Math 505A Homework 6 Neel Gupta

1) a) Given $P(X=(k_1, \dots k_{b+1})) = \frac{b!\, r!}{n!} = \frac{1}{\binom{n}{r}}$, find ~~P(~~ $E(X_1) \dots E(X_{b+1})$

Given $k_1 + \dots + k_{b+1} = r$, & $X_1 = k_1$,

So ~~$E(X_1) = E(k_1)$~~

$E(X_1 + X_2 + \dots X_{b+1}) = E(r)$

$\because k_1 + \dots k_{b+1} = r$, and

$E(X_1) \dots E(X_{b+1})$ are symmetric, so they equal eachother, so

$\Rightarrow E(X_1) + E(X_2) + \dots E(X_{b+1}) = E(r)$

Due to symmetry,

$\underbrace{E(X_1) + \dots E(X_1)}_{b+1} = E(r)$

$(b+1)\, E(X_1) = r \quad \because E(r) = r.$

$E(X_1) = \frac{r}{b+1}$, due to symmetry

$\boxed{E(X_1), \dots, E(X_{b+1}) = \frac{r}{b+1}}$

b) Given $Y_1 = X_1 + 1$ and symmetry of pmfs

$E(Y_1) = E(X_1 + 1) \Rightarrow$ by symmetry

$E(Y_1) = E(X_1) + E(1)$

$E(Y_1) = \frac{r}{b+1} + 1$ by part (a)

$E(Y_1) = \frac{r}{b+1} + \frac{b+1}{b+1} = \boxed{\frac{r+b+1}{b+1}} = E(Y_i)$

b) $Y_2 = x_2 + x_1 + 1 = 2x_1 + 1$

$Y_i = i(x) + 1 \Rightarrow E(Y_i) = (i)E(x_i) + 1$

$= \frac{i(r)}{(b+1)} + \frac{b+1}{b+1} = \frac{ir + b + 1}{\cancel{bi + i + b + 1}}$

$\Rightarrow \frac{ir + b + 1}{\cancel{(1+i)(b+1)}} = \frac{ir + b + 1}{\cancel{(b+1)(i+1)}} = E(Y_i)$

$\cancel{Y_2 = \frac{2r + b + 1}{2b + 2 + b + 1}} \Rightarrow E(Y_i) = \frac{b + 1 + (r)(i)}{b+1}$

When $i$ ranges from 0 to b

$$\boxed{E(Y_b) = \frac{b + 1 + rb}{b+1}}$$

Problem 2

c) Find pmf of $X_1$ and $Y_1$.

$$E(X_1) = \frac{r}{b+1} \;\&\; E(Y_1) = \frac{r+b+1}{b+1}$$

Pmf of all $x$ is given by

$$P(X_1 = k_1, \dots, X_{b+1} = k_{b+1}) = f(k_1, \dots k_{b+1})$$

$$= \frac{b!\, r!}{n!}, \text{ so } P(X_1 = x) = f(k_1)$$

$$Y_1 = X_1 + 1 \implies X_1 = Y_1 - 1$$

because $X_1$ is binomial

$$E(Y_1) = \sum_{y=1}^{b+1} y\, P(Y = y) \implies \text{pmf } Y_1 \text{ is binomial } \left(n, \frac{br}{b+1}\right)$$

Math 505A HW 6 Neel Gupta

2) A biased coin is tossed $n$ times with $P(H) = p$. Let $I_j = \begin{cases} 1 & \text{if } j\text{th} \text{ \& } (j+1)\text{th are diff.} \\ 0 & \text{otherwise} \end{cases}$

When $I_j$ is 1, the # of runs increases.

When you have 1 toss, $R = 1$.

From $j=0$, to $j = n-1$, the indicator will tell us $R-1$, since it doesn't count the case of single toss that is a single run.

Let $R$ be # of runs, then

$$R = 1 + \sum_{j=0}^{n-1} I_j \text{, then}$$

$$E(R) = E\left[1 + \sum_{j=0}^{n-1} I_j\right] = 1 + \sum_{j=0}^{n-1} E(I_j)$$

So $E(I_j) = \sum_{j\geq 0} P(I_j)_j = \underset{H}{p}(1-\underset{T}{p}) + \underset{T}{(1-p)}\underset{H}{p}$

$= 2p(1-p)$ is the probability that $j$th and $j-1$th toss differ, so

$$E(R) = 1 + \sum_{j=0}^{n-1} 2p(1-p) = 2(n-1)p(1-p)+1,$$

So the expected # of runs is $\underline{1+2(n-1)pq}$.

Then the $Var(R) = V\left(\sum_{j=0}^{n-1} I_j\right)$

$$= \sum_{j=0}^{n-1} V(I_j) + 2\sum_{j\neq k} cov(I_j, I_k)$$

Since the successive indicator $I_k = I_{j+1}$, the sum can be reduced down to

$$V(R) = \sum_{j=0}^{n-1} V(I_j) + 2\sum_{j=0}^{n-2} cov(I_j, I_{j+1})$$

$$cov(I_j, I_{j+1}) = E(I_j I_{j+1}) - E(I_j)E(I_{j+1})$$

Since $I_j$ & $I_{j+1}$ are dependent.

$$E(I_j I_{j+1}) = P(\underset{pqp}{HTH}) + P(\underset{qpq}{THT})$$

$$= p^2q + pq^2 = pq(p+q) = pq(1) = pq$$

Given $E(I_j) = 2pq$ and $E(I_j I_{j+1}) = pq$.

Then $cov(I_j, I_{j+1}) = pq - 2pq \cdot 2pq$

$= pq - 4p^2q^2 = pq(1-4pq)$. (1)

$E(I_j^2) - (E(I_j))^2 = V(I_j)$

$V(I_j) = 2pq - (2pq)^2$

$= 2pq - 4p^2q^2 = 2pq(1-2pq)$ (2)

Substituting (1) & (2) back into $V(R)$.

$$V(R) = \sum_{j=0}^{n-1} 2pq(1-2pq) + 2\sum_{j=0}^{n-2} pq(1-4pq)$$

$= 2(n-1)pq(1-2pq) + 2(n-2)pq(1-4pq)$

$= (2n-2)(pq - 2p^2q^2) + (2n-4)(pq - 4p^2q^2)$

$= 2npq - 4np^2q^2 - 2pq + 4p^2q^2 + 2npq - 8np^2q^2$
$- 4pq + 16p^2q^2$

$\Rightarrow 4npq - 12np^2q^2 - 6pq + 20p^2q^2$

$= 2pq(2n - 6npq - 3 + 10pq) = V(R)$

3) Let $X$ := # of floors that someone wants to get on.

Let $I_j$ be an indicator s.t.

$$I_j = \begin{cases} 1 & \text{if someone gets out on } j^{th} \text{ floor} \\ 0 & \text{otherwise} \end{cases}$$

Then, $X = \sum_{j=1}^{10} I_j$, so $E(X) = E\left[\sum_{j=1}^{10} I_j\right]$.

$= \sum_{j=1}^{10} E(I_j)$, so $E(I_j) = P(\text{"someone gets off"})$

Given 20 people and 10 floors,

$P(\text{"someone gets off"}) = 1 - P(\text{"no one gets off"})$

$P(\text{"no one gets off"}) = \left(\frac{9}{10}\right)^{20}$.

3) So $E(X) = \sum_{j=1}^{10} \left[1 - \left(\frac{9}{10}\right)^{20}\right]$, so

$$E(X) = 10 \times \left(1 - \left(\frac{9}{10}\right)^{20}\right) = 10 \times .87842$$

$$= \boxed{8.784 \text{ floors}}$$

The expected # of floors the elevator stops on is $\sim 9$.

$$Var(X) = \underset{I}{E(X^2)} - (E(X))^2$$

$$E(X^2) = \sum_{j=1}^{10} E(I_j^2) + 2 \sum_{j=1}^{10} \sum_{k \geq j+1}^{10} E(I_j X_k)$$

Let $X_j = I_j$ and be a Bernoulli indicator variable, then $E(I_j^2) = E(I_j)$ over all $j$. Since $k = j+1$,

$$2 \sum_{j=1}^{10} \sum_{k=j+1}^{10} E(I_j I_{j+1}) = 2 \sum_{j=0}^{10-1} cov(I_j, I_{j+1})$$

$$\Rightarrow V(X) = \sum_{j=1}^{10} E(I_j) + 2 \sum_{j=1}^{9} cov(I_j, I_{j+1})$$

~~$cov(I_j, I_{j+1}) = E(I_j I_{j+1}) - E(I_j)E(I_{j+1})$~~

$$= n\left(1 - \left(\frac{n-1}{n}\right)^m\right) + (n)(n-1)\left(1 - \left(\frac{n-2}{n}\right)\right)^m - E(X)^2$$

$$= 10\left(1 - \left(\frac{9}{10}\right)^{20}\right) + (10)(9)\left(1 - \left(\frac{8}{10}\right)^{20}\right)$$

$$- (8.784)^2 = \boxed{20.5879}$$

$$Var(X) = 20.588$$

4) There are $n$ different types of objects. You buy 1 each day.

a) Let $X_j$ be # of days ~~which~~ between getting the $j$th new object and the ~~$(j+1)$~~ $(j+1)$th new item. Recognise the distribution and Find $E(X_j)$.

4) a) given $n$ objects and $j$ distinct objects out of $n$, we know the remaining overlapping / duplicate object # is $(n-j)$.

So the propability the next package is different

$$= \frac{\text{\# remaining objects}}{\text{\# total}} = \frac{n-j}{n}$$

So P("$j^{th}$ item contains no new type")

$$= 1 - \frac{n-j}{n} = \frac{n-n+j}{n} = \boxed{\frac{j}{n}}$$

Since the packet can either have a new type or not, we want to find the probability of the next package given distinct object #.

We observe this is a geometric distribution because the geometric distribution is the chance something happens first.

$$P(X_j) = \begin{cases} \frac{n-j}{n}, & j^{th} \text{ package is new} \\ j/n, & \text{otherwise} \end{cases}$$

Since $X_j$ is geometric, $E(X_j) = \frac{1}{P(X_j)}$

$$E[X_j] = \frac{1}{\left(\frac{n-j}{n}\right)} = \boxed{\frac{n}{n-j}}$$

$$E[X_j] = \frac{n}{n-j}.$$

b) Let X = # of days to get all ~~x~~ objects.

So, the time required to collect all types of objects is the sum of the time for every new type. $n$ distinct objects, so after the first drawing, $n-1$ distinct left to collect.

4)b) ~~$E(X) = E[$~~

The number of days is given by

$$X = \sum_{j=0}^{n-1} X_j \text{ , so}$$

$$E(X) = E\left[\sum_{j=0}^{n-1} X_j\right] = \sum_{j=0}^{n-1} E(X_j)$$

from (a), $E(X_j) = \cancel{\frac{n-j}{n}}$, ~~so~~ $\frac{n}{n-j}$.

$$E(X) = \sum_{j=0}^{n-1} \frac{n}{n-j} \cancel{\frac{n-j}{k}} = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} \cdots + \frac{n}{1}$$

$$E(X) = n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + 1\right)$$

$$E(X) = n\left(1 + \frac{1}{n-(n-2)} + \cdots + \frac{1}{n-1} + \frac{1}{n}\right)$$

$$\Rightarrow E(X) = n \sum_{k=1}^{n} \frac{1}{k} \text{ , so}$$

$$E(X) = \boxed{n \sum_{k=1}^{n} \frac{1}{k}}$$

4) c) Let $A_i$ be the event that none of the first $k$ objects hold the $i^{th}$ object.
Find $P(A_1 \cup A_2 \cup A_3 \cup A_4)$.

~~$P(A_1) = 1 - P($"first obj has ith k objects have first item"$)$~~

$P(A_1)$ is the probability that object 1 is not amongst the first $k$ boxes.
The total way to arrange $k$ ~~all the~~ boxes out of $n$ $= \binom{n}{k}$.
The total ways to arange box 1 such that it has not been chosen is $\binom{n}{1}(n-1)^k$, Since the box can go anywhere/even after $k$.
Since we're hoping the $i^{th}$ item is not in the $k^{th}$ tries we hope for ~~$\frac{n-1}{n}$~~ tries 0 to $k$ to never reach item 1, so

$$P(A_1) = \frac{\binom{n}{1}(n-1)^k}{\binom{n}{k}} \quad \ldots \text{ and so on.}$$

Since $A_1 \longrightarrow A_k$ are independent as defined by each item having an equal chance to be in every $n$ boxes,
so $P(A_1 \cap A_2) = 0$ ... and so on.
$P(A_1 \cap A_2 \cap A_3) = 0$ ... and so on.
$P(A_1 \cap A_2 \cap A_3 \cap A_4) = 0$.
No box can contain two items or no day can have 2 boxes, so

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = \frac{\binom{n}{1}(n-1)^k + \binom{n}{2}(n-2)^k + \binom{n}{3}(n-3)^k + \binom{n}{4}(n-4)^k}{\binom{n}{k}}$$

4c) By independence and inclusion/exclusion

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = \frac{\binom{n}{1}(n-1)^k + \binom{n}{2}(n-2)^k + \binom{n}{3}(n-3)^k + \binom{n}{4}(n-4)^k}{\binom{n}{k}}$$

5) Rolling a fair die where 1 makes you stop with score 1 or you can stop prior for a higher score.

a) P("k or greater" before 1)

P("1 before k or greater")

When $k=6$, only 2 numbers 1, 6

$P(1 \text{ before } 6) = P(6 \text{ before } 1) = \frac{1}{2}$

When $k=4$, 1, 4, 5, 6 which means 4 numbs

$P(\text{"}4\geq \text{ before } 1\text{"}) = \frac{3}{4}$, but $P(1 \text{ before } 4\geq) = \frac{1}{4}$,

so $P(k \text{ or greater before }) = \frac{6-k+1}{6}$

$\Rightarrow \frac{7-k}{6} = P(\text{"}k \geq \text{ before } 1\text{"})$

$\frac{1}{8-k} = P(\text{"}1 \text{ before } k \geq\text{"})$

5) a) $P(A_k) = P(\text{"k or greater after 1"})$

By conditioning and Bayes' Law,

$$P(A_k) = P(A_k \mid \text{Roll } 1)\,P(\text{Roll } 1) +$$
$$+ P(A_k \mid 1 < c < k)\,(P(1 < c < k))$$
$$+ P(A_k \mid \text{Roll} \geq k)\,P(\text{Roll} \geq k)$$

where $c \in \{(1, k)\}$

The probability a roll $\geq k$ is given by $\frac{6-(k-1)}{6} = \frac{7-k}{6}$

$P(1 < c < k) = \frac{k-2}{6}$ since $k \neq 1$ & $k \neq k$

$$\Rightarrow P(A_k) = 0 + P(A_k)\left(\frac{k-2}{6}\right) + \frac{7-k}{6}$$

$$\Rightarrow P(A_k) = \left(\frac{k-2}{6}\right)P(A_k) + \frac{7-k}{6}$$

$$P(A_k) - P(A_k)\left(\frac{k-2}{6}\right) = \frac{7-k}{6}$$

$$P(A_k)\left(\frac{8-k}{6}\right) = \frac{7-k}{6} \Rightarrow$$

$$\boxed{P(A_k) = \cancel{8}\,\frac{7-k}{8-k}}$$

b) Let $X_k$ be the score when strategy $S(k)$ was used. Find $E(X_k)$, $k=4$, $k=6$.

We will stop the first time that the die shows r or greater, so if we roll a 6 we could either have score 6 from there not being a 1 before or score 1 from the 1 coming before the 6. By Bayes' Law

$$S(6) = 6 \cdot P(6 \text{ before } 1) + 1\, P(1 \text{ before } 6)$$

5) b) from (a) $P(6 \text{ before } 1) = \frac{7-6}{2} = \frac{1}{8-6}$

So $P(1 \text{ before } 6) = \frac{1}{2} = P(6 \text{ before } 1)$, so

$$S(6) = 6 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{7}{2} = 3.5$$

By similar reasing and knowing

$P(4 \geq \text{ before } 1) = \frac{3}{4}$ and

$P(1 \text{ before } 4 \geq) = \frac{1}{4}$, so

$$S(4) = 4 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 5\left(\frac{1}{4}\right) + 6\left(\frac{1}{4}\right)$$

$$\Rightarrow S(4) = \frac{4}{4} + \frac{1}{4} + \frac{5}{4} + \frac{6}{4} = \frac{16}{4}$$

$$\Rightarrow S(4) = 4.$$

When $S(k) = \cancel{E(X_k)}\ X_k$, $E(X_k)$ for $k=4$, $k=6$ are 4 and 3.5 respectively.

The better strategy is to stop when rolling a number ~~over 4~~ 4 or over.

$$S(5) = 5 \cdot \frac{1}{3} + 6 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{12}{3} = 4.$$

Stopping only at 5 or 6 results in the same expectation of score, so the strategy of stopping at $k=4$ or greater expects a larger out come then stopping only on 6. Stopping on 4 or greater expects the same score as stopping on 5 or greater.