Lecture-16

Hypergeometric Distribution

General discussion:

Suppose total number of items in a a bag $= N$

Total number of defective items (out of N) $= k$

Total number of items selected $= n$

lets discuss the probability that x out of n $(x \leq n)$ items selected is defective.

Now, total number of ways n items can be selected out of N items $= \begin{pmatrix} N \\ n \end{pmatrix}$.

Our requirement is x out of n are defective i.e. remaining $(n - x)$ are non-defective.

Number of ways x defective items can be selected from k defective items $= \begin{pmatrix} k \\ x \end{pmatrix}$.

Number of ways $n - x$ defective items can be selected from $N - k$ defective items $= \begin{pmatrix} N-k \\ n-x \end{pmatrix}$.

Probability of selecting x defectives

$$= \frac{all\ favorable\ cases}{all\ possible\ cases}$$

$$= \frac{\begin{pmatrix} k \\ x \end{pmatrix} \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

Definition:

Let $X =$ The number of successes in a random sample size n selected from N items of which k are labeled success and $(N - k)$ labeled failure.Then probability distribution of the above hypergeometric random variable is

$$f(x) = h(x; N, n, k) = \frac{\begin{pmatrix} k \\ x \end{pmatrix} \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

Such that $\max\{0, n-(N-k)\} \le x \le \min(n, k)$

(Q.30) A random committee of size 3 is selected from 4 doctors and 2 nurses. Write a formula for the probability distribution of the random variable X representing the number of doctors on the committee. Find $P(2 \le X \le 3)$.

Ans:

There are 4 doctors and 2 nurses

3 persons will be selected out of 4+2=6 persons

Let X: number of doctors in the committee which consists of 3 persons

So $x = 1, 2, 3$ ($x \neq 0$ why?)

$$P(X = x) = f(x) = \frac{\begin{pmatrix} 4 \\ x \end{pmatrix} \begin{pmatrix} 2 \\ 3-x \end{pmatrix}}{\begin{pmatrix} 6 \\ 3 \end{pmatrix}}$$

Now,

$P(2 \le X \le 3) = P(X = 2) + P(X = 3)$

$= \frac{\begin{pmatrix} 4 \\ 2 \end{pmatrix} \begin{pmatrix} 2 \\ 3-2 \end{pmatrix}}{\begin{pmatrix} 6 \\ 3 \end{pmatrix}} + \frac{\begin{pmatrix} 4 \\ 3 \end{pmatrix} \begin{pmatrix} 2 \\ 3-3 \end{pmatrix}}{\begin{pmatrix} 6 \\ 3 \end{pmatrix}} = \frac{4}{5}$

(Q.32) From a lot of 10 missiles, 4 are selected at random and fired. If the lot contains 3 defective missiles that will not fire, what is the probability that (a) all 4 will fire? (b) at most 2 will not fire?

Ans:

Total number of missiles = 10

Total number of defective missiles = 3

Hence, total number of non-defective missiles = 7

4 missiles will be fired

Let X: number of non-defective missiles fired

(a) $P(X=0) = \dfrac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \frac{1}{6}$

(b) At most 2 will not fire means 2 or more will fire

$$P(X \geq 2) = P(X=2) + P(X=3) + P(X=4)$$
$$= \frac{\binom{7}{2}\binom{3}{2}}{\binom{10}{4}} + \frac{\binom{7}{3}\binom{3}{1}}{\binom{10}{4}} + \frac{\binom{7}{4}\binom{3}{0}}{\binom{10}{4}} = \frac{29}{30}$$

Multivariate Hyper-geometric Distribution:

If N items can be partitioned into k cells $A_1, A_2, ...A_k$ with $a_1, a_2, ...a_k$elements, respectively then probability distribution of the random variables $X_1, X_2, ...X_k$representing the number of elements selected from $A_1, A_2, ...A_k$in a random sample of size n is

$$f(x_1, x_2, ...x_k; a_1, a_2, ...a_k, N, n) = \frac{\binom{a_1}{x_1}\binom{a_2}{x_2}.......\binom{a_k}{x_k}}{\binom{N}{n}}$$

With $\sum_{i=1}^{n} x_i = n, \ \sum_{i=1}^{n} a_i = N$

(Q.43) A foreign student club lists as its members 2 Canadians, 3 Japanese, 5 Italians, and 2 Germans. If a committee of 4 is selected at random, ?nd the probability that (a) all nationalities are represented; (b) all nationalities except Italian are represented.

Ans:

Total number of members = 2+3+5+2 12

Total number of members selected = 4

(a) All nationalities represented means one from each country.

One person can be selected from 2 Canadians in $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ways

One person can be selected from 3 Japanies in $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ ways

One person can be selected from 5 Italians in $\begin{pmatrix} 5 \\ 1 \end{pmatrix}$ ways

One person can be selected from 2 Germans in $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ways

Four persons can be selected from 12 persons in $\begin{pmatrix} 12 \\ 4 \end{pmatrix}$ ways

$$P(all\ nationalities\ are\ represented)$$
$$= \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}\begin{pmatrix} 5 \\ 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} = \frac{4}{33}$$

(b) All nationalities except Italians are represented, then 3 cases arise;

Case-I: 2 Canadians + 1 Japanies+ 0 Italian + 1 German

Case-2: 1 Canadians + 2 Japanies+ 0 Italian + 1 German

Case-3: 1 Canadians + 1 Japanies+ 0 Italian + 2 German

$$P(all\ nationalities\ are\ represented)$$
$$= \frac{\begin{pmatrix} 2 \\ 2 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} + \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} + \frac{\begin{pmatrix} 2 \\ 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix}\begin{pmatrix} 5 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix}}{\begin{pmatrix} 12 \\ 4 \end{pmatrix}} = \frac{8}{165}$$

(Q.44) An urn contains 3 green balls, 2 blue balls, and 4 red balls. In a random sample of 5 balls, find the probability that both blue balls and at least 1 red ball are selected.

Ans:

Total number of balls = 3+2+4 = 9

Total number of balls selected = 5

Number of blue balls = 2, green balls = 3, red balls = 4

$$P(2\,blue\ balls\ and\ atleast\ 1\ red\ ball)$$
$$= \frac{\binom{2}{2}\binom{4}{1}\binom{3}{2}}{\binom{9}{5}} + \frac{\binom{2}{2}\binom{4}{2}\binom{3}{1}}{\binom{9}{5}} + \frac{\binom{2}{2}\binom{4}{3}\binom{3}{0}}{\binom{9}{5}} = \frac{17}{63}$$

Negative Binomial Distribution:

Let repeated independent trials results in a success with probability p and a failure with probability $q = 1 - p$.

Where X: number of trials in which the kth success occurs,

Then,

$$P(X = x) = f(x) = b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}$$

Where $x = k, k + 1, k + 2....$

Geometric Distribution:

A particular case of negative binomial distribution for $k = 1$ is known as geometric distribution.

Here, X= the number of trials on which the first success occurs.

$$P(X = x) = f(x) = g(x; p) = pq^{x-1}; \quad x = 1, 2, 3........$$

Where $q = 1 - p$

(Q.49) The probability that a person living in a certain city owns a dog is estimated to be 0.3. Find the probability that the tenth person randomly interviewed in that city is the fifth one to own a dog.

Ans:

Here, $p = 0.3, q = 1 - p = 0.7$

X= The number of persons interviewed in which kth person own a dog.

Given $x = 10, k = 5$

$$b^*(10; 5, 0.3) = \begin{pmatrix} 10-1 \\ 5-1 \end{pmatrix} p^5 q^{10-5}$$
$$= \begin{pmatrix} 9 \\ 4 \end{pmatrix} (0.3)^5 (0.7)^{10-5} = 0.0515$$

(Q.50) Find the probability that a person flipping a coin gets (a) the third head on the seventh flip; (b) the first head on the fourth flip.

Ans:

Here, $p = 0.5, q = 1 - p = 0.5$

X= The number of trials in which kth head occurs.

(a)Third head in 7^{th} flip means $x = 7, k = 3$

Hence,

$$b^*(7; 3, 0.5) = \begin{pmatrix} 7-1 \\ 3-1 \end{pmatrix} (0.5)^3 (0.5)^4 = 0.1172$$

(b)First head in the fourth flip means$x = 4, k = 1$

Using negative binomial or geometric distribution

$$g(x; p) = g(4, 0.5) = 0.5(1 - 0.5)^{4-1} = (0.5)^4$$