Discrete Random Variables

January 14, 2021

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

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

Denote $X \sim Ber(p)$

Property of $X \sim Ber(p)$

- $E(X) = p$
- $Var(X) = p(1 - p)$

Use Bernoulli RV to

model generic probabilistic situations with just two outcomes:

- The state of a telephone at a given time that can be either free or busy.
- A person who can be either healthy or sick with a certain disease.

Use Bernoulli RV to

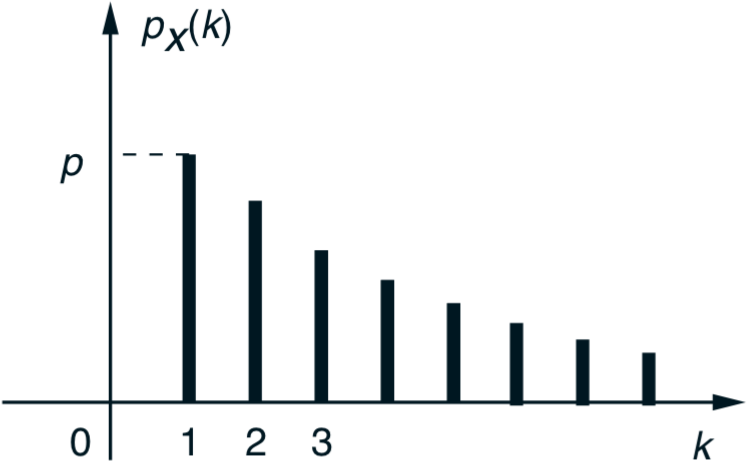
construct more complicated RV by combining multiple Bernoulli RV

- toss a biased coin n time
- $P(Head) = p$, $P(Tail) = 1 - p$
- X: number of tosses until a head comes up for the first time
- pmf of X

$$p(k) = (1-p)^{k-1}p,\ k \geq 1$$

- X is Geometric with parameters p, denoted by $X \sim Geo(p)$

Repeat independent Bernoulli trials until the first success



$p_X(k)$
p
0
1
2
3
k

- $E(X) = \frac{1}{p}$
- $Var(X) = \frac{1-p}{p^2}$

The chance that a bit transmitted through a digital transmission channel is received in error is .1. Also, assume that the transmission trials are independent. X denote the number of bits transmitted until the first error. Determine $P(X = 5)$.

- $X \sim Geo(.1)$
-
$$P(X = 5) = (.9)^4(.1) \approx .066$$

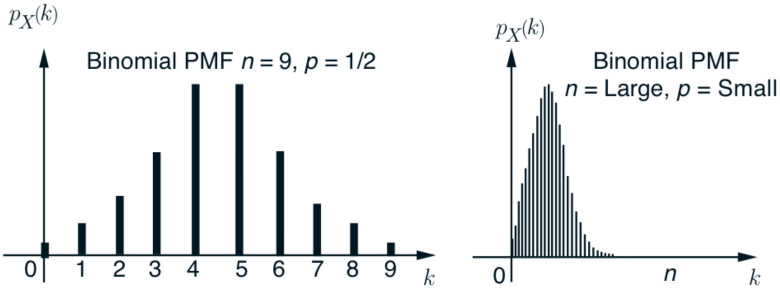
At a “busy time,” a telephone exchange is very near capacity, so callers have difficulty placing their calls. It may be of interest to know the number of attempts necessary in order to make a connection. Suppose that we let $p = 0.05$ be the probability of a connection during a busy time. We are interested in knowing the probability that 5 attempts are necessary for a successful call

- toss a biased coin n time
- $P(Head) = p$, $P(Tail) = 1 - p$
- X: number of heads in the n-toss sequence
- pmf of X

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k},\ 0 \le k \le n$$

- X is Binomial with parameters (n, p), denoted by $X \sim Bino(n, p)$

- Counting the number of success in an experiment consisting of n independent Bernoulli trials
- sum of n independent and identical Bernouilli RV



If $p = 1/2$, the pmf is symmetric around $n/2$. Otherwise, the pmf is skewed towards 0 if $p < 1/2$, and towards n if $p > 1/2$.

- $E(X) = np$
- $Var(X) = np(1-p)$

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that

1. exactly 5 survive?
2. from 3 to 8 survive
3. at least 10 survive

Solution

X: number of survie among 15 patients $\Rightarrow X \hookrightarrow Bino(15, .4)$

1. $P(X = 5) = \binom{15}{5}(.4)^5(.6)^{15-5}$
2. $P(3 \le X \le 8) = \sum_{k=3}^{8} \binom{15}{5}(.4)^k(.6)^{15-k}$
3. $P(X \ge 10) = \sum_{k=10}^{15} P(X = k) = \sum_{k=10}^{15} \binom{15}{5}(.4)^k(.6)^{15-k}$

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that

1. exactly 5 survive?
2. from 3 to 8 survive
3. at least 10 survive

Solution

X: number of survie among 15 patients $\Rightarrow X \hookrightarrow Bino(15, .4)$

1. $P(X=5) = \binom{15}{5}(.4)^5(.6)^{15-5}$
2. $P(3 \leq X \leq 8) = \sum_{k=3}^{8} \binom{15}{5}(.4)^k(.6)^{15-k}$
3. $P(X \geq 10) = \sum_{k=10}^{15} P(X=k) = \sum_{k=10}^{15} \binom{15}{5}(.4)^k(.6)^{15-k}$

A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%.

1. The inspector randomly picks 20 items from a shipment. What is the probability that there will be at least one defective item among these 20?
2. Suppose that the retailer receives 10 shipments in a month and the inspector randomly tests 20 devices per shipment. What is the probability that there will be exactly 3 shipments each containing at least one defective device among the 20 that are selected and tested from the shipment?

The chance that a bit transmitted through a digital transmission channel is received in error is .1. Also, assume that the transmission trials are independent. Let X the number of bits in error in the next four bits transmitted. Determine $P(X = 2)$.

Solution

- $X \sim Bino(4, .1)$
-
$$P(X = 2) = \binom{4}{2}(.1)^2(.9)^2 \approx .0486$$

- error in transmission a bit: p
- transmit n bits
- X: number of bits in error
- pmf

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

- $n \to \infty$, painful to compute $p(k)$
- Solution?
- $\lambda = np$

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

- $n \to \infty$, painful to compute $p(k)$
- Solution?
- $\lambda = np$

$$
\begin{aligned}
P(X = k) &= \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\
&= \frac{n!}{(n-k)!k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)\ldots(n-k+1)}{n^i} \frac{\lambda^k}{k!} \frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k} \\
&\xrightarrow{n \to \infty} e^{-\lambda} \frac{\lambda^k}{k!}
\end{aligned}
$$

For n very large

- $\frac{n(n-1)\ldots(n-k+1)}{n^i} \approx 1$
- $\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}$
- $\left(1 - \frac{\lambda}{n}\right)^k \approx 1$

- $n = 100, p = .01$
- Prob of 5 error in 100 transmission
  - exact value
$$\frac{100!}{95!5!}(.01)^5(.99)^{95} \approx .00290.$$
  - approximation by Poisson RV: $\lambda = np = 1$
$$e^{-1}\frac{1^5}{1!} \approx 00306$$

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 $X \sim$Poisson(λ)

- approximate $Bin(n, p)$ when n is large and p is small
- The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region -*no memory property*.

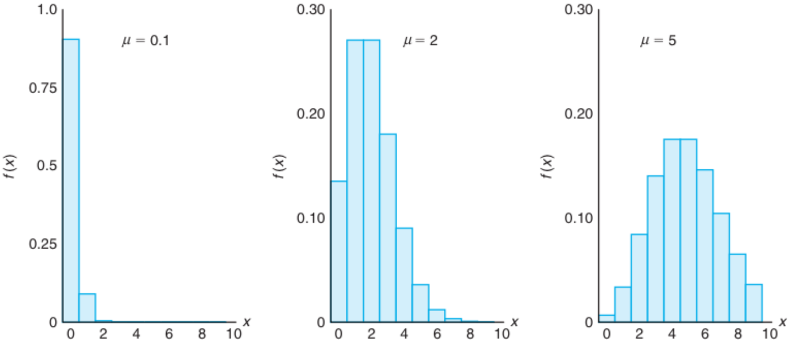
- The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval or the size of the region.
- The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

- Number of people stopping by a small shop on a crowded street in one day
- Number of students forgetting the exam date
- Number of flights having accident in one year
- Number of students in Probability class do their homework

- Number of people stopping by a small shop on a crowded street in one day
- Number of students forgetting the exam date
- Number of flights having accident in one year
- Number of students in Probability class do their homework

- Number of people stopping by a small shop on a crowded street in one day
- Number of students forgetting the exam date
- Number of flights having accident in one year
- Number of students in Probability class do their homework

- Number of people stopping by a small shop on a crowded street in one day
- Number of students forgetting the exam date
- Number of flights having accident in one year
- Number of students in Probability class do their homework



the form of the Poisson distribution becomes more and more symmetric, even bell-shaped, as the λ grows large

Property of $X \sim Poiss(\lambda)$

- $E(X) = \lambda$
- $Var(X) = \lambda$

parameter λ represents the expected number of events which occur per unit of time or area

During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution

- X: number of particles entering the counter in a millisecond
- $X \sim Pois(4)$
- $P(X = 6) = e^{-4}\frac{4^6}{6!}$

During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution

- X: number of particles entering the counter in a millisecond
- $X \sim Pois(4)$
- $P(X = 6) = e^{-4}\frac{4^6}{6!}$

Ten is the average number of oil tankers arriving each day at a certain port. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

- average number of flaws per millimeter of wire is 3.4, then the
- average number of flaws in 10 millimeters of wire is 34, and the
- average number of flaws in 100 millimeters of wire is 340

Example

The number of failures of a testing instrument from contamination particles on the product is a Poisson random variable with a mean of 0.02 failure per hour. What is the probability that the instrument does not fail in an eight-hour shift?

Solution

- X: number of failures in an eight - hour shift
- Average number of failures in an eight - hour shift: $(0.02) \times 8 = 0.16$
- $X \sim Pois(.16)$
- $P(X = 0) = e^{.16} \frac{(.16)^0}{0!}$

The number of failures of a testing instrument from contamination particles on the product is a Poisson random variable with a mean of 0.02 failure per hour. What is the probability that the instrument does not fail in an eight-hour shift?

Solution

- X: number of failures in an eight - hour shift
- Average number of failures in an eight - hour shift: $(0.02) \times 8 = 0.16$
- $X \sim Pois(.16)$
- $P(X = 0) = e^{.16} \frac{(.16)^0}{0!}$

What is the probability of at least one failure in a 24-hour day?

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.

- 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.

- 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.

- 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.

- 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.

A homeowner plants 6 bulbs selected at random from a box containing 5 tulip bulbs and 4 daffodil bulbs. What is the probability that he planted 2 daffodil bulbs and 4 tulip bulbs?

Solution

- X: number of tulip bulbs in a sample of 6 bulbs
- X is hyper geometric with parameters $(5, 4, 6)$
- $P(X = 4) = \frac{\binom{5}{4}\binom{4}{2}}{\binom{9}{6}}$

A homeowner plants 6 bulbs selected at random from a box containing 5 tulip bulbs and 4 daffodil bulbs. What is the probability that he planted 2 daffodil bulbs and 4 tulip bulbs?

Solution

- X: number of tulip bulbs in a sample of 6 bulbs
- X is hyper geometric with parameters $(5, 4, 6)$
- $P(X = 4) = \frac{\binom{5}{4}\binom{4}{2}}{\binom{9}{6}}$