Process Modeling by Statistical Methods (0905331)
03- Discrete Probability Distributions

Dr. Ali Khalaf Al-matar
Chemical Engineering Department
University of Jordan
aalmatar@yahoo.com

Random Variables

- **Random variable**: a quantity resulting from an experiment, by chance can assume different values
  - Discrete
  - Continuous
- **Any quantity that is a function of a random variable is a random variable itself.**

Probability distribution Function

- **Probability distribution function** (pdf)
  - an exhaustive list of the sample space (all outcomes) of an experiment and the probability associated with each outcome.
  - Must satisfy the following two conditions

$$p(x) \geq 0$$
$$\sum_{\text{all x}} p(x) = 1$$

  - Intuitively, the probability of any outcome that does not belong to the sample space is defined to be zero since it is an impossible event.

Cumulative Distribution Function

- **Cumulative distribution function** (cdf) gives the probability that a random variable X will assume a value less than or equal to a stipulated value x

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

  - Satisfies the following properties

$$0 \leq F(X) \leq 1$$
$$\text{If } x \leq y, \text{ then } F(x) \leq F(y)$$

Example 3-8

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?

The question can be answered by first finding the probability mass function of X.

$$P(X=0) = \frac{800}{850} \cdot \frac{799}{849} = 0.886$$

$$P(X=1) = 2 \cdot \frac{800}{850} \cdot \frac{50}{849} = 0.111$$

$$P(X=2) = \frac{50}{850} \cdot \frac{49}{849} = 0.003$$

Therefore,

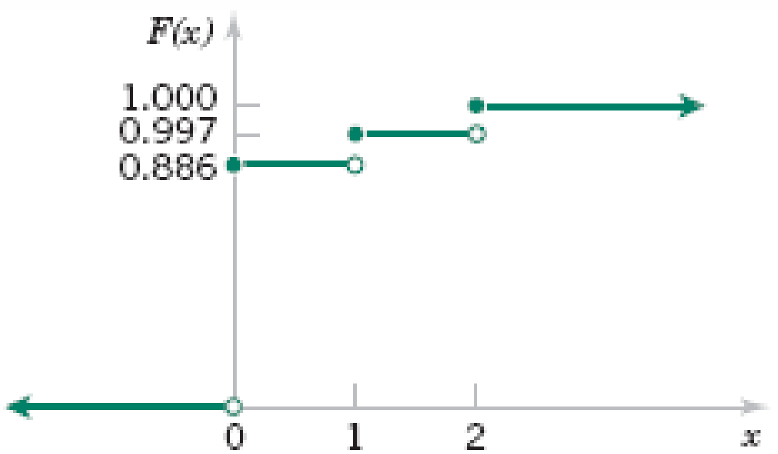
$$F(0) = P(X \le 0) = 0.886$$

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

$$F(2) = P(X \le 2) = 1$$

The cumulative distribution function for this example is graphed in Fig. 3-4. Note that $F(x)$ is defined for all x from $-\infty < x < \infty$ and not only for 0, 1, and 2.

Example 3-8



Expected Value and Variance

- The **expected value** of a random variable is a weighted average of the possible values, using the respective probabilities as weights

$$\mu = E(X) = \sum_{\text{all x}} x f(x)$$

- The variance is a weighted average. Wherein the quantities involved indicate how much individual values differ from the center of the distribution

$$\sigma^2 = V(X) = E(x-\mu)^2 = \sum_{\text{all x}} (x-\mu)^2 f(x) = \sum_{\text{all x}} x^2 f(x) - \mu^2$$

3-4 Mean and Variance of a Discrete Random Variable

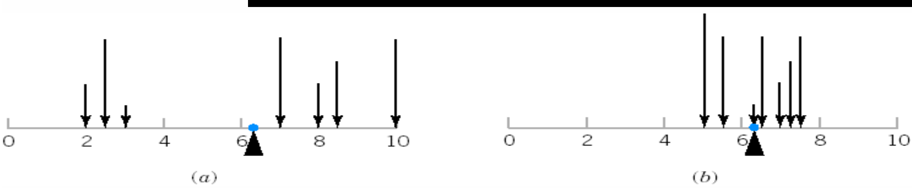


Figure 3-5 A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

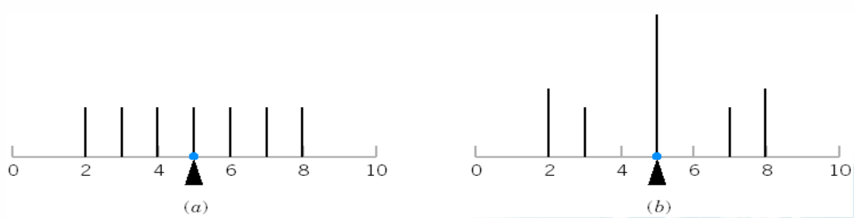


Figure 3-6 The probability distribution illustrated in Parts (a) and (b) differ even though they have equal means and equal variances.

Example 3-11

The number of messages sent per hour over a computer network 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 sent per hour.

$$E(X) = 10(0.08) + 11(0.15) + \cdots + 15(0.07) = 12.5$$

$$V(X) = 10^2(0.08) + 11^2(0.15) + \cdots + 15^2(0.07) - 12.5^2 = 1.85$$

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

Properties of $E(X)$ and $V(X)$

$$E[h(x)] = \sum_{\text{all x}} h(x) f(x)$$

$$E(c) = c$$
$$E(cX) = cE(X)$$
$$E(a + bX) = a + bE(X)$$
$$E[g(X)] \neq g[E(X)]$$

$$V(X) = \sum_{\text{all x}} [X - E(X)]^2 f(x) = E([X - E(X)]^2)$$

$$V(c) = 0$$
$$V(a + bX) = b^2 V(X)$$
$$V(X) = E(X^2) - E(X)^2$$

Example 1

Given: an auto company sells its cars with the daily distribution to the right

Wanted:

1. What type of distribution is this?
2. How many cars should the dealer expect to sell in a typical day?
3. What is the variance?

Cars Sold	$p(x)$
0	0.10
1	0.20
2	0.30
3	0.30
4	0.10

Solution

1. A discrete probability distribution
2. The mean number of cars sold is computed using the expectation value = 2.10 cars.
3. Variance

$$V(X) = \sum_{\text{all x}} x^2 f(x) - \mu^2 = 5.70 - 2.10^2$$
$$= 1.29$$

x	$p(x)$	$x\ p(x)$	$x^2\ p(x)$
0	0.10	0.00	0.00
1	0.20	0.20	0.20
2	0.30	0.60	1.20
3	0.30	0.90	2.70
4	0.10	0.40	1.60
Σ	1.00	2.10	5.70

Uniform Probability Distribution I

- A random variable X has a discrete uniform distribution if each of the n values in its range has equal probability

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

- Suppose X is a discrete uniform variable on the consecutive integers a, $a+1$, ..., b for $a \le b$ then

$$\mu = E(X) = \frac{b+a}{2}$$

$$\sigma^2 = \frac{(b-a+1)^2 - 1}{12}$$

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 $R = \{0, 1, 2, \dots, 9\}$. That is,

$$f(x) = 0.1$$

for each value in R. The probability mass function of X is shown in Fig. 3-7.

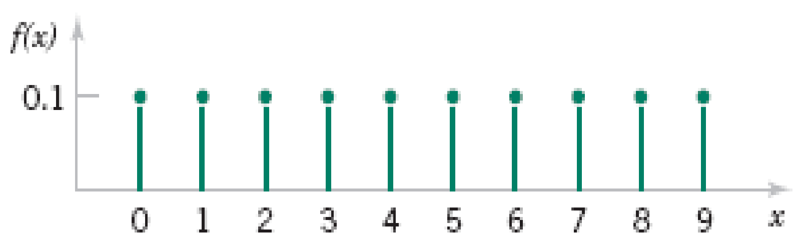


Figure 3-7 Probability mass function for a discrete uniform random variable.

Binomial Distribution: Random experiments and random variables

1. Flip a coin 10 times. Let X = number of heads obtained.
2. A worn machine tool produces 1% defective parts. Let X = number of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. Let X = the number of air samples that contain the rare molecule in the next 18 samples analyzed.
4. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X = the number of bits in error in the next five bits transmitted.
5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X = the number of questions answered correctly.
6. In the next 20 births at a hospital, let X = the number of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X = the number of patients who experience improvement.

Binomial Probability Distribution I

- **Bernoulli trial**: A trial with only two possible **mutually exclusive** outcomes (success or failure).
- A random experiment consists of n Bernoulli trials such that
  - The trials are **statistically independent**, the outcome of one trial does not affect any other trial.
  - Each trial results in only two possible outcomes (success or failure).
  - The probability of a success (p) in each trial remains constant.
- The random variable X that equals the number of trials that result in a success has a **binomial random variable** with parameters $0 < p < 1$ and $n = 1,2,....$

Binomial Probability Distribution II

The probability mass function of X is

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

n: the number of trials

x: the number of successes

p: the probability of success on each trial

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

Example

Given: based on recent experience**, 5%** of the worm gears produced by an automatic, high speed Carter-bell milling machine are defective.

Wanted: what is the probability that out of **6** gears selected at random, exactly 0, 1, 2, 3, 4, 5, 6 out of 6 will be defective?

Solution

- The binomial distribution can be used since the underlying conditions are met.

x	$f(x)$
0	0.7351
1	0.2321
2	0.03054
3	0.002143
4	8.46E-5
5	1.78E-6
6	1.56E-8

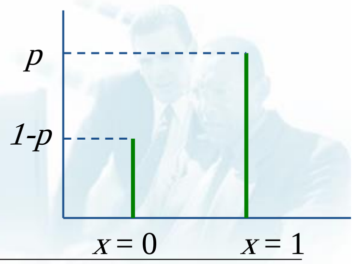
Bernoulli Probability Distribution

When n in the **Binomial is set to 1** the special distribution obtained is called **Bernoulli** distribution

$$f(n=1; x; p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \Rightarrow f(x;p) = p^x(1-p)^{1-x}$$

x: the number of successes $x \in [0,1]$

p: the probability of success on each trial



Geometric Distribution

- In a series of Bernoulli trials (independent trials with constant probability of success), let the random variable *X* denote the **number of trials until the first success**.
- Then *X* is a geometric random variable with parameter $0 < p < 1$ and

$$f(x; p) = p(1-p)^{x-1}, \quad x = 1, 2, \cdots$$

$$\mu = E(X) = 1/p$$

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

Geometric Distribution Example

Suppose that a primary device has failed as a result of a high temperature environment. The probability that an electronic switch will successfully activate a backup device is 0.6. If switch failures are statistically independent and the switches are tried one at a time, how many parallel switches are required to achieve at least 95% probability of successful switching?

Solution

- This is a **countably infinite** sample space
- The geometric distribution applies in this case. We want to evaluate

$$\sum_{x=1}^{?} p(1-p)^{x-1} > 0.95 \Rightarrow \sum_{x=1}^{?} 0.6(0.4)^{x-1} > 0.95$$

x	1	2	3	4	…
$F(x)$	0.6	0.84	0.936	0.9744	…

- Four or less is 0.9744, so we use 4 switches to get a probability in excess of 95%.

- A chemical plant has a pump that will operate with a probability of 0.9. How many pumps, in parallel, do you need to have a probability of the pump(s) operating greater than 0.99?

Negative Binomial Distribution

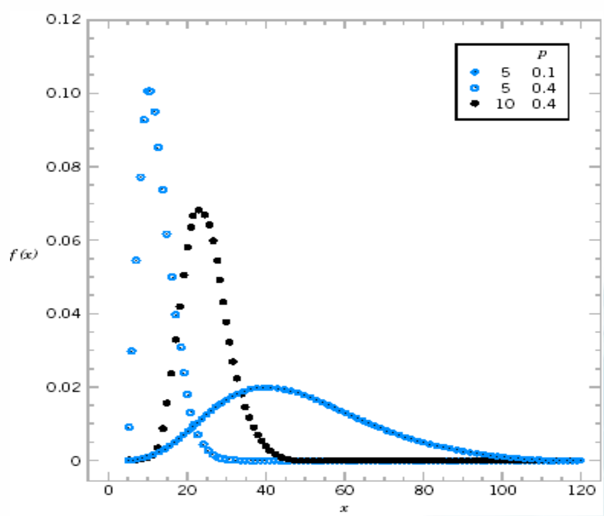
- In a series of Bernoulli trials (independent trials with constant probability of success), let the random variable X denote the **number of trials until r successes occur**. Then X is a negative binomial random variable with parameter 0 < p < 1 and $r = 1, 2, \ldots$.
- The geometric distribution is a a special case where $r = 1$.

$$f(x;r;p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, r+2, \cdots$$

$$\mu = E(X) = r/p$$

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

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



Example Negative Binomial

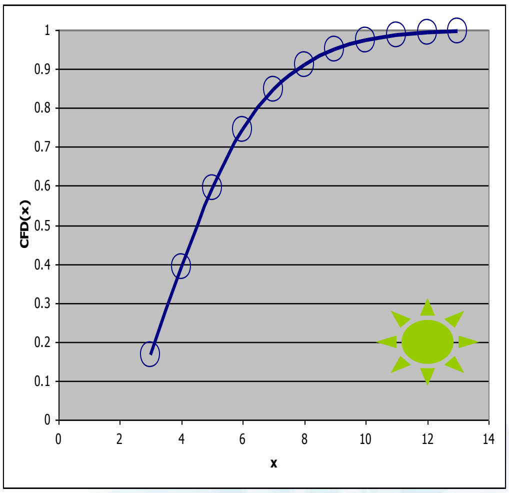
The probability that a car will pass EPA standards is 0.45. How many such cars will need to be checked before the probability is greater than 0.95 for finding 3 cars that pass the EPA standard?

Solution

- Negative binomial; solve

$$\sum_{x=3}^{\cdots} \binom{x-1}{r-1} p^r (1-p)^{x-r} > 0.95$$

- 9 cars has to be inspected



Hypergeometric Distribution - I

- A **finite population** is a population consisting of a small number of individuals, objects, or measurements.
- When sampling from small or (finite) populations **without replacement** the probability does not remain the same
  - preaching the premise of binomial applicability.
- Hypergeometric distribution is used instead

Hypergeometric Distribution - II

- To apply the Hypergeometric distribution
  - The sample of size $n \leq N$ is selected from a finite population (size N objects) **without replacement**
    - K ($K \leq N$) objects classified as success
    - N - K objects classified as failure
  - Typically, sample size is greater than 5% of the size of the population, N.

Hypergeometric Distribution - III

- Let the random variable X denote the number of successes in the sample. Then X is a hypergeometric random variable with

$$f(n;N;K;x)=\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}=\frac{\left({}_{K}C_{x}\right)\cdot\left({}_{N-K}C_{n-x}\right)}{\left({}_{N}C_{n}\right)}$$

$$x=\max\{0,n+K-N\}\text{ to }\min\{K,n\}$$

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

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

Proportion of successes $p = K/N$

Finite population correction factor $\frac{N-n}{N-1}$

Example Hypergeometric

Fifty ($N = 50$) computers were manufactured during a certain week. Forty ($K = 40$) operated perfectly, and 10 had at least one defect. A sample of 5 is selected at random ($n = 5$). What is the probability that 4 out of the 5 will operate perfectly?

Solution-I

- Note that sampling is done without replacement, and the sample size is 5/50 or 10% which is greater than the 5% requirement to apply the hypergeometric distribution.

N = 50, n = 5, K = 40, x = 4.

$$f(4) = \frac{\left({}_{40}C_4\right)\left({}_{50-40}C_{5-4}\right)}{\left({}_{50}C_5\right)}$$

$$= \frac{(91390)(10)}{2118760} = 0.431$$

Solution-II

x	*p*(*x*) *Hypergeometric*	*p*(*x*) *Binomial (n=5,p=0.8)*
0	0.000	0.000
1	0.004	0.006
2	0.044	0.051
3	0.210	0.205
4	0.431	0.410
5	0.311	0.328

Poisson Distribution - I

- Binomial distribution for small p and large n is time consuming computationally
  - The **limiting** form of the binomial distribution for **very small** p **and high** n is called the **Poisson distribution**.
  - Often it is referred to as the "**law of improbable events**"

$$\lim_{n\to\infty} P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{n}{x}\left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0,1,2,\cdots$$

Poisson Distribution - II

- Usually, used to characterize physical situations in which the number of events during a specific period of time is of interest
  - Number of customers arriving at a bank or a restaurant during one hour (waiting line applications)
  - Number of cars crossing a certain bridge during a certain period of time
  - Baggage lost occurrences on a certain airline
- A Poisson process has **no memory**. The number of events occurring in one segment of time or space is independent of the number of events in non-overlapping segment
- The mean process rate λ **must remain constant** for the entire time span or space considered
- The smaller the segment of time or space, the less likely it is for more than one event to occur in that segment.
  - As the segment size tends to 0, the probability of 2 or more occurrences approaches 0.

Poisson Distribution - III

- Given an interval of real numbers, assume counts occur at random **throughout the interval**. If the interval can be partitioned into subintervals of small enough length such that
  1. the probability of more than one count in a subinterval is zero,
  2. the probability of one count in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and
  3. the count in each subinterval is independent of other subintervals, the random experiment is called a **Poisson process.**
- The random variable X that equals the number of counts in the interval is a Poisson random variable with parameter $0 < \lambda$, and the probability mass function of X is

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

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

$$\sigma^2 = V(X) = \lambda = np$$

Poisson Example I

Baggage is rarely lost by Poisson Airlines. A random sample of 1000 flights showed that a total of 300 bags were lost. Find the probability of not losing any bag, exactly one bag is lost.

Poisson distribution is assumed

λ=average of lost baggage per flight; λ=300/1000 = 0.3

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

$$p(0) = \frac{0.3^0 e^{-0.3}}{0!} = 0.7408$$

$$p(1) = \frac{0.3^1 e^{-0.3}}{1!} = 0.2222$$

Poisson Example II

The white blood-cell count of a healthy individual can average as low as 6000/mL. To detect white-cell deficiency, a 0.001 mL drop of blood is taken and the number of white cells is found. How many white cells are expected in a healthy individual? If at most two are found, is there an evidence of white-cell deficiency?

$\lambda=(0.001)(6000)=6$

In a healthy individual we would expect to see 6 white-cells on average. This is a Poisson distribution

$$p(X \le 2) = \sum_{x=0}^{2} \frac{\lambda^x e^{-\lambda}}{x!}$$
$$= 0.062$$