Distributions for Discrete Random Variables

Botao Jin

University of California, Santa Barbara — October 26, 2024

Binomial Distribution

1. Exercise 2.21: Suppose X be the number of problems she correctly answer, so $X \sim Bin(4, .8)$.
   (a) $P(\text{She got an A}) = P(X \geq 3) = \binom{4}{3}(.8)^3(.2) + \binom{4}{4}(.8)^4 = .8192$
   (b) X_1 be the indicator of whether she gets the first problem correct, and Y be the number of the remaining three problems she correctly answer, so $X \sim Bern(.8)$, $Y \sim Bin(3, .8)$, X and Y are independent, and $X = X_1 + Y$

$$\begin{aligned} P(X \geq 3 | X_1 = 1) &= \frac{P(X \geq 3, X_1 = 1)}{P(X_1 = 1)} \\ &= \frac{P(Y \geq 2, X_1 = 1)}{P(X_1 = 1)} \\ &= \frac{P(Y \geq 2)P(X_1 = 1)}{P(X_1 = 1)} \\ &= P(Y \geq 2) = 3(.8)^2(.2) + (.8)^3 = .896 \end{aligned}$$

2. Exercise 2.62: Note that

$$p = P(\text{Three marble are blue}) = \frac{\binom{9}{3}}{\binom{9+4}{3}} = \frac{42}{143}$$

Let X be the number of times we got three blue marbles, then $X \sim Bin(20, p)$ with p.m.f

$$P(X = k) = \binom{20}{k} p^k (1-p)^{20-k}$$

for $p = \frac{42}{143}$ and $k = 0, 1, \ldots, 20$.

Geometric Distribution

1. Exercise 2.20: A fair dice is rolled repeatedly.
   (a) Let X be the number of threes in the first five rolls, then

$$P(\text{get a three for at most two times}) = P(X \leq 2) = \sum_{k=0}^{2} \binom{5}{k} (1/6)^k (5/6)^{5-k}$$

   (b) Let N be the number of times we needed to see the first three, then $N \sim Geo(1/6)$

$$\begin{aligned} P(\text{No three before the fifth roll}) &= P(\text{No three in the first four rolls}) \\ &= P(N \geq 5) = P(N > 4) = (5/6)^4 \end{aligned}$$

   (c) Consider the event that {The first three appears before the twentieth but not in the fifth roll}, which is equivalent to {The first three appears between the fifth roll and the nineteenth roll}, hence

$$P(5 \leq N < 20) = P(N \geq 5) - P(N \geq 20) = P(N > 4) - P(N > 19) = (5/6)^4 - (5/6)^{19}$$

2. Exercise 2.22: Let X be the number of rounds we needed to see the first success, then $X \sim Geo(p)$

   (a) By symmetric, $P(\text{Anne wins}) = P(\text{Anne loses}) = P(\text{tie}) = 1/3$, so $p = 1/3$
   Remark. You can also check it by the law of total probability

   (b) $P(\text{The first win in the fourth round}) = P(X = 4) = (2/3)^3(1/3) = 8/81$

   (c) $P(\text{The first win comes after the fourth round}) = P(\text{She didn't win in the first four round}) = (2/3)^4 = 16/81$

Hyper-geometric Distribution

1. Exercise 2.24: Three people are randomly chosen from a group of two men and four women. Let X be the number of women, then

   (a) $X \sim hypergeo(6, 4, 3)$

   (b) The p.m.f of X is
   $$P(X = k) = \frac{\binom{4}{k}\binom{2}{3-k}}{\binom{6}{3}}$$
   for $k = 1, 2, 3$

2. Exercise 2.28:

   (a) Hyper-geometric distribution with parameters $(52, 4, 13)$

   (b) Binomial distribution with parameters $n = 50$ and
   $$p = 1 - \frac{\binom{4}{0}\binom{48}{13}}{\binom{52}{13}}$$
   You can directly leave it as the fraction.

   (c) Binomial distribution with parameters $n = 50$ and
   $$p = \frac{\binom{4}{1}\binom{13}{13}}{\binom{52}{13}} = \frac{4}{\binom{52}{13}}$$

   (d) Hyper-geometric distribution with parameters $(52, 13, 13)$

Poisson distribution

1. Exercise 4.10: Suppose X is the score, then $X \sim Poi(\lambda)$. Given the information, we have $.5 = P(X \geq 1)$, and in other words, $P(X = 0) = e^{-\lambda} = .5$. Thus $\lambda = \log 2$ and $P(X = 3) = e^{-\lambda}\lambda^3/3! = .028$

2. Exercise 4.33: Let X be the number of calls received with $X \sim Poi(\lambda)$. Then $0.5\% = P(X = 0) = e^{-\lambda}$, thus average number of calls $\lambda = -\log(.005) = 5.298$

3. Exercise 4.34: Suppose X be the number of accidents with $X \sim Poi(\lambda)$. Average $\lambda = 3$ implies that $P(X \leq 2) = e^{-3}(1 + 3/1! + 3^2/2!) = 8.5e^{-3}$

Note: This study guide is used for Botao Jin's sections only. Comments, bug reports: b_jin@ucsb.edu