Math 10B with Professor Stankova
Worksheet, Discussion #13; Thursday, 3/7/2019
GSI name: Roy Zhao

1 Discrete Probability Distributions

1.1 Concepts

1.

Distribution	PMF	Example
Uniform	If $\#R(X) = n$, then $f(x) = \frac{1}{n}$ for all $x \in R(X)$.	Dice roll, $f(1) = f(2) = \cdots = f(6) = \frac{1}{6}$.
Bernoulli Trial	$f(0) = 1 - p$, $f(1) = p$	Flipping a biased coin
Binomial	$f(k) = \binom{n}{k} p^k (1-p)^{n-k}$	p is probability of success. Repeat n Bernoulli trials. Number of 6's rolled when rolling 10 die is $f(k) = \binom{10}{k}(1/6)^k(5/6)^{10-k}$.
Geometric	$f(k) = (1-p)^k p$	k failures and then a success.
Hyper-Geometric	$f(k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$	Counting the number of red balls I pick out of n balls drawn if there are m red balls out of N balls total.
Poisson	$f(k) = \frac{\lambda^k e^{-\lambda}}{k!}$	Count the number of babies born today if on average there are 3 babies born a day.

1.2 Examples

2. I am picking cards out of a deck. What is the probability that I pull out 1 heart out of 5 cards if I pull with replacement? If I pull 5 cards at once?

Solution: With replacement is repeated Bernoulli trials which means binomial distribution. The probability of a success or pulling out a heart is $\frac{1}{4}$. Therefore, the probability of pulling 1 heart out of 5 is

$$\binom{5}{1}\frac{1}{4^1} \cdot \frac{3^{5-1}}{4^{5-1}}.$$

If do not have replacement, then this is a hyper-geometric distribution with $N = 52, n = 5, m = 13$, the the answer is

$$\frac{\binom{13}{1}\binom{39}{4}}{\binom{52}{5}}.$$

3. What is the probability that first heart is the third card I draw (with replacement)?

Solution: We want to know the probability of the first success, which is geometric since we are doing with replacement. The probability of a success is $p = \frac{1}{4}$ and we have two failures before we have a success so $k = 2$. Hence the answer is $f(2) = (3/4)^2(1/4) = \frac{9}{64}$.

1.3 Problems

4. True **FALSE** We cannot talk about Bernoulli trials for rolling a 5 because there are 6 outputs and we need 2 for a Bernoulli trial.

Solution: We can change this into a Bernoulli trial by saying the probability of success is 1/6 and all other outputs are failures, which makes it into success and failure.

5. True **FALSE** When flipping a fair ($p = 1/2$) coin, the probability of flipping 100 heads on 200 flips is 50%.

Solution: If you calculate it, getting exactly 100 heads is very small (5.6%).

6. True **FALSE** The geometric distribution, like the hyper-geometric distribution, assumes that the trials are dependent (without replacement).

Solution: The geometric distribution has independent Bernoulli trials but the hypergeometric distribution has dependent ones because we pick without replacement.

7. In a class of 50 males and 80 females, I give out 3 awards randomly. What is the probability that females will win all 3 awards if the awards must go to different people? What about if the same person can win all three awards?

Solution: This is like the probability of picking 3 females out of 3 people chosen. If the awards must go to different people, there is no replacement so it is the hypergeometric distribution where a success is picking a female. So we have $N = 130$

students total, there are $m = 80$ females, and I am picking $n = 3$ students and I want $k = 3$ females. This gives

$$f(3) = \frac{\binom{80}{3}\binom{50}{0}}{\binom{130}{3}} = \frac{\binom{80}{3}}{\binom{130}{3}}.$$

If the same person can will all the awards, then we are choosing with replacement. So, this is a binomial distribution where the probability of success is $p = \frac{80}{130}$. Thus, we have that the answer is

$$f(3) = \binom{3}{3}\left(\frac{80}{130}\right)^3\left(\frac{50}{130}\right)^0 = \frac{80^3}{130^3}.$$

8. Suppose I am baking souffles and each souffle has a 10% chance of not rising. What is the probability that when baking 50 souffles for class, at most 2 of them will not rise? Suppose that I keep baking them until a souffle doesn't rise. Let X be the number of souffles I bake before one doesn't rise. What is $P(X \geq 20)$?

Solution: For the first part, this is a binomial distribution since we are repeating a Bernoulli trial independently. The probability of success is $p = 0.9$ and we want $P(X \geq 48) = P(X = 48) + P(X = 49) + P(X = 50)$. There are $n = 50$ total trials so the answer is

$$\binom{50}{48}(0.9)^{48}(0.1)^2 + \binom{50}{49}(0.9)^{49}(0.1)^1 + \binom{50}{50}(0.9)^{50} + (0.1)^0.$$

Now for the second part, this is a geometric distribution because we are repeating something until success and the probability of ending is the probability of failure which is $p = 0.1$. Then we want $P(X \geq 20) = (1-p)^{20} = (0.9)^{20}$.

9. At Berkeley, there is an equal number of people aged $18, 19, \ldots, 27$. I cold call someone at random and ask for their age. What is the PMF for their age? Suppose that undergraduates are aged 18 though 21 inclusive. What is the probability that I have to call 10 people until I call an undergraduate (the undergraduate is the 10th person I call)? What is the probability that I call 4 undergraduates out of 10 people I call (if I can call someone more than once)?

Solution: This is a uniform distribution on $[18, 27]$. There are 10 numbers in between and hence

$$f(k) = \begin{cases} \frac{1}{10} & 18 \leq x \leq 27 \\ 0 & \text{otherwise} \end{cases}.$$

The probability of calling an undergrad is $f(18)+f(19)+f(20)+f(21)=\frac{4}{10}=\frac{2}{5}$. The probability that the first undergrad I call is the 10th person is given by the geometric distribution since we are talking about times until a success. I have 9 failures before and so this is

$$f(9)=(1-2/5)^9(2/5)=\frac{3^9\cdot 2}{5^{10}}.$$

The probability that I call 4 undergraduates out of 10 people is given by a binomial distribution since I can call someone more than once and hence there is replacement. So plugging this into the binomial distribution gives

$$f(4)=\binom{10}{4}(2/5)^4(3/5)^6.$$

10. (Challenge) For a lottery, 6 distinct numbers are drawn out of 60 and to win, you need to match all 6 numbers. What is the probability that I win? If I buy 100 different tickets, what is the probability that I win?

Solution: This can be thought of as there are a total of $\binom{60}{6}$ different lottery tickets and only one of them is a winner, or only one of them is tagged. When I buy different tickets, I am picking without replacement so this is the hyper-geometric distribution. There are $N=\binom{60}{6}$ total tickets and $m=1$ of them is a success. Then out of the n balls I draw, I want $k=1$ success. For the first case, if I only pick one ticket, then $n=1$ and we get

$$f(1)=\frac{\binom{1}{1}\binom{\binom{60}{6}-1}{0}}{\binom{\binom{60}{6}}{1}}=\frac{1}{\binom{60}{6}}.$$

If I pick 100 tickets, then $n=100$ and we get

$$f(1)=\frac{\binom{1}{1}\binom{\binom{60}{6}-1}{99}}{\binom{\binom{60}{6}}{100}}.$$

1.4 Examples

11. On average, there are 20 rainy days in Berkeley per year. What is the probability that this year, there are 30? What is the probability that over 5 years it will rain 100 times?

Solution: This is a Poisson distribution since we are talking about the average number of times something occurs in a place or over some period of time. The average is $\lambda = 20$ and hence $P(X = 30) = f(30) = \frac{20^{30}e^{-20}}{30!}$.

For the second part, our new time span is 5 years and the average amount it will rain over 5 years is $\lambda = 5 \cdot 20 = 100$. So with this new λ we calculate the probability of $k = 100$ rainy days is

$$f(100) = \frac{\lambda^{100}e^{-\lambda}}{100!} = \frac{100^{100}e^{-100}}{100!}.$$

12. The probability of seeing a shiny Pokemon is approximately 1 in $10000 = 10^4$. What is the probability that I don't see any in my playthrough if I see 10^5 Pokemon total? (calculate both exactly and an approximation)

Solution: Assuming each Pokemon sighting is independent, this is binomial process with $p = 1/10^4$. We see 10^5 Pokemon and want to see $k = 0$ shiny Pokemon so the answer is $\binom{10^5}{0}p^0(1-p)^{10^5} = (1 - \frac{1}{10^4})^{10^5}$.

Now a Poisson distribution approximates a binomial one whenever n is very large and p is very small. The average number of Pokemon we can expect to see when seeing 10^5 Pokemon is $np = \lambda = 10$ shiny Pokemon. Thus, the probability that we see 0 is $f(0) = \frac{\lambda^0 e^{-\lambda}}{0!} = e^{-10}$.

1.5 Problems

13. **TRUE** False We can use the Poisson distribution to show that $\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = 1$.

Solution: This is a consequence of the fact that the sum of all the possibilities of the Poisson distribution must equal 1.

14. When a cell undergoes mitosis, the number of mutations that occurs is Poisson distributed and an average of 11 mutations occur. What is the probability that no more than 1 mutation occurs when a cell divides?

Solution: This is a Poisson process with $\lambda = 11$. We want to know $P(X \leq 1) = P(X = 0) + P(X = 1) = f(0) + f(1) = \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} = e^{-11} + 11e^{-11} = 12e^{-11}$.

15. The number of chocolate chips in a cookie is Poisson distributed with an average of 15 chocolate chips. What is the probability that you pick up a cookie with only 10 chocolate chips in it?

> **Solution:** This is a Poisson distribution with $\lambda = 15$. We want to calculate $f(10) = \frac{\lambda^{10}e^{-\lambda}}{10!} = \frac{15^{10}e^{-15}}{10!}$.

16. The number of errors on a page is Poisson distributed with approximately 1 error per 100 pages of a book. What is the probability that a novel of 300 pages contains no errors?

> **Solution:** This is a Poisson distribution and the probability that we have 300 pages with no errors means that the first 100 pages has no errors, the second 100 and the last 100 all have no errors. So the answer is $f(0)^3 = e^{-3\lambda} = e^{-3}$.
>
> Another way to do this is noting that if we average 1 error per 100 pages, then over a novel of 300 pages, we should expect $300/100 \cdot 1 = 3$ errors. Thus, the probability of having no errors with $\lambda = 3$ is $e^{-\lambda} = e^{-3}$.

17. (Challenge) Approximately 4 people are born every second. What is the probability that in a minute, there are 240 people born?

> **Solution:** Since approximately 4 people are born every second and there are 60 seconds in a minute, approximately $4 \cdot 60 = 240$ people are born every minute. Thus, this is Poisson distribution with $\lambda = 240$ and we want to calculate $f(240) = \frac{\lambda^{240}e^{-\lambda}}{240!} = \frac{240^{240}e^{-240}}{240!}$.