Chapter 3
Discrete Random Variables & Probability Distributions

MAS291 - STATISTICS & PROBABILITY

Ly Anh Duong

duongla3@fe.edu.vn

Table of Contents

Random variable

1 Probability Distributions and Probability Mass Functions

A random variable is a function that assigns a real number to each outcome in the sample space of a random experiment.

$$X : S \to \mathbb{R}$$

$$X(\omega) \in \mathbb{R}$$

If $X(S) = \{x_1, x_2, \dots, x_n\}$ or $X(S) = \{x_1, x_2, \dots, x_n, \dots\}$
X is called **Discrete random variable**.

Example:

Flipping a coin twice. $\to$ The sample space is $S = \{HH, HT, TH, TT\}$.
$X : S \to \mathbb{R}$
$X(\omega) =$ number of heads in each outcome ω

$\omega \in S$	HH	HT	TH	TT
$X(\omega)$	2	1	1	0

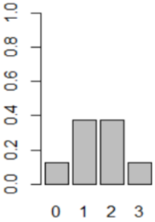
The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X.

Example:

Toss a fair coin three times and let **X be the number of Heads** observed, $X(\omega) \in \{0, 1, 2, 3\}$. Then we have the following probabilities

	TTT	HTT THT TTH	HHT HTH THH	HHH
X	0	1	2	3
P(X = x)	1/8	3/8	3/8	1/8

$$\boxed{\Sigma_x P(x) = 1}$$



Probability distribution for number of heads.

Probability Mass Functions (Hàm khối XS)

For a discrete random variable X with possible values $x_1, x_2, .., x_n$, a probability mass function (pmf) is a function such that

- $f(x_i) \geq 0$
- $\displaystyle\sum_{i=1}^{n} f(x_i) = 1$
- $f(x_i) = P(X = x_i)$

Example: Verify that the following function is a pmf

$$f(x) = \frac{2x+1}{25}, x = 0, 1, 2, 3, 4$$

(1) $f(x) \geqslant 0$
(2) $f(x) = P(X = x)$, and $P(X = 4) = f(4) = 9/25$
(3) $\Sigma_i f(x_i) = f(0) + f(1) + f(2) + f(3) + f(4) = 1$

Given the $pmf\, f(x)$, determine the probabilities.

$$f(x) = \frac{2x+1}{25},\ x = 0, 1, 2, 3, 4$$

$$\begin{array}{ll} \text{a/ } P(X = 4) & b/P(X \leqslant 3) \\ \text{c/ } P(2 \leqslant X \leqslant 4) & d/P(X > -3) \end{array}$$

Given the $pmf\, f(x)$, determine the probabilities.

$$f(x) = \frac{2x+1}{25},\ x = 0, 1, 2, 3, 4$$

$$\begin{array}{ll} \text{a/ } P(X=4) & b/P(X \leqslant 3) \\ \text{c/ } P(2 \leqslant X \leqslant 4) & d/P(X > -3) \end{array}$$

Solution

c/ $P(2 \leqslant X \leqslant 4) = P(X=2) + P(X=3) + P(X=4)$

$$= \frac{(2 \cdot 2 + 1)}{25} + \frac{(2 \cdot 3 + 1)}{25} + \frac{(2 \cdot 4 + 1)}{25} = \frac{21}{25}$$

Example

Let the random variable X denote the number of semiconductor wafers that need to be analyzed in order to detect a large particle of contamination. Assume that the probability that a wafer contains a large particle is 0.01 and that the wafers are independent.
Determine the probability distribution of X.

Solution

Let p denote a wafer in which a large particle is present, and let a denote a wafer in which it is absent. The sample space of the experiment is infinite, and it can be represented as all possible sequences that start with a string of a's and end with p. That is,

$$S = \{p, ap, aap, aaap, aaaap, \text{ and so forth}\}$$

$P(X = 1) = P(p) = 0.01$
$P(X = 2) = P(ap) = 0.99(0.01)$
$P(X = 3) = P(aap) = (0.99)^2(0.01)$
$P(X = x) = (0.99)^{x-1}(0.01)$

Table of Contents

2 Cumulative Distribution Functions

The cumulative distribution function of a discrete random variable X, denoted as $F(x)$ is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

- $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$
- $0 \leq F(x) \leq 1$
- If $x \leq y$ then $F(x) \leq F(y)$.

Example: Find $F(-1), F(1), F(1.9)$

x	**−1**	**0**	**1**	**2**	otherwise
f(x)	0.2	0.5	0	0.3	0

Cumulative Distribution Functions

The cumulative distribution function of a discrete random variable X, denoted as $F(x)$ is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

For a discrete random variable X, $F(x)$ satisfies the following properties.

- $F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$
- $0 \leq F(x) \leq 1$
- If $x \leq y$ then $F(x) \leq F(y)$.

Example: Find $F(-1), F(1), F(1.9)$

x	**−1**	**0**	**1**	**2**	otherwise
f(x)	0.2	0.5	0	0.3	0

$F(-1) = f(-1) = 0.2$
$F(1) = f(-1) + f(0) + f(1) = 0.7$
$F(1.9) = f(-1) + f(0) + f(1) = 0.7$

Exercise

2 Cumulative Distribution Functions

Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement, from the batch. Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

Exercise

2 Cumulative Distribution Functions

Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement, from the batch. Let the random variable X equal the number of nonconforming parts in the sample. What is the cumulative distribution function of X?

Solution

$$P(X=0)=\frac{800}{850}.\frac{799}{849}=0.886$$
$$P(X=1)=2.\frac{800}{850}.\frac{50}{849}=0.111$$
$$P(X=2)=\frac{50}{850}.\frac{49}{849}=0.003$$

Therefore

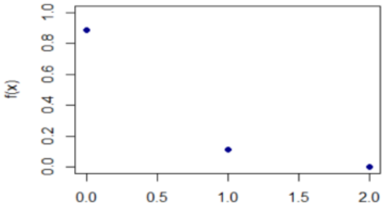
$$F(0)=P(X\le 0)=0.866$$
$$F(1)=P(X\le 1)=0.886+0.111=0.997$$
$$F(2)=P(X\le 2)=1$$

cmf and pmf

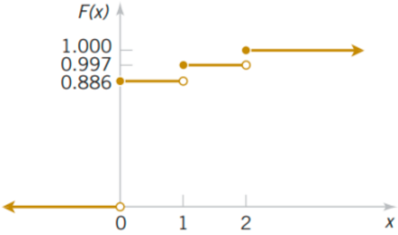
2 Cumulative Distribution Functions

x	0	1	2
f(x)	0.886	0.111	0.003

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.886 & 0 \le x < 1 \\ 0.997 & 1 \le x < 2 \\ 1 & 2 \le x \end{cases}$$



Probability mass function pmf

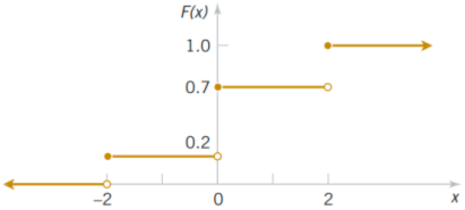


Cumulative distribution function cdf

Example

Determine the pmf of X from the following cdf.

$$F(x)=\begin{cases}0 & x<-2\\ 0.2 & -2\le x<0\\ 0.7 & 0\le x<2\\ 1 & 2\le x\end{cases}$$



Find f(x) from F(x):
f(x) = F(x) – F(x⁻)

x	-3	-2	-1	0	1	2	3
F(x)	0	0.2	0.2	0.7	0.7	1	1
f(x)	0	0.2	0	0.5	0	0.3	0

3-52. The thickness of wood paneling (in inches) that a customer orders is a random variable with the following cumulative distribution function:

$$F(x)=\begin{cases} 0 & x<1/8 \\ 0.2 & 1/8\le x<1/4 \\ 0.9 & 1/4\le x<3/8 \\ 1 & 3/8\le x \end{cases}$$

Determine the following probabilities:

(a) $P(X\le 1/18)$ (b) $P(X\le 1/4)$ (c) $P(X\le 5/16)$

(d) $P(X>1/4)$ (e) $P(X\le 1/2)$

Table of Contents

3 Mean and Variance of a discrete random variable

Mean and Variance

3 Mean and Variance of a discrete random variable

The mean (trung bình) or expected value of X, denoted as μ or $E(X)$ is

$$\mu = E(X) = \sum_x xP(X = x) = \sum_x xf(x)$$

The variance (phương sai) of X, denoted as σ^2 or $V(X)$ is

$$\sigma^2 = V(X) = E(X - \mu)^2 = \sum_x (x - \mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

The standard deviation (độ lệch chuẩn) of X is $\sigma = \sqrt{\sigma^2}$

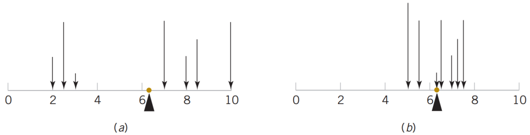


FIGURE 3-5 A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but part (a) illustrates a larger variance.

Example

3 Mean and Variance of a discrete random variable

The number of e-mail messages per hour has the following distribution:

x = number of messages	10	11	12	13	14	15
$f(x)$	0.08	0.15	0.30	0.20	0.20	0.07

Determine the mean and standard deviation of the number of messages received per hour.

Example

3 Mean and Variance of a discrete random variable

The number of e-mail messages per hour has the following distribution:

x = number of messages	10	11	12	13	14	15
$f(x)$	0.08	0.15	0.30	0.20	0.20	0.07

Determine the mean and standard deviation of the number of messages received per hour.

Solution

$E(X) = \sum_x xf(x) = 12.5$

$V(X) = \sum_x x^2 f(x) - E^2(X) = 1.85$

$\sigma = \sqrt{V(X)} = \sqrt{1.85} = 1.36$

Expected Value of a Function of a Discrete Random Variable

If X is a discrete random variable with probability mass function of $f(x)$, then

$$E(h(X)) = \sum_{x} h(x)f(x)$$

Example

- Putting $h(x) = x^2$ then
$$E(X^2) = \sum_{x} x^2 f(x)$$
- Putting $h(x) = e^x$ then
$$E(e^X) = \sum_{x} e^x f(x)$$

Some Useful properties

3 Mean and Variance of a discrete random variable

$$E(aX+b) = aE(X)+b, \quad V(aX+b) = a^2V(X)$$
$$\sigma^2 = \mathrm{V}(\mathrm{X}) = \Sigma_{\mathrm{x}}(\mathrm{X}-\mu)^2\mathrm{f}(\mathrm{x}) = \mathrm{E}\left[(\mathrm{X}-\mu)^2\right] = \mathrm{E}\left[\mathrm{X}^2-2\mu\mathrm{X}+\mu^2\right]$$
$$= \mathrm{E}\left(\mathrm{X}^2\right)-\mu^2 = \mathrm{E}\left(\mathrm{X}^2\right)-\mathrm{E}(\mathrm{X})^2$$

Example

- Putting $h(x) = x^2$ then

$$E(X^2) = \sum_x x^2 f(x)$$

- Putting $h(x) = e^x$ then

$$E(e^X) = \sum_x e^x f(x)$$

Table of Contents

Discrete Uniform Distribution

A random variable X has a **discrete uniform distribution** if each of the n values in its range, say, $x_1, x_2, ..., x_n$ has **equal probability**. Then

$$f(x_i) = P(X = x_i) = \frac{1}{n}$$

Mean and Variance

Suppose that X is a discrete uniform random variable on the consecutive integers $a, a+1, a+2, ..., b$ for $a \leq b$. Then

$$\mu = E(X) = \frac{b+a}{2}, \ \sigma^2 = V(X) = \frac{(b-a+1)^2 - 1}{12}$$

Example

4 Discrete Uniform Distribution

The first digit of a part's serial number is equally likely to be any one of the digits 0 through 9. If one part is selected from a large batch and X is the first digit of the serial number, X has a discrete uniform distribution with probability 0.1 for each value in R = {0,1,2,...,9}. That is,

$$f(x) = 0.1,\ \ x = 0, 1, 2, ..., 9$$

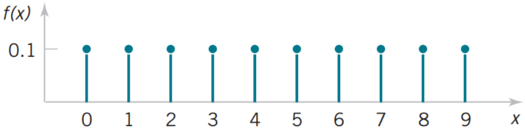


FIGURE 3.6

Probability mass function for a discrete uniform random variable.

Example

Let the random variable X denote the number of the voice lines that are in use at a particular time. Assume that X is a discrete uniform random variable with a range of 0 to 48. Find $E(X)$, $V(X)$

Example

Let the random variable X denote the number of the voice lines that are in use at a particular time. Assume that X is a discrete uniform random variable with a range of 0 to 48. Find $E(X)$, $V(X)$

Solution

$$E(X) = \frac{48+0}{2} = 24$$

$$V(X) = \sigma^2 = \frac{(48-0+1)^2-1}{12} = 200$$

Table of Contents

5 Binomial Distribution

Bernoulli trials

Bernoulli trial: A trial with only two possible outcomes (success or failure).

Example: The following random experiments are series of Bernoulli trials:

- Flip a coin 10 times.
- Guess each question of a multiple-choice exam with 50 questions, each with four choices.

Independence: The outcome from one trial has no effect on the outcome to be obtained from any other trial.

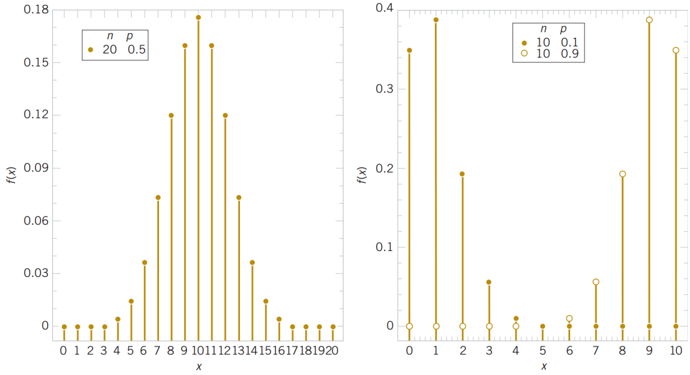
Binomial Distribution

Definition: A random experiment consists of n Bernoulli trials such that
(1) The trials are independent.
(2) Each trial results in only two possible outcomes, labeled as "success" and "failure".
(3) The probability of a success in each trial, denoted as p, remains constant.
The random variable X that equals the number of trials that result in a success is a binomial random variable with parameters $0 < p < 1$ and n $= 1\ 2\ ,\ ,\ldots$. The probability mass function of X is

$$f(x) = C_n^x p^x (1-p)^{n-x},\ x = 0, 1, 2, ..., n$$

Mean and Variance If X is a binomial random variable with parameters p and n,

$$\mu = E(X) = np$$
$$\sigma^2 = V(X) = np(1-p)$$



Hình: Binomial distributions for selected values of n and p

Example

5 Binomial Distribution

Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant.

a) Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

b) Determine the probability that at least four samples contain the pollutant.

Solution

Let X be the number of samples that contain the pollutant in the next 18 samples analyzed. Then X is a binomial random variable with p = 0.1 and n = 18.

a) $P(X=2) = C_{18}^{2}(0.1)^2(0.9)^{16} = 153.(0.1)^2.(0.9)^{16} = 0.284$

b)

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) \\ &= 1 - \sum_{x=0}^{3} P(X = x) \\ &= 1 - \sum_{x=0}^{3} C_{18}^{x}(0.1)^x(0.9)^{18-x} \\ &= 1 - (0.150 + 0.3 + 0.284 + 0.168) = 0.098 \end{aligned}$$

Table of Contents

6 Geometric and Negative Binomial Distribution

Geometric Random Variable

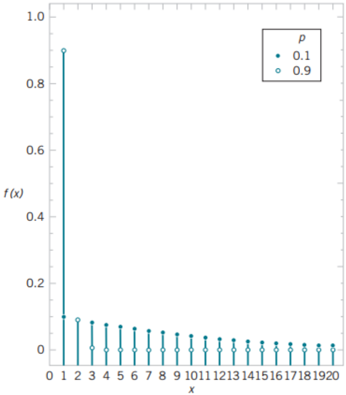
Definition: In a series of Bernoulli trials (independent trials with constant probability p of a success), the random variable X that equals the number of trials until the first success is a geometric random variable with parameter $0 < p < 1$ and

$$f(x) = (1-p)^{x-1}p$$

Mean and Variance

If X is a geometric random variable with parameter p,

$$\mu = E(X) = \frac{1}{p}, \ \sigma^2 = V(X) = \frac{1-p}{p^2}$$



Hình: Geometric distributions for selected values of the parameter p.

Example: Digital Channel

The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume that the transmissions are independent events. Compute probability that the first four bits are transmitted correctly and the fifth bit is in error.

Example: Digital Channel

The probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume that the transmissions are independent events. Compute probability that the first four bits are transmitted correctly and the fifth bit is in error.

Solution

Let the random variable X denote the number of bits transmitted until the first error.

$P(X = 5) = P(okay\ okay\ okay\ okay\ fail) = P(okay)^4.P(fail) = (0.9)^4(0.1)$

Example

The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent.
what is the probability that exactly 125 wafers need to be analyzed before a large particle is detected?

Example

The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent.
what is the probability that exactly 125 wafers need to be analyzed before a large particle is detected?

Solution

Let X denote the number of samples analyzed until a large particle is detected.
Then X is a geometric random variable with $p = 001$. The requested probability is

$$P(X = 125) = (0.99)^{124}0.01 = 0.0029$$

Table of Contents

7 Negative Binomial Distributions

Negative Binomial Distributions

(Generalization of a geometric distribution)

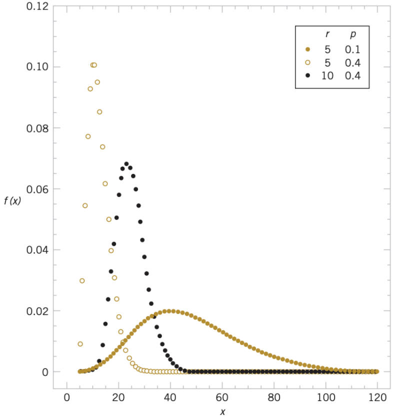
Definition: In a series of Bernoulli trials (independent trials with constant probability p of a success), the random variable X that equals the number of trials until r successes occur is a negative binomial random variable with parameters $0 < p < 1$ and $x = r, r+1, r+2, \ldots$

$$f(x) = C_{x-1}^{r-1}(1-p)^{x-r}p^r$$

Mean and Variance

If X is a negative binomial distribution with parameters p and r, then

$$\mu = E(X) = \frac{r}{p},\ \sigma^2 = V(X) = \frac{r(1-p)}{p^2}$$



Hình: Negative binomial distributions for selected values of the parameters r and p.

Example

A website contains three identical computer servers. Only one is used to operate the site, and the other two are spares that can be activated in case the primary system fails. The probability of a failure in the primary computer (or any activated spare system) from a request for service is 0.0005. Assuming that each request represents an independent trial.

a) Compute the probability that all three servers fail within five requests?

b) What is the mean number of requests until failure of all three servers?

a) Let X denotes the number of requests until all three servers fail. Then X has a negative binomial distributions with $p = 0.0005$ and $r = 3$.

$$
\begin{aligned}
P(X \leq 5) &= P(X = 3) + P(X = 4) + P(X = 5) \\
&= 0.0005^3 + C_3^2 0.0005^3 0.9995 + C_4^2 0.0005^3 (0.9995) \\
&= 1.249 \times 10^{-9}
\end{aligned}
$$

b) The mean number of requests until failure of all three servers is

$$
E(X) = \frac{3}{0.0005} = 6000 \text{ (requests)}
$$

Table of Contents

Hypergeometric Distribution

Definition: A set of N objects contains

- K objects classified as success
- $N - K$ objects classified as failures

A sample of size n objects is selected randomly (without replacement) from the N objects, where $K \leq N$ and $n \leq N$.
Let the random variable X denote the number of successes in the sample. Then X is a hypergeometric random variable and

$$f(x) = \frac{C_K^x C_{N-K}^{n-x}}{C_N^n}, \ x = \max\{0, n+K-N\} \text{ to } \min\{K, n\}$$

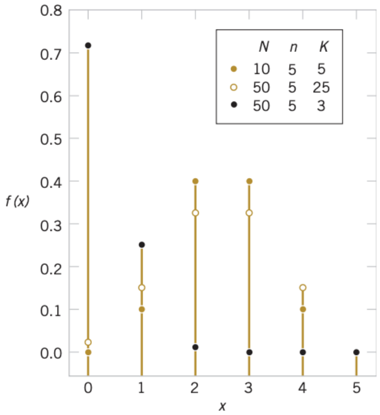
Mean and Variance of Hypergeometric Distribution

If X is a hypergeometric random variable with parameters N, K, n then

$$\mu = E(X) = np$$

$$\sigma^2 = V(X) = np(1-p)\frac{N-n}{N-1}, \quad p := \frac{K}{N}$$

Here p is interpreted as the proportion of successes in the set of N objects.



Hình: Hypergeometric distributions for selected values of parameters N, K, and n.

Example

8 Hypergeometric Distribution

A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in the next state. Four parts are selected randomly and without replacement.

a) What is the probability they are all from the local supplier?

b) What is the probability that two or more parts in the sample are from the local supplier?

Let X equal the number of parts in the sample from the local supplier. Then X has a hypergeometric distribution.

a) The requested probability is $P(X = 4)$. Consequently,

$$P(X = 4) = \frac{C_{100}^{4} C_{200}^{0}}{C_{300}^{4}} = 0.0119$$

b)

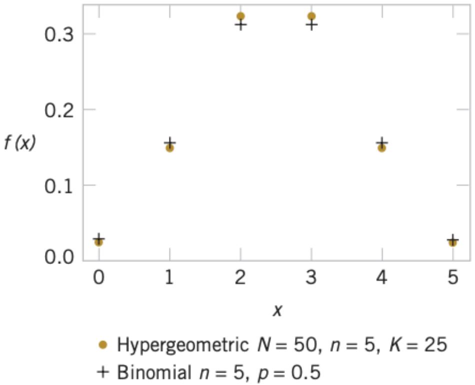
$$\begin{aligned}
P(X \geq 2) &= P(X = 2) + P(X = 3) + P(X = 4) \\
&= \frac{C_{100}^{2} C_{200}^{2}}{C_{300}^{4}} + \frac{C_{100}^{3} C_{200}^{1}}{C_{300}^{4}} + \frac{C_{100}^{4} C_{200}^{0}}{C_{300}^{4}} \\
&= 0.298 + 0.098 + 0.0119 = 0.408
\end{aligned}$$

Finite Population Correction Factor

The term in the variance of a hypergeometric random variable

$$\frac{N-n}{N-1}$$

is called the finite population correction factor.



Hình: Comparison of hypergeometric and binomial distributions.

Table of Contents

9 Poisson Distribution

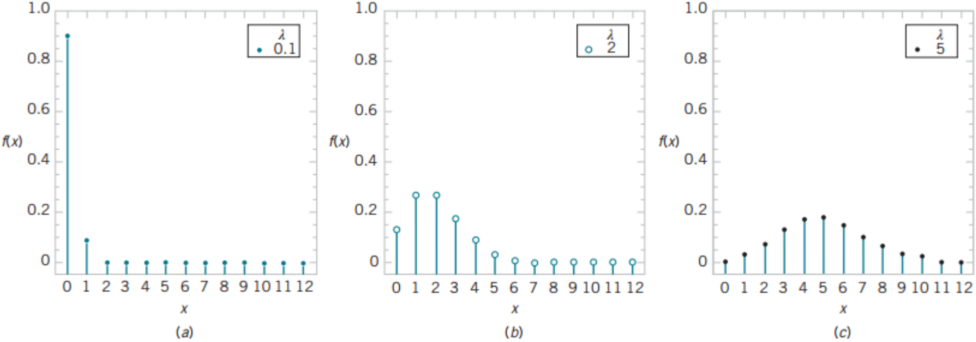
Definition: The random variable X that equals the number of events in the interval is a Poisson random variable with parameter $\lambda > 0$, and the probability mass function of X is

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \; x = 0, 1, 2, \ldots$$

Mean and Variance

If X is a Poisson Random Variable then

$$\mu = E(X) = \lambda, \; \sigma^2 = V(X) = \lambda$$



Hình: Poisson distributions for selected values of the parameters.

Contamination is a problem in the manufacture of optical storage disks. The number of particles per centimeter squared of media surface is 0.1. The area of a disk under study is 100 cm^2.

a) Compute the probability that 12 particles occur in the area of a disk under study.

b) Compute the probability that zero particles in the area of the disk under study.

Contamination is a problem in the manufacture of optical storage disks. The number of particles per centimeter squared of media surface is 0.1. The area of a disk under study is 100 cm^2.

a) Compute the probability that 12 particles occur in the area of a disk under study.

b) Compute the probability that zero particles in the area of the disk under study.

Solution

Let X denote the number of particles in the area of under study.

a)

$$E(X) = 100 \times 0.1 = 10 \quad \text{(particles)}$$

$$\longrightarrow P(X = 12) = \frac{e^{-10}10^{12}}{12!} = 0.095$$

b) $P(X = 0) = e^{-10} = 4.54 \times 10^{-5}$

Summary

Binomial Distribution

n	=	number of trials
X	=	number of successes
p	=	probability of success
P(x)	=	$\binom{n}{x}p^x(1-p)^{n-x}$
E(X)	=	np
V(X)	=	np(1 – p)

Negative Binomial Distribution

X	=	number of trials until r successes
p	=	probability of success
P(x)	=	$\binom{x-1}{r-1}p^r(1-p)^{x-r}$, x = r, r + 1, …
E(X)	=	r/p
V(X)	=	$r(1-p)/p^2$

Poisson Distribution

λ: frequency, average number of events

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots$$

$\mu = \lambda$

$\sigma^2 = \lambda$

Geometric Distribution

X	=	number of trials
p	=	probability of success
P(x)	=	$p(1-p)^{x-1}$, x = 1, 2, …
E(X)	=	1/p
V(X)	=	$(1-p)/p^2$

Hypergeometric Distribution

N	=	number of objects
K	=	number of success-objects
n	=	sample size
P(x)	=	$\dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$
E(X)	=	np, where p = K/N
V(X)	=	$np(1-p)\dfrac{N-n}{N-1}$

Q&A

Thank you for listening!