23/11

HW 3

HW3/1. Write

$$RHS = \sum_{n=1}^{\infty} P(X \geq n)$$

$$= P(X \geq 1)$$

$$+ P(X \geq 2)$$

$$+$$

$$\vdots$$

$$
\begin{aligned}
&= P(X=1) + P(X=2) + P(X=3) + \cdots \\
&\qquad\qquad\quad + P(X=2) + P(X=3) + \cdots \\
&\qquad\qquad\qquad\qquad\quad + P(X=3) + \cdots \\
&\qquad\qquad\qquad\qquad\qquad\quad \vdots \qquad\qquad \vdots \\
\left(\begin{array}{l}\text{Since the}\\ \text{summands} \geq 0\end{array}\right) &= P(X=1) + 2P(X=2) + 3P(X=3) + \cdots \\
&= E(X) \\
&= \text{LHS}
\end{aligned}
$$

$$
\left[\begin{array}{cccccc|c}
1 & -1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & -1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & -1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
\hline
1 & 0 & 0 & \cdots & & \cdots & 1 \quad 0
\end{array}\right]
$$

The above would not have been possible

if either the numbers being added are all non-negative or the sum of the absolute values of all the numbers is finite.

Fact: If $\{a_{ij}, i,j \in \mathbb{N}\}$ is a collection of real numbers such that either $a_{ij} \geq 0$ for all i,j or $\sum_{i\geq 1} \sum_{j\geq 1} |a_{ij}| < \infty$, then $\sum_{i\geq 1} \sum_{j\geq 1} a_{ij} = \sum_{j\geq 1} \sum_{i\geq 1} a_{ij}$.

HW3|2 What is the distn of X_n?

For $n \geq 1$, $X_n \sim \mathrm{Bin}\left(n, \frac{1}{n}\right)$.

$\boxed{\begin{array}{l} \underline{\text{Thm.}} \text{ If } X_n \sim \mathrm{Bin}(n, p_n) \text{ and} \\ \qquad \lim_{n\to\infty} n p_n = \lambda \in (0, \infty), \\ \text{then } \lim_{n\to\infty} P(X_n = k) = \dfrac{e^{-\lambda} \lambda^k}{k!} \text{ for all } k \in \{0, 1, 2, \dots\}. \end{array}}$

Using the above thm, it follows that

$$\lim_{n\to\infty} P(X_n = k) = \frac{e^{-1}}{k!} \text{ for } k \in \{0, 1, 2, \ldots\}.$$

<u>HW3/3</u> The distribution of X_n is $\text{Bin}\left(\binom{n}{2}, \frac{1}{n^2}\right)$.

For $k = 0, 1, 2, \ldots$,

$$\lim_{n\to\infty} P(X_n = k) = \frac{e^{-1/2}}{2^k\, k!}.$$

HW3|4 If X_n denotes the # broken connections,

then $X_n \sim \text{Bin}\left(n-1, \frac{\lambda}{n}\right)$.

Hence $\lim_{n\to\infty} P(X_n=k) = \frac{e^{-\lambda}\lambda^k}{k!}$, $k \in \{0,1,2,\ldots\}$.

Taking $k=0$,

$$p_n \to e^{-\lambda}, \quad n \to \infty.$$

HW3/5. For a fixed n, and $1 \leq k \leq n$,

$$P(X_n \geq k) = P\left(\text{The minimum of the } n \text{ chosen units is } \geq k\right)$$

$$= P\left(\text{All the chosen units belong to } \{k, k+1, \ldots, n\}\right)$$

$$= \left(\frac{n-k+1}{n}\right)^n$$

The PMF of X_n is

$$P(X_n = k) = P(X_n \geq k) - P(X_n \geq k+1)$$

$$= \left(\frac{n-k+1}{n}\right)^n - \left(\frac{n-k}{n}\right)^n, \quad k = 1, \dots, n.$$

$$\boxed{\lim_{n\to\infty}\left(1+\frac{\lambda}{n}\right)^n = e^{\lambda}, \quad \lambda \in \mathbb{R}.}$$

$$= \left(1 - \frac{k-1}{n}\right)^n - \left(1 - \frac{k}{n}\right)^n.$$

Hence, $\displaystyle\lim_{n\to\infty} P(X_n = k) = e^{-(k-1)} - e^{-k}, \quad k \in \{1, 2, 3, \dots\}.$

$$= e^{-(k-1)}\left(1-e^{-1}\right).$$

The RHS is the PMF of Geo$\left(1-e^{-1}\right)$.

Recall that if $X \sim \text{Geo}(p)$, then $P(X=k) = (1-p)^{k-1}p$.

Consider a coin with probability of Heads p. Toss the coin till the first Head is obtained. The # tosses required follows Geometric(p).

Caution: Sometimes the # Tails before the first Head is also taken as Geo(p).

HW3/6

(a) The PMF of X is

$$P(X=k) = P(X \geq k) - P(X \geq k+1)$$

$$= \left(\frac{N-k+1}{N}\right)^n - \left(\frac{N-k}{N}\right)^n, \quad k=1,\dots,N.$$

Using HW3/1, we get

$$E(X) = \sum_{i=1}^{\infty} P(X \geq i) = \sum_{i=1}^{N} P(X \geq i)$$

$$= \sum_{i=1}^{N} \left(\frac{N-i+1}{N} \right)^n$$

$$= N^{-n} \sum_{j=1}^{N} j^n . \qquad (j = N-i+1)$$

For example, if $n = 2$, then

$$E(X) = N^{-2} \frac{N(N+1)(2N+1)}{6} = \frac{(N+1)(2N+1)}{6N} .$$

(b) For fixed N,

$$P(X=1) = 1 - \left(\frac{N-1}{N}\right)^{n} \longrightarrow 1 \text{ as } n \to \infty,$$

and

$$E(X) = \sum_{j=1}^{N} \left(\frac{j}{N}\right)^{n} \longrightarrow 1 \text{ as } n \to \infty$$

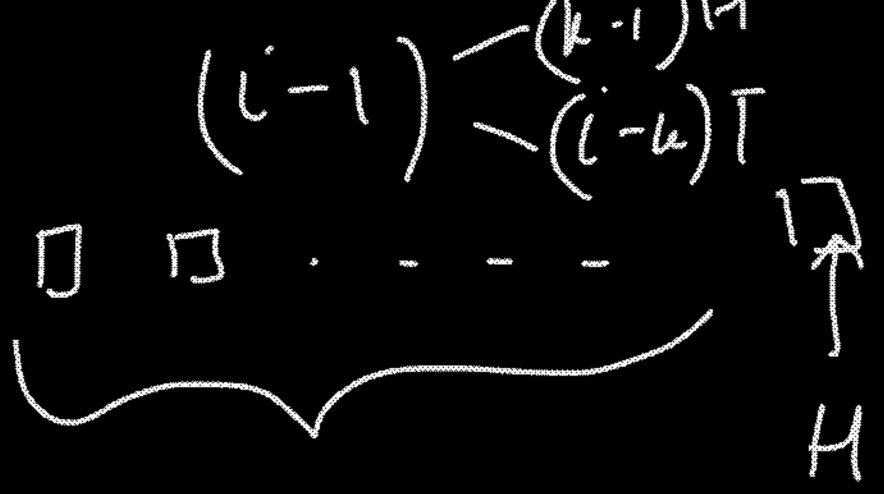
because $\lim_{n\to\infty} \left(\frac{j}{N}\right)^{n} = \begin{cases} 1, & \text{if } j = N, \\ 0, & \text{if } 1 \le j \le N-1. \end{cases}$

The sum and limit can be interchanged when the sum is over a fixed finite number.

HW3/7. The possible values of X is $\{k, k+1, k+2, \dots\}$.

The PMF of X is

$$P(X=i) = \binom{i-1}{k-1} p^{k} (1-p)^{i-k}, \quad i \in \{k, k+1, \dots\}.$$



$$\sum_{i=k}^{\infty} P(X=i) = \sum_{i=k}^{\infty} \binom{i-1}{k-1} p^k q^{i-k} \qquad (\text{denote } q = 1-p)$$

$$= p^k \sum_{i=k}^{\infty} \binom{i-1}{k-1} q^{i-k} \qquad (j = i-k)$$

$$= p^k \sum_{j=0}^{\infty} \binom{j+k-1}{k-1} q^j$$

$$= p^k \left(1 + kq + \frac{k(k+1)}{2} q^2 + \frac{k(k+1)(k+2)}{3!} q^3 + \cdots \right)$$

$$(*) \quad = p^k (1-q)^{-k} \qquad (\text{because } 0 < q < 1)$$

$$= 1.$$

Due to $(*)$, this is called the "Negative Binomial" distn.

Denote by Y_1 the # tosses needed to get the first H.

" " Y_2 " " " " after the first H for getting the second H.

Denote by Y_i, for $1 \le i \le k$, the # tosses needed to get the i-th H after the (i-1)-th H.

For example if $k = 4$, and the outcomes ar T T H H T T T H H

then $Y_1 = 3$, $Y_2 = 1$, $Y_3 = 4$, $Y_4 = 1$.

Clearly,

$$X = \sum_{i=1}^{k} Y_i ,$$

and $Y_1, \ldots, Y_k$ follow Geometric (p).

Recall that $E(Y_i) = \frac{1}{p}$, for $i = 1, \ldots, k$.

Hence $E(X) = \frac{k}{p}$.

Let us calculate the variance of Geometric.

Denote $q = 1-p$ and recall that

$$\sum_{i=0}^{\infty} q^{i} = (1-q)^{-1}.$$

Differentiating both sides w.r.t. q, we get

$$\sum_{i=1}^{\infty} i q^{i-1} = (1-q)^{-2}.$$

Differentiating once again,

$$\sum_{i=2}^{\infty} i(i-1)\, q^{i-2} = 2(1-q)^{-3}.$$

Rewrite the above as (putting $j = i-1$)

$$\sum_{j=1}^{\infty} (j+1)\, j\, q^{j-1} p = 2(1-q)^{-3} p = 2p^{-2},$$

and note that $\text{LHS} = \sum_{j=1}^{\infty} (j+1)\, j\, P(Y_1 = j)$

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

Hence, $\operatorname{Var}(Y_1) = E(Y_1^2) - E^2(Y_1)$

$$= E\big(Y_1(Y_1+1)\big) - E(Y_1) - E^2(Y_1)$$

$$= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{1}{p^2} - \frac{1}{p} = \frac{1-p}{p^2}.$$

In other words, variance of $\text{Geo}(p)$ is $\frac{1-p}{p^2}$.

Since $X = \sum_{i=1}^{k} Y_i$,

$$\mathrm{Var}(X) = \sum_{i=1}^{k} \mathrm{Var}(Y_i) + 2 \sum_{1 \le i < j \le k} \mathrm{Cov}(Y_i, Y_j).$$

<u>Exc.</u> Check that for $i \neq j$, Y_i and Y_j are independent and hence $\mathrm{Cov}(Y_i, Y_j) = 0$.

Thus, $\mathrm{Var}(X) = \sum_{i=1}^{k} \mathrm{Var}(Y_i) = k \frac{1-p}{p^2}$.

HW3/8. Clearly X takes values $\{0, 1, \ldots, N-M\}$.

For $i \in \{0, 1, \ldots, N-M\}$

$P(X = i) = P($In the first i draws, black is obtained and the $(i+1)$-th draw yields a white$)$

$$= \frac{(N-M)(N-M-1)\cdots(N-M-i+1)\,M}{N(N-1)\cdots(N-i)}.$$

$$= \frac{M\prod_{j=0}^{i-1}(N-M-j)}{\prod_{j=0}^{i}(N-j)}.$$

(Convention: Empty product $= 1$, Empty sum $= 0$.)

Let us label all the black balls as $1, 2, \ldots, N-M$.

Notice that

$$X = \sum_{i=1}^{N-M} \mathbb{1}\left(\begin{array}{l}\text{The black ball with label } i \text{ is drawn} \\ \text{before the first white ball is drawn}\end{array}\right).$$

Hence,

$$E(X) = \sum_{i=1}^{N-M} P\left(\begin{array}{l}\text{The black ball with label } i \text{ is drawn} \\ \text{before the firs white ball is drawn}\end{array}\right)$$

$$= \sum_{i=1}^{N-M} P\left(\begin{array}{c}\text{The black ball with label } i \text{ is drawn before} \\ \text{all the } M \text{ white balls}\end{array}\right)$$

$$= \sum_{i=1}^{N-M} \frac{1}{M+1} \qquad (\text{Exc.})$$

$$= \frac{N-M}{M+1}.$$

<u>Moral</u>: Expressing a count random variable as sum of indicators is often helpful.

HW3|9. The possible values of X are $\{1, 2, \ldots, 6\}$.

For $1 \leq i \leq 5$,

$$P(X > i) = P(\text{The sum of the first } i \text{ rolls} \leq 5)$$

$$= 6^{-i} \# \{\text{Ways in which the sum of first } i \text{ rolls} \leq 5\}$$

$$= 6^{-i} \# \{(x_1, \ldots, x_i) : x_1, \ldots, x_i \in \{1, \ldots, 6\} \text{ and } x_1 + \cdots + x_i \leq 5\}$$

$$= 6^{-i} \# \{(x_1, \ldots, x_i) : x_1, \ldots, x_i \in \mathbb{N} \text{ and } x_1 + \cdots + x_i \leq 5\}$$

$$(J_j = I_j - 1) = 6^{-i} \# \{ (y_1, \dots, y_i) : y_1, \dots, y_i \in \{0, 1, 2, \dots\} \text{ and } y_1 + \dots + y_i \le 5 - i \}$$

$$= 6^{-i} \# \{ (y_1, \dots, y_i, y_{i+1}) : y_1, \dots, y_{i+1} \in \{0, 1, 2, \dots\} \text{ and } y_1 + \dots + y_{i+1} = 5 - i \}$$

$$= 6^{-i} \# \{ \text{Ways of throwing } (5-i) \text{ identical balls into } (i+1) \text{ distinct boxes} \}$$

$$= 6^{-i} \binom{5}{i}$$

Thus, the PMF of X is

$$P(X=i) = \begin{cases} 6^{-(i-1)} \binom{5}{i-1} - 6^{-i} \binom{5}{i}, & i=1,\dots,6, \\ 0, & \text{o.w.} \end{cases}$$

Convention: $\binom{a}{b} = 0$ if $b > a$ or $b < 0$.

Recalling HW3/1, write

$$E(X) = \sum_{i=1}^{\infty} P(X \geq i)$$

$$= 1 + \sum_{i=2}^{6} P(X \geq i)$$

$$= 1 + \sum_{j=1}^{5} P(X > j) \qquad (j = i-1)$$

$$= 1 + \sum_{j=1}^{5} 6^{-j} \binom{5}{j}$$

$$= \sum_{j=0}^{5} 6^{-j} \binom{5}{j}$$

$$= \left(1 + \frac{1}{6}\right)^5$$

Probability theory 1

Homework 3 (Solutions of 10 - 15)

November 23, 2021

10. A man has a key ring with n keys, exactly one of which would unlock a door. He has forgotten the correct key, and keeps picking keys from the ring, with replacement, and tries them. If X denotes the number of trials needed to unlock the door, calculate $\mathrm{E}(X)$.

    Soln.: Clearly, X follows Geometric$(1/n)$ and hence

$$\mathrm{E}(X) = n\,.$$

11. The wife of the man in the above problem tells him that a better strategy is to try the keys without replacement. Justify or contradict her claim by recalculating $\mathrm{E}(X)$ in this scheme, and comparing the same with your answer to the above problem.

    Soln.: In this case, X is distributed uniformly over $\{1, \ldots, n\}$ and hence

$$\mathrm{E}(X) = \frac{n+1}{2} \le n\,,$$

    with the inequality being strict for $n \ge 2$. Thus, trying without replacement is better.

12. Apply the following tweak to the above problem. Suppose that the ring has exactly two keys which would unlock the door. Recalculate the expected number of trials needed to unlock the door, when the keys are tried

    (a) with replacement,

    (b) without replacement.

    Soln.:

    (a) Now X follows Geometric $(2/n)$ and hence

$$\mathrm{E}(X) = \frac{n}{2}\,.$$

(b) This situation is very similar to that of HW3/8, from which it follows that

$$\mathrm{E}(X) = 1 + \frac{n-2}{3} = \frac{n+1}{3}.$$

It is easy to see that

$$\frac{n+1}{3} < \frac{n}{2}, n \geq 3.$$

13. A set is chosen at random from the collection of all subsets of $\{1, \ldots, n\}$. Let X denote the cardinality of the chosen set. Calculate the PMF and expectation of X.

Soln.: Since the total number of subsets is 2^n, of which, $\binom{n}{k}$ are of size k for $k = 0, 1, \ldots, n$, it follows that X follows Binomial$(n, 1/2)$. Hence,

$$\mathrm{E}(X) = \frac{n}{2}.$$

14. For events $A_1, A_2, \ldots$, show that

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n).$$

Soln.: For finite n, it can be shown by induction that

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i). \tag{1}$$

As $n \to \infty$, the LHS goes to

$$P\left(\bigcup_{i=1}^{\infty} A_i\right),$$

because

$$\bigcup_{i=1}^{n} A_i \uparrow \bigcup_{i=1}^{\infty} A_i.$$

Thus, letting $n \to \infty$ in (1) completes the solution because the RHS therein goes to

$$\sum_{i=1}^{\infty} P(A_i).$$

15. Calculate the expectation and variance of the the following distributions. Even if some of these are done in class, redo the calculations for your own practice.

(a) Bernoulli (p) (Bernoulli (p) is the distribution of a random variable taking values 1 and 0 with probabilities p and $1-p$, respectively.)

(b) Binomial (n, p)

(c) Geometric (p)

(d) Poisson (λ)

(e) Uniform on $\{1, \ldots, n\}$

Soln.:

Distribution	Mean	Variance
Bernoulli(p)	p	$p - p^2$
Binomial(n, p)	np	$np(1-p)$
Geometric(p)	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson(λ)	λ	λ
Uniform on $\{1, \ldots, n\}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$