memoryless** because of this. If X follows $\text{Geo}(p)$,

$$\begin{aligned} P(X = n + k | X > n) &= \frac{P(X = n + k, X > n)}{P(X > n)} = \frac{p(1-p)^{n+k-1}}{p(1-p)^{n\frac{1}{p}}} \\ &= p(1-p)^{k-1} \\ &= P(X = k) \end{aligned}$$

□

5. (a) Let X and Y be discrete random variables with the joint probability mass function $f(x, y)$. Assume that $f(x, y) = g(x)k(y)$ for some probability mass functions $g(x)$ and $k(y)$. Prove that X and Y are independent.

Solution.

By the definition in the handout, a collection of random variables X, Y is called independent if for any x in the range of X (call it S_X), and y in the range of Y (call it S_Y) we have

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

It will be great to have $g(x)$ as p.m.f of X , and $h(y)$ as p.m.f of Y . Because if this is true, we have

$$P(X = x, Y = y) = f(x, y) = g(x)h(y) = P(X = x)P(Y = y),$$

To confirm the guess, take any $x \in S_X$

$$P(X = x) = \sum_{y \in S_Y} P(X = x, Y = y) = g(x) \underbrace{\sum_{y \in S_Y} h(y)}_{=1} = g(x)$$

The same goes for $y \in S_Y$, the guess is correct that $g(x), h(y)$ is the p.m.f of X, Y , respectively. □

- (b) In a sequence of tosses of a p-coin, let X be the number of tosses required to get the first head and Y be the number of tosses to get the second head after getting the first head. Show that X and Y are independent $\text{Geo}(p)$ random variables.

Solution.

6. (De Moivre trials) Each trial may result in any of t given outcomes (t is a fixed positive integer). The i th outcomes having probability p_i , and $\sum_{i=1}^t p_i = 1$. Let the

number of occurrences of the i th outcome in n independent trials. The probability is as the following

$$P(N_i = n_i, \text{ for all } 1 \leq i \leq t) = \frac{n!}{n_1!n_2!\dots n_t!} p_1^{n_1} p_2^{n_2} \dots p_t^{n_t}$$

for any collection n_1, n_2, \dots, n_t of non-negative integers with sum n . The vector N is said to have the multinomial distribution. Calculate the marginal probability of N_1 , i.e calculate $P(N_1 = n_1), n_1 = 0, 1, \dots, n$.

Solution. The question seems to be a bit complicated, but the marginal is actually a binomial. TBD

$$\begin{aligned} P(N_1 = n_1) &= \sum_{\substack{n_2, n_3, \dots, n_t \\ n_1 + n_2 + \dots + n_t = n}} \frac{n!}{n_1!n_2!\dots n_t!} p_1^{n_1} p_2^{n_2} \dots p_t^{n_t} \\ &= \frac{n!}{n_1!(n - n_1)!} p_1^{n_1} \sum_{\substack{n_2, n_3, \dots, n_t \\ n_1 + n_2 + \dots + n_t = n}} \frac{(n - n_1)!}{n_2! \dots n_t!} p_2^{n_2} \dots p_t^{n_t} \\ &= \frac{n!}{n_1!(n - n_1)!} p_1^{n_1} \\ &\quad \cdot \sum_{n_2=0}^{n-n_1} \frac{(n - n_1)!}{n_2!(n - n_1 - n_2)!} p_2^{n_2} \sum_{n_3=0}^{n-n_1-n_2} \dots \underbrace{\sum_{n_{t-1}=0}^{n-n_1-\dots-n_{t-2}} \frac{(n - n_1 - \dots - n_{t-2})!}{n_{t-1}!n_t!} p_{t-1}^{n_{t-1}} p_t^{n_t}}_{(p_{t-1} + p_t)^{n-n_1-\dots-n_{t-2}}} \\ &= \frac{n!}{n_1!(n - n_1)!} p_1^{n_1} \\ &\quad \cdot \sum_{n_2=0}^{n-n_1} \frac{(n - n_1)!}{n_2!(\textcolor{red}{n} - \textcolor{red}{n}_1 - \textcolor{red}{n}_2)!} p_2^{n_2} (p_3 + p_4 + \dots + p_t)^{\textcolor{red}{n} - \textcolor{red}{n}_1 - \textcolor{red}{n}_2} \\ &= \frac{n!}{n_1!(n - n_1)!} p_1^{n_1} \underbrace{(p_2 + \dots + p_t)^{n-n_1}}_{(1-p_1)^{n-n_1}} = \binom{n}{n_1} p_1^{n_1} (1 - p_1)^{n-n_1} \end{aligned}$$

Hence, the marginal of N_1 follows binomial distribution \square