Lecture
Discrete Distributions

Probability and Statistics

Instructor: Dr. Shabbir Ahmad

Discrete Distribution

DEFINITION:

A discrete distribution is a distribution of data in statistics that has discrete values. Discrete values are countable, finite, non-negative integers, such as 1, 10, 15, etc.

UNDERSTANDING DISCRETE DISTRIBUTION:

Distribution is a statistical concept used in data research. Those seeking to identify the outcomes and probabilities of a particular study will chart measurable data points from a data set, resulting in a probability distribution diagram. There are many types of probability distribution diagram shapes that can result from a distribution study, such as the normal distribution ("bell curve").

Types of discrete distribution

The most common discrete distributions used by statisticians or analysts include the

- Binomial distribution
- Poisson distributions
- Bernoulli distribution
- Multinomial distribution
- Geometric distribution
- Hypergeometric distribution

The Binomial Distribution

- The binomial distribution is based on a Bernoulli trial, which is a random experiment in which
- there are only two possible outcomes: success (S) and failure (F). We conduct the Bernoulli Trial.

  p q
- If the probability of success is p then the probability of failure must be $1 - p = q$

Understanding Binomial Distribution

The binomial distribution is a common discrete distribution used in statistics, as opposed to a continuous distribution, such as the normal distribution. This is because the binomial distribution only counts two states, typically represented as 1 (for a success) or 0 (for a failure) given a number of trials in the data. The binomial distribution, therefore, represents the probability for x successes in n trials, given a success probability p for each trial.

Real life examples:

The binomial distribution is often used in social science statistics as:

- Whether a Republican or Democrat will win an upcoming election
- An individual will die within a specified period of time

Many instances of binomial distributions can be found in real life:

- If a new drug is introduced to cure a disease, it either cures the disease (it’s successful) or it doesn’t cure the disease (it’s a failure).
- If you purchase a lottery ticket, you’re either going to win money, or you aren’t. Basically, anything you can think of that can only be a success or a failure can be represented by a binomial distribution.

PROPERTIES

The Binomial model has three defining properties:

- The number of observations n is fixed.
- Each observation is independent.
- Each observation represents one of two outcomes ("success" or "failure").
- The probability of "success" p is the same for each outcome.

PROBAILITY DENSITY FUNCTION

- The formula for the binomial probability density function is

$N(\mu, \sigma)$

$B(n, p)$

Parameter

$$P(\ x\ ;p\ ,n\ ) = \binom{n}{x}(p)^{x}(1-p)^{(n-x)}$$

For x=0,1,2, … ,n

${}^{n}C_{x}$

Conditions for binomial distribution usage

It is used under the condition:

- The variable is discrete
- Statistical independence is assumed
- The exponent power is finite and small
- For symmetrical distribution p=q and for asymmetrical $p \neq q$

EXAMPLES

QUESTION NO 1:

A (blindfold) marksman finds that on the average he hits the target 4 times out of 5.if he fires 4 shots, what is the probability of More than 2 hits?

$p = 4/5 = 0.8$

n

$q = 1 - p$
$= 0.2$

$(x=3)$ $\binom{4}{3}(0.8)^3(0.2)^{4-3}$

$(x=4)$ $\binom{4}{4}(0.8)^4(0.2)^{4-4}$

SOLUTION:

Here, $n = 4$, $p = 0.8$, $q = 0.2$.

Let X = number of hits.

Let x_0 = no hits, $x_1 = 1$ hit, $x_2 = 2$ hits, etc.

(a) $P(X) = P(x_3) + P(x_4)$

$$= C_3^4 (0.8)^3 (0.2)^1 + C_4^4 (0.8)^4 (0.2)^0$$

$$= 4(0.8)^3 (0.2) + (0.8)^4$$

$$= 0.8192$$

Question no 2

In the old days there was a probability of 0.8 of success in ~~any~~ 10 attempt 5 to make a telephone call (This often depended on the importance of the person making the call, or the operator's curiosity!)

Calculate the probability of having 7 successes in attempts

$P = 0.8$

$q = 0.2$

$\binom{10}{7}(0.8)^7(0.2)^{10-7}$

SOLUTION:

Probability of success $p = 0.8$, so $q = 0.2$.

$X =$ success in getting through.

n

Probability of 7 successes in 10 attempts:

$$\text{Probability} = P(X = 7)$$

$$= C_7^{10}(0.8)^7(0.2)^{10-7}$$

$$= 0.20133$$

POISSON DISTRIBUTION

- In statistics, a Poisson distribution is a probability distribution that can be used to show how many times an event is likely to occur within a specified period of time.
- In other words, it is a count distribution. Poisson distributions are often used to understand independent events that occur at a constant rate within a given interval of time. It was named after French mathematician *Siméon Denis Poisson*.

REAL LIFE EXAMPLES

- the number of hungry persons entering McDonald's restaurant per day.
- Estimating the number of car crashes in a city of a given size.
- The number of phone calls received at switchboard per minute
- The number of defective parts per batch shipped from the factory for assembly.
- The number of machine breaking down any one day.
- the number of network failures per day.
- the number of birth, deaths, marriages, suicides over a given period of time.

PROPERTIES

- The number of outcomes of a trial occurring in one time interval
- The probability of a single outcome during a short time interval does not depend on the number of outcomes occurring outside this time interval or region
- The probability that more than one outcome will occur in such a short time interval or small region is negligible.

PROBABILITY DENSITY FUNCTION

The Poisson distribution is used to model the number of events occurring within a given time interval so the formula for the Poisson probability density function is:

mean

$$p(x;\lambda) = \frac{e^{-\lambda}\ \lambda^{x}}{x!}$$

For $x = 0,1,2,\ldots$

CHARACTERISTICS

- It is the limiting form of binomial distribution when n is large and p (or q) is small.
- Here p (or q) is very close to zero or unity
- If p is very close to zero the distribution is j shaped or unimodal
- As it consists of a single parameter the entire distribution can be obtained by knowing the mean only

EXAMPLE 1

The average number of major storms in your city is 2 per year. What is the probability that exactly 3 storms will hit your city next year?

Step 1: Figure out the components you need to put into the equation.

- $\mu = 2$ (average number of storms per year, historically)
- $x = 3$ (the number of storms we think might hit next year)
- $e = 2.71828$ (e is Euler's number, a constant)

$$\frac{e^{-\lambda}\lambda^{x}}{x!}$$

Step 2: Plug the values from Step 1 into the Poisson distribution formula:

- $P(x; \mu) = (e^{-\mu}) (\mu^{x}) / x!$
- $= (2.71828^{-2}) (2^{3}) / 3!$
- $= (0.13534) (8) / 6$
- $= 0.180$

The probability of 3 storms happening next year is 0.180, or 18%

ANY QUESTIONS