

Probability & Statistics for EECS:

Homework #05

Due on March 19, 2023 at 23:59

Name: **Zhou Shouchen**
Student ID: 2021533042

Problem 1

(a) Let X be the number of questions until Kratos is sure about the location of the treasure.

So when $X = 1$, Kratos asks the correct location in one time, so $P(X = 1) = \frac{1}{9}$.

When $X = 2$, Kratos asks the wrong position at the first time, and asks the correct location in the second time, so $P(X = 2) = \frac{8}{9} \cdot \frac{1}{8} = \frac{1}{9}$.

Similarly, $P(X = 3) = P(X = 4) = P(X = 5) = P(X = 6) = P(X = 7) = \frac{1}{9}$.

As for $X = 8$, it is special since if Kratos wrongly asks the location for 8 times, he can still get the correct location with using exclusion method, because there are totally 9 realms. So $P(X = 8) = \frac{2}{9}$.

$$\text{So } E(X) = \sum_{x=1}^8 xP(X = x) = \sum_{x=1}^7 x \cdot \frac{1}{9} + 8 \cdot \frac{2}{9} = \frac{44}{9}.$$

So above all, the expected number of questions until Kratos is sure about the location is $\frac{44}{9}$.

(b) Let X be the number of questions until Kratos is sure about the location of the treasure.

Since Kratos eliminates as close to half of the remaining realms as possible, for each subsequence, we eliminate the $\lfloor \frac{\text{length}}{2} \rfloor$, where "length" is the length of the subsequence.

For example, firstly the subsequence is $[1, 9]$, we ask if it is in a realm numbered less than or equal to $1 + \lfloor \frac{9}{2} \rfloor - 1 = 4$.

If so, the second time the subsequence is $[1, 4]$, we will ask if it is in a realm numbered less than or equal to $1 + \lfloor \frac{4}{2} \rfloor - 1 = 2$.

Otherwise, the second time the subsequence is $[5, 9]$, we will ask if it is in a realm numbered less than or equal to $5 + \lfloor \frac{4}{2} \rfloor - 1 = 6$.

And we will continuously do this until the length of the subsequence is equal to 1, then we will be sure that it is in that realm.

So to conclude, if it is in realm 1, 2, 3, 4, 5, 6, 9, then we will ask 3 times.

If it is in realm 7, 8, then we will ask 4 times.

$$\text{So } P(X = 3) = \frac{7}{9}, P(X = 4) = \frac{2}{9}.$$

$$\text{So } E(X) = 3 \cdot \frac{7}{9} + 4 \cdot \frac{2}{9} = \frac{29}{9}.$$

So above all, the expected number of questions until Kratos is sure about the location is $\frac{29}{9}$.

Problem 2

Let N_1 : # of days until first new type of steak appears.

Let N_2 : # of days until the second new type of steak appears.

Let N_3 : # of days until the third new type of steak appears.

Let N_4 : # of days until the fourth new type of steak appears.

Let N_5 : # of days until the fifth new type of steak appears.

So the number of days that we need to watch is $N = N_1 + N_2 + N_3 + N_4 + N_5$.

From the description of the problem, we can get that

$$N_1 = 1, N_2 \sim FS(\frac{4}{5}), N_3 \sim FS(\frac{3}{5}), N_4 \sim FS(\frac{2}{5}), N_5 \sim FS(\frac{1}{5}).$$

$$\text{So } E(N_1) = 1, E(N_2) = \frac{5}{4}, E(N_3) = \frac{5}{3}, E(N_4) = \frac{5}{2}, E(N_5) = \frac{5}{1}.$$

$$\text{So } E(N) = E(N_1 + N_2 + N_3 + N_4 + N_5) = E(N_1) + E(N_2) + E(N_3) + E(N_4) + E(N_5) = 1 + \frac{5}{4} + \frac{5}{3} + \frac{5}{2} + \frac{5}{1} = \frac{137}{12}.$$

So above all, the expectation of the days that we expect to watch is $\frac{137}{12}$.

Problem 3

(a) Let Z be the number of that the first time they simultaneously successful.

Since $X_i \sim \text{Bern}(p_1), Y_i \sim \text{Bern}(p_2)$,

so for one time, they have the possibility of $p_1 p_2$ to success.

So $Z \sim \text{FS}(p_1 p_2)$, and $P(Z = k) = (1 - p_1 p_2)^{k-1} (p_1 p_2)$, and the support is $\{1, 2, \dots\}$.

And from what we have learned, the expectation of $Z \sim \text{FS}(p_1 p_2)$ is $E(Z) = \frac{1}{p_1 p_2}$.

So above all, the distribution of Z is $Z \sim \text{FS}(p_1 p_2)$, and the expectation is $E(Z) = \frac{1}{p_1 p_2}$.

(b) Let A be the first time that at least one has a success.

And the probability that at least one success is that $1 - (1 - p_1)(1 - p_2) = p_1 + p_2 - p_1 p_2$.

So we can say that $A \sim \text{FS}(p_1 + p_2 - p_1 p_2)$.

And $P(A = k) = (1 - p_1 - p_2 + p_1 p_2)^{k-1} (p_1 + p_2 - p_1 p_2)$.

And from what we have learned in class, we could get that the expectation of A is that $E(A) = \frac{1}{p_1 + p_2 - p_1 p_2}$.

So above all, the expectation time until at least one has a success is $E(A) = \frac{1}{p_1 + p_2 - p_1 p_2}$.

(c) Let $p = p_1 = p_2$.

Let B: the first time that Mario is successful.

Let C: the first time that Zelda is successful.

D: their first successes are simultaneous.

So $B \sim \text{FS}(p), Y \sim \text{FS}(p)$. And since B and C are independent, and $P(B = k) = (1 - p)^{k-1} p$, $P(C = k) = (1 - p)^{k-1} p$.

So $P(D) = P(B = C) = \sum_{k=1}^{\infty} P(B = k, C = k) = \sum_{k=1}^{\infty} P(B = k) P(C = k) = \sum_{k=1}^{\infty} (1 - p)^{k-1} p (1 - p)^{k-1} p$
 $= \sum_{k=0}^{\infty} ((1 - p)^2)^k p^2 = p^2 \cdot \frac{1}{1 - (1 - p)^2} = \frac{p^2}{p(2 - p)} = \frac{p}{2 - p}$.

So the probability that their first time success are simultaneous is $\frac{p}{2 - p}$.

And there are total three situations,

1. Mario get the first success earlier than Zelda, i.e. $B < C$
2. Zelda get the first success earlier than Mario, i.e. $C < B$
3. Mario and Zelda get the first success simultaneously i.e. $B = C$ and it is exactly D that we mentioned above.

Since there are total three situations, so $P(B < C) + P(C < B) + P(B = C) = 1$.

Since $p_1 = p_2 = p$, so with symmetry we can get that $P(B < C) = P(C < B)$.

So $2P(B < C) + P(D) = 1$, i.e. $P(B < C) = \frac{1}{2}(1 - P(D)) = \frac{1}{2}(1 - \frac{p}{2 - p}) = \frac{1}{2} \frac{2 - 2p}{2 - p} = \frac{1 - p}{2 - p}$.

So the probability that Mario's first success precedes Zelda's is $\frac{1 - p}{2 - p}$.

So above all, the probability that their first time success are simultaneous is $\frac{p}{2 - p}$.

The probability that Mario's first success precedes Zelda's is $\frac{1 - p}{2 - p}$.

Problem 4

$X \sim \text{Geom}(p), Y \sim \text{Geom}(q)$.

So $P(X = k) = (1-p)^k p, P(Y = k) = (1-q)^k q$.

$$(a) P(X = Y) = \sum_{k=0}^{\infty} P(X = Y = k) = \sum_{k=0}^{\infty} P(X = k)P(Y = k)$$

$$= \sum_{k=0}^{\infty} (1-p)^k p (1-q)^k q = pq \sum_{k=0}^{\infty} [(1-p)(1-q)]^k$$

since $p \in (0, 1), q \in (0, 1)$, so $(1-p)(1-q) \in (0, 1)$.

$$\text{So } P(X = Y) = pq \cdot \frac{1}{1 - (1-p)(1-q)} = \frac{pq}{p + q - pq}.$$

So above all, the probability that $X = Y$ is $P(X = Y) = \frac{pq}{p + q - pq}$.

(b) Let $F(z)$ be the CDF of $\max(X, Y)$, so $F(z) = \sum P(\max(X, Y) \leq z) = P(X \leq z, Y \leq z)$

Since X, Y are independent,

$$\text{so } F(z) = P(X \leq z)P(Y \leq z) = \left(\sum_{k=0}^z (1-p)^k p\right) \left(\sum_{k=0}^z (1-q)^k q\right) = [1 - (1-p)^{z+1}][1 - (1-q)^{z+1}].$$

And the Survival Function $G(z) = 1 - F(z) = (1-p)^{z+1} + (1-q)^{z+1} - [(1-p)(1-q)]^{z+1}$.

$$\text{So } E(\max(X, Y)) = \sum_{z=0}^{\infty} G(z) = \frac{1-p}{p} + \frac{1-q}{q} - \frac{(1-p)(1-q)}{p + q - pq}.$$

$$\text{So above all, } E(\max(X, Y)) = \frac{1-p}{p} + \frac{1-q}{q} - \frac{(1-p)(1-q)}{p + q - pq}.$$

$$(c) P(\min(X, Y) = k) = \sum_{i=k}^{\infty} P(X = k, Y = i) + \sum_{i=k+1}^{\infty} P(X = i, Y = k)$$

Since X, Y are independent, so $P(\min(X, Y) = k) = (1-p)^k p \sum_{i=k}^{\infty} (1-q)^i q + (1-q)^k q \sum_{i=k+1}^{\infty} (1-p)^i p$

$$= (1-p)^k p q \frac{(1-q)^k}{q} + (1-q)^k q p \frac{(1-p)^{k+1}}{p} = (1-p)^k (1-q)^k p + (1-p)^{k+1} (1-q)^k q$$

$$= (1-p)^k (1-q)^k (p + q - pq).$$

So above all, $P(\min(X, Y) = k) = (1-p)^k (1-q)^k (p + q - pq)$.

(d) With the definition of the conditional probability, we have

$$P(X = k | X \leq Y) = \frac{P(X = k, X \leq Y)}{P(X \leq Y)} = \frac{P(X = k)P(Y \geq k)}{P(Y \geq X)}$$

With the LOPT, we have $P(Y \geq X) = \sum_{i=0}^{\infty} P(Y \geq X | X = i)P(X = i) = \sum_{i=0}^{\infty} P(Y \geq i)P(X = i)$

$$\text{And } P(Y \geq k) = \sum_{i=k}^{\infty} (1-q)^i q = q \cdot \frac{(1-q)^k}{1 - (1-q)} = (1-q)^k.$$

$$\text{So } P(Y \geq X) = \sum_{i=0}^{\infty} (1-q)^i (1-p)^i p = p \cdot \frac{1}{1 - (1-p)(1-q)}.$$

$$\text{So } P(X = k | X \leq Y) = \frac{P(X = k)P(Y \geq k)}{P(Y \geq X)} = \frac{(1-p)^k p (1-q)^k}{p \cdot \frac{1}{1 - (1-p)(1-q)}} = (1-p)^k (1-q)^k (p + q - pq).$$

$$\text{So } E[X | X \leq Y] = \sum_{k=0}^{\infty} k \cdot P(X = k | X \leq Y) = (p + q - pq) \sum_{k=0}^{\infty} k [(1-p)(1-q)]^k.$$

Let's construct $S = \sum_{k=0}^{\infty} k q^k$

then $q \cdot S = \sum_{k=1}^{\infty} (k-1)q^k$.

minus the two equation, we can get $(1-q)S = q + q^2 + q^3 + \cdots = \frac{q}{1-q}$.

So $S = \frac{q}{(1-q)^2}$.

With this, we can get that $\sum_{k=0}^{\infty} k[(1-p)(1-q)]^k = \frac{(1-p)(1-q)}{[1 - (1-p)(1-q)]^2}$.

So $E[X|X \leq Y] = (p+q-pq) \cdot \frac{(1-p)(1-q)}{(p+q-pq)^2} = \frac{(1-p)(1-q)}{p+q-pq}$.

So above all, $E[X|X \leq Y] = \frac{(1-p)(1-q)}{p+q-pq}$.

Problem 5

(a) Let A_i be the rank that the i -th dishes we have ordered in the exploration phase.

Let A be the sum of the ranks that all dishes we have eaten.

so $A = A_1 + A_2 + \cdots + A_k + (m - k) \cdot X$. Since during the exploration phase, all food we will only ordered once, so there must exist a single j , s.t. $A_j = X$.

$$\text{So } A = \sum_{i=1, i \neq j}^k A_i + A_j + (m - k) \cdot X = \sum_{i=1, i \neq j}^k A_i + (m - k + 1) \cdot X.$$

$$\text{So for all } i \neq j, \text{ we have } 1 \leq A_i < X, \text{ so } E(A_i) = \frac{1}{E[X] - 1} \cdot (1 + 2 + \cdots + (E[X] - 1)) = \frac{E[X]}{2}.$$

$$\begin{aligned} \text{So } E(A) &= E\left(\sum_{i=1, i \neq j}^k A_i + (m - k + 1) \cdot X\right) = \sum_{i=1, i \neq j}^k E(A_i) + (m - k + 1) \cdot E[X] \\ &= (k - 1) \cdot \frac{E[X]}{2} + (m - k + 1)E[X] = \left(\frac{k + 1}{2} + m - k\right)E[X]. \end{aligned}$$

$$\text{So above all, } E(A) = \left(\frac{k + 1}{2} + m - k\right)E[X].$$

(b) Since we have eaten k dishes at the exploration phase, so the smallest rank we will get is k .

So for a that $1 \leq a < k$, $P(X = a) = 0$.

As for a that $k \leq a \leq n$, since a is the highest rank we get, and we do not eat two same dishes at the exploration phase. So other $k - 1$ dishes have the rank in $\{1, 2, \dots, a - 1\}$, so the number of combination that the dishes with the highest rank a is that $\binom{a-1}{k-1}$, and we can arrange them in different order, which has $k!$ ways. So the number of total number that the dishes with highest rank of a is $k! \binom{a-1}{k-1}$.

As for the number of combinations of all possible choices, we just need to choose k dishes from all n kinds of dishes, which is $\binom{n}{k}$. And we can arrange them in different order, which has $k!$ ways. So the number of all possible choices is $k! \binom{n}{k}$.

$$\text{So the PMF of } X \text{ is that } P(X = a) = \frac{\# \text{ dishes with highest rank of } a}{\# \text{ all possible choices}} = \frac{k! \binom{a-1}{k-1}}{k! \binom{n}{k}} = \frac{\binom{a-1}{k-1}}{\binom{n}{k}}.$$

So above all, the PMF of X is that

$$P(X = a) = \frac{\binom{a-1}{k-1}}{\binom{n}{k}}, \text{ if } k \leq a \leq n,$$

$$P(X = a) = 0, \text{ otherwise.}$$

(c) With the definition of expectation,

$$E[X] = \sum_{a=k}^n a \cdot P(X = a) = \sum_{a=k}^n a \cdot \frac{\binom{a-1}{k-1}}{\binom{n}{k}} = \frac{1}{\binom{n}{k}} \cdot \sum_{a=k}^n a \binom{a-1}{k-1} = \frac{k}{\binom{n}{k}} \cdot \sum_{a=k}^n \binom{a}{k}$$

We can get a lemma that $\sum_{a=k}^n \binom{a}{k} = \binom{n+1}{k+1}$ with the story proof:

Consider a situation that there are a group of $n + 1$ people, and we want to choose $k + 1$ people from the group.

For the left-hand side, let each person has a unique number $1, 2, \dots, n + 1$. So we can discuss about the biggest number of people in the subgroup.

If the biggest number in the subgroup is a , then for this situation, the number total ways to get the subgroup is to choose n people from the number of $\{1, 2, \dots, a-1\}$, i.e. the number of ways is $\binom{a-1}{n}$. Since the subgroup has total $k+1$ people, so $a \geq k+1$. So the number of total ways is that $\sum_{a=k+1}^n \binom{a-1}{n} = \sum_{a=k}^n \binom{a}{n}$

For the right-hand side, we just directly choose the $k+1$ people from all $n+1$ people, so the number of total ways is $\binom{n+1}{k+1}$.

So $\sum_{a=k}^n \binom{a}{n} = \binom{n+1}{k+1}$ has been proved.

Combine the lemma with the $E[X]$ that we have reduced above, we can get that

$$E[X] = \frac{k}{\binom{n}{k}} \cdot \binom{n+1}{k+1} = \frac{k}{\frac{n!}{k!(n-k)!}} \frac{(n+1)!}{(k+1)!((n+1)-(k+1))!} = \frac{k(n+1)}{k+1}$$

So above all, $E[X] = \frac{k(n+1)}{k+1}$.

(d) Combine (a) and (c) we can get that $E(A) = (\frac{k+1}{2} + m - k) \frac{k(n+1)}{k+1}$

To get the optimal value of k , we just need to let the derivatives of $E(A)$ be 0.

So let $f(k) = E(A) = (\frac{k+1}{2} + m - k) \frac{k(n+1)}{k+1}$, then $f'(k) = (n+1)(\frac{1}{2} - \frac{m+k^2+2k}{(k+1)^2})$.

Let $f'(k) = 0$, since $n > 0$, so $n+1 \neq 0$, so $\frac{1}{2} - \frac{m+k^2+2k}{(k+1)^2} = 0$, simplify it we can get $(k+1)^2 = 2(m+1)$.

So $k = -1 \pm \sqrt{2(m+1)}$.

Since $m \geq 1$, so $\sqrt{2(m+1)} \geq 2$, since $k > 0$, so we can only take $k = -1 + \sqrt{2(m+1)}$.

And take it into verify, $k = \sqrt{2(m+1)} - 1$ is the optimal value.

So above all, the optimal value is $k = \sqrt{2(m+1)} - 1$.