Review of Basic Probability and Statistics

Chapter 4

Jiang Li, Ph.D.
Department of Computer Science

4.1 Introduction

- Use of probability and statistics
  - Integral part of a simulation study
  - Every modeling team should include at least one thoroughly trained person
  - Understanding how to model a probabilistic system and validate the model
  - Needed to choose input probability distributions and generate random samples
  - Required to analyze output data

Jiang Li, Ph.D.
Department of Computer Science

4.2 Random Variables and Their Properties

- Experiment
  - Process whose outcome is not known with certainty
- Sample space, S
  - Set of all possible outcomes, called sample points
- Example: Experiment consisting of flipping a coin

$$S = \{H, T\}$$

HOWARD UNIVERSITY
Jiang Li, Ph.D.
Department of Computer Science

Random Variables and Their Properties

- $P(X \leq x)$
  - Probability associated with the event $\{X \leq x\}$
- Discrete random variables
  - Can take on a countable number of values
- Continuous random variables
  - Can take on an uncountably infinite number of different values

Jiang Li, Ph.D.
Department of Computer Science

Discrete Random Variables

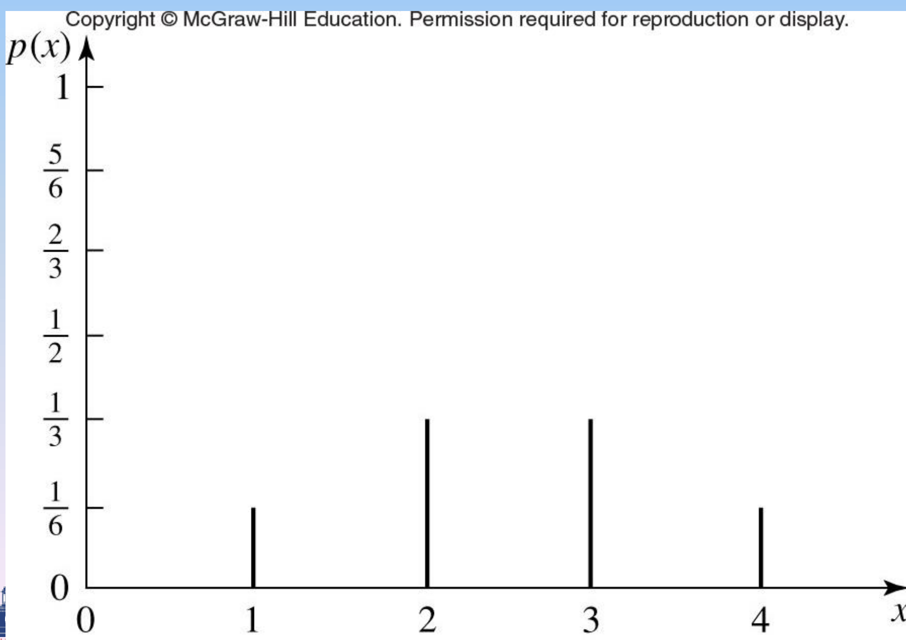
Probability mass function

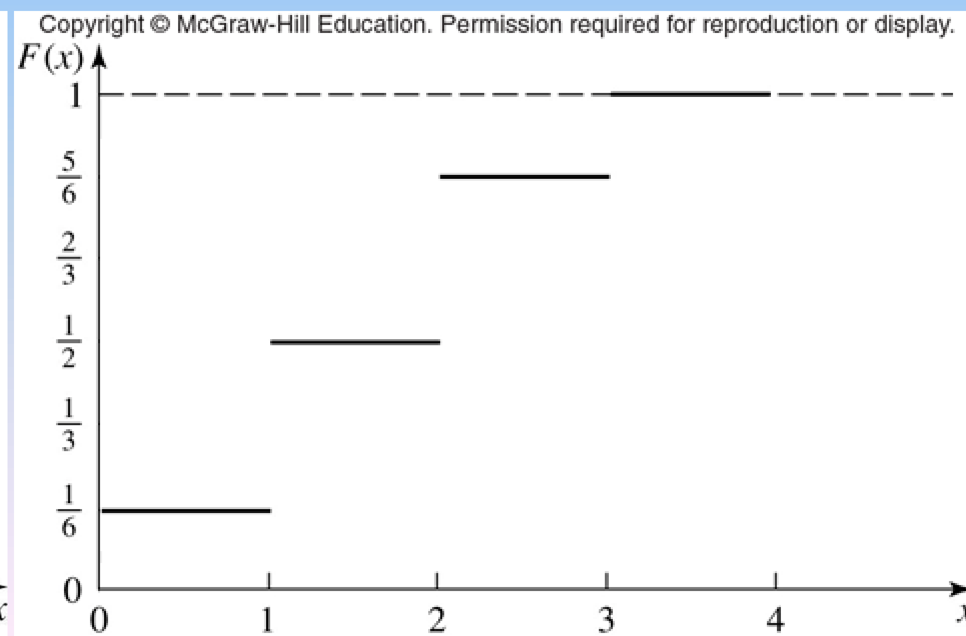
$$\mathrm{p}(x_i) = P(X = x_i) \quad \text{for } i=1,2,\ldots.$$

$$P(X \in I) = \sum_{a \le x \le b} \mathrm{p}(x_i) \text{ for I = [a,b]}$$

Cumulative distribution function

$$\mathrm{F}(x) = P(X \le x) \quad \text{for } -\infty < x < \infty$$





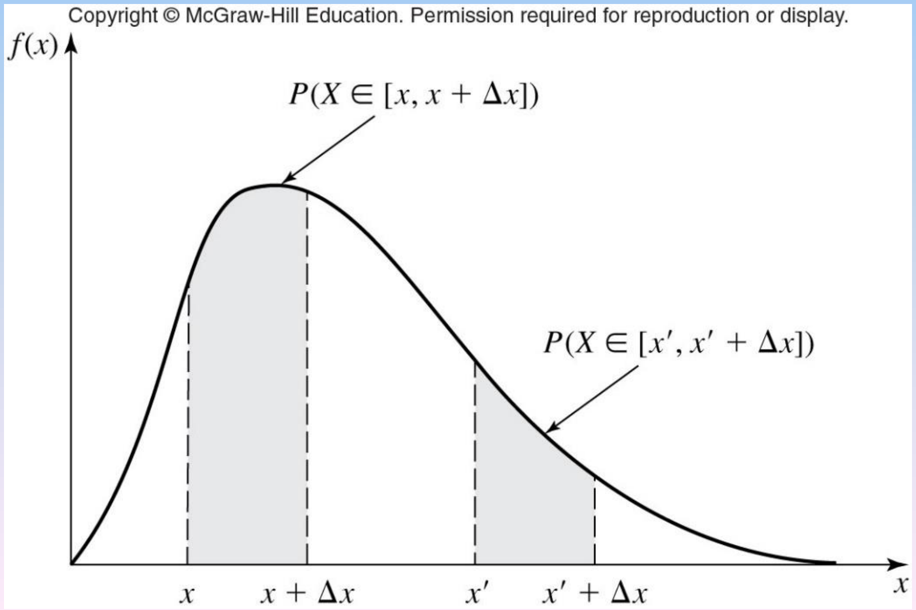
Continuous Random Variables

Probability Mass function

$$P(X = x) = 0$$

$P(X \in B) = \int_B f(x)dx$ f(x):Probability density function

$P(X \in [x, x + \Delta x]) = \int_x^{x+\Delta x} f(y)dy$ for I = [a,b]



Jiang Li, Ph.D.
Department of Computer Science

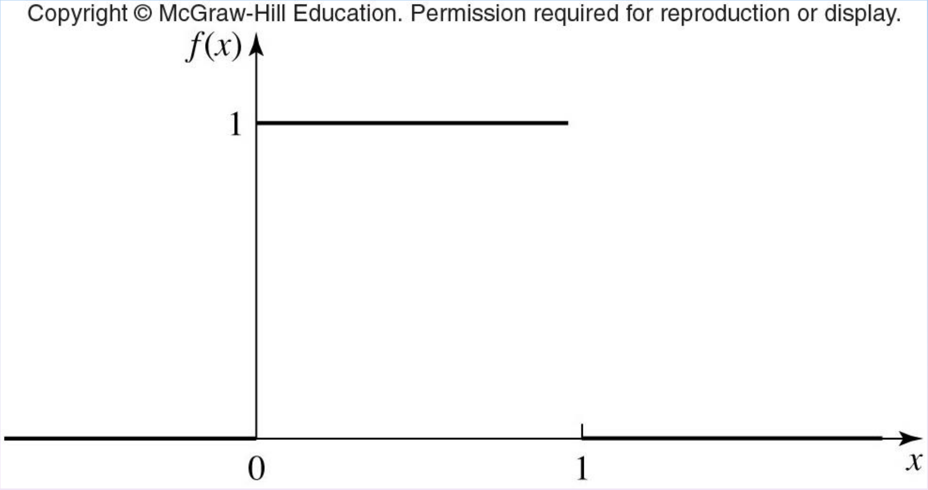
Continuous Random Variables

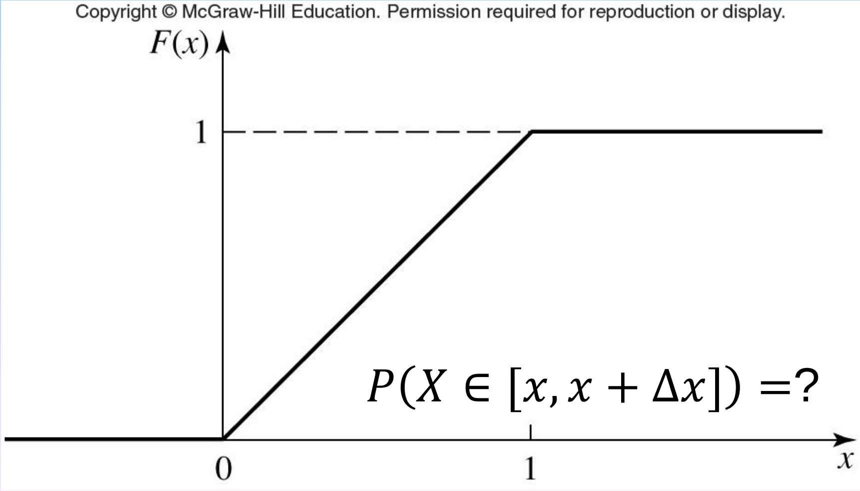
Cumulative Distribution function

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy \quad \text{for } -\infty < x < \infty$$

$$f(x) = F'(x)$$

$$P(X \in I) = \int_{a}^{b} f(y)dy = \text{F(b)} - \text{F(a) for I = [a,b]}$$



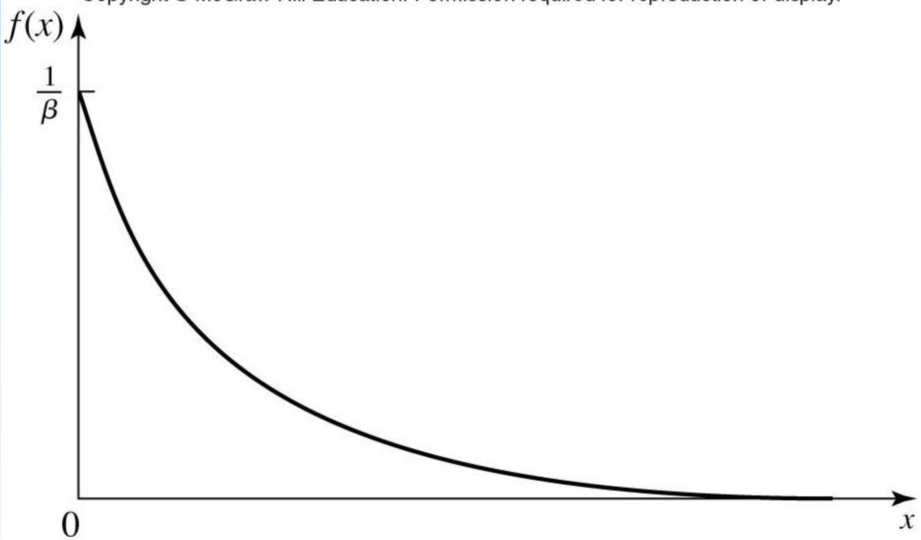


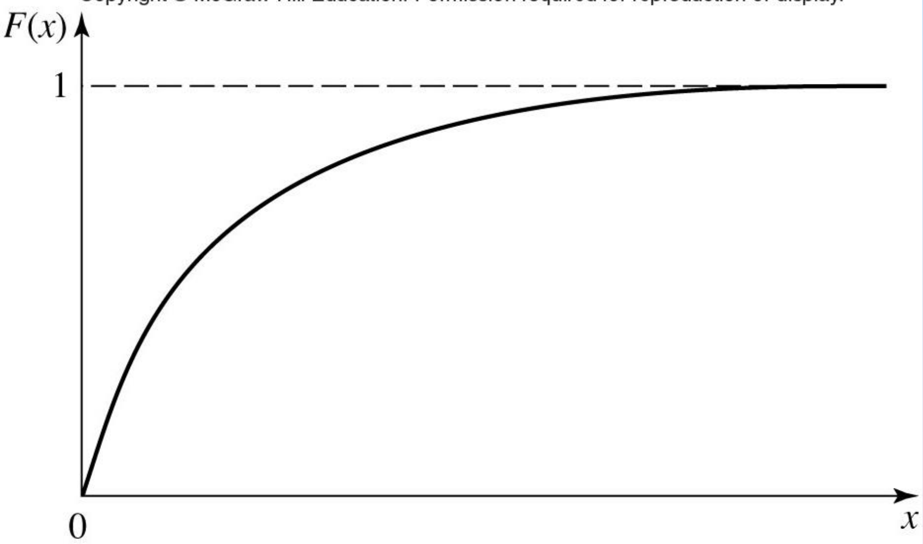
Continuous Random Variables

Exponential distribution

$$f(x) = \frac{1}{\beta} e^{-\frac{1}{\beta}x}$$

$$\mathrm{F}(x) = 1 - e^{-\frac{1}{\beta}x}$$





HOWARD UNIVERSITY

Jiang Li, Ph.D.
Department of Computer Science

CDF Properties

- $0 \leq F(x) \leq 1$
- F(x) is nondecreasing
- $\lim_{x \to \infty} F(x) = 1$
- $\lim_{x \to -\infty} F(x) = 0$

Jiang Li, Ph.D.
Department of Computer Science

Joint Probability Mass Function

- $p(x,y) = P(X = x, Y = y)$ for all x,y
- X and Y are independent if

$$p(x,y) = P_X(x)P_Y(y) \text{ for all x,y}$$

$$P_X(x) = \sum_{\text{all } y} p(x,y)$$

$$P_Y(y) = \sum_{\text{all } x} p(x,y)$$

Jiang Li, Ph.D.
Department of Computer Science

Joint Probability Mass Function

- Are x and y independent?
- $p(x, y) = \begin{cases} \frac{xy}{27} & \text{for x} = 1,2 \text{ and y } = 2,3,4 \\ 0 & \text{otherwise} \end{cases}$
- Draw 2 cards from a deck of 52 w/o replacement. Let R.V. X and Y be the number of aces and kings that occur. $X, Y \in \{0,1,2\}$

Jiang Li, Ph.D.
Department of Computer Science

Jointly Continuous Random Var.

- X and Y are jointly continuous if there is a joint p.d.f. of X and Y

$$P(X \in A, Y \in B) = \int_B \int_A f(x,y)dx\,dy$$

X and Y are independent if

$$f(x,y) = f_X(x)f_Y(y) \quad \text{for all x,y}$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

Joint Continuous Random Var.

- Are x and y independent?

$$f(x, y) = \begin{cases} 24xy & \text{for x,y} \geq 0, x + y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Mean

- *Discrete R.V.*

$$\mu = E(X) = \sum_{j=1}^{\infty} x_j p_X(x_j)$$

- *Continuous R.V.*

$$\mu = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$E(cX) = cE(X)$$

$$E\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i E(X_i)$$

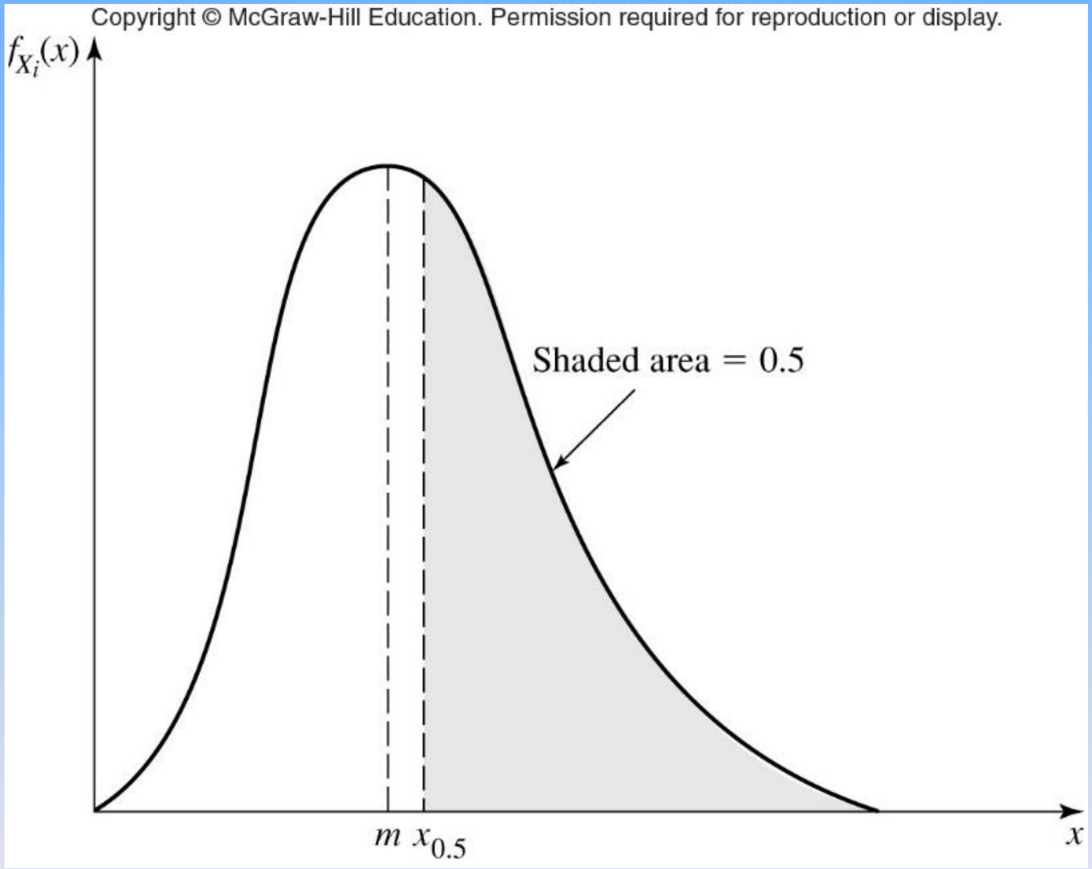
Median

- Median $x_{0.5}$
  - Smallest x s.t. $F_X(x) \geq 0.5$
  - If X is continuous, $F_X(x_{0.5}) = 0.5$
- Example
  - A discrete R.V. X takes on 1,2,3,4,100 with the same probability.
  - Mean?
  - Median?

Jiang Li, Ph.D.
Department of Computer Science

Mode

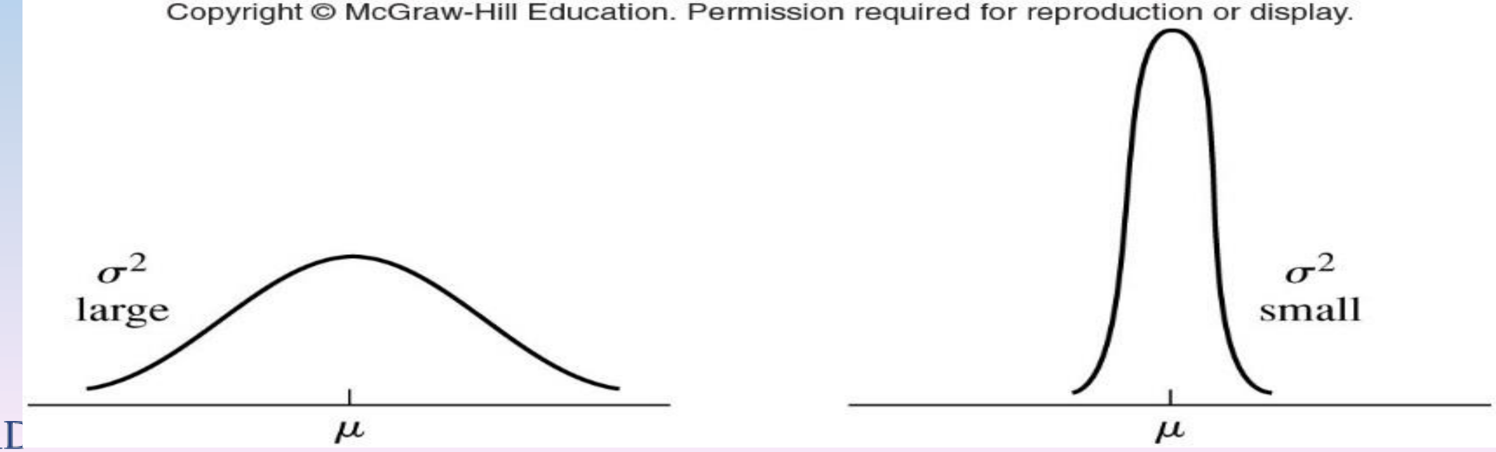
- Mode
  - x that maximize $f_X(x)$ or $p(x)$
  - Mode for prev. example?
  - Mode for uniform distr., expo. distr.?



HOWARD UNIVERSITY
Jiang Li, Ph.D.
Department of Computer Science

Variance

- Variance
  - Measure of the dispersion of a random variable about its mean
  - $\sigma^2 = E[(X-\mu)^2] = \mathrm{E}(X^2) - \mu^2$
  - Variance of a single dice throw? Uniform R.V. on [0,1]?



Jiang Li, Ph.D.
Department of Computer Science

Variance Properties

- $Var(X) \geq 0$
- $Var(cX) = c^2 Var(X)$
- $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$
  - If X_i's are independent or uncorrelated

Jiang Li, Ph.D.
Department of Computer Science

Std. Deviation

- Standard deviation of random variable X_i

$$\sigma_i = \sqrt{\sigma_i^2}$$

- Used more in normal distributions
  - Tell the probability of $X \in [\mu - n\sigma, \mu + n\sigma]$
  - E.g. P = 0.95 when n = 1.96

Dependence Between Two R.V.'s

- Covariance $Cov(X_i, X_j)$ or C_{ij} measures dependence between two random variables X_i and X_j

$$C_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] = E(X_i X_j) - \mu_i \mu_j$$

- $C_{ij} = C_{ji}$
- Example: What is Cov(X,Y)?

$$f(x,y) = \begin{cases} 24xy & \text{for x,y} \geq 0, x + y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Covariance Example

- $E(XY) = \int_0^1 \int_0^{1-x} xyf(x,y)dydx$
$$= \int_0^1 x^2 \left(\int_0^{1-x} 24y^2 dy \right) dx$$
$$= \int_0^1 8x^2(1-x^3)dx$$
$$= \frac{2}{15}$$
- $E(X) = \int_0^1 xf(x)dx = \int_0^1 12x^2(1-x)^2 dx = \frac{2}{5}$
- $Cov(X,Y) = ?$

Jiang Li, Ph.D.
Department of Computer Science

Covariance

- If X and Y are independent, Cov(X,Y) = 0
- If Cov(X,Y) = 0, X and Y are uncorrelated but may be dependant (unless they are jointly normally distributed)
- If Cov(X,Y) > 0, X and Y are positively correlated
  - $X > \mu_X$ and $Y > \mu_Y$ tend to occur together
- If Cov(X,Y) < 0, X and Y are negatively correlated
  - $X > \mu_X$ and $Y < \mu_Y$ tend to occur together

Jiang Li, Ph.D.
Department of Computer Science

Correlation

- Measure of the dependence between X_i and X_j (among X_1, X_2, … X_n, e.g. in simulation output)
- C_{ij} is not dimensionless
- Correlation $\mathrm{Cor}(X_i, X_j) = \rho_{ij} = \frac{C_{ij}}{\sqrt{\sigma_i{}^2 \sigma_j{}^2}}$
  - Same sign as C_{ij}
  - In [-1, 1]
    - Close to -1 => Highly negatively correlated
    - Close to 1 => Highly positively correlated

Jiang Li, Ph.D.
Department of Computer Science

Correlation Example

- $f(x,y) = \begin{cases} 24xy & \text{for x,y} \geq 0, x+y \leq 1 \\ 0 & \text{otherwise} \end{cases}$
- $E(XY) = \frac{2}{15}$
- $E(X) = \int_0^1 xf(x)dx = \int_0^1 12x^2(1-x)^2 dx = \frac{2}{5}$
- $E(X^2) = \int_0^1 x^2 f(x)dx = \int_0^1 12x^3(1-x)^2 dx = \frac{1}{5}$
- $Cor(X,Y) = ?$

Jiang Li, Ph.D.
Department of Computer Science

4.3 Simulation Output Data and Stochastic Processes

- Simulation output data are random
  - Because random variables used as input
  - Take care in drawing conclusions
- Stochastic process
  - Collection of similar random variables ordered over time defined on a common sample space
  - State space

Jiang Li, Ph.D.
Department of Computer Science

Stochastic Processes

- Discrete-time stochastic process
  - Index variable takes a discrete set of values
  - E.g. X_1, X_2, …
  - E.g. Delay observed by customers

    $D_1 = 0$

    $D_{i+1} = \max\{D_i + S_i - A_{i+1}, 0\}$ for $i > 0$

    Why?
- Continuous-time stochastic process
  - Index variable takes values in a continuous range
  - E.g. [X(t), t >= 0]
  - E.g. The number of customers in the queue at time t.

Simulation Output Data and Stochastic Processes

- A discrete-time stochastic process $X_1, X_2, \ldots$ is said to be covariance-stationary if:

$$\mu_i = \mu \qquad \text{for } i = 1, 2, \ldots \text{ and } -\infty < \mu < \infty$$

$$\sigma_i^2 = \sigma^2 \qquad \text{for } i = 1, 2, \ldots \text{ and } \sigma^2 < \infty$$

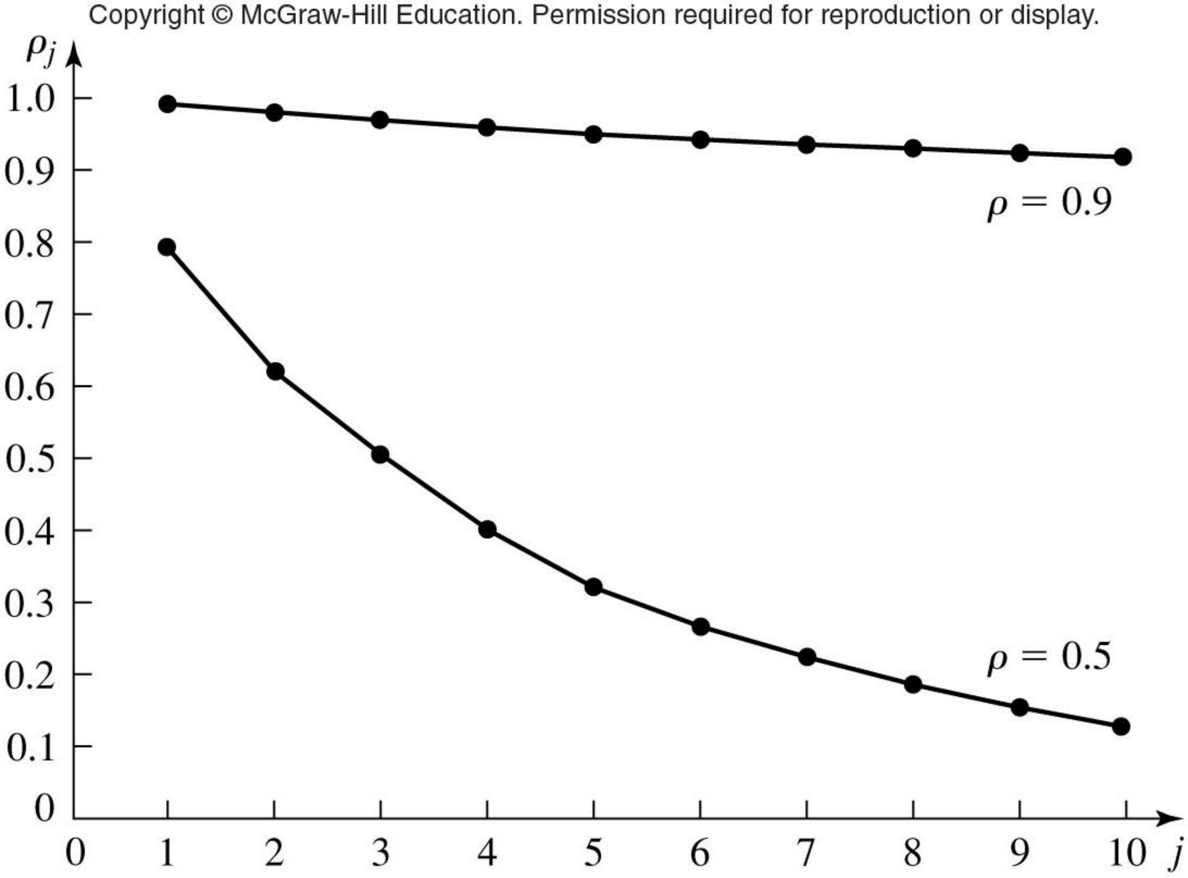
and $C_{i,i+j} = \text{Cov}(X_i, X_{i+j})$ is independent of i for $j = 1, 2, \ldots$.

  - Covariance only depends on the separation (or lag)
  - Denote the covariance and correlation between X_i and X_{i+j} by C_j and ρ_j
    - $\rho_j = \frac{C_{i,i+j}}{\sqrt{\sigma_i^2 \sigma_{i+j}^2}} = \frac{C_j}{\sigma^2} = \frac{C_j}{C_0}$ for j = 0, 1, 2, …
- Definition for continuous-time S.P. is similar

Jiang Li, Ph.D.
Department of Computer Science

Correlation Example 1

- Output process $D_1, D_2, D_3, \ldots$



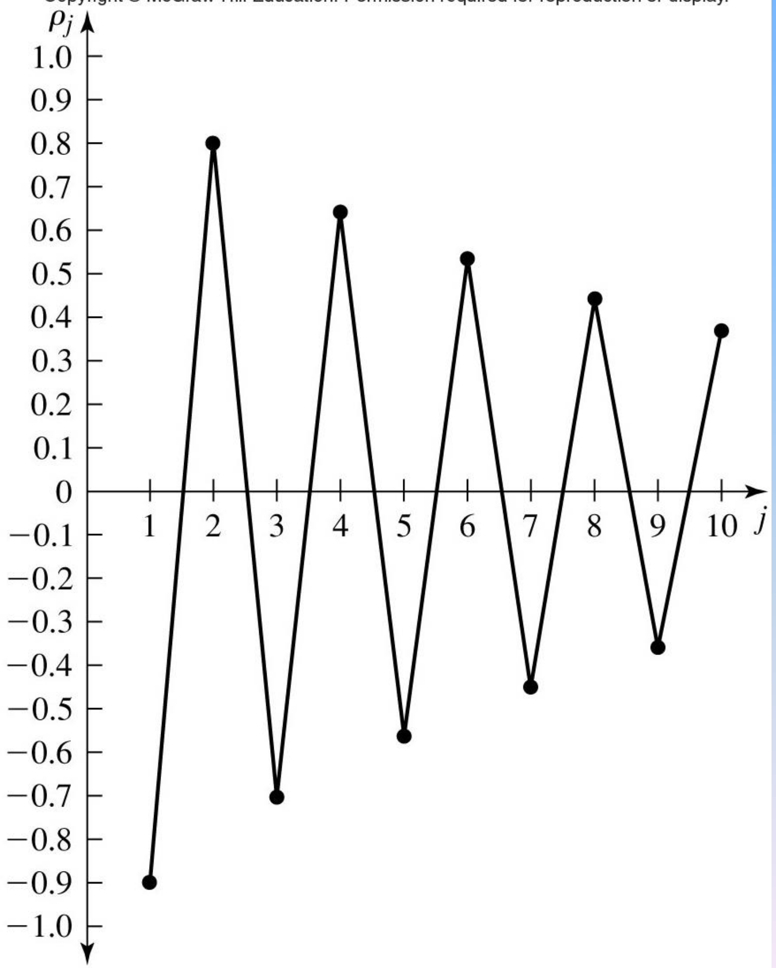
Jiang Li, Ph.D.
Department of Computer Science

Correlation Example 2

- An (s, S) inventory system
- In month i
  - I_i: amount of inventory before ordering
    - $I_i < s$, order $S - I_i$ => ordering cost (major cost)
  - J_i: amount of inventory before ordering
    - $J_i > Q_i$ => holding cost
    - $J_i > Q_i$ => shortage cost
  - Q_i: demand, exponential distribution
  - C_i: total cost

Correlation Example 2 (cont'd)

- Output process $C_1, C_2, C_3, \ldots$
  - $\rho_1 < 0$
  - $\rho_2 > 0$



HOWARD UNIVERSITY
Jiang Li, Ph.D.
Department of Computer Science

Warmup Period in Simulations

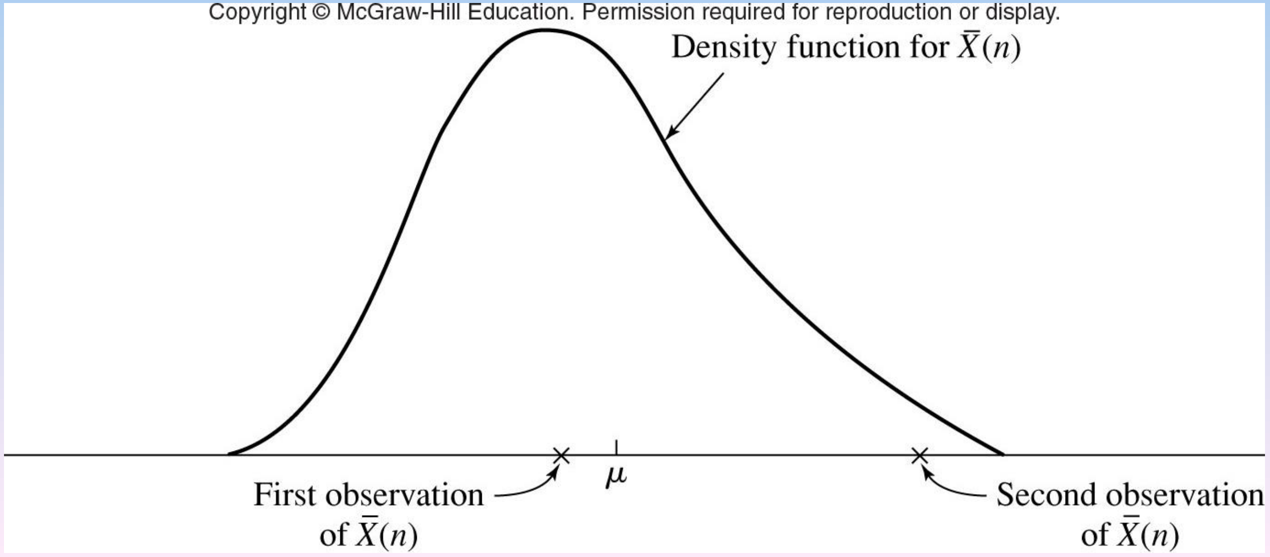
- $X_1, X_2, X_3, \ldots$ is a stochastic process beginning at time 0 in a simulation
  - Likely not covariance-stationary
- $X_{k+1}, X_{k+2}, X_{k+3}, \ldots$ may be approx. covariance-stationary if k is large enough
  - k: the length of warmup period

4.4 Estimation of Means, Variances, and Correlations

- Sample mean ($\hat{\mu}$) of a set of random variables X_1, X_2,…X_n with finite population mean μ and finite population variance σ^2

$$\bar{X}(n) = \frac{\sum_{i=1}^{n} X_i}{n}$$

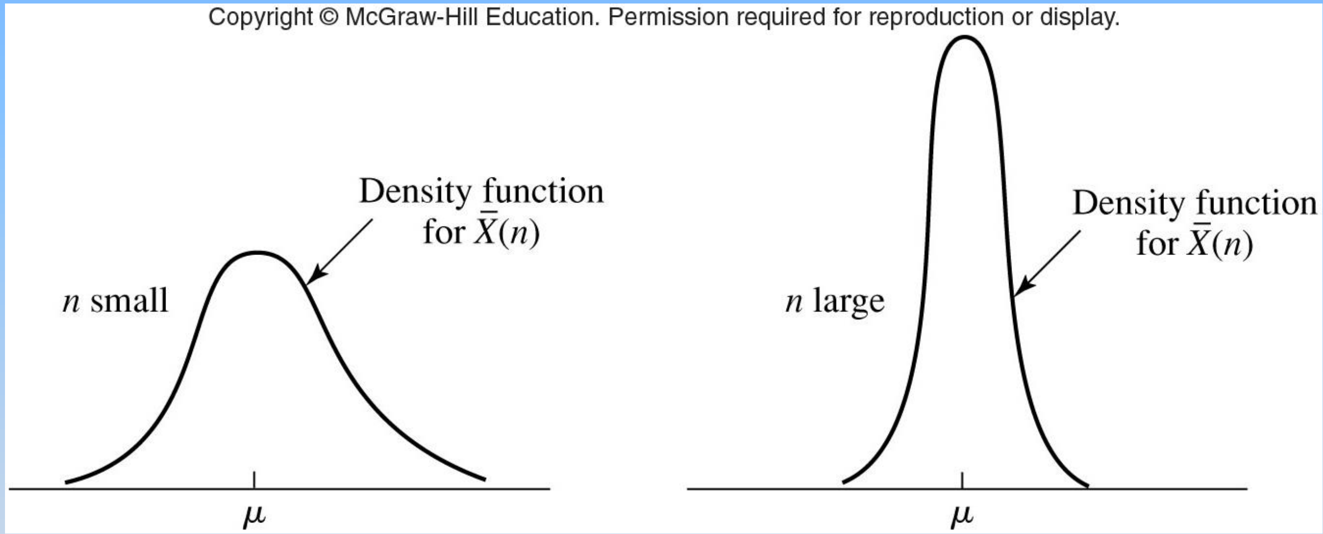
- Unbiased
  - $E(\bar{X}(n)) = \mu$



How Close is $\bar{X}(n)$ to μ? (1)

- If X_i's are independent

$$
\begin{aligned}
&Var(\bar{X}(n)) \\
&= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2} Var(\textstyle\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \\
&= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}
\end{aligned}
$$



- The larger n, the closer $\bar{X}(n)$ is to μ
- But we don't know the exact $Var(\bar{X}(n))$
  - Estimate it, need to estimate σ^2

Estimation of Variances and $Var(\bar{X}(n))$

- If X_i's are IID, sample variance given by $\hat{\sigma}^2$

$$S^2(n) = \frac{\sum_{i=1}^{n} [X_i - \bar{X}(n)]^2}{n-1}$$

- Unbiased

$$E(S^2(n)) = \sigma^2$$

- Estimate $Var(\bar{X}(n))$

$$\widehat{Var}(\bar{X}(n)) = \frac{S^2(n)}{n} = \frac{\sum_{i=1}^{n}[X_i - \bar{X}(n)]^2}{n(n-1)}$$

How Close is $\bar{X}(n)$ to μ? (2)

- Simulation output data are almost always correlated
  - $\rho_j \neq 0$ for j = 1,2,…,n-1
- Assume the stochastic process of $X_1, X_2, \ldots X_n$ is covariance-stationary.
  - $\bar{X}(n)$ is still an unbiased estimator of μ

$$Var(\bar{X}(n)) = \sigma^2 \frac{\left[1 + 2\sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)\rho_j\right]}{n}$$

→ Correlation terms

  - Can we estimate $Var(\bar{X}(n))$ from $\frac{S^2(n)}{n}$?

Jiang Li, Ph.D.
Department of Computer Science

How Close is $\bar{X}(n)$ to μ? (2)

- Can we estimate $Var(\bar{X}(n))$ from $\frac{S^2(n)}{n}$?

$$E(S^2(n)) = \sigma^2 \left(1 - 2\frac{\sum_{j=1}^{n-1}(1-\frac{j}{n})\rho_j}{n-1}\right)$$

  - If $\rho_j > 0$, $E(S^2(n)) < \sigma^2$
  - Don't use $S^2(n)$ to estimate σ^2 (if X_i's are correlated)

$$E\left(\frac{S^2(n)}{n}\right) = \frac{\frac{n}{1+2\sum_{j=1}^{n-1}(1-\frac{j}{n})\rho_j} - 1}{n-1} Var(\bar{X}(n))$$

  - If $\rho_j > 0$, $E\left(\frac{S^2(n)}{n}\right) < Var(\bar{X}(n))$

Example

- $D_1, D_2, \ldots, D_{10}$ from the process of delays for a covariance-stationary M/M/1 queue with $\rho = 0.9$
- Substituting the true ρ_j (j = 1, 2, …, 9) into previous formulas
  - $E(S^2(10)) = 0.0328\sigma^2$
  - $E\left(\frac{S^2(10)}{10}\right) = 0.0034Var(\bar{D}(10))$
  - Underestimate σ^2 and $Var(\bar{D}(10))$

Jiang Li, Ph.D.
Department of Computer Science

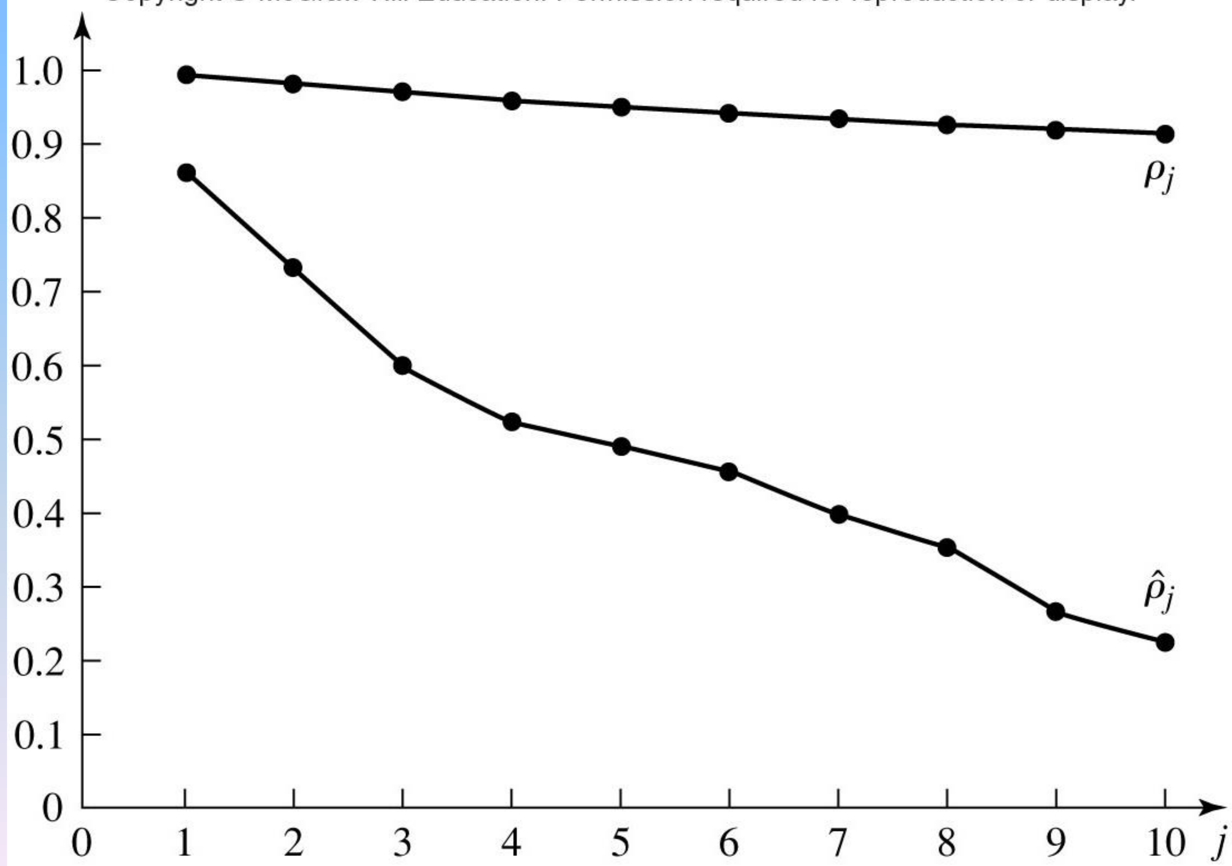
Can we know $Var(\bar{X}(n))$ better?

- Estimate correlations
  - $\hat{C}_{ij} = \frac{\sum_{i=1}^{n-j}[X_i - \bar{X}(n)][X_{i+j} - \bar{X}(n)]}{n-j}$ (denominator can be n)
  - $\widehat{\rho}_j = \frac{\hat{C}_{ij}}{S^2(n)}$
- Can help understand $Var(\bar{X}(n))$ and correlations. However,
  - $\widehat{\rho}_j$ is biased
  - $\widehat{\rho}_j$ has large variance (unless n is very large and j << n)
  - $Cov(\widehat{\rho}_j, \widehat{\rho}_k) \neq 0$, can't consider $\widehat{\rho}_j$ independently
- $\widehat{\rho}_j$ may not be zero even if X_i's are independent. It is random.

Jiang Li, Ph.D.
Department of Computer Science

Correlation Estimate Example

- $D_1, D_2, \ldots, D_{100}$ from the process of delays for a covariance-stationary M/M/1 queue with $\rho = 0.9$
- Estimate ρ_j (j = 1, 2, …, 10)



HOWARD UNIVERSITY
Jiang Li, Ph.D.
Department of Computer Science

Are We Stuck?

- Simulation output data are most likely correlated
- Difficulty dealing with correlations
- It is possible to group simulation output data into new "observations" and we can use results based on IID assumptions
  - Chapter 9

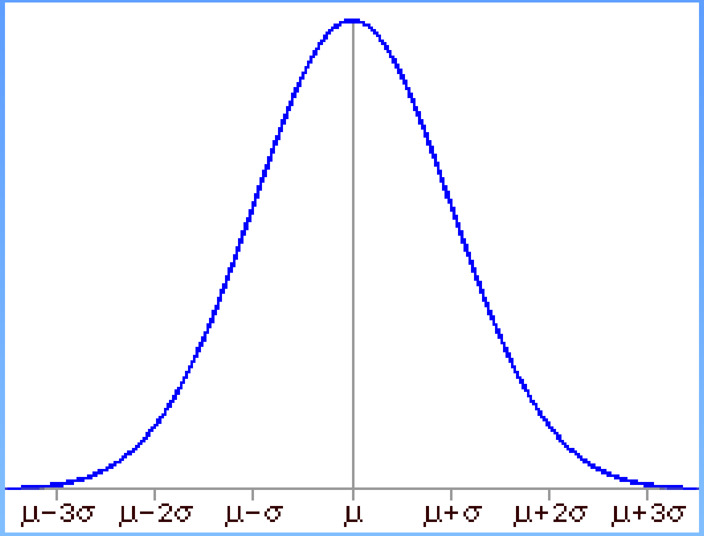
Jiang Li, Ph.D.
Department of Computer Science

Confidence Intervals and Hypothesis Tests for the Mean

- Confidence Intervals
  - Get an idea of the range of the mean
- Hypothesis Tests
  - Get an idea of what the mean exactly is

Jiang Li, Ph.D.
Department of Computer Science

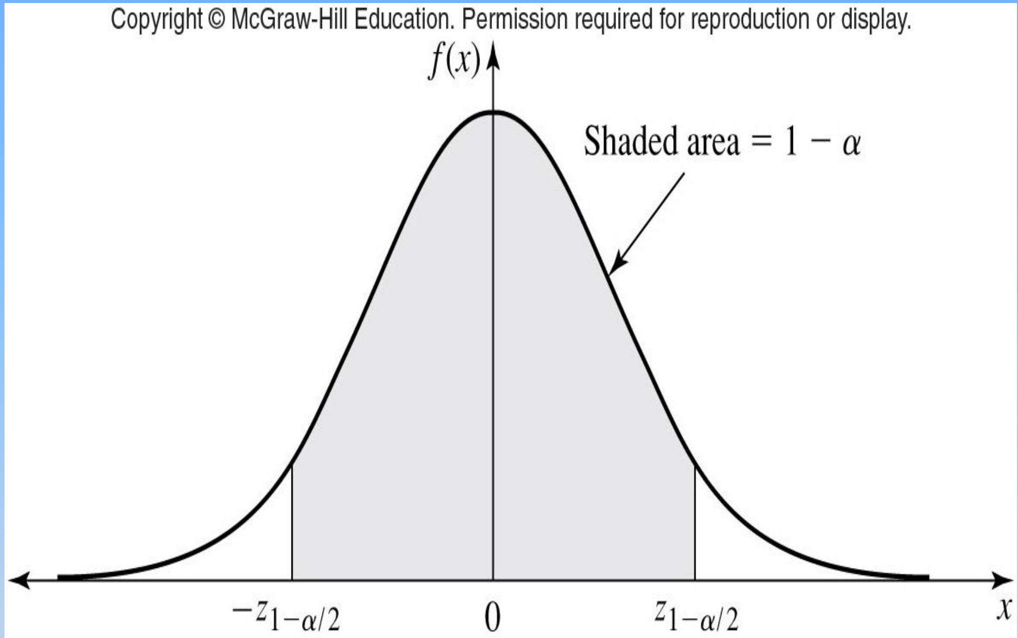
Central Limit Theorem



- $Z_n = \frac{\bar{X}(n)-\mu}{\sqrt{\sigma^2/n}}$
- When n is "sufficiently large", regardless of X_i's distribution (X_i's are IID)
  - Z_n will be approximately distributed as a standard normal distribution variable with mean 0 and variance 1
  - $\bar{X}(n)$ is approximately distributed as a normal distribution variable with mean μ and variance σ^2/n

Confidence Interval for the Mean

- Estimate σ^2 by $S^2(n)$
- $t_n = \frac{\bar{X}(n)-\mu}{\sqrt{S^2(n)/n}}$ is approx. distributed as a std. normal r.v.
- For large n



$$P\left(-z_{1-\alpha/2} \leq \frac{\bar{X}(n)-\mu}{\sqrt{S^2(n)/n}} \leq z_{1-\alpha/2}\right)$$

$$= \mathrm{P}\left[\bar{X}(n) - z_{1-\frac{\alpha}{2}}\sqrt{\frac{S^2(n)}{n}} \leq \mu \leq \bar{X}(n) + z_{1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}\right]$$

Jiang Li, Ph.D.
Department of Computer Science

Confidence Intervals for the Mean

- If n is sufficiently large, an approximate 100(1-α) percent confidence interval for μ is given by

$$\overline{X}(n) \pm z_{1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}$$

- Note that $\bar{X}(n)$, $S^2(n)$ are random variables, not specific values
  - When they are replaced by values, the probability is no longer 1-α as estimates are used.
- Coverage for the confidence interval
  - If one constructs a very large number of independent 100(1- α) percent conf. int., each based on n (>> 1) observations, the proportion of these conf. int. that contains the mean should be 1- α

Jiang Li, Ph.D.
Department of Computer Science

Confidence Interval Example

- 15 independent samples of size n = 10 from a normal distribution with mean 5 and variance 1
- 90 percent confidence interval for mean
- 2 of 13 (15.4%) does not include the mean

Jiang Li, Ph.D.
Department of Computer Science

Confidence Intervals for the Mean

- Problem: The more skewed (non-symmetric) the underlying distribution, the larger the value of *n* is needed
  - If n is small, the actual coverage of a 100(1-α) percent conf. int. will generally be less than 1-α
- Need larger actual coverage -> need larger confidence interval
  - How much larger?

Jiang Li, Ph.D.
Department of Computer Science

t Confidence Intervals for the Mean

- If X_i's are normal random variables
  - n samples: $X_1, X_2, \ldots, X_n$
  - $t_n = \frac{\bar{X}(n) - \mu}{\sqrt{S^2(n)/n}}$ has a *t* distribution with n-1 degree of freedom (n > 1)
  - *t* 100 (1- α) percent confidence interval for μ is given by

$$\overline{X}(n) \pm t_{n-1,1-\alpha/2}\sqrt{\frac{S^2(n)}{n}}$$

  - $t_{n-1,1-\alpha/2} > z_{1-\alpha/2}$
  - $t_{n-1,1-\alpha/2} \to z_{1-\alpha/2}$ as $\mathrm{n} \to \infty$
    - $t_{40.0.95}$ differs from $z_{0.95}$ by less than 3%

t Confidence Interval Example

- 10 observations from a normal distribution
  - 1.20, 1.50, 1.68, 1.89, 0.95, 1.49, 1.58, 1.55, 0.50, 1.09
  - $\bar{X}(10) = 1.34, S^2(10) = 0.17$
  - 90 percent conf. int. for the mean

$$\bar{X}(10) \pm t_{9,0.95}\sqrt{\frac{s^2(10)}{10}} = 1.34 \pm 1.83\sqrt{\frac{0.17}{10}}$$

$$= 1.34 \pm 0.24$$

Distribution Impacts on C.I. Coverage

- Skewness, a measure of symmetry
  - 0 for a symmetric distribution, e.g. normal
  - $v = \frac{E\left[(X-\mu)^3\right]}{(\sigma^2)^{3/2}}$
  - Use n samples to get t conf. int., repeat 500 times, check coverage

Distribution	Skewness v	$n = 5$	$n = 10$	$n = 20$	$n = 40$
Normal	0.00	0.910	0.902	0.898	0.900
Exponential	2.00	0.854	0.878	0.870	0.890
Chi square	2.83	0.810	0.830	0.848	0.890
Lognormal	6.18	0.758	0.768	0.842	0.852
Hyperexponential	6.43	0.584	0.586	0.682	0.774

Jiang Li, Ph.D.
Department of Computer Science

Willlink Confidence Interval

- Use Estimated Skewness
  - $\hat{\mu}_3/[S^2(n)]^{3/2}$
- $\hat{\mu}_3 = \frac{n\sum_{i=1}^{n}[X_i - \bar{X}(n)]^3}{(n-1)(n-2)}$, a $= \frac{\hat{\mu}_3}{6\sqrt{n}(S^2(n))^{3/2}}$
- $G(r) = \frac{[1+6a(r-a)]^{1/3}-1}{2a}$
- $100(1-\alpha)$ percent confidence interval

$$\bar{X}(n) - G(\pm t_{n-1,1-\frac{\alpha}{2}})\sqrt{S^2(n)/n}$$

Jiang Li, Ph.D.
Department of Computer Science

Willlink Confidence Interval Ex.

- 10 observations from a normal distribution
  - 1.20, 1.50, 1.68, 1.89, 0.95, 1.49, 1.58, 1.55, 0.50, 1.09
  - $\hat{\mu}_3 = -0.062, \mathrm{a} = -0.048$
  - $G(r) = \frac{[1-0.288(r+0.048))]^{1/3}-1}{-0.096}$
  - 90 percent conf. int. for the mean

    [1.34 – 0.31, 1.34 + 0.20]
  - For n=10, lognormal dist., 500 experiments, Willink C.I. have a coverage of 0.872 vs. 0.796 of t C.I., with 76% larger avg. half-length.

Jiang Li, Ph.D.
Department of Computer Science

Hypothesis Tests for the Mean

- $X_1, X_2, \ldots, X_n$ are (approx.) normally distributed
- Null hypothesis H_0: $\mu = \mu_0$
- Alternative hypothesis H_1: $\mu \neq \mu_0$
- If H_0 is true, $t_n = \frac{\bar{X}(n) - \mu_0}{\sqrt{S^2(n)/n}}$ has a *t* distribution with n-1 degree of freedom => *t* test
  - If $|t_n| > t_{n-1,1-\alpha/2}$, reject H_0
  - Otherwise, fail to reject H_0

– $\{x \text{ s.t. } |x| > t_{n-1,1-\frac{\alpha}{2}}\}$: rejection (critical) region

– Level (size) of the test = α, usually 0.05 or 0.10
  - Probability to reject H_0 when H_0 is true.

Test Errors

- Type I error
  - Rejecting H_0 when in fact it is true
  - Probability is equal to the level α
- Type II error
  - Failure to reject H_0 when it is false
  - Probability denoted by β

Power of the test.
Can only be increased by increasing n when level is fixed.

Outcome \ H_0	True	False
Reject	α	$\delta = 1 - \beta$
Fail to reject	$1 - \alpha$	β

t Test Example 1

- 10 observations from a normal distribution
  - 1.20, 1.50, 1.68, 1.89, 0.95, 1.49, 1.58, 1.55, 0.50, 1.09
  - $\bar{X}(10) = 1.34, S^2(10) = 0.17$
  - H$_0$: $\mu = 1$, $\alpha = 0.1$
  - $t_{10} = \frac{\bar{X}(10)-1}{\sqrt{S^2(10)/10}} = \frac{0.34}{\sqrt{0.17/10}} = 2.65 > 1.83 = t_{9,0.95}$

HOWARD UNIVERSITY
Jiang Li, Ph.D.
Department of Computer Science

t Test Example 2

- n observations from a normal distribution with $\mu = 1.5$ and $\sigma = 1$
- 1000 observations of $t_{10} = \frac{\bar{X}(10)-1}{\sqrt{S^2(10)/10}}$
  - For 433 of them, $|t_{10}| > 1.83$ => Estimated power is 0.433 for a test at level 0.10
- 1000 observations of t_{25} and t_{100}
  - Estimated power is 0.796 and 0.999

Confidence Interval vs. Hypothesis Test

- Rejection of the null hypothesis H_0: $\mu = \mu_0$ at level α is equivalent to μ_0 not being in the 100(1 - α) confidence interval for μ
- Confidence interval is preferred
  - A range of possible value is provided

Jiang Li, Ph.D.
Department of Computer Science

4.6 The Strong Law of Large Numbers

- If one performs an infinite number of experiments, each resulting in an $\bar{X}(n)$, and n is sufficiently large, then $\bar{X}(n)$ will be arbitrarily close to μ for almost all the experiments

THEOREM 4.2. $\bar{X}(n) \rightarrow \mu$ w.p. 1 as $n \rightarrow \infty$

Jiang Li, Ph.D.
Department of Computer Science

The Strong Law of Large Numbers

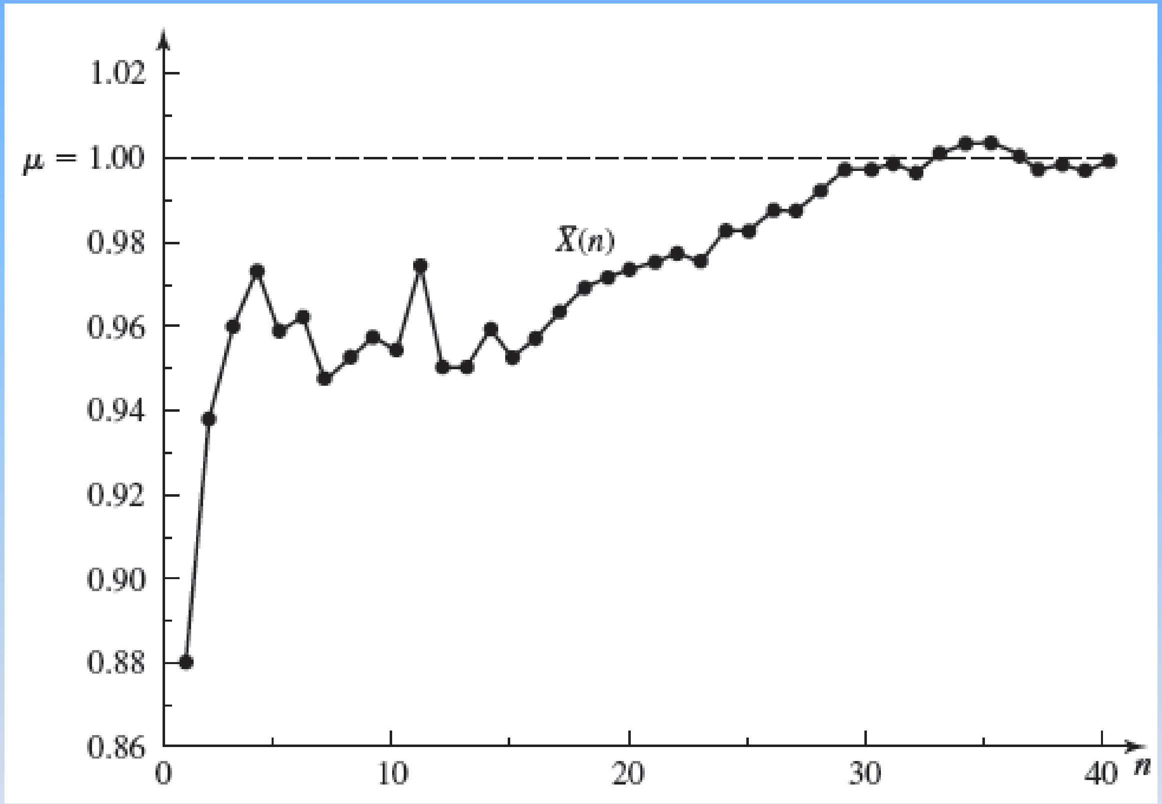


Figure 4.18 $\bar{X}(n)$ for various values of n when the X_i's are normal random variables with μ=1 and σ^2=0.01

4.7 The Danger of Replacing a Probability Distribution by its Mean

- Do not replace an input probability distribution by its mean in a simulation model
- Example: manufacturing system with single machine tool
  - Raw parts arrive at the machine with exponential interarrival times with a mean of one minute

Jiang Li, Ph.D.
Department of Computer Science

The Danger of Replacing a Probability Distribution by its Mean

- Example (cont'd.)
  - Processing times at the machine distributed exponentially with a mean of 0.99 minutes
  - This system is an $M/M/1$ queue with utilization factor $\rho = 0.99$
  - It can be shown that average delay in queue of a part in the long run is 98.01 minutes

Jiang Li, Ph.D.
Department of Computer Science

The Danger of Replacing a Probability Distribution by its Mean

- Example (cont'd.)
  - If we replace each distribution by its corresponding mean (parts arrive at t = 1 min, t = 2 min, etc.)
    - And if each part has a processing time of exactly 0.99 minutes
    - Then no part is ever delayed in the queue
- The variances as well as the means of the input distributions affect the output measures for queuing-type systems

Jiang Li, Ph.D.
Department of Computer Science

Summary

- PMF, PDF, CDF
- Mean, variance, covariance, correlation
  - And their estimates
- Understand how good the estimate of mean is
  - Confidence interval
    - Ordinary, t, Willink
  - Hypotheses test
- Strong law of large numbers
- Don't replace a prob. distribution by its mean