Special Discrete random variables

March 20, 2019

Discrete Random variables

Bernoulli RV

Discrete RV X is called *Bernoulli RV* with parameter p if its pmf is

$$P(X = 0) = 1 - p$$
$$P(X = 1) = p$$

Denote $X \sim \text{Ber}(p)$

Properties

- Bernoulli RV is indicator RV of a trial which has success probability p and failure probability $1 - p$.
- $E(X) = p, \ \mathrm{Var}(X) = p(1-p)$

Properties

- Bernoulli RV is indicator RV of a trial which has success probability p and failure probability $1 - p$.
- $E(X) = p, \ \mathrm{Var}(X) = p(1 - p)$

Binomial RV

- n independent trials, each is Ber(p).
- X is the number of success
- X is called Binomial RV with parameter (n, p)
- Denote $X \sim$ Bino(n, p).

Pmf

For $i = 0, 1, \ldots, n$

- Choose positions for i success
- i successes, $n - i$ failures, all independent

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

Pmf

For $i = 0, 1, \ldots, n$

- Choose positions for i success
- i successes, $n - i$ failures, all independent

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$$

Example

It is known that disks produced by a certain company will be defective with probability .01 independently of each other. The company sells the disks in packages of 10 and offers a money-back guarantee that at most 1 of the 10 disks is defective. What proportion of packages is returned? If someone buys three packages, what is the probability that exactly one of them will be returned?

Solution

- X = number of defective in a random package
- X =Bino$(10, .01)$
- $P(return) = P(X > 1)$
  $= 1 - P(X = 0) - P(X = 1) \approx .005$

Solution

- X = number of defective in a random package
- X =Bino$(10, .01)$
- $P(return) = P(X > 1)$
  $= 1 - P(X = 0) - P(X = 1) \approx .005$

Solution

- X = number of defective in a random package
- X =Bino$(10, .01)$
- $P(return) = P(X > 1)$
  $= 1 - P(X = 0) - P(X = 1) \approx .005$

Solution

- Y = number of returned package in 3 packages
- Y =Bino$(3, .005)$
- $P(Y = 1) = .015$

Solution

- Y = number of returned package in 3 packages
- Y =Bino$(3, .005)$
- $P(Y = 1) = .015$

Solution

- Y = number of returned package in 3 packages
- Y =Bino$(3, .005)$
- $P(Y = 1) = .015$

Properties

- $X_1, \ldots, X_n$ independent $\mathrm{Ber}(p)$
- $X = \sum_{i=1}^{n} X_i \rightarrow X$ is $\mathrm{Bino}(n, p)$
- $E(X) = \sum_{i=1}^{n} E(X_i) = np$
- $\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = np(1-p)$

Properties

- $X_1, \ldots, X_n$ independent $\text{Ber}(p)$
- $X = \sum_{i=1}^{n} X_i \rightarrow X$ is $\text{Bino}(n, p)$
- $E(X) = \sum_{i=1}^{n} E(X_i) = np$
- $\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(X_i) = np(1-p)$

Properties

- $X_1, \ldots, X_n$ independent $\text{Ber}(p)$
- $X = \sum_{i=1}^{n} X_i \to X$ is $\text{Bino}(n, p)$
- $E(X) = \sum_{i=1}^{n} E(X_i) = np$
- $\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(X_i) = np(1-p)$

Sum of Binomial

- $X \sim \text{Bino}(n_1, p)$, $Y \sim \text{Bino}(n_2, p)$
- Think of X as sum of n_1 independent $\text{Ber}(p)$, Y as sum of n_2 independent $\text{Ber}(p)$
- Then $X + Y$ is sum of $n_1 + n_2$ independent $\text{Ber}(p)$
- $X + Y \sim \text{Bino}(n_1 + n_2, p)$.

Sum of Binomial

- $X \sim$ Bino(n_1, p), $Y \sim$ Bino(n_2, p)
- Think of X as sum of n_1 independent Ber(p), Y as sum of n_2 independent Ber(p)
- Then $X + Y$ is sum of $n_1 + n_2$ independent Ber(p)
- $X + Y \sim$ Bino$(n_1 + n_2, p)$.

Sum of Binomial

- $X \sim \text{Bino}(n_1, p)$, $Y \sim \text{Bino}(n_2, p)$
- Think of X as sum of n_1 independent $\text{Ber}(p)$, Y as sum of n_2 independent $\text{Ber}(p)$
- Then $X + Y$ is sum of $n_1 + n_2$ independent $\text{Ber}(p)$
- $X + Y \sim \text{Bino}(n_1 + n_2, p)$.

Computing cdf

- Want to compute the cdf of $X \sim \text{Bino}(n, p)$
- Note that $P(X = k+1) = \frac{p}{1-p}\frac{n-k}{k+1}P(X = k)$
- Start with $P(X = 0) = (1-p)^n$, calculate $P(X = 1), P(X = 2), \ldots$

Computing cdf

- Want to compute the cdf of $X \sim \text{Bino}(n, p)$
- Note that $P(X = k+1) = \frac{p}{1-p}\frac{n-k}{k+1}P(X = k)$
- Start with $P(X = 0) = (1-p)^n$, calculate $P(X = 1), P(X = 2), \ldots$

Computing cdf

- Want to compute the cdf of $X \sim \text{Bino}(n, p)$
- Note that $$P(X = k + 1) = \frac{p}{1-p}\frac{n-k}{k+1}P(X = k)$$
- Start with $P(X = 0) = (1 - p)^n$, calculate $P(X = 1), P(X = 2), \ldots$

$X = \text{Bino}(6, .4)$

$$P\{X=0\} = (.6)^6 = .0467$$

$$P\{X=1\} = \frac{4}{6}\frac{6}{1}P\{X=0\} = .1866$$

$$P\{X=2\} = \frac{4}{6}\frac{5}{2}P\{X=1\} = .3110$$

$$P\{X=3\} = \frac{4}{6}\frac{4}{3}P\{X=2\} = .2765$$

$$P\{X=4\} = \frac{4}{6}\frac{3}{4}P\{X=3\} = .1382$$

$$P\{X=5\} = \frac{4}{6}\frac{2}{5}P\{X=4\} = .0369$$

$$P\{X=6\} = \frac{4}{6}\frac{1}{6}P\{X=5\} = .0041.$$

Poisson Random variables

Definition

Discrete RV X is called **Poisson RV** with parameter λ if the pmf is

$$P(X = i) = e^{-\lambda}\frac{\lambda^i}{i!}, \quad i = 0, 1, \ldots$$

Denote $X \sim \text{Poisson}(\lambda)$

Idea

- Can think of Poisson RV as Binomial RV with large n and small p
- Let $X \sim \text{Bino}(n, p)$ and let $\lambda = np$
- then the pmf of X is approximately the same as pmf of Poisson(λ) if n is large

Idea

- Can think of Poisson RV as Binomial RV with large n and small p
- Let $X \sim$ Bino(n, p) and let $\lambda = np$
- then the pmf of X is approximately the same as pmf of Poisson(λ) if n is large

$$
\begin{aligned}
P\{X = i\} &= \frac{n!}{(n-1)!i!} p^i (1-p)^{n-i} \\
&= \frac{n!}{(n-1)!i!} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} \\
&= \frac{n(n-1)\dots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}
\end{aligned}
$$

For fixed i, large n and small p:

$$\left(1-\frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \quad \left(1-\frac{\lambda}{n}\right)^i \approx 1$$

$$\frac{n(n-1)\ldots(n-i+1)}{n^i} \approx 1$$

then $P(X=i) \approx e^{-\lambda}\frac{\lambda^i}{i!}$

Example

- Let X is the number of people visit a store in one day
- Each person passing by the store has positive chance to visit the store
- Think of each person as one Bernoulli RV: visit = success
- then X is Binomial with n large and p small = Poisson $(\lambda = np)$

Example

- Let X is the number of people visit a store in one day
- Each person passing by the store has positive chance to visit the store
- Think of each person as one Bernoulli RV: visit = success
- then X is Binomial with n large and p small = Poisson $(\lambda = np)$

Example

- Let X is the number of people visit a store in one day
- Each person passing by the store has positive chance to visit the store
- Think of each person as one Bernoulli RV: visit = success
- then X is Binomial with n large and p small = Poisson $(\lambda = np)$

Example

- Let X is the number of people visit a store in one day
- Each person passing by the store has positive chance to visit the store
- Think of each person as one Bernoulli RV: visit = success
- then X is Binomial with n large and p small = Poisson $(\lambda = np)$

Examples

- Number of typos in a page of a book
- Number of airplane accidents in a year
- Number of students in probability class did their homework !!!!!
- Number of people lives longer than 100 years

Examples

- Number of typos in a page of a book
- Number of airplane accidents in a year
- Number of students in probability class did their homework !!!!!
- Number of people lives longer than 100 years

Examples

- Number of typos in a page of a book
- Number of airplane accidents in a year
- Number of students in probability class did their homework !!!!!
- Number of people lives longer than 100 years

Examples

- Number of typos in a page of a book
- Number of airplane accidents in a year
- Number of students in probability class did their homework !!!!!
- Number of people lives longer than 100 years

Properties

- $E(X) = \lambda$
- $\mathrm{Var}(X) = \lambda$
- So think of λ as average number of success

Example

Suppose that the average number of accidents occurring weekly on a particular stretch of a highway equals 3. Calculate the probability that there is at least one accident this week.

Solution

X = number of accident =Pois(3)

$$P(X \geq 1) = 1 - P(X = 0)$$

$$= 1 - e^{-3}\frac{3^0}{0!} \qquad = 1 - e^{-3} = .9502$$

Sum of Poisson

Let $X \sim \text{Pois}(\lambda_1)$ and $Y \sim \text{Pois}(\lambda_2)$. Then

$$X + Y \sim \text{Pois}(\lambda_1 + \lambda_2)$$

Calculating Poisson

$$\frac{P\{X = i+1\}}{P\{X = i\}} = \frac{e^{-\lambda}\lambda^{i+1}/(i+1)!}{e^{-\lambda}\lambda^{i}/i!} = \frac{\lambda}{i+1}$$

- Start with $P(X = 0) = e^{-\lambda}$
- $P(X = 1) = \lambda P(X = 0)$,
  $P(X = 2) = \frac{\lambda}{2} P(X = 1)$,...

Hyper Geometric Random variables

Definition

A box contains N blue balls and M red balls. Choose randomly n balls without replacement. Let X be the number of chosen blue balls. Then pmf of X is

$$P(X = i) = \frac{\binom{N}{i}\binom{M}{n-i}}{\binom{N+M}{n}}$$

X is called **hyper geometric** (N, M, n) RV.

Properties

- Denote $p = \frac{N}{N+M}$
- $E(X) = np$
- $\mathrm{Var}(X) = np(1-p)\left[1 - \dfrac{n-1}{N+M-1}\right]$

Relation with Binomial

Let X and Y be independent binomial random variables having respective parameters (n, p) and (m, p). Then

$$P(X = i | X + Y = k) = \frac{\binom{n}{i}\binom{m}{k-i}}{\binom{n+m}{k}}$$

The *conditional distribution* of X given $X + Y = k$ is hyper geometric (n, m, k).

Application

- Want to know the number N of a certain animal in one area.
- Catch a sample of r individuals, marked them then release
- After a few days, catch another sample of n individuals
- $X =$ number of marked individual in second sample.

Application

- Want to know the number N of a certain animal in one area.
- Catch a sample of r individuals, marked them then release
- After a few days, catch another sample of n individuals
- $X =$ number of marked individual in second sample.

Application

- Want to know the number N of a certain animal in one area.
- Catch a sample of r individuals, marked them then release
- After a few days, catch another sample of n individuals
- X = number of marked individual in second sample.

Application

- Want to know the number N of a certain animal in one area.
- Catch a sample of r individuals, marked them then release
- After a few days, catch another sample of n individuals
- $X =$ number of marked individual in second sample.

- Assuming two sample are random
- then X is hyper geometric $(r, N-r, n)$
- Suppose $X = i$ then estimate $N = nr/i$

$$P(N = nr/i) = P(X = i)$$

- Assuming two sample are random
- then X is hyper geometric $(r, N-r, n)$
- Suppose $X = i$ then estimate $N = nr/i$

$$P(N = nr/i) = P(X = i)$$

Homework 6

Chapter 5: 2, 6, 12, 14, 18