Chapter 5: Some Discrete Probability Distributions:

5.2: Discrete Uniform Distribution:

If the discrete random variable X assumes the values x_1, x_2, …, x_k with equal probabilities, then X has the discrete uniform distribution given by:

$$f(x) = P(X = x) = f(x;k) = \begin{cases} \frac{1}{k} \; ; \; x = x_1, x_2, \cdots, x_k \\ 0 \; ; \; elsewhere \end{cases}$$

Note:

- f(x)=f(x;k)=P(X=x)
- k is called the parameter of the distribution.

Example 5.2:

- Experiment: tossing a balanced die.
- Sample space: S={1,2,3,4,5,6}
- Each sample point of S occurs with the same probability 1/6.
- Let X= the number observed when tossing a balanced die.
- The probability distribution of X is:

$$f(x) = P(X = x) = f(x;6) = \begin{cases} \frac{1}{6} \; ; \; x = 1, 2, \cdots, 6 \\ 0 \; ; \; elsewhere \end{cases}$$

Theorem 5.1:

If the discrete random variable X has a discrete uniform distribution with parameter k, then the mean and the variance of X are:

$$E(X) = \mu = \frac{\sum_{i=1}^{k} x_i}{k}$$

$$Var(X) = \sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \mu)^2}{k}$$

Example 5.3:

Find E(X) and Var(X) in Example 5.2.

Solution:

$$E(X) = \mu = \frac{\sum_{i=1}^{k} x_i}{k} = \frac{1+2+3+4+5+6}{6} = 3.5$$

$$Var(X) = \sigma^2 = \frac{\sum_{i=1}^{k}(x_i-\mu)^2}{k} = \frac{\sum_{i=1}^{k}(x_i-3.5)^2}{6}$$

$$= \frac{(1-3.5)^2+(2-3.5)^2+\cdots+(6-3.5)^2}{6} = \frac{35}{12}$$

5.3 Binomial Distribution:

Bernoulli Trial:

- Bernoulli trial is an experiment with only two possible outcomes.
- The two possible outcomes are labeled:
  success (s) and failure (f)
- The probability of success is P(s)=p and the probability of failure is P(f)= q = 1−p.
- Examples:
  1. Tossing a coin (success=H, failure=T, and p=P(H))
  2. Inspecting an item (success=defective, failure=non-defective, and p=P(defective))

Bernoulli Process:

Bernoulli process is an experiment that must satisfy the following properties:

1. The experiment consists of n repeated Bernoulli trials.
2. The probability of success, P(s)=p, remains constant from trial to trial.
3. The repeated trials are independent; that is the outcome of one trial has no effect on the outcome of any other trial

Binomial Random Variable:

Consider the random variable :

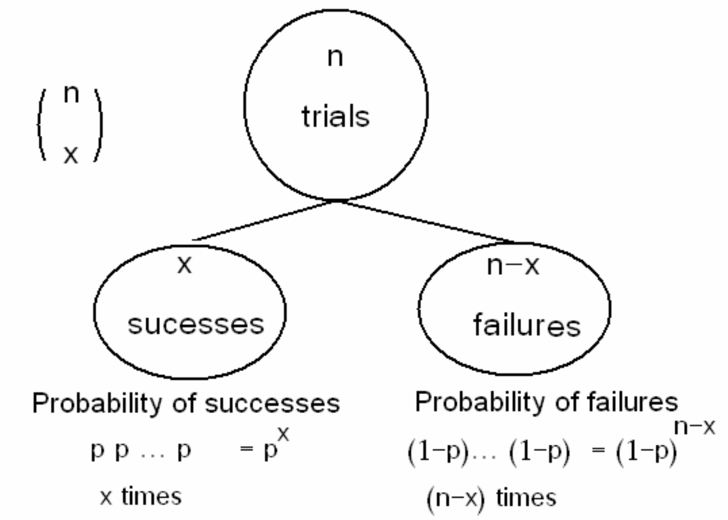
X = The number of successes in the n trials in a Bernoulli process

The random variable X has a binomial distribution with parameters n (number of trials) and p (probability of success), and we write:

$$X \sim \text{Binomial}(n,p) \quad \text{or} \quad X \sim \text{b}(x;n,p)$$

The probability distribution of X is given by:

$$f(x) = P(X = x) = b(x;n,p) = \begin{cases} \binom{n}{x} p^x \,(1-p)^{n-x} \;; & x = 0, 1, 2, \ldots, n \\ 0 \;; & \textit{otherwise} \end{cases}$$



We can write the probability distribution of X as a table as follows.

x	f(x)=P(X=x)=b(x;n,p)
0	$\binom{n}{0} p^0 (1-p)^{n-0} = (1-p)^n$
1	$\binom{n}{1} p^1 (1-p)^{n-1}$
2	$\binom{n}{2} p^2 (1-p)^{n-2}$
$\vdots$	$\vdots$
$n-1$	$\binom{n}{n-1} p^{n-1} (1-p)^1$
n	$\binom{n}{n} p^n (1-p)^0 = p^n$
Total	1.00

Example:

Suppose that 25% of the products of a manufacturing process are defective. Three items are selected at random, inspected, and classified as defective (D) or non-defective (N). Find the probability distribution of the number of defective items.

Solution:

- Experiment: selecting 3 items at random, inspected, and classified as (D) or (N).
- The sample space is
  S={DDD,DDN,DND,DNN,NDD,NDN,NND,NNN}
- Let X = the number of defective items in the sample
- We need to find the probability distribution of X.

(1) First Solution:

Outcome	Probability	x
NNN	$\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{27}{64}$	0
NND	$\frac{3}{4}\times\frac{3}{4}\times\frac{1}{4}=\frac{9}{64}$	1
NDN	$\frac{3}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{9}{64}$	1
NDD	$\frac{3}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{3}{64}$	2
DNN	$\frac{1}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{9}{64}$	1
DND	$\frac{1}{4}\times\frac{3}{4}\times\frac{1}{4}=\frac{3}{64}$	2
DDN	$\frac{1}{4}\times\frac{1}{4}\times\frac{3}{4}=\frac{3}{64}$	2
DDD	$\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{64}$	3

The probability distribution .of X is

.x	.f(x)=P(X=x)
0	$\frac{27}{64}$
1	$\frac{9}{64}+\frac{9}{64}+\frac{9}{64}=\frac{27}{64}$
2	$\frac{3}{64}+\frac{3}{64}+\frac{3}{64}=\frac{9}{64}$
3	$\frac{1}{64}$

(2) Second Solution:

Bernoulli trial is the process of inspecting the item. The results are success=D or failure=N, with probability of success P(s)=25/100=1/4=0.25.

The experiments is a Bernoulli process with:

- number of trials: n=3
- Probability of success: p=1/4=0.25
- X ~ Binomial(n,p)=Binomial(3,1/4)

- The probability distribution of X is given by:

$$f(x) = P(X = x) = b(x;3,\frac{1}{4}) = \begin{cases} \binom{3}{x} (\frac{1}{4})^x (\frac{3}{4})^{3-x}; & x = 0, 1, 2, 3 \\ 0 \ ; & otherwise \end{cases}$$

$$f(0) = P(X = 0) = b(0;3,\frac{1}{4}) = \binom{3}{0}(\frac{1}{4})^0 (\frac{3}{4})^3 = \frac{27}{64}$$

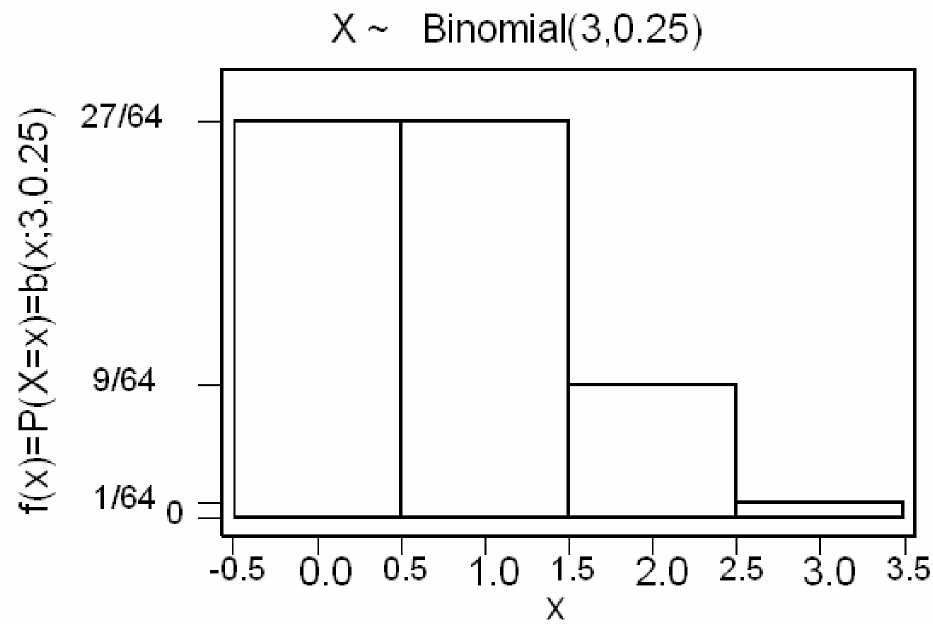
$$f(1) = P(X = 1) = b(1;3,\frac{1}{4}) = \binom{3}{1}(\frac{1}{4})^1 (\frac{3}{4})^2 = \frac{27}{64}$$

$$f(2) = P(X = 2) = b(2;3,\frac{1}{4}) = \binom{3}{2}(\frac{1}{4})^2 (\frac{3}{4})^1 = \frac{9}{64}$$

$$f(3) = P(X = 3) = b(3;3,\frac{1}{4}) = \binom{3}{3}(\frac{1}{4})^3 (\frac{3}{4})^0 = \frac{1}{64}$$

The probability distribution of X is

x	f(x)=P(X=x) =b(x;3,1/4)
0	27/64
1	27/64
2	9/64
3	1/64



Theorem 5.2:

The mean and the variance of the binomial distribution b(x;n,p) are:

$$\mu = n\,p$$
$$\sigma^2 = n\,p\,(1 - p)$$

Example:

In the previous example, find the expected value (mean) and the variance of the number of defective items.

Solution:

- X = number of defective items
- We need to find E(X)=μ and Var(X)=σ^2
- We found that X ~ Binomial(n,p)=Binomial(3,1/4)
- .n=3 and p=1/4

The expected number of defective items is

$$E(X)=\mu = n\,p = (3)\,(1/4) = 3/4 = 0.75$$

The variance of the number of defective items is

$$Var(X)=\sigma^2 = n\,p\,(1-p) = (3)\,(1/4)\,(3/4) = 9/16 = 0.5625$$

Example:

In the previous example, find the following probabilities:
(1) The probability of getting at least two defective items.
(2) The probability of getting at most two defective items.

Solution:

X ~ Binomial(3,1/4)

$$f(x) = P(X = x) = b(x;3,\frac{1}{4}) = \begin{cases} \binom{3}{x} (\frac{1}{4})^x (\frac{3}{4})^{3-x} & for\ \ x = 0, 1, 2, 3 \\ 0 & otherwise \end{cases}$$

.x	.f(x)=P(X=x)=b(x;3,1/4)
0	27/64
1	27/64
2	9/64
3	1/64

(1) The probability of getting at least two defective items:

$$P(X\geq 2)=P(X=2)+P(X=3)= f(2)+f(3)=\frac{9}{64}+\frac{1}{64}=\frac{10}{64}$$

(2) The probability of getting at most two defective item:

$$P(X\leq 2) = P(X=0)+P(X=1)+P(X=2)$$

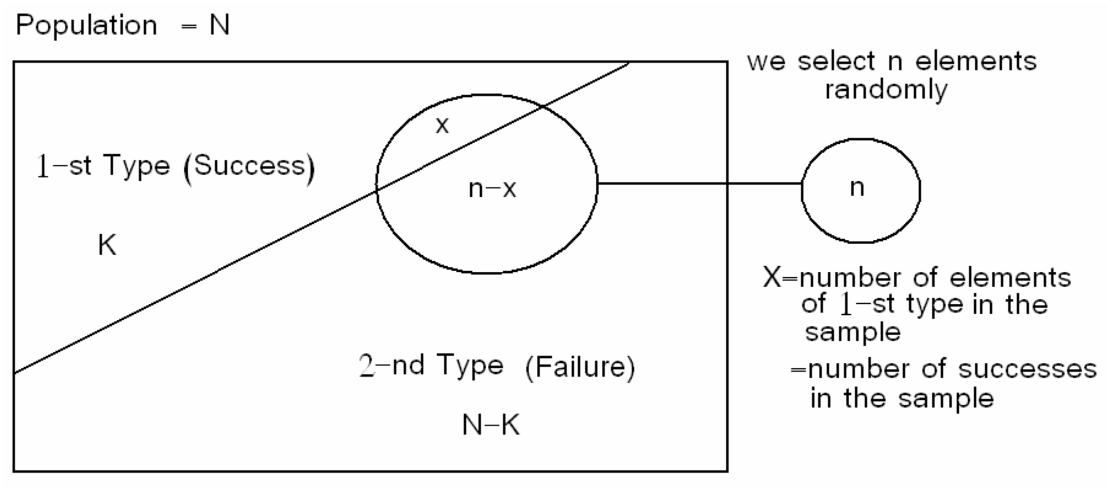
$$= f(0)+f(1)+f(2) = \frac{27}{64}+\frac{27}{64}+\frac{9}{64}=\frac{63}{64}$$

or

$$P(X\leq 2)= 1-P(X>2) = 1-P(X=3) = 1- f(3) = 1-\frac{1}{64}=\frac{63}{64}$$

Example 5.4: Reading assignment
Example 5.5: Reading assignment
Example 5.6: Reading assignment

5.4 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

$$X \sim \text{Binomial}(n,p);\ \text{where } p=\frac{K}{N}$$

(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:

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

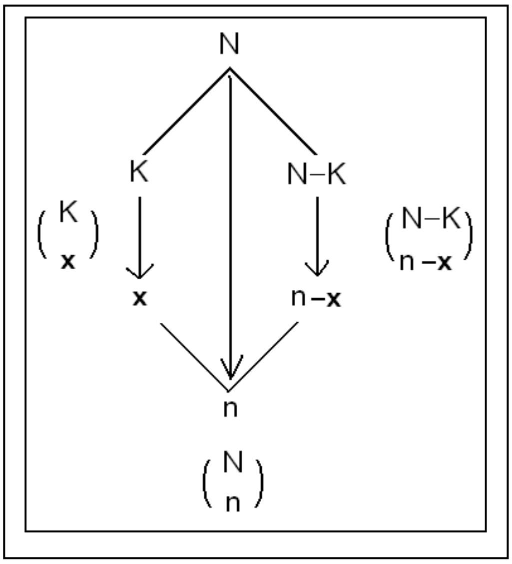
$$= \begin{cases} \dfrac{\binom{K}{x} \times \binom{N-K}{n-x}}{\binom{N}{n}}; & x = 0, 1, 2, \cdots, n \\ 0; \ otherwise \end{cases}$$

Note that the values of X must satisfy:

$$0 \le x \le K \text{ and } 0 \le n - x \le N - K$$

$$\Leftrightarrow$$

$$0 \le x \le K \text{ and } n - N + K \le x \le n$$

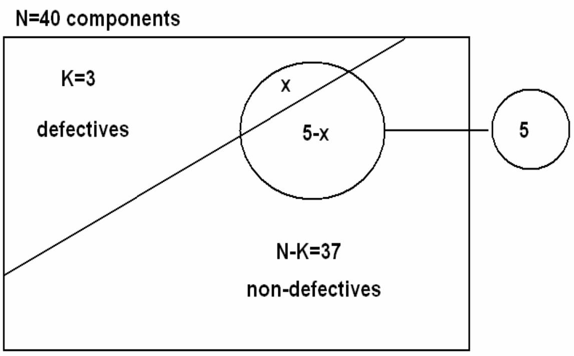


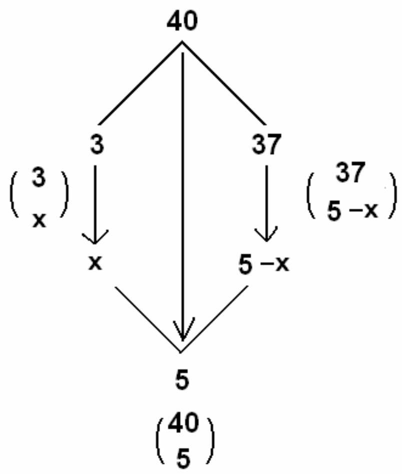
Example 5.8: Reading assignment

Example 5.9:

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.

Solution:





- 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)= $\dfrac{\binom{3}{1} \times \binom{37}{5-1}}{\binom{40}{5}} = \dfrac{\binom{3}{1} \times \binom{37}{4}}{\binom{40}{5}} = 0.3011$

Theorem 5.3:

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

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

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

Example 5.10:

In Example 5.9, 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)=σ^2
- We found that $X \sim$ h(x;40,5,3)
- N=40, n=5, and K=3

The expected number of defective items is

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

The variance of the number of defective items is

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

Relationship to the binomial distribution:

* Binomial distribution: $b(x;n,p)=\binom{n}{x}p^x(1-p)^{n-x};\ x=0,1,\ldots,n$

* Hypergeometric distribution: $h(x;N,n,K)=\dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}};x=0,1,\cdots,n$

If n is small compared to N and K, then the hypergeometric distribution h(x;N,n,K) can be approximated by the binomial distribution b(x;n,p), where $p=\frac{K}{N}$; i.e., for large N and K and small n, we have:

$$h(x;N,n,K)\approx b(x;n,\frac{K}{N})$$

$$\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} \approx \binom{n}{x}\left(\frac{K}{N}\right)^x\left(1-\frac{K}{N}\right)^{n-x};x=0,1,\cdots,n$$

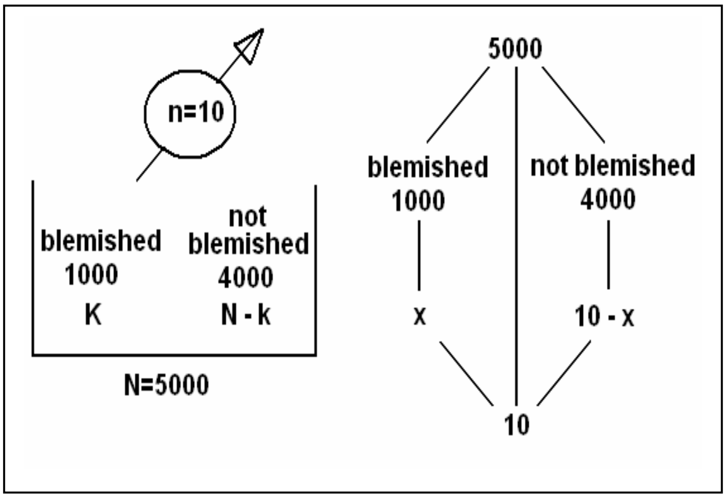
Note:

If n is small compared to N and K, then there will be almost no difference between selection without replacement and selection with replacement ($\frac{K}{N}\approx\frac{K-1}{N-1}\approx\cdots\approx\frac{K-n+1}{N-n+1}$).

Example 5.11:

N=5000 K=1000 n=10

X=Number of blemished tires in the Sample



X~h(x;5000,10,1000)

The exact probability is

$$P(X=3) = \binom{1000}{3}\binom{4000}{7}\Big/\binom{5000}{10}$$

$= \underline{0.201}477715$

$\approx \underline{0.201}$

Since n=10 is small relative to N=5000 and K=4000, we can approximate the hypergeometric probabilities using binomial probabilities as follows:

.n=10 (no. of trials)

.p=K/N=1000/5000=0.2 (probability of success)

X ~ h(x;5000,10,1000) ≈ b(x;10,0.2)

$$P(X=3) \approx \binom{10}{3}(0.2)^3\,(0.8)^7 = \underline{0.201}326592$$

$\approx \underline{0.201}$

5.6 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 (μ=λ*t*), and we write X~Poisson(μ), 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;μ) 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); μ=12λ (t=12)
(ii) W = number of calls received in a day.
W ~ Poisson (μ); μ=λ/30 (t=1/30)

Example 5.16: Reading Assignment

Example 5.17: Reading Assignment

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?

Solution:

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

X ~ Poisson (6) (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≥2) = P(X=2)+ P(X=3)+ . . . = $\sum_{x=2}^{\infty} P(X=x)$

P(X≥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 ~ Poisson(12) (t=2, λ=6, μ=λt=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 ~ Poisson (3) (t=1/2, λ=6, μ=λt=6/2=3)

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

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

Theorem 5.6: (Poisson approximation for binomial distribution:

Let X be a binomial random variable with probability distribution b(x;n,p). If $n\to\infty$, $p\to 0$, and μ=np remains constant, then the binomial distribution b(x;n,p) can approximated by Poisson distribution p(x;μ).

- For large n and small p we have:

b(x;n,p) ≈ Poisson(μ) (μ=np)

$$\binom{n}{x} p^x (1-p)^{n-x} \approx \frac{e^{-\mu}\mu^x}{x!}; x = 0,1,\cdots,n; \quad (\mu = np)$$

Example 5.18:

X = number of items producing bubbles in a random sample of 8000 items

.n=8000 and p=1/1000 = 0.001

X ~ b(x;8000, 0.001)

The exact probability is:

$$P(X<7)=P(X\le 6)=\sum_{x=0}^{6}\binom{8000}{x}(0.001)^x(0.999)^{8000-x} =\ldots=\underline{0.313}252$$

The approximated probability using Poisson approximation:

.n=8000 (n is large, i.e., n→∞)

.p= 0.001 (p is small, i.e. p→0)

μ=np = 8000(0.001)=8

X ≈ Poisson (8)

$$f(x) = P(X = x) = p(x;8) = \frac{e^{-8}\,8^x}{x!} : \quad x = 0, 1, 2\ldots$$

$$P(X<7)=P(X\le 6)=\sum_{x=0}^{6}\frac{e^{-8}\,8^x}{x!} = e^{-8}\sum_{x=0}^{6}\frac{8^x}{x!} =\ldots= \underline{0.313}374$$