Topic:-

Hypergeometric Probability Distribution

Out Line

- Definition
  - Formula
  - Parameters
  - Characteristics
- Property
  - legtimate
  - Mean
  - variance
- Numerical
- Skewness
- Kurtosis
- Real life Example

3
1
4

* Definition:-

There are many experiments in which the condition of independence is violated and the probability of success does not remain constant for all trail. Such experiment are called hypergeometric experiment

Formula:-

$$P(X=x) = \frac{\binom{K}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

N = number of units in the set or population.
n = number of units in the sample.
k = number of success in the set or pop.

Parameters :-

The hypergeometric p.d has three parameter's N , n and k.

01:- a random sample of size "n" is drawn WOR from a finite population of "N" units.

02:- 'K' of the units one of one kind and the remaining "N-k" of another kind.

Characteristics:-

The hypergeometric distribution has the following characteristics.

01:- The are only 2 possible outcomes.

02:- The probability of a success is not the same on each trail without replacement, thus events are not independent.

Scanned with CamScanner

03:- In which population is finite.

04:- Trails are dependent.

* Properties

Legtimate

$$P(X=x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \binom{k}{x}\binom{N-K}{n-x}$$

for using the formula

$$(1+x)^a (1+x)^b = (1+x)^{a+b}$$

$$= \frac{1}{\binom{N}{n}} \binom{\cancel{K}+N-\cancel{K}}{\cancel{x}+n-\cancel{x}}$$

$$= \frac{\binom{N}{n}}{\binom{N}{n}} = 1$$

Hence proved.

Mean

$$E(x) = \sum x f(x)$$

$$= \sum_{x=0}^{n} x \cdot \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \cdot \sum_{x=0}^{n} x \binom{k}{x}\binom{N-k}{n-x}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \cancel{x} \frac{k(k-1)!}{\cancel{x}(x-1)!(k-x)!} \binom{N-k}{n-x}$$

$$= \frac{1}{\binom{N}{n}} \cdot k \binom{k-1}{x-1}\binom{N-k}{n-x}$$

Date:

Mathematical formula.

$$\because (1+x)^a + (1+x)^b = (1+x)^{a+b}$$

$$= \frac{1}{\binom{N}{n}} k \binom{\cancel{k}-1+N-\cancel{k}}{\cancel{x}-1+n-\cancel{x}}$$

$$= \frac{1}{\binom{N}{n}} k \binom{N-1}{n-1}$$

$$= \frac{1}{\frac{N(N-1)!}{n(n-1)!}} \cdot k \binom{N-1}{n-1}$$

$$= k \binom{\cancel{N-1}}{\cancel{n-1}} \frac{n\,\cancel{(n-1)!}}{N\,\cancel{(N-1)!}}$$

$$= \frac{nk}{N}$$

$$= np$$

H

Variance

$$E(X^2) = E[x(x-1) + E(x)]$$

$$E[x(x-1)] = \sum_{x=0}^{n} x(x-1) \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \frac{x(x-1)K!\binom{N-K}{n-x}}{x!(x-1)!\binom{N}{n}}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \frac{x\cancel{(x-1)}K!\binom{N-K}{n-x}}{\cancel{x(x-1)}(x-2)!(K-x)!}$$

$$= \frac{1}{\binom{N}{n}} \sum_{x=0}^{n} \frac{K(K-1)(K-2)!\binom{N-K}{n-x}}{(x-2)!(K-x)!}$$

$$= \frac{K(K-1)}{\binom{N}{n}} \sum_{x=2}^{n} \binom{K-2}{x-2}\binom{N-K}{n-x}$$

$$= \frac{k(k-1)}{\binom{N}{n}} \binom{N-\cancel{k}+\cancel{k}-2}{\cancel{x}-2+n-\cancel{x}}$$

$$= \frac{k(k-1)}{\binom{N}{n}} \binom{N-2}{n-2}$$

$$= \frac{k(k-1)}{\frac{N!}{n!(N-n)!}} \cdot \frac{(N-2)!}{(n-2)!(N-n)!}$$

$$= \frac{k(k-1)\cancel{(N-n)!}\, n(n-1)\cancel{(n-2)!}\,(N-2)!}{N!\,\cancel{(n-2)!}\,\cancel{(N-n)!}}$$

$$= \frac{k(k-1)\,n(n-1)(N-2)!}{N!}$$

$$= \frac{k(k-1)\,n(n-1)\cancel{(N-2)!}}{N(N-1)\cancel{(N-2)!}} = \frac{k(k-1)\,n(n-1)}{N(N-1)}$$

$$E(x^2) = E[x(x-1)] + E(x) \qquad \because E(x) = \frac{nk}{N}.$$

$$= \frac{k(k-1)\cdot n(n-1)}{N(N-1)} + \frac{nk}{N}$$

$$= \frac{nk}{N}\left(\frac{N-k}{N}\right)\left(\frac{N-x}{N-1}\right)$$

Date:

$$E(x^2) = \frac{k(k-1)\cdot n(n-1) + nk(N-1)}{N(N-1)}$$

Numerical

A box contains ten items, seven of which are good and three are defective. two items are selected (WOR); Compute the probability distribution for the number of defectives in the sample of two. Compute the mean and variance of this probability distribution. Is this mean equal to $\frac{nk}{N}$ and variance $\frac{nk(N-k)(N-n)}{N^2(N-1)}$.

Solution

The hypergeometric distribution is

$$P(X=x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \qquad x = 0, 1, 2, 3, \ldots n \text{ or } k$$

Hence $N = 10$, $n = 2$, $k = 3$, $x =$ num of d items then the possible value of x are 0, 1 and 2

Therefore.

$$P(X=x) = \frac{\binom{3}{x}\binom{10-3}{2-x}}{\binom{10}{2}} \quad x = 0, 1, 2$$

$$P(X=0) = \frac{\binom{3}{0}\binom{10-3}{2-0}}{\binom{10}{2}} = \frac{\binom{3}{0}\binom{7}{2}}{\binom{10}{2}} = \frac{7}{15}$$

$$P(X=1) = \frac{\binom{3}{1}\binom{10-3}{2-1}}{\binom{10}{2}} = \frac{\binom{3}{1}\binom{7}{1}}{\binom{10}{2}} = \frac{7}{15}$$

$$P(X=2) = \frac{\binom{3}{2}\binom{10-3}{2-2}}{\binom{10}{2}} = \frac{\binom{3}{2}\binom{7}{0}}{\binom{10}{2}} = \frac{1}{15}$$

Thus the probability distribution in tabular form for the computation of mean and variance of X is given as follows.

Date:

X	$P(X=x)=P(x)$	$xP(x)$	$x^2P(x)$
0	7/15	0	0
1	7/15	7/15	7/15
2	1/15	2/15	4/15
	$\Sigma P(x)=1$	$\Sigma xP(x)=9/15$	$\Sigma x^2P(x)=11/15$

Mean

$$E(X)=\Sigma xP(x)=9/15=0.6$$

$$=\frac{nk}{N}=\frac{(2)(3)}{10}=0.6$$

Variance

$$var(X)=\sigma^2=\Sigma x^2P(x)-[\Sigma xP(x)]^2=\frac{11}{15}-\left[\frac{9}{15}\right]^2=0.373$$

$$\frac{nk(N-k)(N-n)}{N^2(N-1)}=\frac{(2)(3)(7)(8)}{(100)(9)}=0.3733$$

Hence

$$E(X)=\frac{nk}{N}$$

And

$$var(X)=\frac{nk(N-k)(N-n)}{N^2(N-1)}$$

Scanned with CamScanner

Skewness:-

$$= \frac{(N-2K)(N-1)^{1/2}(N-2n)}{[nK(N-K)(N-n)]^{1/2}(N-2)}$$

Kurtosis:-

$$\frac{1}{nK(N-K)(N-n)(N-2)(N-3)} \cdot [\{(N-1)N^2(N(N+1)\} - 6K(N-K) - 6n(N-n) \; 6nK(N-K)(N-n)(5N-6)]$$

Real life Example

01:-

Five individuals from on animal population thought be near extinction in a certain region h been caught, tagged, and released to mix into the population. After they had an opportunity to mix. a random sample of 10 of theses animals is selected let X = the num of tagged animals in the second sample

Ali Ilyas Notes

Scanned with CamScanner

Date:

If there are actually 25 animals of this type in the region what is the $E(X)$. & $V(X)$?

Solutions:-

$n=10$, $M=5$ and $N=25$ so $p=\frac{5}{25}=.2$

and

$$E(X) = (10)(0.2) = 2 \qquad \because E(X) = np$$

$$V(X) = \frac{15}{24}(10)(0.2)(0.8) \qquad \because V(X) = npq$$

$$= 1$$

If the sampling was carried out with replacement

$$V(X) = 1.6$$