Hypergeometric Distribution:

A random variable X has a hypergeometric distribution iff its probability distribution is given by

$$h(x; n, N, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \quad ; \quad \begin{array}{l} x = 0, 1, 2, \dots, k \\ x \le k \; ; \; n-x \le N-k \end{array}$$

Proof:

CD's / infected CD's

Suppose a box of N oranges contains K sweet oranges.

CD's

A sample of n oranges (n < N) is selected. The probability

infected CD's

that the sample contains x sweet oranges (x ≤ K).

Let sample space S contains $\binom{N}{n}$ sample points. The population contains

$$\underbrace{\underbrace{S \cdot S \cdot \dots S}_{\text{K sweet (infected)}} \quad \underbrace{\bar{S} \cdot \bar{S} \cdot \dots \bar{S}}_{\text{N-k not sweet (infected)}}}_{\text{N oranges (CD's)}}$$

and the sample contains:

$$\underbrace{\underbrace{S \cdot S \cdots S}_{\text{x sweet (infected)}} \quad \underbrace{\bar{S} \cdot \bar{S} \cdots \bar{S}}_{\text{n-x not sweet (infected)}}}_{\text{n oranges CD's}}$$

CD's

The number of possible ways of selecting x oranges from K is $\binom{K}{x}$ and the number of ways of selecting n-x not

infected CD's

sweet oranges out of N-K is $\binom{N-K}{n-x}$ ways.

infected

Thus the probability distribution of selecting x sweet out

infected / infected CD's

of K and n-x not sweet out of N-K not sweet oranges by independent law of probability is

$$h(x; n, N, k) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \quad ; \quad \begin{array}{l} x = 0, 1, 2, \dots, k \\ x \le k, \; n-x \le N-k \end{array}$$

which is hypergeometric distribution with parameters n, N and K.

⑧

Moments of Hypergeometric Distribution:

By definition.

$$\mu_r' = E[X^r] = \sum_{x=0}^{n} x^r\, h(x;n,N,k) = \sum_{x=0}^{n} x^r \binom{k}{x}\binom{N-k}{n-x}\Big/\binom{N}{n}$$

at $r=1$

$$\mu_1' = \sum_{x=0}^{n} x\cdot\binom{k}{x}\binom{N-k}{n-x}\Big/\binom{N}{n} = \frac{1}{\binom{N}{n}}\cdot\sum_{x=1}^{n} x\cdot\binom{k}{x}\binom{N-k}{n-x}$$

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

$$= \frac{k}{\binom{N}{n}}\sum_{x=1}^{n}\frac{(k-1)!}{(x-1)!\,(k-x)!}\cdot\frac{(N-k)!}{[N-k-(n-x)]!\,(n-x)!}$$

Let $z = x-1 \Rightarrow x = z+1$ at $x=1 \Rightarrow z = 1-1 = 0$.
at $x=n \Rightarrow z = n-1$.

$$\therefore \mu_1' = \frac{k}{\binom{N}{n}}\sum_{z=0}^{n-1}\frac{(k-1)!}{z!\,(k-z-1)!}\cdot\frac{(N-k)!}{(N-k-\overline{n-z-1})!\,(n-z-1)!}$$

$$= \frac{k\cdot n(n-1)!\,(N-n)!}{N(N-1)!}\cdot\sum_{z=0}^{n-1}\frac{(k-1)!}{z!\,(k-z-1)!}\,\frac{(N-k)!}{(N-k-\overline{n-z-1})!\,(n-z-1)!}$$

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

$$= k\cdot\frac{n}{N}\cdot(1) \Rightarrow \boxed{\mu_1' = \frac{kn}{N}}$$

at $r=2$

$$\mu_2' = E(X^2) = E[X(X-1)+X] = E[X(X-1)] + E(X) = E[X(X-1)] + \frac{kn}{N}$$

$$\therefore E[X(X-1)] = \sum_{x=0}^{n} x(x-1)\cdot\binom{k}{x}\binom{N-k}{n-x}\Big/\binom{N}{n}$$

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

$$= \frac{n(n-1)(n-2)!\,(N-n)!}{N(N-1)(N-2)!}\sum_{x=2}^{n} x(x-1)\cdot\frac{k(k-1)(k-2)!}{(k-x)!\,x(x-1)(x-2)!}\cdot\frac{(N-k)!}{(n-x)!\cdot(N-k-\overline{n-x})!}$$

$$= \frac{n(n-1)}{N(N-1)}\cdot k(k-1)\frac{\sum_{x=2}^{n}\frac{(k-2)!}{(x-2)!\,(k-x)!}\cdot\frac{(N-k)!}{(n-x)!\,(N-k-\overline{n-x})!}}{\frac{(N-2)!}{(n-2)!\,(N-n)!}}$$

Let $z = x-2 \Rightarrow x = z+2$ at $x=2 \Rightarrow z = 2-2 = 0$.
at $x=n \Rightarrow z = n-2$

$$E[X(X-1)] = E[(z+2)(z+1)] = \frac{n(n-1)}{N(N-1)}\cdot k(k-1)\frac{\sum_{z=0}^{n-2}\frac{(k-2)!}{z!\,(k-z-2)!}\cdot\frac{(N-k)!}{(n-z-2)!\,(N-k-n)!}}{\frac{(N-2)!}{(n-2)!\,(N-n)!}}$$

9

Since the expression $\binom{k-2}{z}\binom{N-k}{n-z-2}/\binom{N-2}{n-2}$ is p.d.f for hypergeometric distribution for $z = 0, 1, 2, \dots, n-2$

Using the property of p.d.f $\sum_{z=0}^{n-2} \binom{k-2}{z}\binom{N-k}{n-z-2}/\binom{N-2}{n-2} = 1.$

$$\Rightarrow E[x(x-1)] = E[(z+1)(z+2)] = k \cdot (k-1) \cdot \frac{n(n-1)}{N(N-1)} \cdot (1)$$

$$\therefore \mu_2' = k(k-1)\frac{n(n-1)}{N(N-1)} + k\frac{n}{N}$$

$$= \frac{kn}{N}\left[(k-1)\frac{(n-1)}{(N-1)} + 1\right]$$

$$\text{Variance} = \mu_2' = E[x^2] - \{E[x]\}^2$$

$$= \frac{nk}{N}\left[\frac{(k-1)(n-1)}{(N-1)} + 1\right] - \frac{k^2 n^2}{N^2}$$

$$\boxed{Var(X) = \mu_2' = \frac{nk}{N}\left[\frac{(k-1)(n-1)}{(N-1)} + 1 - \frac{kn}{N}\right]}$$

Higher order moments can be obtained by proceeding the same manner.

Hypergeometric Experiment:

An experiment is called as hypergeometric experiment iff

a) The outcomes of the experiment classified into two categories, successes and failures

b) The probability of successes changes from trial to trial.

c) Successive trials are dependent

d) The experiment is repeated upto a fixed number of times.

9

Since the expression $\binom{k-2}{z}\binom{N-k}{n-z-2}\Big/\binom{N-2}{n-2}$ is p.d.f for hypergeometric distribution for $z = 0, 1, 2, \dots, n-2$

Using the property of p.d.f $\sum_{z=0}^{n-2} \binom{k-2}{z}\binom{N-k}{n-z-2}\Big/\binom{N-2}{n-2} = 1.$

$$\Rightarrow E[x(x-1)] = E[(z+1)(z+2)] = k\cdot(k-1)\cdot\frac{n(n-1)}{N(N-1)} \cdot (1)$$

$$\therefore \mu_2' = k(k-1)\frac{n(n-1)}{N(N-1)} + k\frac{n}{N}$$

$$= \frac{kn}{N}\left[(k-1)\frac{(n-1)}{(N-1)} + 1\right]$$

$$\text{Variance} = \mu_2' = E[x^2] - \{E[x]\}^2$$

$$= \frac{nk}{N}\left[\frac{(k-1)(n-1)}{(N-1)} + 1\right] - \frac{k^2 n^2}{N^2}$$

$$\boxed{Var(X) = \mu_2' = \frac{nk}{N}\left[\frac{(k-1)(n-1)}{(N-1)} + 1 - \frac{kn}{N}\right]}$$

Higher order moments can be obtained by proceeding the same manner

Hypergeometric Experiment:

An experiment is called as hypergeometric experiment iff

a) The outcomes of the experiment classified into two categories, successes and failures

b) The probability of successes changes from trial to trial.

c) Successive trials are dependent

d) The experiment is repeated upto a fixed number of times.

urrence relation of probabilities for hypergeometric dist

The p.d. of exactly x successes is

$$P(x) = \binom{K}{x}\binom{N-K}{n-x} \Big/ \binom{N}{n}$$

and the p.d. of x+1 successes is

$$P(x+1) = \binom{K}{x+1}\binom{N-K}{n-\overline{x+1}} \Big/ \binom{N}{n}$$

The ratio

$$\frac{P(x+1)}{P(x)} = \frac{\binom{K}{x+1}\binom{N-K}{n-\overline{x+1}}}{\binom{N}{n}} \cdot \frac{\binom{N}{n}}{\binom{K}{x}\binom{N-K}{n-x}}$$

simplifying we get

$$\boxed{P(x+1) = \frac{(n-x)(K-x)}{(x+1)(N-K-n+x+1)} P(x)}$$

Problem:

the post of Computer Programmer

Among the 120 applicants for a job only 80 are actually qualified. If 5 of these applicants are randomly selected for an "indepth interview"; find the probability that only 2 of the 5 will be qualified for the job.

Sol:

Let $x = 2$, $n = 5$, $N = 120$, $K = 80$

$$h(2; 5, 120, 80) = \frac{\binom{80}{2}\binom{40}{3}}{\binom{120}{5}} = 0.164$$

using Binomial distribution:

$$b(2; 5, \tfrac{2}{3}) = \binom{5}{2}\left(\tfrac{2}{3}\right)^2\left(1-\tfrac{2}{3}\right)^3 = 0.165$$

$p = \frac{80}{120} = \frac{2}{3}$

Walpole 3rd edition.

Solve
- Exercise page 165
- Exercise page 173
- Exercise page 180

What is the probability that a ~~waitress~~ waiter will refuse to serve alcoholic beverages to only 2 minors if she checks the I.D.s of 6 students from among 9 students of which 4 are not of legal age?

Sol.

There are 9 students out of which 6 does'nt reach the particular age. So sample space is $\binom{9}{6}$.

$$\frac{\binom{4}{2}\binom{5}{4}}{\binom{9}{6}} = 5/4$$ //.