Introduction to Probability
Chapter:5 Discrete Distributions

Dr Nitin Gupta
Department of Mathematics
Indian Institute of Technology Kharagpur,
Kharagpur - 721 302, INDIA.

Outline

References

1. Probability and statistics in engineering by Hines et al (2003) Wiley.
2. Mathematical Statistics by Richard J. Rossi (2018) Wiley.
3. Probability and Statistics with reliability, queuing and computer science applications by K. S. Trivedi (1982) Prentice Hall of India Pvt. Ltd.

Bernoulli trail $X \sim Ber(p)$

- Consider a trail in which we can have success with probability p and failure with probability q, such that $p + q = 1$
- Let X counts the number of successess. Then X can take values $0, 1$.
- PMF is $P(X = 1) = p$ and $P(X = 0) = q$.
- MGF is $M(t) = q + pe^t$.
- $E(X) = p$ and $Var(X) = E(X^2) - (E(X))^2 = p - p^2 = pq$

Binomial distribution $X \sim Bin(n, p)$

- Let n independent bernoulli trails are conducted.
- Let X counts the number of successess in n trails. Then X can take values $0, 1, \ldots, n$.
- PMF is $P(X = x) = \binom{n}{x}\ p^x\ q^{n-x},\ x = 0, 1, \ldots, n;\ 0,$ otherwise.
- MGF is $M(t) = (q + pe^t)^n$.
- $E(X) = np$; $E(X^2) = n(n-1)p^2 + np$; and $Var(X) = npq$.

Binomial distribution

Example

Suppose that we are throwing pair of dice 10 times and asking for the probability of double five. Here the probability of getting double five is $p = \frac{1}{36}$. If X denote the number of times getting a double five from rolling two dices. Then $X \sim Bin(n, p)$, $n = 10$, $p = \frac{1}{36}$ Then Required probability $P(X > 0) = 1 - P(X = 0) = 1 - \binom{10}{0} \left(\frac{1}{36}\right)^0 \left(\frac{35}{36}\right)^{10}$

Binomial distribution

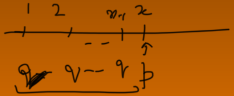
Example

Suppose that a football player makes 30% of his shot attempts. If the players shoots 10 shots in a game, and X is the number of shots made. Assuming that the shots are independent, find the mean of random variable X. Clearly $X \sim Bin(n, p)$, $n = 10$, $p = 0.30$. Then mean of random variable X is

$$E(X) = np = 10 \times 0.30 = 3.$$

Geometric distribution $X \sim Geo(p)$

- Let independent bernoulli trails are conducted till we get a success.
- Let X counts the number of trails to get a success. Then X can take values $1, 2, \dots \infty$.
- PMF is $P(X = x) = q^{x-1}p, \ x = 1, 2, \dots \infty; \ 0, \ $ otherwise. $p+q=1$
- MGF is $M(t) = \frac{pe^t}{1-qe^t}$.
- $E(X) = \frac{1}{p}$; and $Var(X) = \frac{q}{p^2}$.



Geometric distribution

Example

In each trial the probability of success is 0.25, in how many trails we expect first success?

Solution: Let X denote the number of trails to have first success.
$X \sim Geo(p)$, where $p = 0.25$.Then expected number of trials are
$E(X) = \frac{1}{p} = 4.$ ✓

$X \sim Geo(p)$, $p = 0.25$

Geometric distribution

- Geometric distribution has memoryless property, i.e.,

$$P(X > x + y | X > y) = P(X > x)$$

Example

Let getting a '1' is a success in roll of a die. We will roll a fair die until we observe success. Let we have already rolled the die ten times without a success. The probability that more than two additional tosses are required to have a success is

$X \sim Geo(p)$, $p = \frac{1}{6}$, $q = \frac{5}{6}$

$X \in \{1, 2, 3, \ldots\}$

$$\begin{aligned} P(X > 12 | X > 10) &= P(X > 2) = 1 - P(X \le 2) \\ &= 1 - P(X = 1) - P(X = 2) \\ &= 1 - \frac{1}{6} - \frac{1}{6} \times \frac{5}{6} \\ &= \frac{25}{36}. \end{aligned}$$

$\because P(X = x) = pq^{x-1}$

Negative Binomial distribution $X \sim NB(r, p)$

- Let independent bernoulli trails are conducted till we get r successes.
- Let X counts the number of trails to get r success. Then X can take values $r, r+1, \ldots \infty$. $X \in \{r, r+1, r+2, \ldots\}$
- PMF is $P(X = x) = \binom{x-1}{r-1} p^r q^{x-r}$, $x = r, r+1, \ldots \infty$; 0, otherwise.
- MGF is $M(t) = \left(\frac{pe^t}{1-qe^t}\right)^r$.
- $E(X) = \frac{r}{p}$; and $Var(X) = \frac{rq}{p^2}$.

x

$x-1$

p

$r-1 \quad x-r \quad 1$

$p(x) = \binom{x-1}{r-1} \quad p^{r-1} \quad q^{x-r} \quad p$

$X \sim NB(r, p)$

Negative Binomial distribution

Example

In ODI cricket match series between two teams A and B, the team who wins three games will be the winner. Suppose that the team A has probability 0.60 of winning over team B. We want to find the probability that team A will win the series in 5 games. Let X counts the number of games required by team A to win the series. Then $X \sim NB(r, p)$, $r = 3$, $p = 0.60$. Now the required probability is $P(X = 5) = \binom{5-1}{3-1}$ $(0.60)^3(1 - 0.60)^{5-3}$.

$X \in \{3, 4, 5, \ldots\}$

$P(X = x) = \binom{x-1}{r-1} p^r q^{x-r}$

Negative Binomial distribution

Example

A factory produces components for computers. Let 5% of components are defective. We need to find 3 non-defective components for our 3 new computers. Components are tested until 3 non-defectives are found. What is the probability that more than 5 components will be tested?

Solution: Let X be the number of components tested for getting r non-defectives. Then $X \sim NB(r, p)$, $r = 3$, $p = 0.95, q = 1 - p$. Then required probability is

$X \in \{3, 4, 5, \ldots\}$ $q = 0.05$

$$
\begin{aligned}
P(X > 5) &= 1 - P(X \leq 5) \\
&= 1 - [P(X = 3) + P(X = 4) + P(X = 5)] \\
&= 1 - [p^3 + 3p^3 q + 6p^3 q^2] \\
&= 0.0012.
\end{aligned}
$$

$P(X=x) = \binom{x-1}{r-1} p^r q^{x-r}$

Hypergeometric distribution

A box contains N balls. M are drawn at random, marked and returned to the box. Next n balls are drawn at random from box and then the marked balls are counted. Let X denote the number of marked balls. Then PMF is

$$P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \ \max(0, M+n-N) \leq x \leq \min(M, n).$$

$x \geq 0$, $N - M \geq n - x$

$x \geq M + n - N$

$E(X) = \frac{n}{N}M$, $Var(X) = \frac{nM}{N^2(N-1)}(N-M)(N-n)$.

Summary

The discrete distributions were presented in this chapter. These distributions has wide use in the field of engineering, science and management. Particulaly we intoduced the binomial, geometric, negative Binomial and hypergeometric distribution.