Engineering Probability & Statistics (AGE 1150)

Chapter 5: Some Discrete Probability Distributions – Part 2

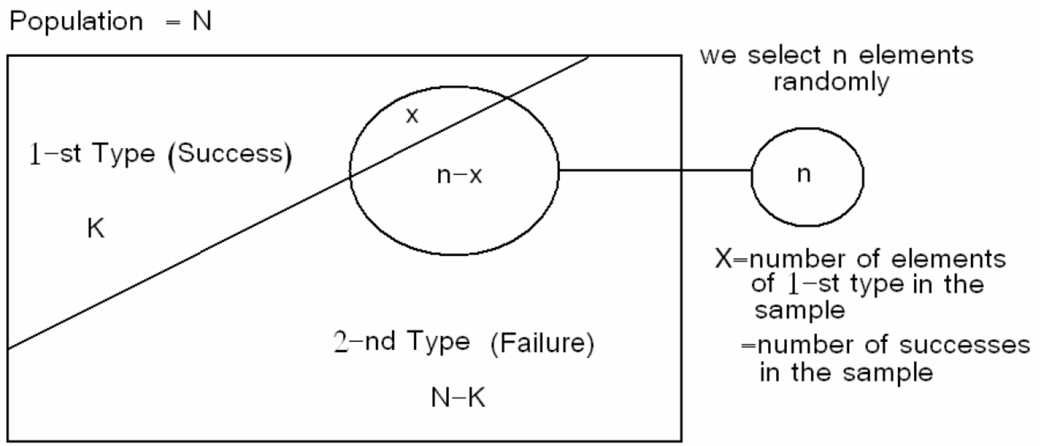
Hypergeometric Distribution

- Suppose there is a population with 2 types of elements:

1-st Type = success

2-nd Type = failure

- N= population size
- K= number of elements of the 1-st type
- N −K = number of elements of the 2-nd type
- We select a sample of n elements at random from the population
- Let X = number of elements of 1-st type (number of successes) in the sample
- We need to find the probability distribution of X.



- There are to two methods of selection:

1. selection with replacement
2. selection without replacement

- (1) If we select the elements of the sample at random and **with** replacement, then

X* ~ Binomial(*n*,*p*); where *p* = *K* / *N

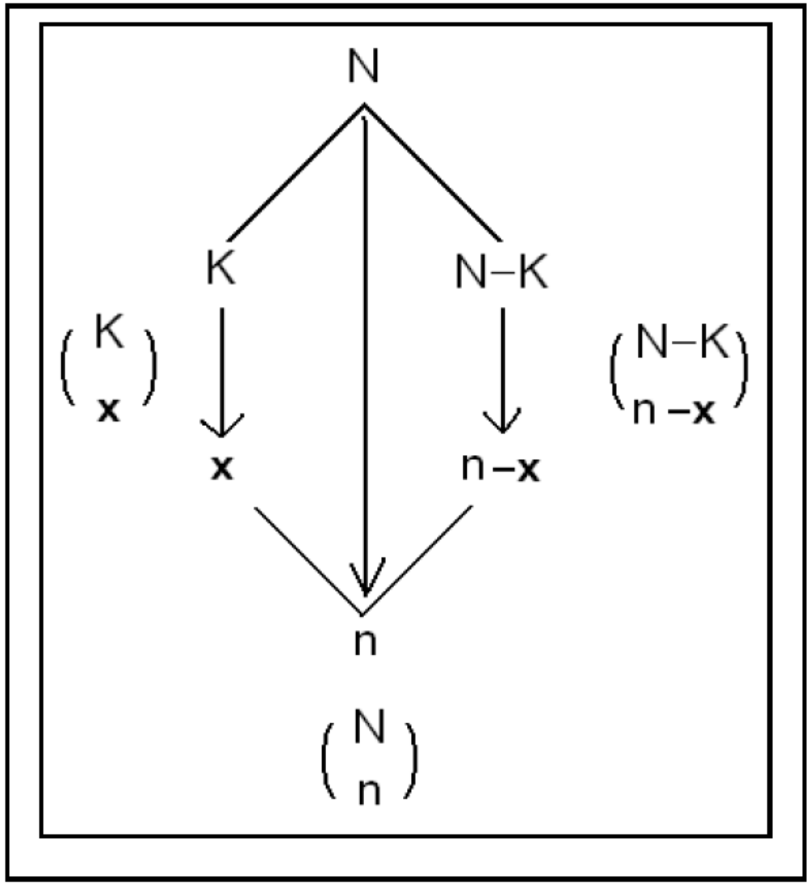
- (2) Now, suppose we select the elements of the sample at random and **without** replacement. When the selection is made without replacement, the random variable X has a hypergeometric distribution with parameters *N*, *n*, and *K*. and we write **X~h(x;*N*,*n*,*K*).**
- The probability distribution of *X* is given by:

$$f(x)=P(X=x)=h(x;N,n,K)$$

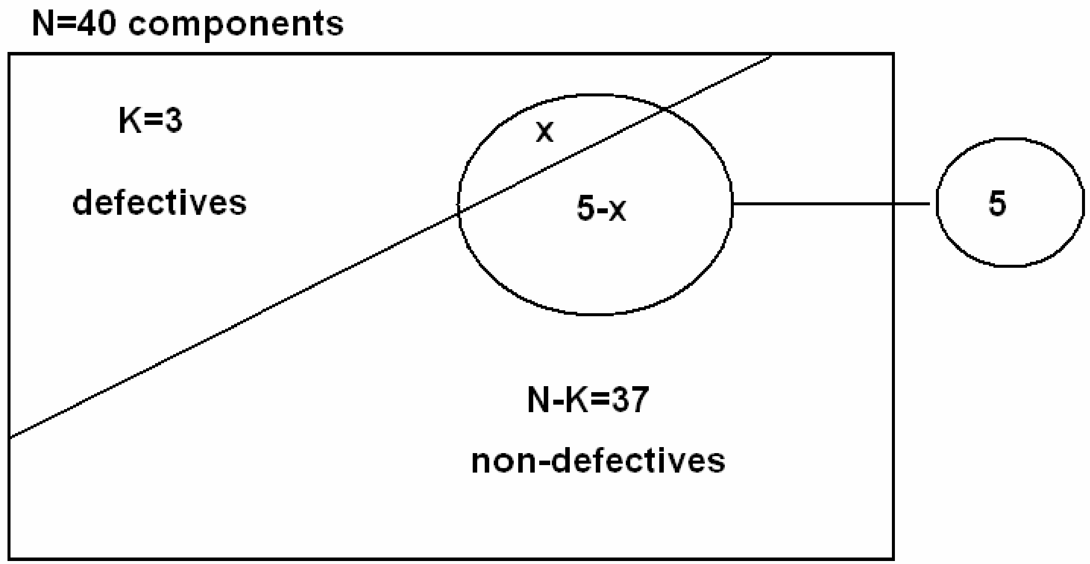
$$=\begin{cases}\dfrac{\dbinom{K}{x}\times\dbinom{N-K}{n-x}}{\dbinom{N}{n}}; & x=0,1,2,\cdots,n\\ 0;\ otherwise\end{cases}$$

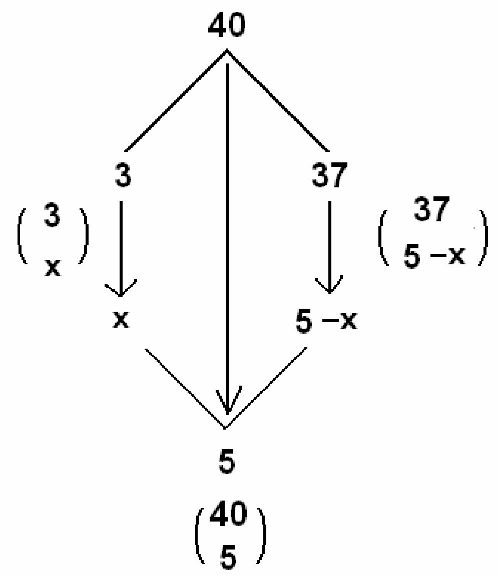
Note that the values of X must satisfy:

$0\le x\le K$ and $n-N+K\le x\le n$



Example: Lots of 40 components each are called acceptable if they contain no more than 3 defectives. The procedure for sampling the lot is to select 5 components at random (without replacement) and to reject the lot if a defective is found. What is the probability that exactly one defective is found in the sample if there are 3 defectives in the entire lot.





- Let X= number of defectives in the sample
- *N*=40, *K*=3, and *n*=5
- X has a hypergeometric distribution with parameters *N*=40, *n*=5, and *K*=3.
- X~h(x;*N*,*n*,*K*)=h(x;40,5,3).
- The probability distribution of *X*

is given by:

$$f(x) = P(X = x) = h(x;40,5,3) = \begin{cases} \dfrac{\binom{3}{x} \times \binom{37}{5-x}}{\binom{40}{5}}; & x = 0, 1, 2, \cdots, 5 \\ 0; \ otherwise \end{cases}$$

But the values of X must satisfy:

$0 \le x \le K$ and $n - N + K \le x \le n \Leftrightarrow 0 \le x \le 3$ and $-32 \le x \le 5$

Therefore, the probability distribution of X is given by:

$$f(x) = P(X = x) = h(x;40,5,3) = \begin{cases} \dfrac{\binom{3}{x} \times \binom{37}{5-x}}{\binom{40}{5}}; & x = 0, 1, 2, 3 \\ 0; \ otherwise \end{cases}$$

Now, the probability that exactly one defective is found in the sample is

$$.f(1)=P(X=1)=h(1;40,5,3)=\frac{\binom{3}{1}\times\binom{37}{5-1}}{\binom{40}{5}}=\frac{\binom{3}{1}\times\binom{37}{4}}{\binom{40}{5}}=0.3011$$

Theorem 5.3:

- The mean and the variance of the hypergeometric distribution h(x;N,n,K) are:

$$\mu = n\frac{K}{N}$$

$$\sigma^2 = n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}$$

- In previous Example, find the expected value (mean) and the variance of the number of defectives in the sample.

Solution:

- X = number of defectives in the sample
- We need to find E(X)=μ and Var(X)=σ2
- We found that X ~ h(x;40,5,3)
- N=40, n=5, and K=3
- The expected number of defective items is

$$\mathrm{E(X)}=\mu = n\frac{K}{N} = 5 \times \frac{3}{40} = 0.375$$

The variance of the number of defective items is

$$\mathrm{Var(X)}=\sigma^2 = n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1} = 5 \times \frac{3}{40}\left(1-\frac{3}{40}\right)\frac{40-5}{40-1} = 0.311298$$

Poisson Distribution

- Poisson experiment is an experiment yielding numerical values of a random variable that count the number of outcomes occurring in a given time interval or a specified region denoted by t.
- X = The number of outcomes occurring in a given time
- interval or a specified region denoted by t.
- Example:

1. X = number of field mice per acre (t= 1 acre)
2. X= number of typing errors per page (t=1 page)
3. X=number of telephone calls received every day (t=1 day)
4. X=number of telephone calls received every 5 days (t=5 days)

Let λ be the average (mean) number of outcomes per unit time or unit region (t=1).

- The average (mean) number of outcomes (mean of X) in the time interval or region t is:

$\mu = \lambda t$

- The random variable X is called a Poisson random variable with parameter μ ($\mu=\lambda t$), and we write $X \sim Poisson(\mu)$, if its probability distribution is given by:

$$f(x) = P(X = x) = p(x;\mu) = \begin{cases} \dfrac{e^{-\mu}\mu^{x}}{x!} & ; \quad x = 0, 1, 2, 3, \ldots \\ 0 & ; \quad otherwise \end{cases}$$

Theorem 5.5:

- The mean and the variance of the Poisson distribution Poisson($x;\mu$) are:

$$\mu = \lambda t$$
$$\sigma^2 = \mu = \lambda t$$

- Note:
- λ is the average (mean) of the distribution in the unit time ($t=1$).
- If X=The number of calls received in a month (unit time $t=1$ month) and X~Poisson(λ), then:

(i) Y = number of calls received in a year.

Y ~ Poisson (μ); $\mu=12\lambda$ ($t=12$)

(ii) W = number of calls received in a day.

W ~ Poisson (μ); $\mu=\lambda/30$ ($t=1/30$)

Example:

Suppose that the number of typing errors per page has a Poisson distribution with average 6 typing errors.

(1) What is the probability that in a given page:

(i) The number of typing errors will be 7?

(ii) The number of typing errors will be at least 2?

(2) What is the probability that in 2 pages there will be 10 typing errors?

(3) What is the probability that in a half page there will be no typing errors?

(1) X = number of typing errors per page.

X ~ Poisson (6) $\quad\quad$ (t=1, λ=6, μ=λt=6)

$$f(x) = P(X = x) = p(x;6) = \frac{e^{-6}6^x}{x!}\;;\quad x = 0, 1, 2, \ldots$$

(i) $$f(7) = P(X = 7) = p(7;6) = \frac{e^{-6}6^7}{7!} = 0.13768$$

(ii) $P(X\geq 2) = P(X=2) + P(X=3) + \ldots = \sum_{x=2}^{\infty} P(X = x)$

$P(X\geq 2) = 1 - P(X<2) = 1 - [P(X=0) + P(X=1)]$

$$= 1 - [f(0) + f(1)] = 1 - [\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!}]$$

$= 1 - [0.00248+0.01487]$

$= 1 - 0.01735 = 0.982650$

(2) X = number of typing errors in 2 pages

$X \sim \text{Poisson}(12)$ $\quad\quad$ $(t=2,\ \lambda=6,\ \mu=\lambda t=12)$

$$f(x) = P(X = x) = p(x;12) = \frac{e^{-12}12^{x}}{x!}: \quad x = 0, 1, 2\ldots$$

$$f(10) = P(X = 10) = \frac{e^{-12}12^{10}}{10!} = 0.1048$$

(3) X = number of typing errors in a half page.

$X \sim \text{Poisson}(3)$ $\quad$ $(t=1/2,\ \lambda=6,\ \mu=\lambda t=6/2=3)$

$$f(x) = P(X = x) = p(x;3) = \frac{e^{-3}\,3^{x}}{x!}: \quad x = 0, 1, 2\ldots$$

$$P(X = 0) = \frac{e^{-3}(3)^{0}}{0!} = 0.0497871$$