Statistical Models In Simulation

Discrete Random Variables

- X is a discrete random variable if the number of possible values of X is finite, or countable infinite.
- Example: Consider packets arriving at a router.
  - Let X be the number of packets arriving each second at a router.

    R_X = possible values of X (range space of X) = $\{0,1,2,\ldots\}$

    $p(x_i)$ = probability the random variable X is x_i , $p(x_i) = P(X = x_i)$
  - $p(x_i)$, $i = 1,2, \ldots$ must satisfy:

    1. $p(x_i) \geq 0, \quad \text{for all } i$

    2. $\sum_{i=1}^{\infty} p(x_i) = 1$

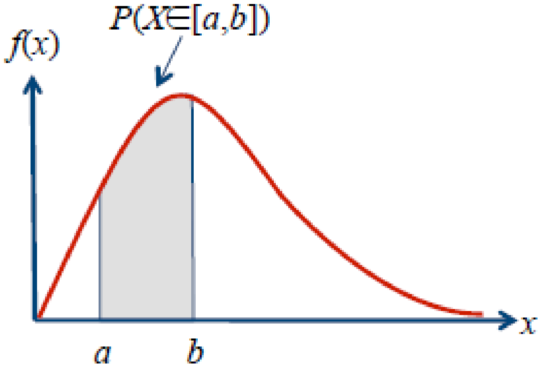
  - The collection of pairs $(x_i, p(x_i))$, $i = 1,2,\ldots$, is called the **probability distribution** of X, and
  - $p(x_i)$ is called the **probability mass function** (**PMF**) of X.

Continuous Random Variables

- X is a continuous random variable if its range space R_X is an interval or a collection of intervals.
- The probability that X lies in the interval $[a, b]$ is given by:

$$P(a \le X \le b) = \int_a^b f(x)dx$$

- $f(x)$ is called the **probability density function** (PDF) of X, and satisfies:



$$1.\ f(x) \ge 0,\ \text{for all } x \text{ in } R_X$$

$$2.\ \int_{R_X} f(x)dx = 1$$

$$3.\ f(x) = 0,\ \text{if } x \text{ is not in } R_X$$

- Properties

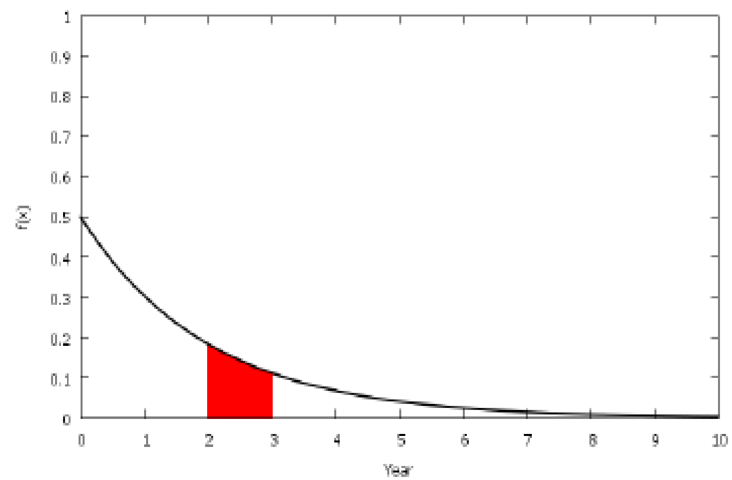
$$1.\ P(X = x_0) = 0,\ \text{because} \int_{x_0}^{x_0} f(x)dx = 0$$

$$2.\ P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$

Continuous Random Variables

- Example: Life of an inspection device is given by X, a continuous random variable with PDF:

$$f(x) = \begin{cases} \frac{1}{2} e^{-x/2} & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$



- X has exponential distribution with mean 2 years
- Probability that the device's life is between 2 and 3 years is:

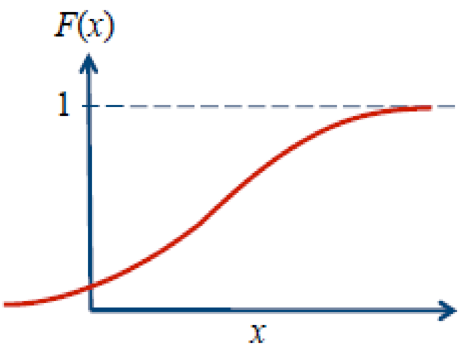
$$P(2 \leq x \leq 3) = \frac{1}{2} \int_2^3 e^{-x/2} dx = 0.145$$

Cumulative Distribution Function

- Cumulative Distribution Function (**CDF**) is denoted by $F(x)$, where

$$F(x) = P(X \le x)$$

  - If X is discrete, then $$F(x) = \sum_{x_i \le x} p(x_i)$$
  - If X is continuous, then $$F(x) = \int_{-\infty}^{x} f(t)dt$$



- Properties
  1. F is nondecreasing function. If $a \le b$, then $F(a) \le F(b)$
  2. $\lim_{x \to \infty} F(x) = 1$
  3. $\lim_{x \to -\infty} F(x) = 0$

- All probability questions about X can be answered in terms of the CDF:

$$P(a \le X \le b) = F(b) - F(a), \text{ for all } a \le b$$

Cumulative Distribution Function

- Example: The inspection device has CDF:

$$F(x) = \frac{1}{2}\int_0^x e^{-t/2}dt = 1 - e^{-x/2}$$

  - The probability that the device lasts for less than 2 years:

$$P(0 \le X \le 2) = F(2) - F(0) = F(2) = 1 - e^{-1} = 0.632$$

  - The probability that it lasts between 2 and 3 years:

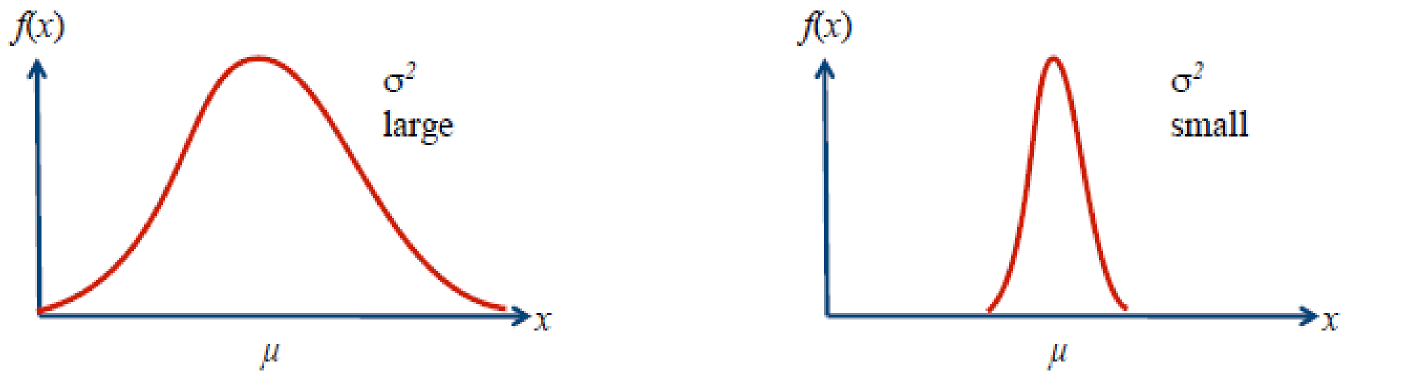
$$P(2 \le X \le 3) = F(3) - F(2) = \left(1 - e^{-\frac{3}{2}}\right) - \left(1 - e^{-1}\right) = 0.145$$

Expected value

- The expected value of X is denoted by $E(X)$
  - If X is discrete $E(X) = \sum_{\text{all } i} x_i p(x_i)$
  - If X is continuous $E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$
  - a.k.a the mean, m, μ, or the 1st moment of X
  - A measure of the central tendency

Variance

- The variance of X is denoted by $V(X)$ or $Var(X)$ or σ^2
  - Definition: $V(X) = E(\,(X - E[X])^2\,)$
  - Also $V(X) = E(X^2) - (\,E(X)\,)^2$
  - A measure of the spread or variation of the possible values of X around the mean



Standard deviation

- The standard deviation (SD) of X is denoted by σ
  - Definition: $\sigma = \sqrt{V(x)}$
  - The standard deviation is **expressed** in the **same units** as the **mean**
  - Interprete σ always together with the mean
- Attention:
  - The standard deviation of two different data sets may be difficult to compare

Expected value and variance: Example

- Example: The mean of life of the previous inspection device is:

$$E(X) = \frac{1}{2}\int_0^{\infty} xe^{-x/2}dx = -xe^{-x/2}\Big|_0^{\infty} + \int_0^{\infty} e^{-x/2}dx = 2$$

- To compute the variance of X, we first compute $E(X^2)$:

$$E(X^2) = \frac{1}{2}\int_0^{\infty} x^2e^{-x/2}dx = -x^2e^{-x/2}\Big|_0^{\infty} + \int_0^{\infty} e^{-x/2}dx = 8$$

- Hence, the variance and standard deviation of the device's life are: $V(X) = 8 - 2^2 = 4$

$$\sigma = \sqrt{V(X)} = 2$$

Expected value and variance: Example

$$E(X) = \frac{1}{2}\int_0^\infty xe^{-x/2}dx = -xe^{-x/2}\Big|_0^\infty + \int_0^\infty e^{-x/2}dx = 2$$

Partial Integration

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx$$

Set

$$u(x) = x$$

$$v'(x) = e^{-x/2}$$

$$\Rightarrow$$

$$u'(x) = 1$$

$$v(x) = -2e^{-x/2}$$

$$E(X) = \frac{1}{2}\int_0^\infty xe^{-x/2}dx = \frac{1}{2}\left(x\cdot(-2e^{-x/2})\Big|_0^\infty - \int_0^\infty 1\cdot(-2e^{-x/2})dx\right)$$

Mean and variance of sums

- If $x_1, x_2, \ldots, x_k$ are k random variables and if $a_1, a_2, \ldots, a_k$ are k constants, then

$$E(a_1x_1+a_2x_2+\ldots+a_kx_k) = a_1E(x_1)+a_2E(x_2)+\ldots+a_kE(x_k)$$

- For independent variables

$$\mathrm{Var}(a_1x_1 + a_2x_2 + \cdots + a_kx_k) = a_1^2\,\mathrm{Var}(x_1) + a_2^2\,\mathrm{Var}(x_2) + \ldots + a_k^2\,\mathrm{Var}(x_k)$$

Coefficient of variation

- The ratio of the standard deviation to the mean is called coefficient of variation (C.O.V.)
  - Dimensionless
  - Normalized measure of dispersion

$$C.O.V = \frac{\text{standard deviation}}{\text{mean}} = \frac{\sigma}{\mu} \qquad , \mu > 0$$

  - Can be used to compare different datasets, instead the standard deviation.

Covariance

- Given two random variables x and y with μ_x and μ_y, their covariance is defined as

$$\mathrm{Cov}(x, y) = \sigma^2_{xy} = E[(x-\mu_x)(y-\mu_y)] = E(xy) - E(x)\,E(y)$$

  - $\mathrm{Cov}(x, y)$ measures the dependency of x and y, i.e., how x and y vary together.

- For independent variables, the covariance is **zero**, since

$$E(xy) = E(x)E(y)$$

Correlation coefficient

- The normalized value of covariance is called the correlation coefficient or simply correlation

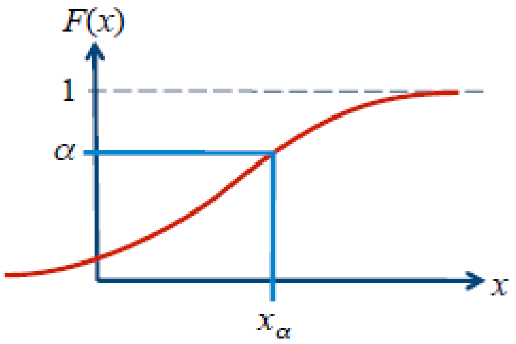
$$\text{Correlation}(x, y) = \rho_{x,y} = \frac{\sigma_{xy}^2}{\sigma_x \sigma_y}$$

- The correlation lies between -1 and +1

Quantile

- The x value at which the CDF takes a value α is called the **α-quantile** or **100α-percentile**. It is denoted by x_α.

$$P(X \le x_\alpha) = F(x_\alpha) = \alpha \quad , \ \alpha \in [0,1]$$

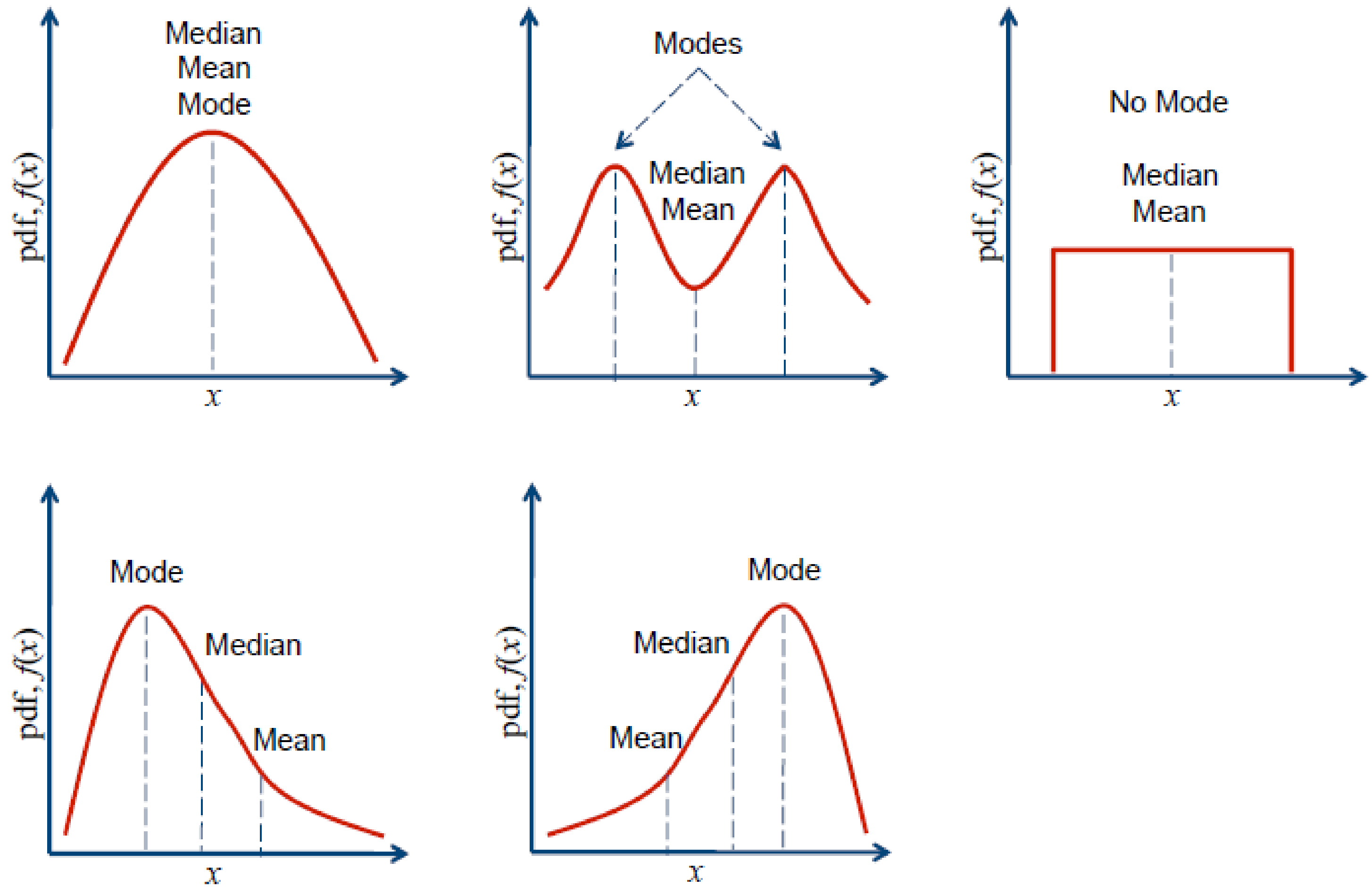


- Relationship:
  - The **median** is the **50-percentile** or **0.5-quantile**

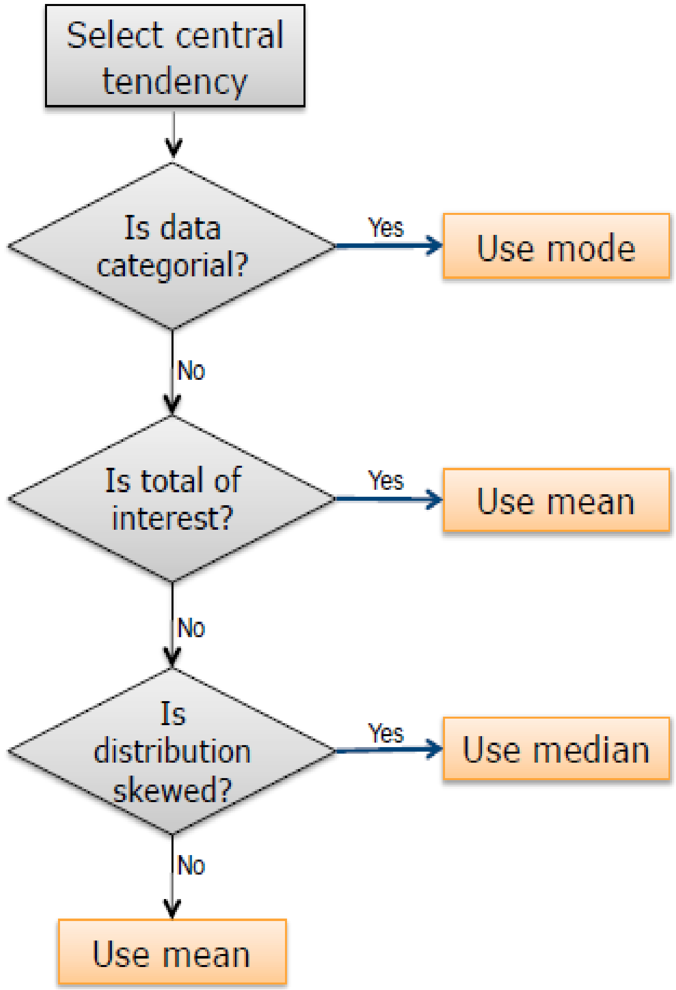
Mean, median, and mode

- Three different indices for the central tendency of a distribution:
  - Mean: $$E(X) = \mu = \sum_{i=1}^{n} p_i x_i = \int_{-\infty}^{\infty} x \cdot f(x) dx$$
  - Median: The 0.5-quantile, i.e., the x_i for that half of the values are smaller and the other half is larger.
  - Mode: The most likely value, i.e., the x_i that has the highest probabiliy p_i or the x at which the PDF is maximum.

Mean, median, and mode



Selecting among mean, median, and mode



Useful Statistical Models

- Queueing systems
- Inventory and supply-chain systems
- Reliability and maintainability
- Limited data

Queueing Systems

In a queueing system, inter arrival and service-time patterns can be probabilistic

Sample statistical models for inter arrival or service time distribution:

- *Exponential distribution:* if service times are completely random
- *Normal distribution:* fairly constant but with some random variability (either positive or negative)
- *Truncated normal distribution:* similar to normal distribution but with restricted value.
- *Gamma and Weibull distribution:* more general than exponential (involving location of the modes of pdf's and the shapes of tails.)

Inventory and supply chain

In realistic inventory and supply-chain systems, there are at least three random variables:

- The number of units demanded per order or per time period
- The time between demands
- The lead time (time between the placing of an order for stocking the inventory system and the receipt of that order)

Sample statistical models for lead time distribution:

- *Gamma*

Sample statistical models for demand distribution:

- *Poisson:* simple and extensively tabulated.
- *Negative binomial distribution*: longer tail than Poisson (more large demands).
- *Geometric:* special case of negative binomial given at least one demand has occurred.

Reliability and maintainability

Time to failure (TTF)

- *Exponential:* failures are random
- *Gamma:* for standby redundancy where each component has an exponential TTF
- *Weibull:* failure is due to the most serious of a large number of defects in a system of components
- *Normal:* failures are due to wear

Other areas

For cases with limited data, some useful distributions are:

- *Uniform, triangular and beta*

Other distribution: *Bernoulli, binomial and hyper-exponential.*

Discrete Distributions

Discrete random variables are used to describe random phenomena in which only *integer values* can occur.

In this section, we will learn about:

- Bernoulli trials and Bernoulli distribution
- Binomial distribution
- Geometric and negative binomial distribution
- Poisson distribution

Bernoulli Trials and Bernoulli Distribution

The **Bernoulli distribution** is a special case of the binomial distribution where there is only a single trial. It is the simplest type of probability distribution, representing a single experiment that has exactly two possible outcomes: success or failure.

Key Characteristics of the Bernoulli Distribution:

Binary Outcomes:

- The experiment or trial has only two possible outcomes, typically labeled as "success" (often represented by 1) and "failure" (often represented by 0).

Single Trial:

- There is only one trial or experiment, meaning the Bernoulli distribution describes the outcome of just one event.

Probability of Success (p):

- The probability of success in the trial is denoted by p. Consequently, the probability of failure is $1-p$.

Bernoulli Trials and Bernoulli Distribution

Bernoulli Trials:

- Consider an experiment consisting of n trials, each can be a success or a failure.
  - Let $X_j = 1$ if the j^{th} trial is a success with probability p
  - and $X_j = 0$ if the j^{th} trial is a failure

$$p_j(x_j) = p(x_j) = \begin{cases} p, & x_j = 1, \quad j = 1,2,\ldots,n \\ 1-p = q, & x_j = 0, \quad j = 1,2,\ldots,n \\ 0, & \text{otherwise} \end{cases}$$

- For one trial, it is called the Bernoulli distribution where $E(X_j) = p$ *and* $V(X_j) = p(1-p) = pq$

Bernoulli process:

- The n Bernoulli trials where trails are independent:

$$p(x_1, x_2, \ldots, x_n) = p_1(x_1)p_2(x_2) \ldots p_n(x_n)$$

Example of a Bernoulli Distribution:

Consider flipping a fair coin once. The possible outcomes are:

- Success (getting heads) with probability $p = 0.5$
- Failure (getting tails) with probability $1 - p = 0.5$

Here, X is a Bernoulli random variable that can take the value 1 (for heads) or 0 (for tails).

- The expected value $E(X) = 0.5$, representing the probability of getting heads.
- The variance $\mathrm{Var}(X) = 0.5 \times 0.5 = 0.25$.

Relation to the Binomial Distribution:

The Bernoulli distribution is a special case of the binomial distribution with n=1. In other words, a binomial distribution with only one trial is equivalent to a Bernoulli distribution.

Applications of the Bernoulli Distribution:

Coin Tosses: Representing the probability of getting heads (success) or tails (failure) in a single coin toss.

Pass/Fail Tests: Determining the probability of passing or failing a single test or trial.

Customer Behavior: Modeling whether a customer makes a purchase (success) or does not make a purchase (failure) in a single interaction.

Used to generate discrete random variate such as binomial and geometric

The Bernoulli distribution is fundamental in probability and statistics, serving as the building block for more complex distributions like the binomial distribution. It is used in situations where there is a simple, binary outcome in a single trial or experiment.

Binomial Distribution

- Gives probability of number of successes in n independent trials, when probability of success p on single trial is a constant.
- To determine the probability of a particular outcome with all the success.
- E.g. what is the probability of 8 or more "tails" in 10 tosses of a fair coin?
- Can be sometimes approximated by normal or by Poisson distribution.

Applications: used frequently in quality control, reliability, survey sampling and other industrial problems.

- To classify defective or non defective items in a batch of size n.
- Find out demand (no of items) placed by a customer in case of inventory problem.

The **binomial distribution** is a discrete probability distribution that describes the number of successes in a fixed number of independent trials of a binary (yes/no or success/failure) experiment. Each trial is identical, and the probability of success remains constant across all trials.

Key Characteristics of the Binomial Distribution:

Binary Outcomes:

- Each trial in a binomial experiment has only two possible outcomes: "success" or "failure."

Fixed Number of Trials (n):

- The number of trials, denoted by n, is fixed in advance.

Constant Probability of Success (p):

- The probability of success in each trial is the same and is denoted by p. Consequently, the probability of failure is 1−p.

Independent Trials:

- The trials are independent, meaning the outcome of one trial does not affect the outcome of another.

Binomial Distribution

The number of successes in n Bernoulli trials, X, has a binomial distribution.

$$p(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0,1,2,\ldots,n \\ 0, & \text{otherwise} \end{cases}$$

The number of outcomes having the required number of successes and failures

Probability that there are x successes and (n-x) failures

- Easy approach is to consider the binomial distribution X as a sum of n independent Bernoulli Random variables *(X=X1+X2+…+Xn)*
- The mean, $E(X) = p + p + \ldots + p = n*p$
- The variance, $V(X) = pq + pq + \ldots + pq = n*pq$

Example of a Binomial Distribution:

Suppose you flip a fair coin 10 times (so $n = 10$), and you want to find the probability of getting exactly 4 heads. Here, the probability of getting a head on any single flip is $p = 0.5$, and the probability of getting tails is $1 - p = 0.5$.

Using the binomial distribution formula:

$$P(X = 4) = \binom{10}{4} (0.5)^4 (0.5)^6$$

First, calculate the binomial coefficient $\binom{10}{4}$:

$$\binom{10}{4} = \frac{10!}{4!(10-4)!} = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} = 210$$

Then, calculate the probability:

$$P(X = 4) = 210 \times (0.5)^{10} = 210 \times \frac{1}{1024} \approx 0.205$$

So, the probability of getting exactly 4 heads in 10 flips is approximately 0.205 (or 20.5%).

Example: The probability of a chip failure is 0.05. everyday a random sample of size 14 is taken. What is the probability that

i) At most 3 will fail

ii) At least 3 will fail.

Applications of the Binomial Distribution:

Quality Control: For example, determining the probability of finding a certain number of defective products in a batch.

Clinical Trials: Estimating the likelihood of a certain number of patients responding to a treatment.

Sports: Calculating the probability of a player achieving a certain number of hits in a given number of attempts.

The binomial distribution is widely used in statistics to model scenarios where there are fixed numbers of independent trials, each with the same probability of success.

Geometric Distribution

- Gives probability of requiring exactly x Bernoulli trials before the first success is achieved.
- E.g. determination of probability of requiring exactly five tests firings before first success is achieved.
- A doctor is seeking an anti-depressant for a newly diagnosed patient. Suppose that, of the available anti-depressant drugs, the probability that any particular drug will be effective for a particular patient is p=0.6.
- What is the probability that the first drug found to be effective for this patient is the first drug tried, the second drug tried, and so on? What is the expected number of drugs that will be tried to find one that is effective?
- A patient is waiting for a suitable matching kidney donor for a transplant. If the probability that a randomly selected donor is a suitable match is p=0.1, what is the expected number of donors who will be tested before a matching donor is found?

The **geometric distribution** is a discrete probability distribution that models the number of trials required to get the first success in a sequence of independent and identically distributed Bernoulli trials (each with the same probability of success). It is used to determine the probability that the first occurrence of success requires a specific number of trials.

Key Characteristics of the Geometric Distribution:

Binary Outcomes:

- Each trial in a geometric experiment has two possible outcomes: success or failure.

Independent Trials:

- The trials are independent, meaning the outcome of one trial does not affect the outcome of another.

Constant Probability of Success (p):

- The probability of success p is constant across all trials.

Number of Trials Until First Success:

- The geometric distribution is concerned with the number of trials, X, until the first success occurs.

The **negative binomial distribution** is a discrete probability distribution that models the number of trials required to achieve a specified number of successes in a sequence of independent and identically distributed Bernoulli trials. It generalizes the geometric distribution, which is a special case where the number of successes required is 1.

Key Characteristics of the Negative Binomial Distribution:

Binary Outcomes:

- Each trial in the negative binomial experiment has two possible outcomes: success or failure.

Independent Trials:

- The trials are independent, meaning the outcome of one trial does not affect the outcome of another.

Constant Probability of Success (p):

- The probability of success p is constant across all trials.

Number of Successes (r):

- The distribution is concerned with the number of trials needed to achieve a specified number of successes r.

The probability mass function (PMF) of the negative binomial distribution is given by:

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

where:

- X is the random variable representing the number of trials needed to achieve r successes.
- k is the number of trials (where $k \geq r$).
- p is the probability of success on each trial.
- $1 - p$ is the probability of failure on each trial.
- $\binom{k-1}{r-1} = \frac{(k-1)!}{(r-1)!(k-r)!}$ is the binomial coefficient, representing the number of ways to arrange $r - 1$ successes in the first $k - 1$ trials.

Special Case - Geometric Distribution:

The geometric distribution is a special case of the negative binomial distribution when r=1. In this case, the distribution models the number of trials needed to achieve the first success.

Mean (Expected Value):

The mean (expected value) of a negative binomial distribution is given by:

$$E(X) = \frac{r}{p}$$

This represents the average number of trials needed to achieve r successes.

Variance:

The variance of a negative binomial distribution is given by:

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

This variance measures the variability in the number of trials required to achieve r successes.

Example of a Negative Binomial Distribution:

Suppose you are rolling a fair six-sided die, and you want to know the probability that the 3rd time you roll a 6 occurs on the 7th roll.

Here, the probability of success (rolling a 6) is $p = \frac{1}{6}$, and the probability of failure (not rolling a 6) is $1 - p = \frac{5}{6}$. You are interested in the number of trials $k = 7$ required to achieve $r = 3$ successes.

Using the negative binomial distribution formula:

$$P(X = 7) = \binom{7-1}{3-1} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{7-3}$$

First, calculate the binomial coefficient:

$$\binom{6}{2} = \frac{6!}{2!(6-2)!} = \frac{6 \times 5}{2 \times 1} = 15$$

Then, calculate the probability:

$$P(X = 7) = 15 \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^4 = 15 \times \frac{1}{216} \times \frac{625}{1296} \approx 0.032$$

So, the probability that the 3rd 6 occurs on the 7th roll is approximately 0.032, or 3.2%.

Geometric & Negative Binomial Distribution

Geometric distribution (Used frequently in data networks)

- The number of Bernoulli trials, X, to achieve the 1st success:

$$P(FFF....FS) = p(x) = \begin{cases} q^{x-1}p, & x = 0,1,2,\ldots,n \\ 0, & \text{otherwise} \end{cases}$$

- *$E(x) = 1/p$, and $V(X) = q/p^2$*

Negative binomial distribution

- The number of Bernoulli trials, X, until the kth success
- If Y is a negative binomial distribution with parameters p and k, then:

$$p(x) = \begin{cases} \binom{y-1}{k-1} q^{y-k} p^k, & y = k, k+1, k+2, \ldots \\ 0, & \text{otherwise} \end{cases}$$

- *$E(Y) = k/p$, and $V(X) = kq/p^2$*
- *Y is the sum of k independent geometric RVs*

Suppose you are rolling a fair six-sided die, and you want to know the probability that the first time you roll a 6 occurs on the 3rd roll.

Here, the probability of success (rolling a 6) is $p = \frac{1}{6}$, and the probability of failure (not rolling a 6) is $1 - p = \frac{5}{6}$.

Using the geometric distribution formula:

$$P(X = 3) = \left(\frac{5}{6}\right)^{3-1} \times \frac{1}{6} = \left(\frac{5}{6}\right)^{2} \times \frac{1}{6} \approx 0.1157$$

So, the probability that you first roll a 6 on the 3rd roll is approximately 0.1157, or 11.57%.

Properties of the Geometric Distribution:

1. **Memorylessness:**
   - The geometric distribution has the "memoryless" property, meaning the probability of success on the next trial is independent of the number of failures that have already occurred. Mathematically, $P(X > k + m | X > k) = P(X > m)$.
2. **Skewness:**
   - The geometric distribution is typically right-skewed, meaning there are more small values (fewer trials) and fewer large values (more trials) in the distribution.

Applications of the Geometric Distribution:

- **Quality Control:** Determining the number of items inspected before finding the first defective item.
- **Customer Behavior:** Modeling the number of customer interactions before the first purchase is made.
- **Reliability Testing:** Estimating the number of trials needed before the first system failure occurs.

Geometric & Negative Binomial Distribution

Example: 40% of the assembled ink-jet printers are rejected at the inspection station. Find the probability that the first acceptable ink-jet printer is the third one inspected. Considering each inspection as a Bernoulli trial with $q=0.4$ *and* $p=0.6$,

$$p(3) = 0.4^2(0.6) = 0.096$$

Thus, in only about 10% of the cases is the first acceptable printer is the third one from any arbitrary starting point

What is the probability that the third printer inspected is the second acceptable printer?

Use Negative Binomial Distribution with $y=3$ *and* $k=2$

$$p(3) = \binom{3-1}{2-1} 0.4^{3-2} (0.6)^2 = 0.288$$

Poisson Distribution

- Gives probability of exactly x independent occurrences during a given period of time. If events take place independently and at constant rate.
- May also represent number of occurrences over constant areas or volumes.
- E.g. used to represent distribution of number of defects in a piece of material, customer arrivals, insurance claims, incoming telephone calls etc..
- Frequently used as approximation to binomial distribution.

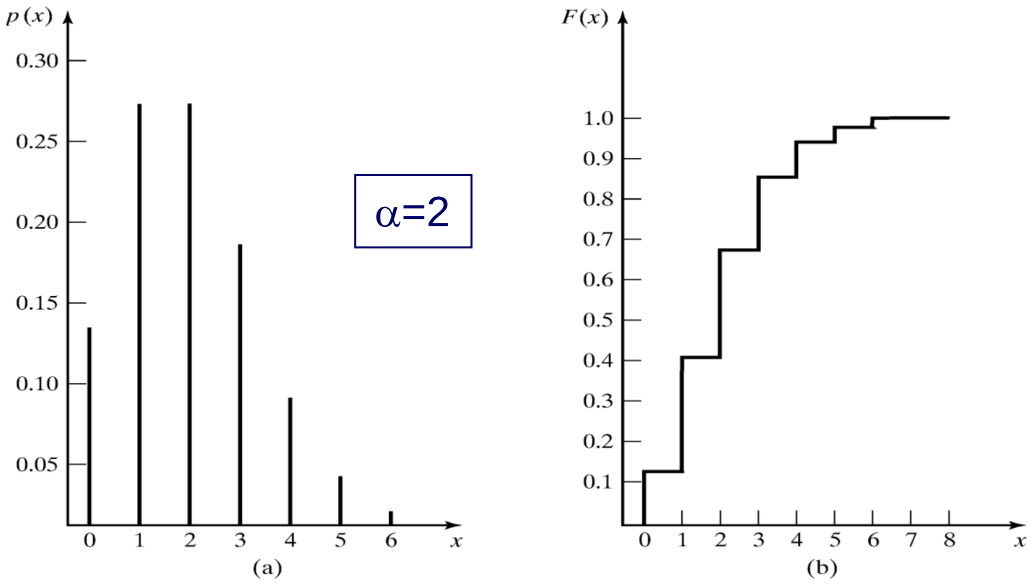
Poisson Distribution

Poisson distribution describes many random processes quite well and is mathematically quite simple. The pmf and cdf are:

$$p(x) = \begin{cases} \dfrac{e^{-\alpha}\alpha^x}{x!}, & x = 0,1,\ldots \\ 0, & \text{otherwise} \end{cases} \qquad F(x) = \sum_{i=0}^{x} \frac{e^{-\alpha}\alpha^i}{i!}$$

where $\alpha > 0$

- ***E(X) = α = V(X)***



Poisson Distribution

Example: A computer repair person is "beeped" each time there is a call for service. The number of beeps per hour ~ Poisson(α = 2 per hour).

- The probability of three beeps in the next hour:

  $p(3) = e^{-2} 2^3 / 3! = 0.18$

  also, $p(3) = F(3) – F(2) = 0.857-0.677=0.18$

- The probability of two or more beeps in a 1-hour period:

  $p(2 \text{ or more}) = 1 – p(0) – p(1)$

  $= 1 – F(1)$

  $= 0.594$

The number of accidents in a year to taxi driver in Mumbai follows a Poisson distribution with mean equal to 3. out of 100 taxi drivers, find approximately the number of drivers with:

i) No accident in a year

ii) More than 3 accidents in a year

Continuous Distributions

Continuous random variables can be used to describe random phenomena in which the variable can take on any value in some interval.

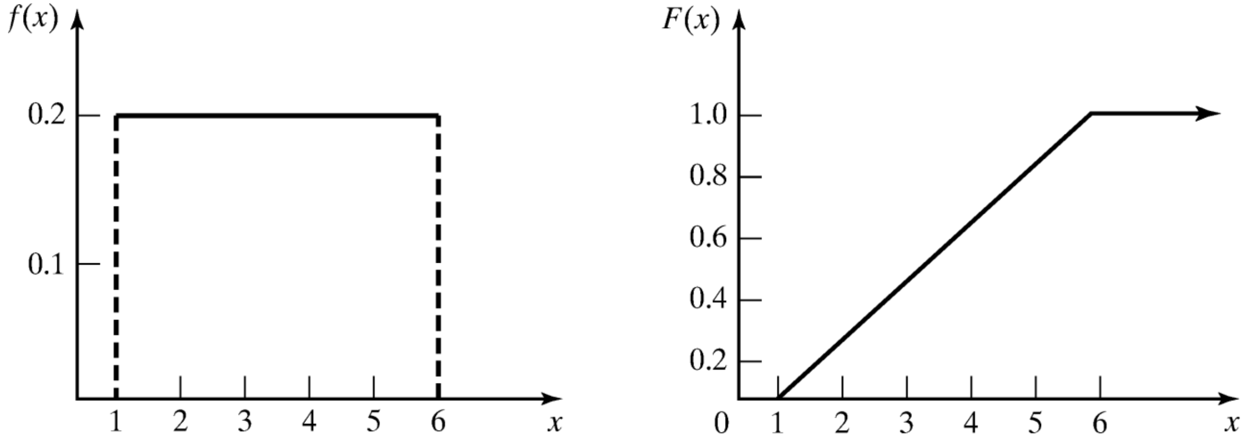
In this section, the distributions studied are:

- Uniform
- Exponential
- Normal
- Weibull
- Lognormal

Uniform Distribution

A random variable X is uniformly distributed on the interval (a,b), $U(a,b)$, if its pdf and cdf are:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases} \qquad F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}$$



Example with $a = 1$ and $b = 6$

Uniform Distribution

Properties

- $P(x_1 \leq X < x_2)$ is proportional to the length of the interval $[F(x_2) - F(x_1) = (x_2\text{-}x_1)/(b\text{-}a)]$
- $E(X) = (a+b)/2$ $V(X) = (b\text{-}a)^2/12$

U(0,1) provides the means to generate random numbers, from which random variates can be generated.

Example: In a warehouse simulation, a call comes to a forklift operator about every 4 minutes. With such a limited data, it is assumed that time between calls is uniformly distributed with a mean of 4 minutes with ($a=0$ and $b=8$)

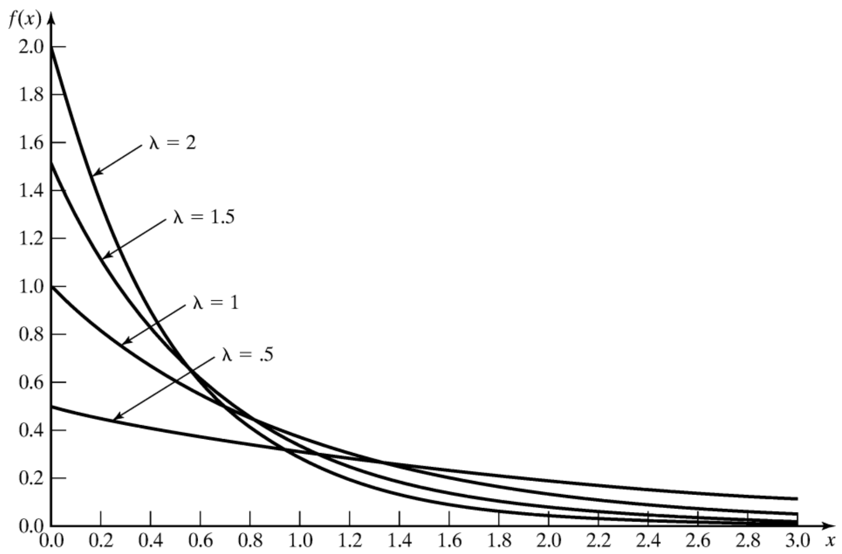
Exponential Distribution

A random variable X is exponentially distributed with parameter λ > 0 if its pdf and cdf are:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

$$F(x) = \begin{cases} 0, & x \prec 0 \\ \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

- $E(X) = 1/\lambda \quad V(X) = 1/\lambda^2$
- Used to model interarrival times when arrivals are completely random, and to model service times that are highly variable
- For several different exponential pdf's (see figure), the value of intercept on the vertical axis is λ, and all pdf's eventually intersect.



Exponential Distribution

Example: A lamp life (in thousands of hours) is exponentially distributed with failure rate ($\lambda = 1/3$), hence, on average, 1 failure per 3000 hours.

- The probability that the lamp lasts longer than its "mean life" is: $P(X > 3) = 1\text{-}(1\text{-}e^{-3/3}) = e^{-1} = 0.368$
- **This is independent of λ. That is, the probability that an exponential random variable is greater than it's mean is 0.368 for any λ**
- The probability that the lamp lasts between 2000 to 3000 hours is:

$$P(2 \le X \le 3) = F(3) - F(2) = 0.145$$

Exponential Distribution

Memoryless property is one of the important properties of exponential distribution

- For all $s \geq 0$ and $t \geq 0$:

$$P(X > s+t \mid X > s) = P(X > t) = P(X > s+t)/P(s) = e^{-\lambda t}$$

- Let X represent the life of a component and is exponentially distributed. Then, the above equation states that the probability that the component lives for at least $s+t$ hours, given that it survived s hours is the same as the probability that it lives for at least t hours. That is, the component doesn't remember that it has been already in use for a time s. ***A used component is as good as new!!!***

- **Light bulb example:** The probability that it lasts for another 1000 hours given it is operating for 2500 hours is the same as the new bulb will have a life greater than 1000 hours

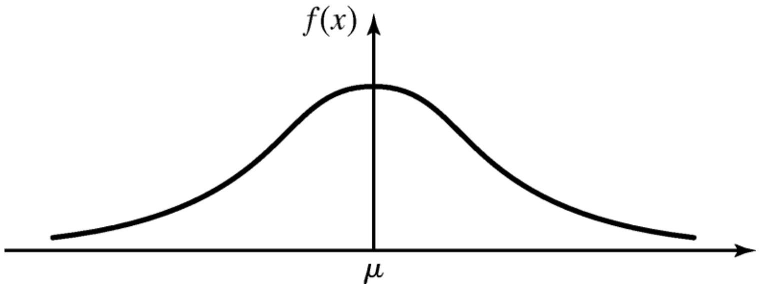
$$P(X > 3.5 \mid X > 2.5) = P(X > 1) = e^{-1/3} = 0.717$$

Normal Distribution

A random variable *X* is normally distributed has the pdf:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty$$

- Mean: $-\infty < \mu < \infty$
- Variance: $\sigma^2 > 0$



- Denoted as $X \sim N(\mu, \sigma^2)$

Special properties:

$$\lim_{x\to-\infty} f(x) = 0, \text{ and } \lim_{x\to\infty} f(x) = 0$$

- $f(\mu-x)=f(\mu+x)$; the pdf is symmetric about μ.
- The maximum value of the pdf occurs at $x = \mu$; the mean and mode are equal.

Normal Distribution

The CDF of Normal distribution is given by

$$F(x) = P(X \le x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] \mathrm{dt},$$

It is not possible to evaluate this in closed form

Numerical methods can be used but it would be necessary to evaluate the integral for each pair (μ, σ^2).

A transformation of variable allows the evaluation to be independent of μ and σ.

Normal Distribution

Evaluating the distribution:

- Independent of μ and σ, using the standard normal distribution:

$$Z \sim N(0,1)$$

- Transformation of variables: let Z = (X - μ) / σ,

$$F(x) = P(X \le x) = P\left(Z \le \frac{x-\mu}{\sigma} \right)$$

$$= \int_{-\infty}^{(x-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$= \int_{-\infty}^{(x-\mu)/\sigma} \phi(z) dz = \Phi(\tfrac{x-\mu}{\sigma})$$

where $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$ is very well tabulated.

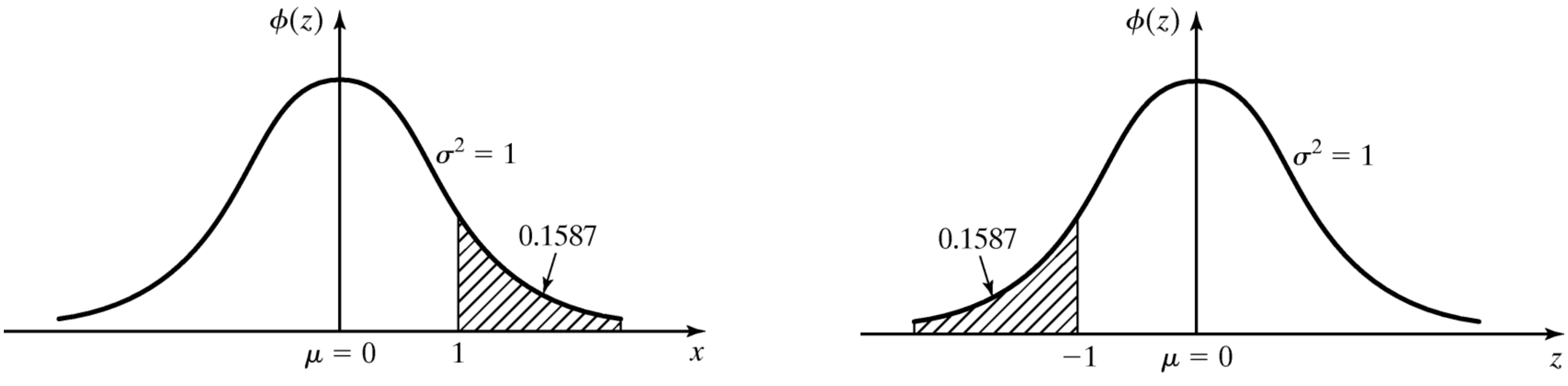
Exponential Distribution

Example: The time required to load an ocean going vessel, X, is distributed as N(12,4)

- The probability that the vessel is loaded in less than 10 hours:

$$F(10) = \Phi\left(\frac{10-12}{2}\right) = \Phi(-1) = 0.1587$$

  – Using the symmetry property, Φ(1) is the complement of Φ (-1), i.e., Φ (-1) = 1- Φ(1)



(a) (b)

Normal Distribution

Example: The time to pass through a queue to begin self-service at a cafeteria is found to be $N(15,9)$. The probability that an arriving customer waits between 14 and 17 minutes is:

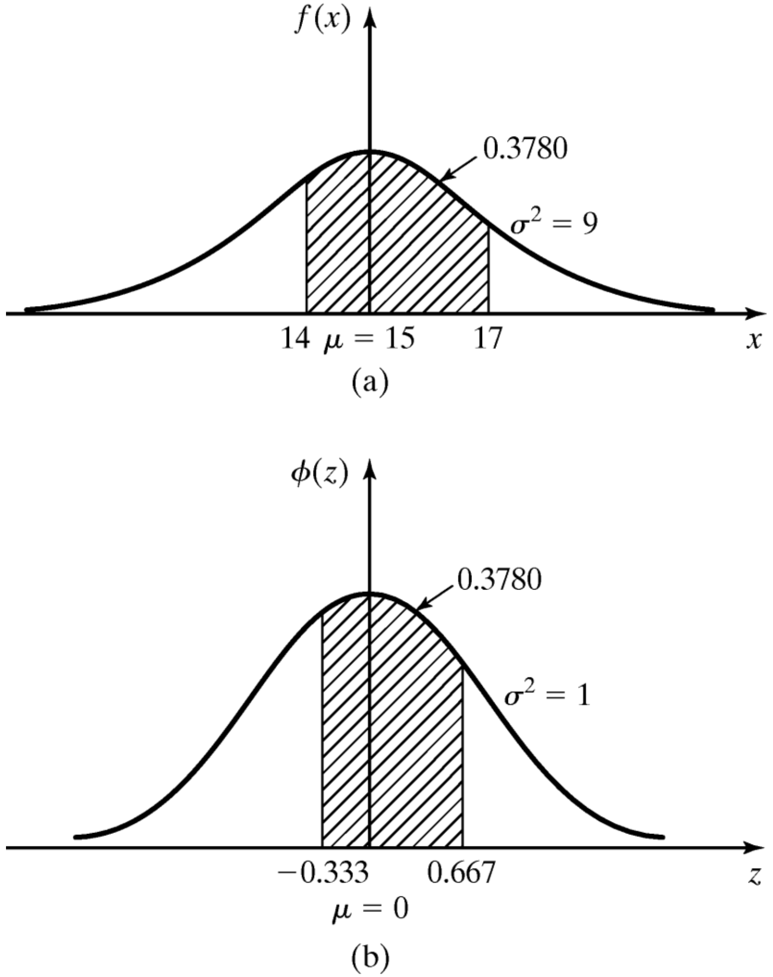
$$P(14 \leq X \leq 17) = F(17) - F(14)$$

$$= \phi((17-15)/3) - \phi((14-15)/3)$$

$$= \phi(0.667) - \phi(-0.333) = 0.3780$$

Normal Distribution

Transformation of pdf for the queue example is shown here



Weibull Distribution

A random variable X has a Weibull distribution if its pdf has the form:

$$f(x) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x-\nu}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x-\nu}{\alpha}\right)^{\beta}\right], & x \geq \nu \\ 0, & \text{otherwise} \end{cases}$$

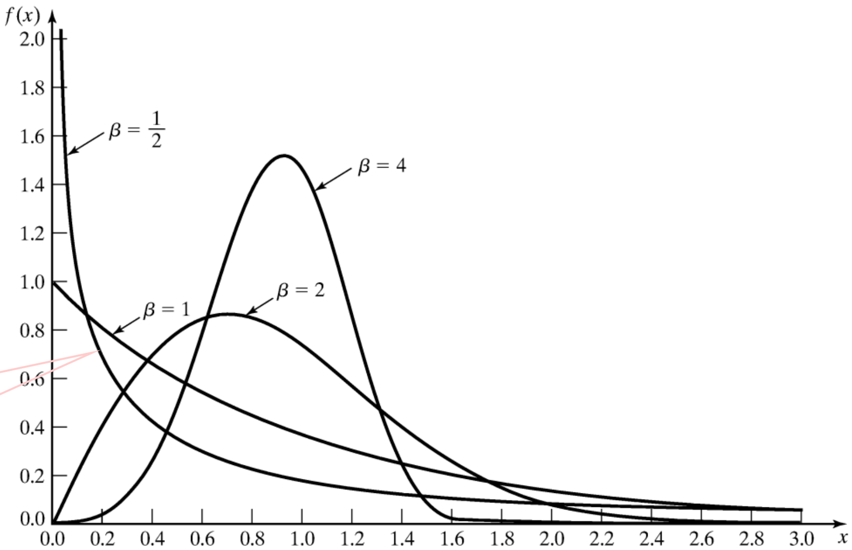
3 parameters:

- Location parameter: υ, $(-\infty < \nu < \infty)$
- Scale parameter: β, $(\beta > 0)$
- Shape parameter. α, (> 0)

Example: $\upsilon = 0$ and $\alpha = 1$:

Exponential Distribution

When $\beta = 1$,
$X \sim exp(\lambda = 1/\alpha)$



Weibull Distribution

The mean and variance of Weibull is given by

$$E(X) = \nu + \alpha\Gamma\left(\frac{1}{\beta} + 1\right)$$

$$V(X) = \alpha^2\left[\Gamma\left(\frac{2}{\beta} + 1\right) - \left[\Gamma\left(\frac{1}{\beta} + 1\right)\right]^2\right]$$

where $\Gamma(.)$ is a Gamma function defined as $\Gamma(\beta) = \int_0^\infty x^{\beta-1}e^{-x}dx$

If β is an integer, $\Gamma(\beta) = (\beta - 1)!$

The CDF is given by

$$F(x) = \begin{cases} 0, & x < \nu \\ 1 - \exp\left[-\left(\frac{x-\nu}{\alpha}\right)^\beta\right], & x \geq \nu \end{cases}$$

Weibull Distribution

Example: The time it takes for an aircraft to land and clear the runway at a major international airport has a Weilbull distribution with ν=1.35 minutes, β=0.5, α=0.04 minute. Find the probability that an incoming aircraft will take more than 1.5 minute to land and clear the runway.

$$P(X > 1.5) = 1 - P(X \leq 1.5)$$

$$P(X \leq 1.5) = F(1.5) = 1 - \exp\left[-\left(\frac{1.5 - 1.34}{0.04}\right)^{0.5}\right] = 0.865$$

Therefore, the probability that an aircraft will require more than 1.5 minutes to land and clear runway is

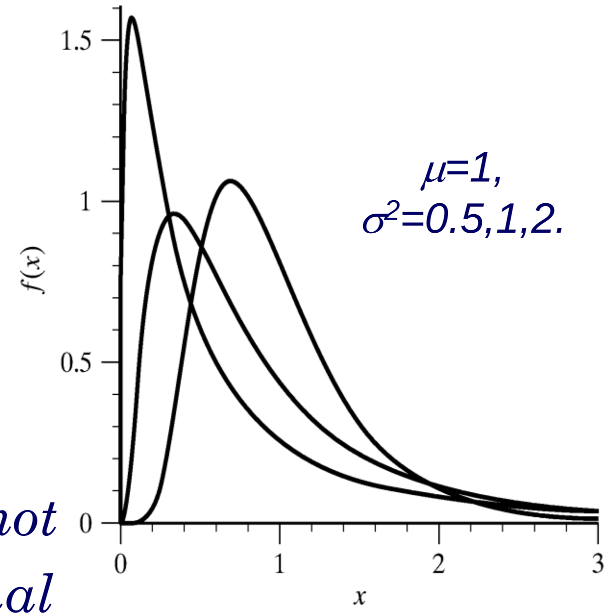
$$1 - 0.865 = 0.135 \text{ minutes}$$

Lognormal Distribution

A random variable X has a lognormal distribution if its pdf has the form:

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\dfrac{(\ln x - \mu)^2}{2\sigma^2}\right], & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

- Mean $E(X) = e^{\mu+\sigma^2/2}$
- Variance $V(X) = e^{2\mu+\sigma^2/2}(e^{\sigma^2} - 1)$
- *Note that* ***parameters μ and σ²*** *are not the mean and variance of the lognormal*



Relationship with normal distribution

- When $Y \sim N(\mu, \sigma^2)$, then $X = e^Y \sim \text{lognormal}(\mu, \sigma^2)$

Lognormal Distribution

If the mean and variance of lognormal are known to be μ_{L} and σ_L^2 respectively, then the parameters μ and σ^2 is given by:

$$\mu = \ln\left(\frac{\mu_{\mathrm{L}}^2}{\sqrt{\mu_{\mathrm{L}}^2 + \sigma_L^2}}\right)$$

$$\sigma^2 = \ln\left(\frac{\mu_L^2 + \sigma_L^2}{\mu_L^2}\right)$$

Example: The rate of return on a volatile investment is modeled as lognormal with mean 20% (=μ_L) and standard deviation 5% (=σ_L^2). What are the *parameters* for lognormal?

- $\mu = 2.9654$; $\sigma^2 = 0.06$

Poisson Process

Definition: $N(t)$, $t \geq 0$ is a counting function that represents the number of events occurred in $[0,t]$.

- *e.g.*, arrival of jobs, e-mails to a server, boats to a dock, calls to a call center

A counting process $\{N(t), t \geq 0\}$ is a **Poisson process** with mean rate λ if:

- **Arrivals occur one at a time**
- **$\{N(t), t \geq 0\}$ has stationary increments**: The distribution of number of arrivals between t and $t+s$ depends only on the length of interval s and not on starting point t. Arrivals are completely random without rush or slack periods.
- **$\{N(t), t \geq 0\}$ has independent increments**: The number of arrivals during non-overlapping time intervals are independent random variables.

Poisson Process

Properties

$$P[N(t)=n]=\frac{e^{-\lambda t}(\lambda t)^n}{n!}, \qquad \text{for } t \geq 0 \text{ and } n=0,1,2,\ldots$$

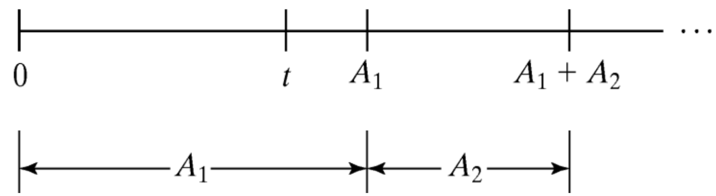
- Equal mean and variance: $E[N(t)] = V[N(t)] = \lambda t$
- Stationary increment: For any s and t, such that $s < t$, the number of arrivals in time s to t is also Poisson-distributed with mean $\lambda(t\text{-}s)$

$$P[N(s)-N(t)=n]=\frac{e^{-\lambda(t-s)}\left[\lambda(t-s)\right]^n}{n!}, \qquad \text{for } \; n=0,1,2,\ldots$$

$$\text{and } E\left[N(t)-N(s)\right]=\lambda(t-s)=V\left[N(t)-N(s)\right]$$

Interarrival Times

Consider the inter-arrival times of a Possion process (A_1, A_2, …), where A_i is the elapsed time between arrival *i* and arrival *i+1*

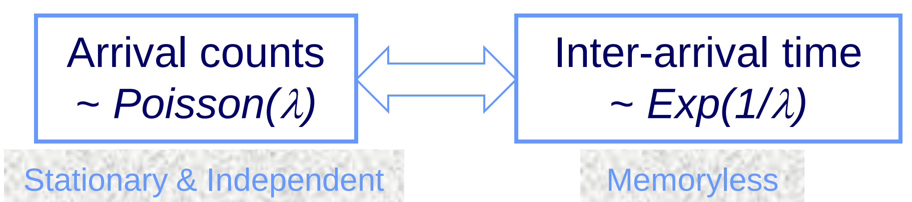


- The *1st arrival* occurs after time t *iff* there are no arrivals in the interval [0,t], hence:

$$P\{A_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}$$

$$P\{A_1 \le t\} = 1 - e^{-\lambda t} \qquad \text{[cdf of exp}(\lambda)]$$

- Inter-arrival times, A_1, A_2, …, are exponentially distributed and independent with mean $1/\lambda$



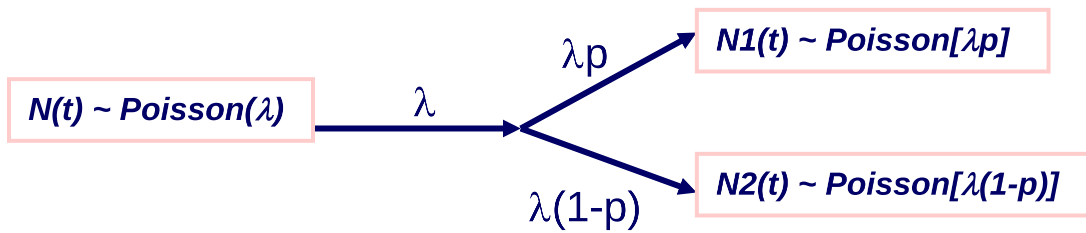
Poisson Process

The jobs at a machine shop arrive according to a Poisson process with a mean of $\lambda = 2$ jobs per hour. Therefore, the inter-arrival times are distributed exponentially with the expected time between arrivals being $E(A)=1/\lambda=0.5$ hour

Other Properties

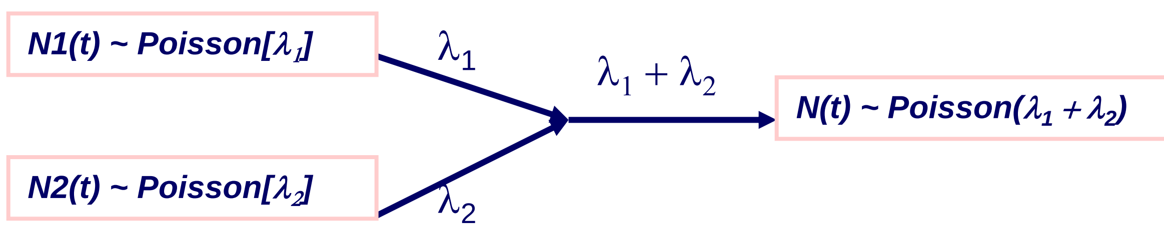
Splitting:

- Suppose each event of a Poisson process can be classified as Type I, with probability p and Type II, with probability $1\text{-}p$.
- $N(t) = N_1(t) + N_2(t)$, where $N_1(t)$ and $N_2(t)$ are both Poisson processes with rates λp and $\lambda(1\text{-}p)$



Pooling:

- Suppose two Poisson processes are pooled together
- $N_1(t) + N_2(t) = N(t)$, where $N(t)$ is a Poisson processes with rates $\lambda_1 + \lambda_2$



Poisson Process

Example: Suppose jobs arrive at a shop with a Poisson process of rate λ. Suppose further that each arrival is marked "high priority" with probability 1/3 (Type I event) and "low priority" with probability 2/3 (Type II event). Then $N_1(t)$ and $N_2(t)$ will be Poisson with rates $\lambda/3$ and $2\lambda/3$.

Non-stationary Poisson Process (NSPP)

Poisson Process without the stationary increments, characterized by $\lambda(t)$, the arrival rate at time t. (Drop assumption 2 of Poisson process, stationary increments)

The *expected number* of arrivals by time t, $\Lambda(t)$:

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

Relating stationary Poisson process $N(t)$ *with rate* $\lambda=1$ and NSPP $N(t)$ with rate $\lambda(t)$:

- Let arrival times of a stationary process with rate $\lambda = 1$ be t_1, t_2, …, and arrival times of a NSPP with rate $\lambda(t)$ be T_1, T_2, …, we know:

$$t_i = \Lambda(T_i) \qquad \textit{[Expected \# of arrivals]}$$

$$T_i = \Lambda^{-1}(t_i)$$

- An NSPP can be transformed into a stationary Poisson process with arrival rate 1 and vice versa.

Non-stationary Poisson Process (NSPP)

Example: Suppose arrivals to a Post Office have rates 2 per minute from 8 am until 12 pm, and then 0.5 per minute until 4 pm.

Let $t = 0$ correspond to 8 am, NSPP $N(t)$ has rate function:

$$\lambda(t) = \begin{cases} 2, & 0 \le t < 4 \\ 0.5, & 4 \le t < 8 \end{cases}$$

Expected number of arrivals by time t:

$$\Lambda(t) = \begin{cases} 2t, & 0 \le t < 4 \\ \int_0^4 2ds + \int_4^t 0.5ds = \frac{t}{2} + 6, & 4 \le t < 8 \end{cases}$$

Hence, the probability distribution of the number of arrivals between 11 am and 2 pm, corresponds to times 3 and 6 respectively.

$$P[N_{ns}(6) - N_{ns}(3) = k] \quad = P[N(\Lambda(6)) - N(\Lambda(3)) = k]$$

$$= P[N(9) - N(6) = k]$$

$$= e^{-(9-6)}(9-6)^k / k! \qquad = e^{-3}(3)^k / k!$$

Empirical Distributions

A distribution whose parameters are the observed values in a sample of data.

- May be used when it is impossible or unnecessary to establish that a random variable has any particular parametric distribution.
- Advantage: no assumption beyond the observed values in the sample.
- Disadvantage: sample might not cover the entire range of possible values.

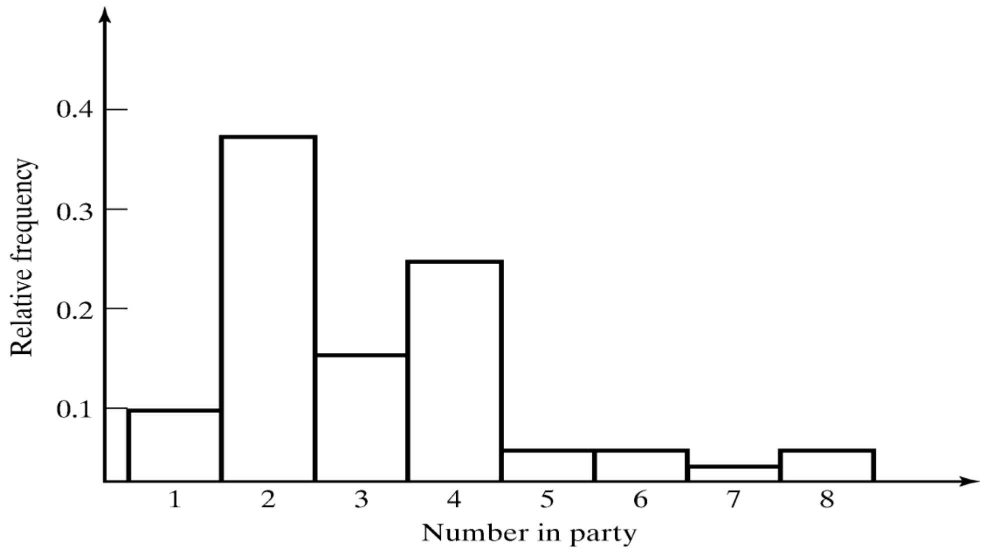
Empirical Example – Discrete

Customers at a local restaurant arrive at lunch time in groups of eight from one to eight persons. The number of persons per party in the last 300 groups has been observed. The results are summarized in Table 5.3. A histogram of the data is plotted and a CDF is constructed. The CDF is called the empirical distribution

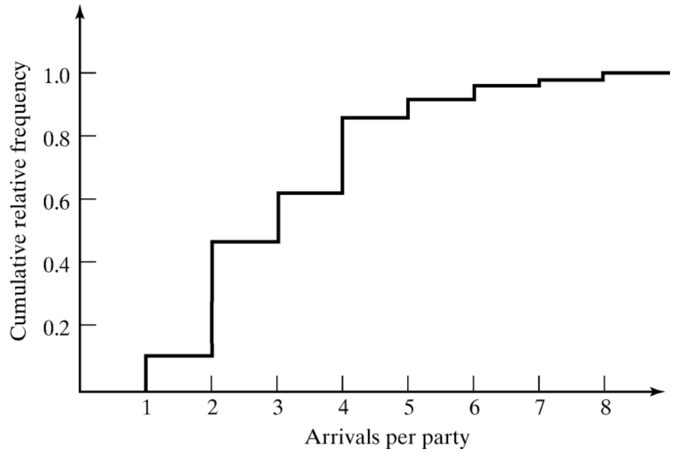
Table 5.3 Arrivals per Party Distribution

Arrivals per Party	*Frequency*	*Relative Frequency*	*Cumulative Relative Frequency*
1	30	0.10	0.10
2	110	0.37	0.47
3	45	0.15	0.62
4	71	0.24	0.86
5	12	0.04	0.90
6	13	0.04	0.94
7	7	0.02	0.96
8	12	0.04	1.00

Empirical Example – Discrete



Histogram



CDF

Empirical Example - Continuous

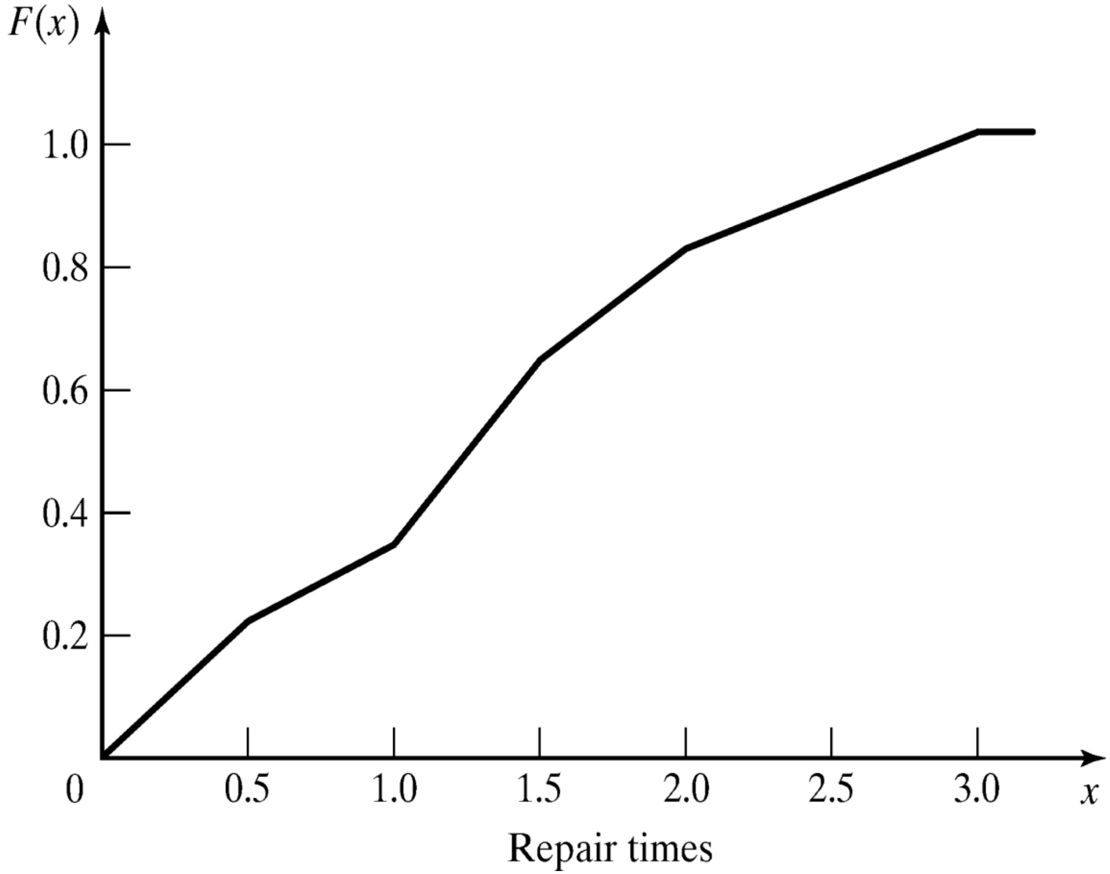
The time required to repair a conveyor system that has suffered a failure has been collected for the last 100 instances; the results are shown in Table 5.4. There were 21 instances in which the repair took between 0 and 0.5 hour, and so on. The empirical cdf is shown in Figure 5.29. A piecewise linear curve is formed by the connection of the points of the form *[x,F(x)]*. The points are connected by a straight line. The first connected pair is (0, 0) and (0.5, 0.21); then the points (0.5, 0.21) and (1.0, 0.33) are connected; and so on.

Empirical Example – Continuous

Table 5.4 Repair Times for Conveyor

Interval (Hours)	*Frequency*	*Relative Frequency*	*Cumulative Frequency*
$0 < x \leq 0.5$	21	0.21	0.21
$0.5 < x \leq 1.0$	12	0.12	0.33
$1.0 < x \leq 1.5$	29	0.29	0.62
$1.5 < x \leq 2.0$	19	0.19	0.81
$2.0 < x \leq 2.5$	8	0.08	0.89
$2.5 < x \leq 3.0$	11	0.11	1.00

Empirical Example – Continuous



- A college professor of electrical engg is leaving home for the summer, but would like to have a light burning at all times to discourage burglars. The professor rigs up a device that will hold two light bulbs. The device will switch the current to the second bulb if the first bulb fails. The box in which the light bulbs are packaged says, ”Average life 1000 hrs, exponentially distributed.” the professor will be gone 90 days (2160 hrs). What is the probability that a light will be burning when the summer is over and the professor returns?