Chapter 5 Summary and Review (draft: 2019/11/10-15:47:08)

Bernoulli, binomial, geometric, negative binomial

Summary of topics and terminology:

- Bernoulli: $X \sim \text{Bernoulli}(\theta)$. Represents an simple random experiment with two possible outcomes. $X = 1$ is a success with probability θ, and $X = 0$ is a failure with probability $1 - \theta$.
  - $f_X(x) = \theta^x(1-\theta)^{1-x}$. pmf, discrete, $x = 0, 1$.
  - $\text{E}(X) = \theta$
  - $\text{Var}(X) = \theta(1-\theta)$
- Binomial: $X \sim \text{Bin}(n, \theta)$ is the number of successes in n independent Bernoulli trials with probability of success θ.
  - $f_X(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$. pmf, discrete, $x = 0, 1, 2, \ldots, n$.
  - $\text{E}(X) = n\theta$
  - $\text{Var}(X) = n\theta(1-\theta)$
- Geometric: $X \sim \text{Geom}(\theta)$ is the number of trials up to and including the first success
  - $f_X(x) = \theta(1-\theta)^{x-1}$. pmf, discrete, $x = 1, 2, 3, \ldots$
  - $\text{E}(X) = \frac{1}{\theta}$
  - $\text{Var}(X) = \frac{1-\theta}{\theta^2}$
- Negative binomial: $X \sim \text{NB}(k, \theta)$ is the number of trials to get k total successes
  - $f_X(x) = \binom{x-1}{k-1}\theta^k(1-\theta)^{x-k}$. pmf, discrete, $x = k, k+1, k+2, \ldots$.
  - $\text{E}(X) = \frac{k}{\theta}$
  - $\text{Var}(X) = \frac{k(1-\theta)}{\theta^2}$

Example problems:

1. Calculate the probability of getting 4 heads when flipping a fair coin ten times.

   Solution:

   X =number of heads out of 10 flips, $X \sim \text{Bin}(n = 10, \theta = 0.5)$.

   $P(X = 4) = f_X(4) = \binom{10}{4}(0.5)^4(1-0.5)^6 = \frac{10\cdot 9\cdot 8\cdot 7}{4\cdot 3\cdot 2\cdot 2^10} \approx 0.2050781$

2. If a factory is producing items with a probability of an item being defective is 1%, how many items are expected to be produced before the first defective item?

   Solution:

   X = the number of items produced up to and including the first defective one.

   $X \sim \text{Geom}(\theta = 0.01)$. $\text{E}(X) = 1/\theta = 100$. So we expect 100 items to be produced, where the first 99 are good and the 100th is defective.

   *note that in R the Geometric RV is defined differently, it counts the number of failures instead of the total number of trials.

3. A consumer satisfaction researcher needs to recruit 10 individuals with a particular preference for a particular study. It is known that 27% of the general population has the particular preference of interest. The researcher is to conduct random phone calls until the 10 individuals are found. What is the probability that this takes less than 20 phone calls?

   Solution:

   Let X = the total number of phone calls required to get the desired 10 individuals

   $X \sim \text{NB}(k = 10, \theta = 0.27)$.

   $P(X < 20) = \sum_{x=10}^{19} \binom{x-1}{9} (0.27)^{10}(0.73)^{x-10} \approx 0.01561759$

   *note that in R the Negative binomial RV is defined differently, it counts the number of failures instead of the total number of trials. So for this problem it is `pnbinom(9,size=10,prob=0.27)`.

4. Challenge: Consider an experiment where Bernoulli trials are performed until the first success. The total number of trials is noted and then the same number of trials is performed again. The total number of successes is then counted after this second stage. What is the expected number of successes? *(note: this is a challenging problem, but try it and see what you think.)*

   Solution:

   Let $X \sim \text{Geom}(\theta)$ and $Y \sim \text{Binom}(n = X, \theta)$.

   X represents the total number of trials to be performed during the first stage. Y represents the number of successes in the second stage. Note that the total number of trials performed during the second stage depends on the outcome of the first stage. So we cannot know exactly how many trials to perform until completing the first stage and getting and X value.

   We know that for a binomial random variable, the expected value is $n\theta$. Since Y is binomial with $n = X$ trials, we can calculate a conditional expectation: $\text{E}(Y \mid X = x) = x\theta$.

   Now we sum over all possible x-values and multiply by their probabilities:
   $\text{E}(Y) = \sum_{x=1}^{\infty} \text{E}(Y \mid X = x) \cdot P(X = x) = \theta \sum_{x=1}^{\infty} xP(X = x) = \theta\text{E}(X) = 1$.

   Another way to conceptualize this is to realize $\text{E}(Y \mid X) = X\theta$ is a random variable if we leave X undetermined, and then we get $\text{E}(Y) = \text{E}[\text{E}(Y \mid X)] = \text{E}[X\theta] = \text{E}(X) \cdot \theta = \frac{1}{\theta} \cdot \theta = 1$.

   Another possible approach would be to write the joint pmf for X, Y as

   $$f(x, y) = f_{Y|X=x}(y|x) \cdot f_X(x) = \binom{x}{y} \theta^y (1-\theta)^{x-y} \cdot \theta(1-\theta)^{x-1}$$

   And calculate

   $$\begin{aligned} \text{E}(Y) &= \sum_{x=1}^{\infty} \sum_{y=0}^{x} y \binom{x}{y} \theta^y (1-\theta)^{x-y} \cdot \theta(1-\theta)^{x-1} \\ &= \sum_{x=1}^{\infty} x\theta \cdot \theta(1-\theta)^{x-1} \\ &= \theta \sum_{x=1}^{\infty} x\theta(1-\theta)^{x-1} \\ &= \theta\text{E}(X) = 1 \end{aligned}$$

Poisson

Summary of topics and terminology:

- The Poisson random variable is for counting how many events occur over a given physical extent. This could be number of events over a given interval of time, length, or even area of volume. Many different examples exist.
- $X \sim \text{Pois}(\lambda)$
  - ➠ $f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$. pmf, discrete, $x = 0, 1, 2, 3, \ldots$.
  - ➠ $\text{E}(X) = \lambda$
  - ➠ $\text{Var}(X) = \lambda$
- λ can be called the rate parameter. It is the number of events per given unit of physical extent. But we have to be careful as to what the base unit of physical extent is.
- Scaling the Poisson parameter: if $X \sim \text{Pois}(\lambda)$ where λ is the number of events to expect over a given physical extent, then if we scale the size of that physical extent by a factor of t, then Y the number of events over this new scaled physical extent is $Y \sim \text{Pois}(\lambda t)$. I.e. that $\text{E}(X) = \lambda$ but $\text{E}(Y) = \lambda t$.
- Poisson approximation to binomial: $X \sim \text{Bin}(n, \theta)$ when n is large enough, then $n(x; n, \theta) \approx p(x,\ \lambda = n\theta)$.

Example problems:

1. A textile manufacturer makes fabric with on average 1 imperfection for every 20 square meters. A particular customer places an order to 100 square meters. What is the expected number of imperfections? What is the probability of no imperfections? What is the probability of at least 3 imperfections?

   Solution:

   The rate of imperfection events is 1 per 20m^2 there are 5 such regions in 100 m^2 so we expect 5 imperfections. Thus we will use $\lambda = 5$. $X \sim \text{Pois}(\lambda = 5)$.

   The probability of zero imperfections is $e^{-\lambda} = e^{-5}$ which is a very tiny probability!

   The probability of at least 3 imperfections is
   $P(X \geq 3) = e^{-5} \sum_{x=3}^{\infty} \frac{5^x}{x!} = e^{-5}(e^5 - 1 - 5 - \frac{5^2}{2}) \approx 0.875348$

   We can calculate this in R as `1-ppois(2,lambda=5)`.

2. Suppose a biologist is taking lake water samples and testing for presence of a particular microbe. If we assume the microbe population is large and uniformly distributed throughout the lake, then it can be modeled well by a Poisson. The biologist takes a 1 milliliter sample and estimates that there are 5 microbes in the sample. Calculate the probability that there are more than 5,100 microbes in a 1L sample.

   Solution:

   The number of microbes in a t milliliter sample will be Poisson distributed with parameter $\lambda = 5t$. With 1,000 mL = 1 L, we will use $\lambda = 5000$.

   $P(X > 5100) = \sum_{x=5101}^{\infty} \frac{e^{-5000}(5000)^x}{x!} \approx 0.07796233$

   We can calculate this in R as `1-ppois(5100,lambda=5000)`.

3. If a certain high energy particle enters the Earth's atmosphere at rate 280 per day, calculate the probability that there are at most 1 in a given 10 minute interval.

   Solution:

   280 events per day means 280/24/60 per minute, then times 10 for a 10min interval.

   We use: $X \sim \text{Pois}(\lambda = 35/18)$.

   $P(X \leq 1) = P(X = 0) + P(X = 1) = e^{-\lambda} + e^{-\lambda}\lambda = e^{-35/18} + e^{-35/18}\frac{35}{18} \approx 0.4212519$

4. If Earth only encounters a particular high energy particle type on average once per century, Calculate the probability that there is at least one such particle in the next year.

   Solution:

   We'll use $\lambda = 0.01$ since once per century would scale down by a factor of 100 to calculate probabilities for a single year.

   $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-0.01} \approx 0.009950166$.

   This is fairly close to a 1% chance. This is generally true for a very small rate parameter. If λ is very small, say less than 0.1 or so, then the probability of at least one event is approximately λ.