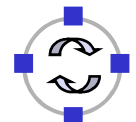


Simulative Evaluation of Internet Protocol Functions

Chapter 9 Obtaining Data From Simulations & Short Probability Primer

(Acknowledgement: these slides have mostly been compiled from [Kar04, Rob01, BLK02])



- ❑ So far, we have mainly used the OMNet++ simulation environment as a tool to set up a kind of “playground” for instructive illustration of and experimentation with networking algorithms:
 - ❑ While this is a worthwhile way of using a simulation environment, this is not the original purpose for which simulations are usually build
 - ❑ Most of the times, a simulation program is build in order to help answer a series of (quantitative) questions
 - ❑ This leads us to an inevitable excursion into the topics of model parameters, metrics and measurements... :o)

- ❑ Parameters, metrics, and measurements:
 - ❑ Parameters and metrics
 - ❑ How to measure typical types of metrics
 - ❑ On the meaning of measurements

(compiled from [Kar04])

System, Load, and Fault Models (1)

- ❑ System model:
 - ❑ Describes the composition of an entire system out of simpler subsystems
 - ❑ Represents the behavior, the way a system works
- ❑ In communication networks:
 - ❑ Entities communicate by exchanging messages over links
 - ❑ Protocols implemented in entities, can have parameters like processing delays, limited queue length
 - ❑ Links can have parameters like bandwidth and delay

System, Load, and Fault Models (2)

- ❑ Load model:
 - ❑ Describes the pattern with which requests are made to the system to perform different kinds of activities
 - When do such requests occur?
 - What is the time between requests?
 - What are the parameters of such a request?

- ❑ In communication networks:
 - ❑ Load is the desire of a “user” to send a packet to another user
 - Example parameter: how big is the packet
 - ❑ On a coarser abstraction level:
 - how often are connections established,
 - how much data is transmitted within a connection,
 - what is the mix between different types of connections (QoS) within a session, etc.

System, Load, and Fault Models (3)

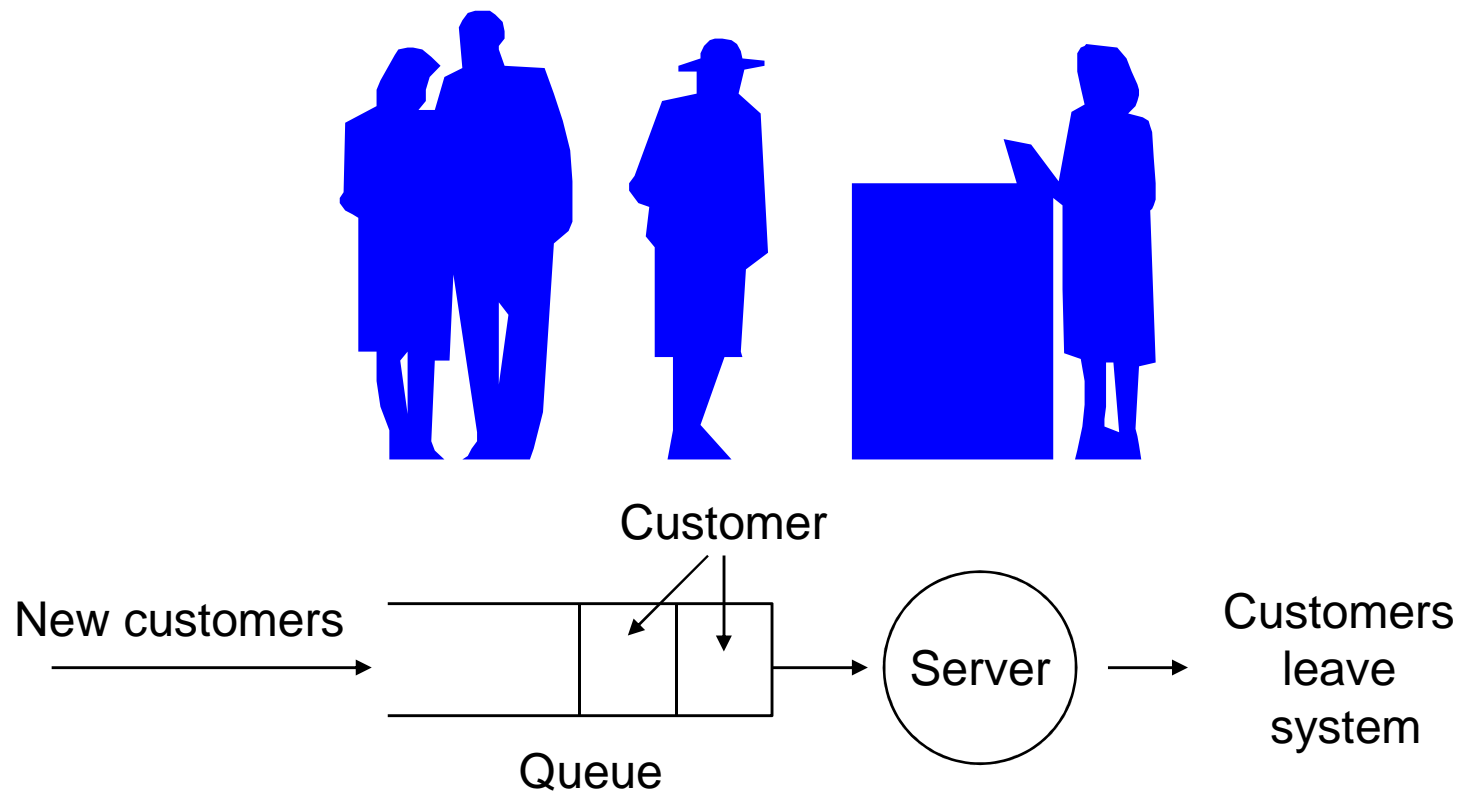
- ❑ Fault model:
 - ❑ Describes which parts of a system can deviate from their proscribed/desired behavior, and in which form
 - When do such deviations occur?
 - Can they be repeated (are faulty entities repaired)?
 - How long do deviations last?
 - Which kind of faulty behavior occurs?
- ❑ In communication networks:
 - ❑ Entities can be faulty (e.g., a node can crash, often considered a permanent error)
 - ❑ Communication links can be faulty (e.g., some bits are not transmitted correctly, due to electromagnetic noise, usually considered a transient error)

System, Load, and Fault Models (4)

- ❑ Neither the arrival of requests nor the occurrence of faults can be described deterministically
- ❑ Random distributions needed to model these events along with their parameters:
 - ❑ What distributions are available, appropriate and easy to use in simulations?
 - ❑ How can random numbers be generated such that the simulated events occur according to these distributions? (Based on general-purpose random number generators?)
 - ❑ How well do standard distributions match observed behavior of real (communication) systems? How to choose parameters for distributions to model real systems?
 - ❑ What to do if no simple standard distributions can be found that match a real system's behavior?
 - ❑ Is it sufficient to just look at distributions?
- ❑ And what exactly is a “*distribution*”???
 - ❑ Later in this chapter, we will review some basics of probability

Back to Basics: A Simple Model

- In order to discuss parameters and metrics in more detail, let us recall our simple model of the introduction chapter:



(Simulation) Model Parameters

- ❑ In our introductory manual simulation example, parameters have not been taken into account
- ❑ Parameters are:
 - ❑ Pattern of customer arrival
 - Deterministic modeling impossible
 - Stochastic modeling: consider the time between customers as a random variable, choose a suitable distribution for this variable (by the way: what are *random variables* and *distributions*? → see later)
 - Common case: interarrival time of customers is *exponentially distributed* (explained later), characterized by the distribution's *mean*
 - ❑ Pattern of service times
 - Similar to arrival patterns
 - Service time modeled as a random variable, e.g. also with an exponential distribution
 - ❑ Speed of the server
 - Can arbitrarily be set to 1 by scaling the time
 - Direct representation is simple, too

And What About Metrics?

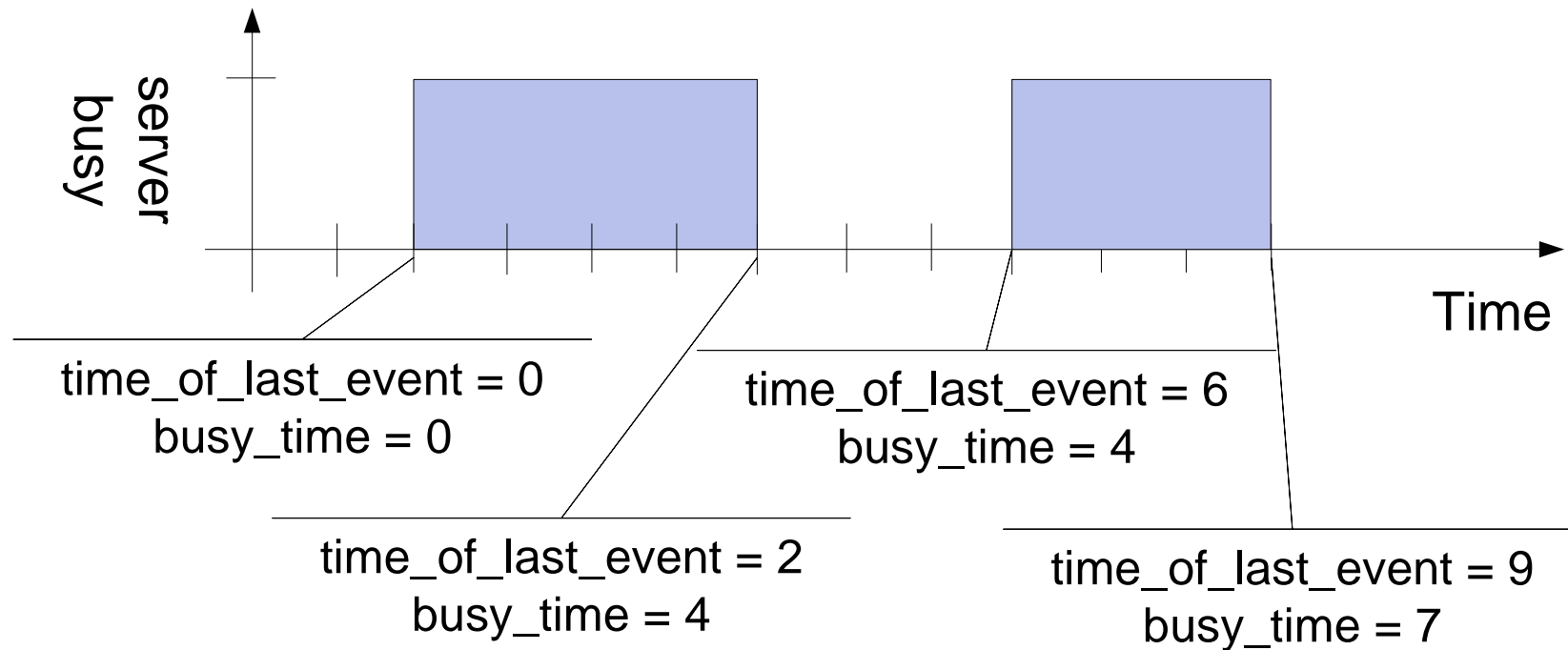
- ❑ The initial goal was to investigate:
 - ❑ Utilization of the server (“what percentage of time is the server busy”)
 - ❑ Length of the queue (“how much space do we need in the store?”)
 - What does this mean? At most? On average?
 - ❑ Waiting time of customers
 - Again: At most? On average? Fridays?

- ❑ How to capture such information from a simulation?

Measuring Server Utilization (1)

- ❑ Directly measuring utilization is difficult
- ❑ Simple to measure the absolute time the server has been busy, and at the end divide it by the total time that has been simulated
 - ❑ Introduce a counter “busytime”, initialize to zero
 - ❑ At every event: if the server has been busy in the time before this event, add the time since the last event to “busytime”
 - Easy to check the “old” server state if this check is done before the state is updated according to the current event
 - ❑ Additional counter “time_of_last_event” necessary
 - Trivial to compute time *since* last event from this
 - Before setting the simulation clock to the new time, store it in “time_of_last_event”

Measuring Server Utilization (2)

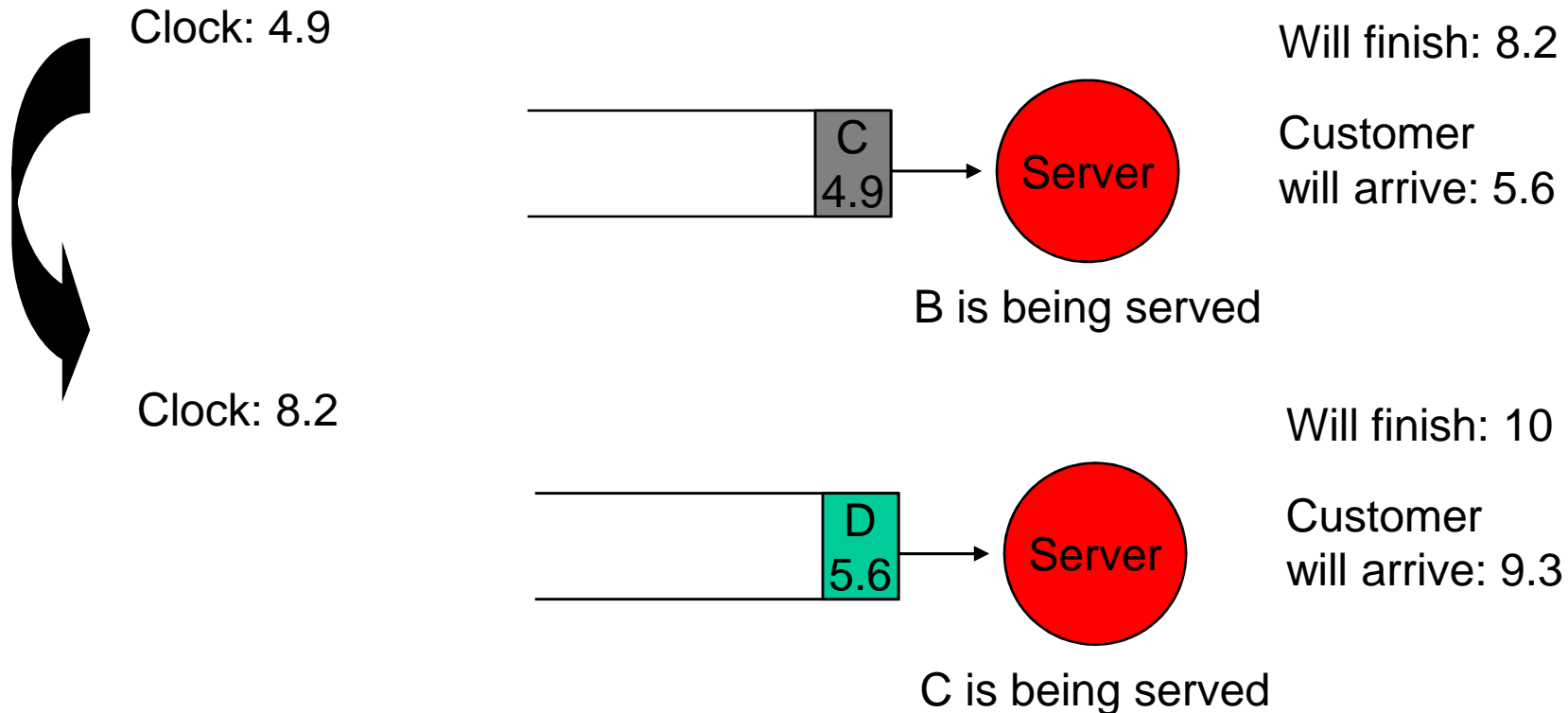


Measuring Customer Waiting Time (1)

- ❑ How much time does a customer spend in the queue?
- ❑ When putting customer in the queue, mark it with the time this was done
- ❑ When retrieving a customer from the queue, take difference between current simulation clock and time of entry into queue
- ❑ If customer does not enter queue, waiting time is 0

Measuring Customer Waiting Time (2)

- Example from above: waiting time for customer C



- Waiting time for C is current simulation clock – time of enqueueing = $8.2 - 4.9 = 3.3$

Measuring Customer Waiting Time (3)

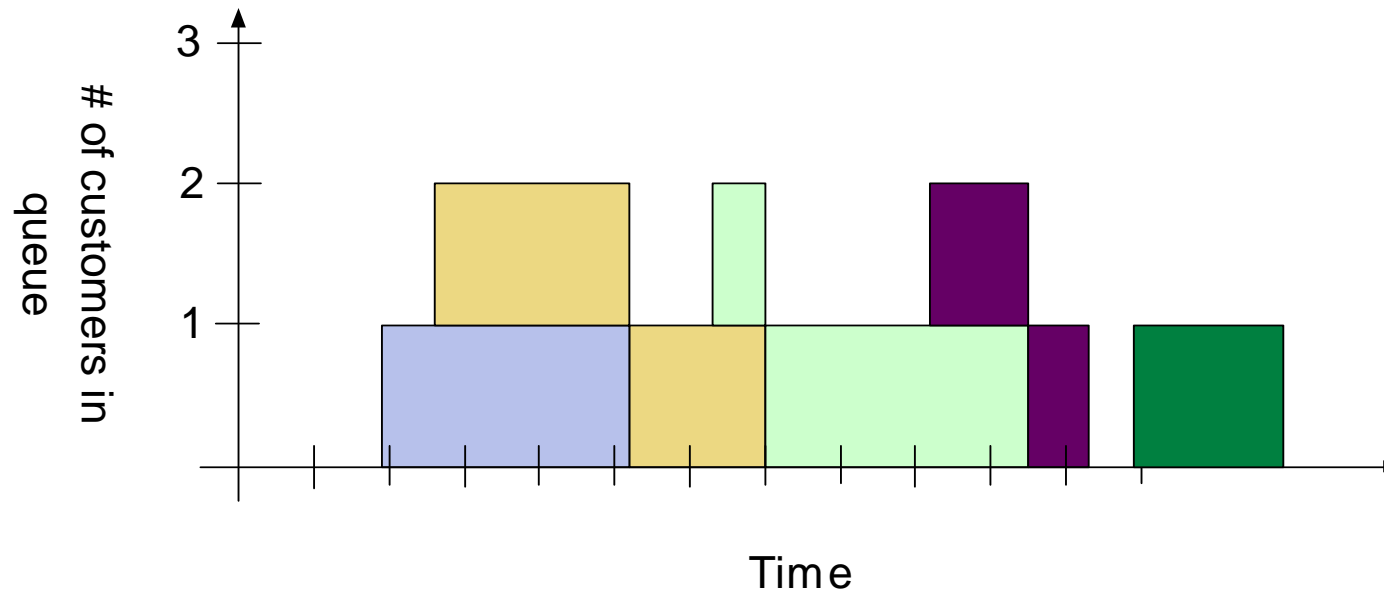
- ❑ This yields waiting times for each customer
- ❑ How to aggregate this information?
- ❑ Build the average!
 - ❑ Let D_i stand for the delay of customer i (possibly 0)
 - ❑ Compute the usual arithmetic average of the D_i s

$$1/n \sum_{i=1}^n D_i$$

- ❑ Note: (Arithmetic) average over a discrete number of data

Measuring Queue Length (1)

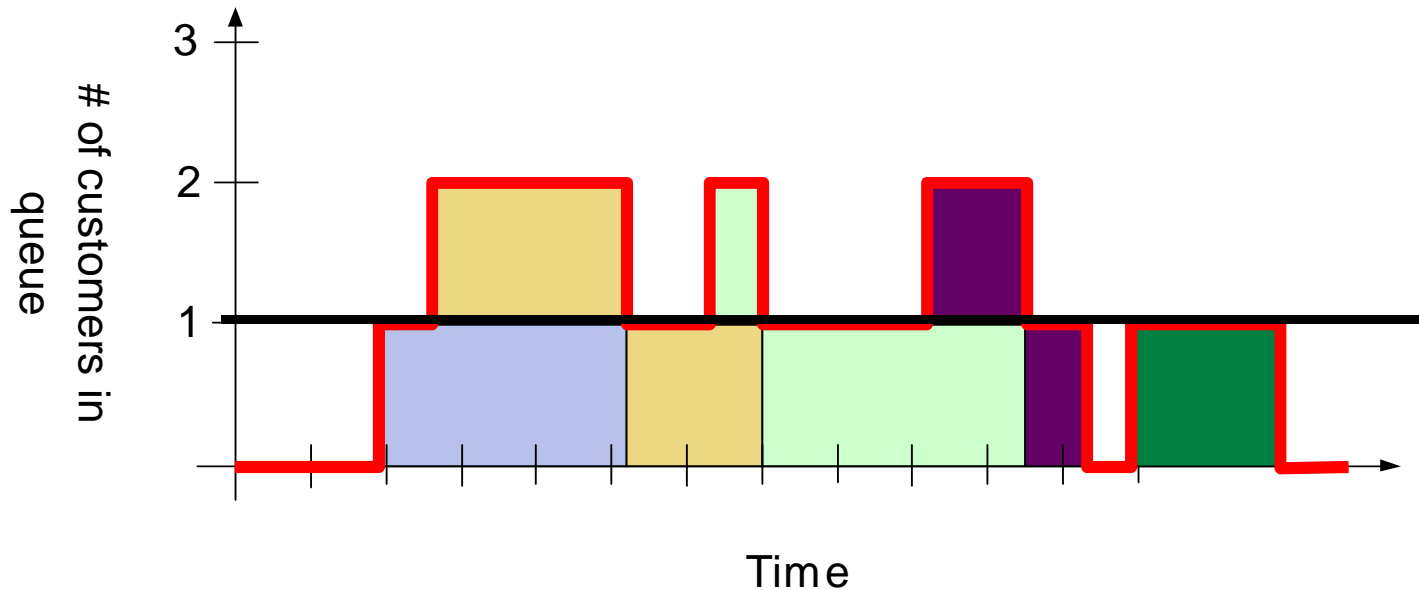
- How does the number of customers in the queue behave?



- Discrete-valued function over time, changing at arbitrary points in time (customers joining and leaving)

Measuring Queue Length (2)

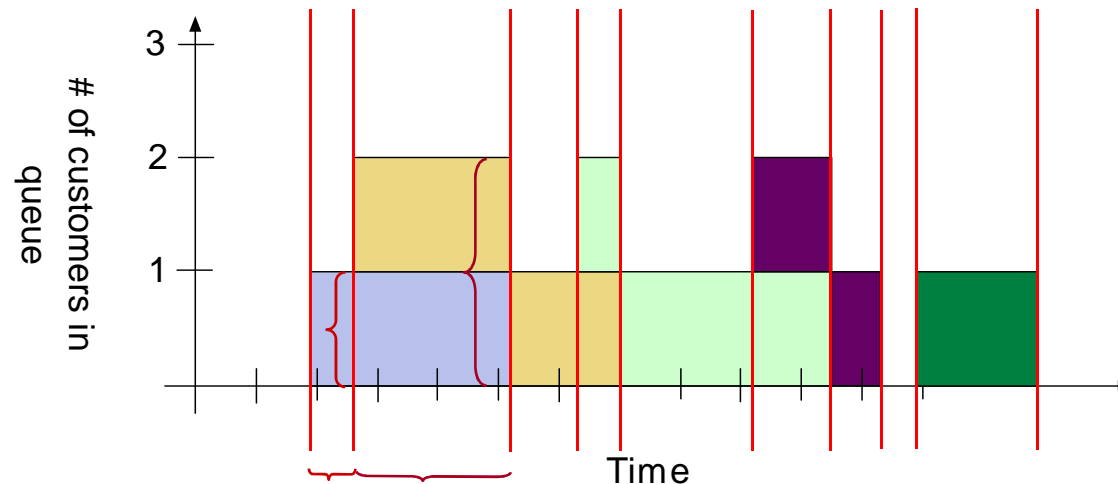
- Appropriate representation: average of this function over time



- Area of this curve = 17 between time 0 and 16.9
- Average length of the queue is hence roughly 1.006

Measuring Queue Length (3)

- ❑ How to easily compute this average?
 - ❑ Many complicated approaches possible 😊
- ❑ Simplest way
 - ❑ Compute the total area under the curve, and at the end of the simulation, divide by the total simulated time
 - ❑ At every event that manipulates the queue: add the area since the last event to the total area



Different Kinds of Metrics

- ❑ Average waiting time of customers
 - ❑ Average is taken over a discrete set of individuals
 - ❑ Measured value is time
 - ❑ Example for a *discrete-time* metric (or statistic)

- ❑ Average queue length
 - ❑ Average is continuously taken over time
 - ❑ Measured value is a number, a discrete metric
 - ❑ Example for a *continuous-time* metric (or statistic)

Meaning of Measurements (1)

- ❑ What is the correct interpretation of such simulation-based measurements?
- ❑ Look e.g. at waiting time in queue
 - ❑ Let D_i represent the waiting time as measured for customer i
 - ❑ This refers to one particular simulation run – or to observations on one particular day (for the real system)!
 - ❑ For different simulations/on different days, the D_i s will be different
 - ❑ D_i s are averaged over n customers to obtain aggregate information, here: *average* waiting time in queue
 - ❑ Other possible aggregations: max, min, proportions, ...
- ❑ Both D_i s and their aggregate will change from one set of observations to the next!

Meaning of Measurements (2)

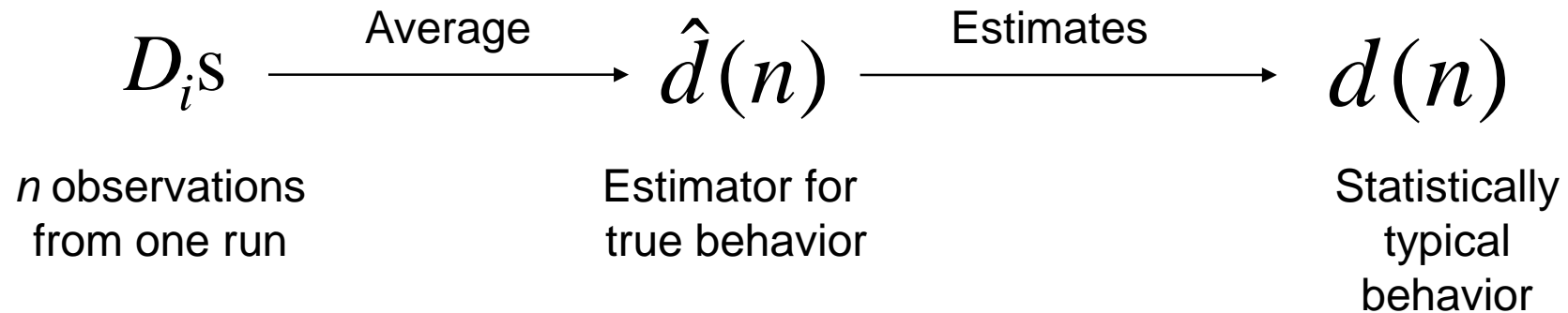
- ❑ Why look at aggregated information of measurements?
- ❑ Aggregated information gives a more concise representation/description of the system under study
- ❑ It is easier to compare aggregated information from different system/simulation runs
 - ❑ The average waiting times in a supermarket on weekdays as opposed to Saturdays is more informative than waiting times of individual customers (which are not really important)
- ❑ So what about looking at distributions instead of averages?
- ❑ Successive values might not be distributed in the same way
 - ❑ In the queuing example, D_1 is always 0, whereas D_2 , D_3 , etc. are not
 - ❑ Hence, the average of such values are not the usual statistical average which is usually computed over independent, but identically distributed observations

Meaning of Measurements (3)

- ❑ The “truly typical” (in a statistical sense) behavior of the model can be considered to be the “average” over all possible behaviors the model can exhibit
- ❑ Weighted by the probabilities of these behaviors
- ❑ Sometimes possible to analytically derive closed-form descriptions for such behavior

Meaning of Measurements (4)

- What is the relationship between truly typical behavior and the result of a simulation run?



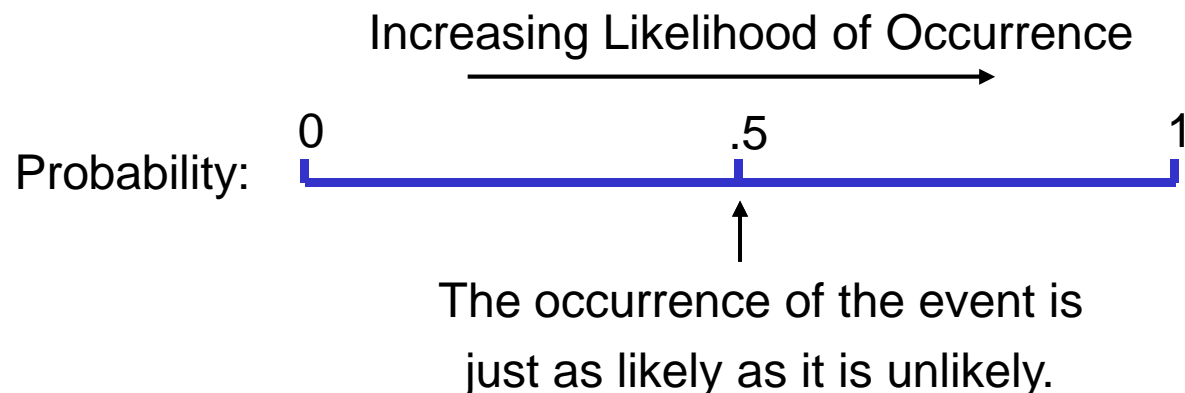
- How good is such an estimator? How to improve the estimation? What are typical sources of errors?
 - There is a course on the topic of simulation which studies these questions in much more detail (“Leistungsbewertung”, held every winter term)

How Long to Run Simulations?

- ❑ When have sufficiently many observations been collected?
- ❑ Depends on the purpose (again)
- ❑ Sometimes (rarely): behavior of system for a well-defined finite amount of time is of interest
 - ❑ Example: supermarket closes at 8pm, no matter what
 - ❑ Simulate for a certain fixed amount of simulated time, or a fixed number of events
 - ❑ Problem: quality of estimations is variable
- ❑ Often: certain metrics (depending on input parameters) are of interest
 - ❑ Simulate as long as it takes for the *estimation* of these metrics to have reached a desired “quality level” (we will come back to this later)
 - ❑ Correctly computing the quality level is difficult
 - ❑ Amount of necessary simulated time can vary
 - ❑ “Stopping rules” regulate when to stop a simulation

Short Review of Some Probability Basics

- ❑ *Probability* is a numerical measure of the likelihood that an event will occur.
- ❑ Probability values are always assigned on a scale from 0 to 1.
- ❑ A probability near 0 indicates an event is very unlikely to occur.
- ❑ A probability near 1 indicates an event is almost certain to occur.
- ❑ A probability of 0.5 indicates the occurrence of the event is just as likely as it is unlikely.



(compiled from [Rob01])

An Experiment and Its Sample Space

- ❑ An *experiment* is any process that generates well-defined outcomes.
- ❑ The *sample space* for an experiment is the set of all experimental outcomes.
- ❑ A *sample point* is an element of the sample space, any one particular experimental outcome.
- ❑ Examples:

Experiment

Outcomes

- ❑ Draw a card from a pack
- ❑ Telephone sales call
- ❑ First number drawn in National Lottery

- ❑ {Ace hearts, 2 hearts, ..., King of spades}
- ❑ Sale or no sale
- ❑ {1,2,3,...,49}

Constructing Sample Spaces

- ❑ A good way to construct the sample space is to write down examples of typical outcomes and try to identify the complete set

- ❑ Example:

Toss a coin four times

One typical outcome is four consecutive heads (H,H,H,H), another is a head, tail, head and head (H,T,H,H)

A little thought results in identifying the sample space as the set of all such 4-tuples

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

A Counting Rule for Multiple-Step Experiments

- ❑ If an experiment consists of a sequence of k steps in which there are n_1 possible results for the first step, n_2 possible results for the second step, and so on, then the total number of experimental outcomes is given by $(n_1)(n_2) \dots (n_k)$
- ❑ A helpful graphical representation of a multiple-step experiment is a tree diagram

Counting Rule for Combinations

Another useful counting rule enables us to count the number of experimental outcomes when n objects are to be selected from a set of N objects

- Number of combinations of N objects taken n at a time

$${}^N C_n = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

where

$$N! = N(N-1)(N-2) \dots (2)(1)$$

$$n! = n(n-1)(n-2) \dots (2)(1)$$

$$0! = 1$$

Counting Rule for Permutations

A third useful counting rule enables us to count the number of experimental outcomes when n objects are to be selected from a set of N objects where the order of selection is important

- Number of permutations of N objects taken n at a time:

$${}^N P_n = n! \binom{N}{n} = \frac{N!}{(N-n)!}$$

Assigning Probabilities

- ❑ **Classical Method:**
 - ❑ Assigning probabilities based on the assumption of *equally likely outcomes*
- ❑ **Relative Frequency Method:**
 - ❑ Assigning probabilities based on *experimentation or historical data*
- ❑ **Subjective Method:**
 - ❑ Assigning probabilities based on the *assignor's judgment*
 - ❑ Applied in economics and related sciences

Classical Method

If an experiment has n possible outcomes, this method would assign a probability of $1/n$ to each outcome.

□ **Example:**

Experiment: Rolling a die

Sample Space: $S = \{1, 2, 3, 4, 5, 6\}$

Probabilities: Each sample point has a $1/6$ chance of occurring.

Relative Frequency of an Outcome

- Suppose that, in a large number of repetitions, N , of the experiment, the outcome O , occurs n_O times. The relative frequency of O is,

$$\frac{n_O}{N}$$

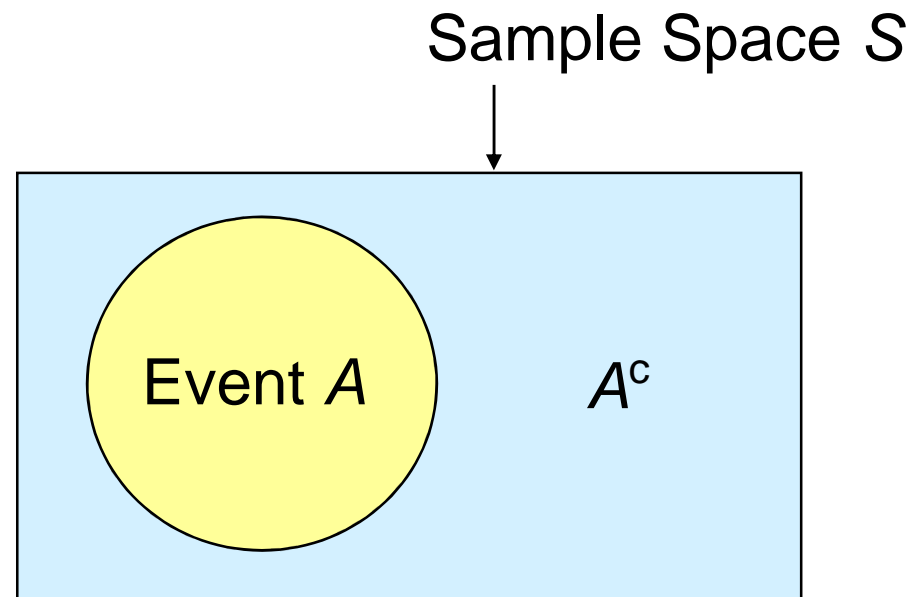
We can think of the probability of O as the value to which the relative frequency settles down as N gets larger and larger.

Events and Their Probability

- ❑ An *event* is a collection of sample points
- ❑ The *probability of any event* is equal to the sum of the probabilities of the sample points in the event
- ❑ If we can identify all the sample points of an experiment and assign a probability to each, we can compute the probability of an event
- ❑ There are some *basic probability relationships* that can be used to compute the probability of an event without knowledge of all the sample point probabilities:
 - ❑ Complement of an Event
 - ❑ Union of Two Events
 - ❑ Intersection of Two Events
 - ❑ Mutually Exclusive Events

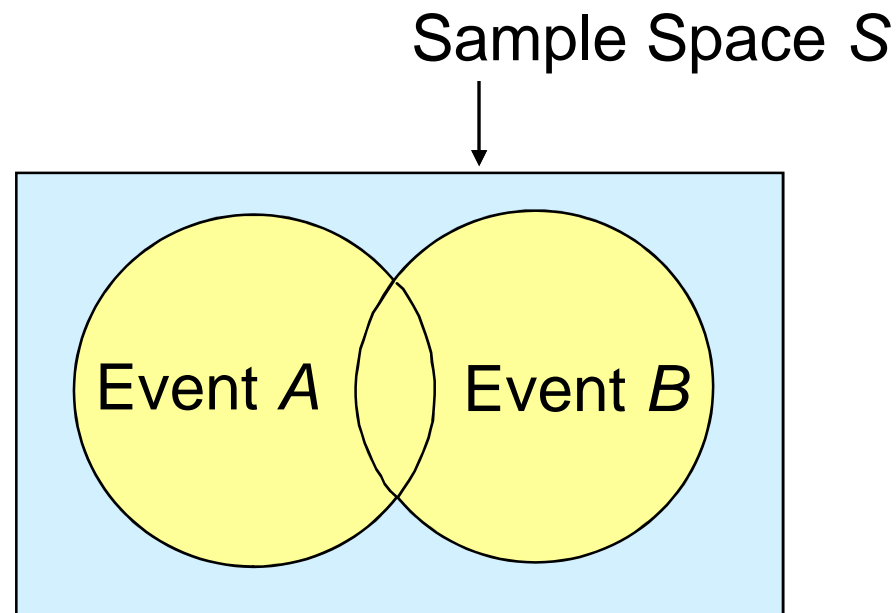
Complement of an Event

- ❑ The *complement* of event A is defined to be the event consisting of all sample points that are not in A
- ❑ The complement of A is denoted by A^c
- ❑ The *Venn diagram* below illustrates the concept of a complement



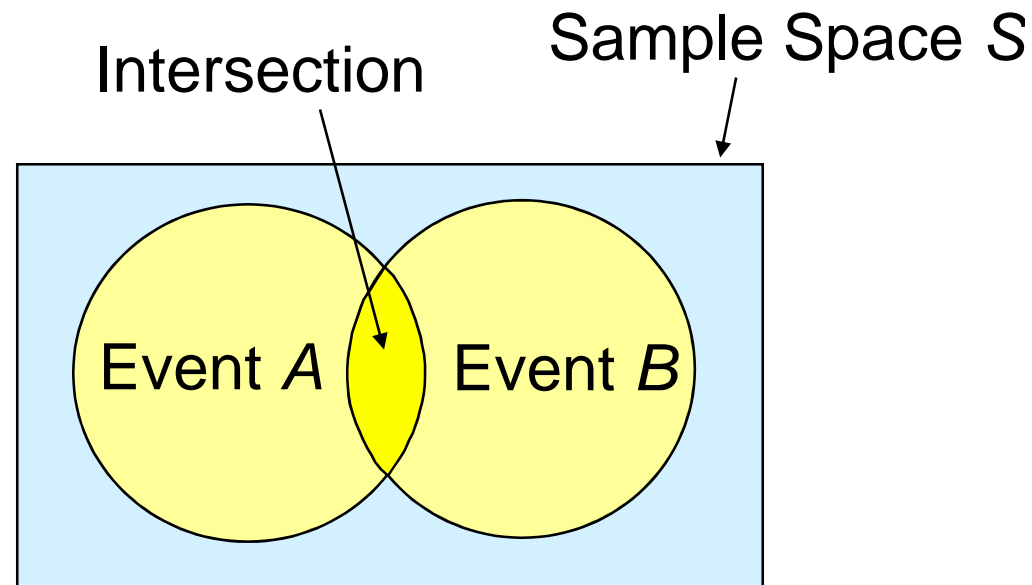
Union of Two Events

- ❑ The *union* of events A and B is the event containing all sample points that are in A or B or both
- ❑ The union is denoted by $A \cup B$
- ❑ The union of A and B is illustrated below



Intersection of Two Events

- ❑ The *intersection* of events A and B is the set of all sample points that are in both A and B
- ❑ The intersection is denoted by $A \cap B$
- ❑ The intersection of A and B is the area of overlap in the illustration below



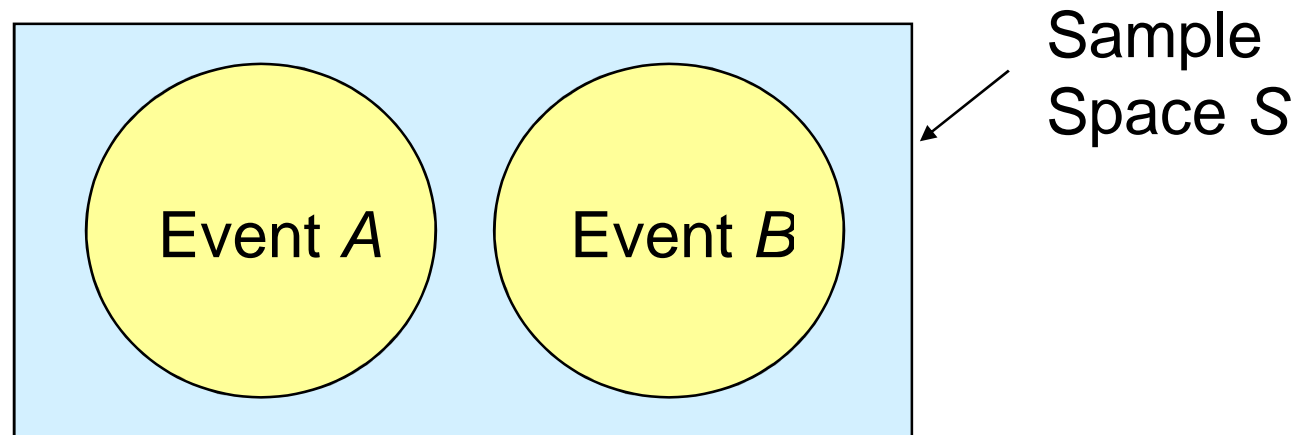
Addition Law

- ❑ The *addition law* provides a way to compute the probability of event A , or B , or both A and B occurring
- ❑ It is written as:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Mutually Exclusive Events

- Two events are said to be *mutually exclusive* if the events have no sample points in common
 - That is, two events are mutually exclusive if, when one event occurs, the other cannot occur



- Addition Law for Mutually Exclusive Events:

$$P(A \cup B) = P(A) + P(B)$$

Conditional Probability

- ❑ The probability of an event given that another event has occurred is called a *conditional probability*
- ❑ The conditional probability of *A given B* is denoted by $P(A|B)$
- ❑ A conditional probability is computed as follows:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Multiplication Law

- ❑ The *multiplication law* provides a way to compute the probability of an intersection of two events
- ❑ The law is written as:

$$P(A \cap B) = P(B)P(A|B)$$

Independent Events

- ❑ Events A and B are *independent* if $P(A|B) = P(A)$
- ❑ Multiplication Law for Independent Events

$$P(A \cap B) = P(A)P(B)$$

- ❑ The multiplication law also can be used as a test to see if two events are independent

Random Variables

- ❑ A *random variable* is a numerical expression of the outcome of an experiment
 - ❑ Example 1: toss a die twice; count the number of times the number 4 appears (0, 1 or 2 times)
 - ❑ Example 2: toss a coin; assign \$10 to head and -\$30 to a tail
- ❑ Discrete random variable:
 - ❑ Can only take discrete values
 - ❑ Obtained by counting (0, 1, 2, 3, etc.)
 - ❑ Usually a finite number of different values
 - ❑ E.g., toss a coin 5 times; count the number of tails (0, 1, 2, 3, 4, or 5 times)

(compiled from [BLK02])

Discrete Probability Distribution

- ❑ A discrete probability distribution is the list of all possible $[X_j, P(X_j)]$ pairs, with:
 - ❑ X_j = Value of random variable
 - ❑ $P(X_j)$ = Probability associated with value
 - ❑ This list can be visualized with a histogram
- ❑ Mutually Exclusive (Nothing in Common)
- ❑ Collective Exhaustive (Nothing Left Out)

$$0 \leq P(X_j) \leq 1 \quad \sum P(X_j) = 1$$

Basic Summary Measures (1)

- *Expected Value (The Mean):*
 - Weighted average of the probability distribution

$$\mu = E(X) = \sum_j X_j P(X_j)$$

- E.g., toss 2 coins, count the number of tails, compute expected value:

$$\begin{aligned} \mu &= \sum_j X_j P(X_j) \\ &= (0)(.25) + (1)(.5) + (2)(.25) = 1 \end{aligned}$$

Basic Summary Measures (2)

- *Variance:*

- Weighted average squared deviation about the mean

$$\sigma^2 = E\left[(X - \mu)^2\right] = \sum (X_j - \mu)^2 P(X_j)$$

- E.g., Toss 2 coins, count number of tails, compute variance:

$$\begin{aligned} \sigma^2 &= \sum (X_j - \mu)^2 P(X_j) \\ &= (0 - 1)^2 (.25) + (1 - 1)^2 (.5) + (2 - 1)^2 (.25) \\ &= .5 \end{aligned}$$

- The *standard deviation* is the square root of the variance:

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

Binomial Probability Distribution

- ❑ 'n' Identical Trials
 - ❑ E.g., 15 tosses of a coin; 10 light bulbs taken from a warehouse
- ❑ 2 Mutually Exclusive Outcomes on Each Trial
 - ❑ E.g., Heads or tails in each toss of a coin; defective or not defective light bulb
- ❑ Trials are Independent
 - ❑ The outcome of one trial does not affect the outcome of the other
- ❑ Constant Probability for Each Trial
 - ❑ E.g., Probability of getting a tail is the same each time we toss the coin
- ❑ 2 Sampling Methods
 - ❑ Infinite population without replacement
 - ❑ Finite population with replacement

Binomial Probability Distribution Function

$$P(X) = \frac{n!}{X!(n-X)!} p^X (1-p)^{n-X}$$

$P(X)$: probability of X successes given n and p

X : number of "successes" in sample ($X = 0, 1, \dots, n$)

p : the probability of each "success"

n : sample size

Tails in 2 Tosses of Coin

<u>X</u>	<u>$P(X)$</u>
0	$1/4 = .25$
1	$2/4 = .50$
2	$1/4 = .25$

Binomial Distribution Characteristics

- Mean

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

- E.g.,

$$\mu = np = 5(.1) = .5$$

- Variance and Standard Deviation

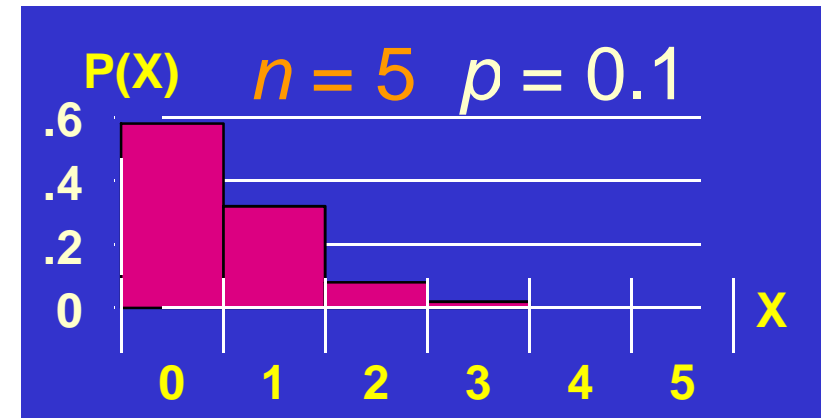
-

$$\sigma^2 = np(1-p)$$

$$\sigma = \sqrt{np(1-p)}$$

- E.g.,

$$\sigma = \sqrt{np(1-p)} = \sqrt{5(.1)(1-.1)} = .6708$$





Siméon Poisson

Poisson Distribution

- ❑ Discrete events (“successes”) occurring in a given area of opportunity (“interval”)
 - ❑ “Interval” can be time, length, surface area, etc.
- ❑ The probability of a “success” in a given “interval” is the same for all the “intervals”
- ❑ The number of “successes” in one “interval” is independent of the number of “successes” in other “intervals”
- ❑ The probability of two or more “successes” occurring in an “interval” approaches zero as the “interval” becomes smaller
 - ❑ E.g., # customers arriving in 15 minutes
 - ❑ E.g., # defects per case of light bulbs

Poisson Probability Distribution Function

$$P(X) = \frac{e^{-\lambda} \lambda^X}{X!}$$

$P(X)$: probability of X "successes" given λ

X : number of "successes" per unit

λ : expected (average) number of "successes"

e : 2.71828 (base of natural logs)

E.g., Find the probability of 4 customers arriving in 3 minutes when the mean is 3.6.

$$P(X) = \frac{e^{-3.6} 3.6^4}{4!} = .1912$$

Poisson Distribution Characteristics

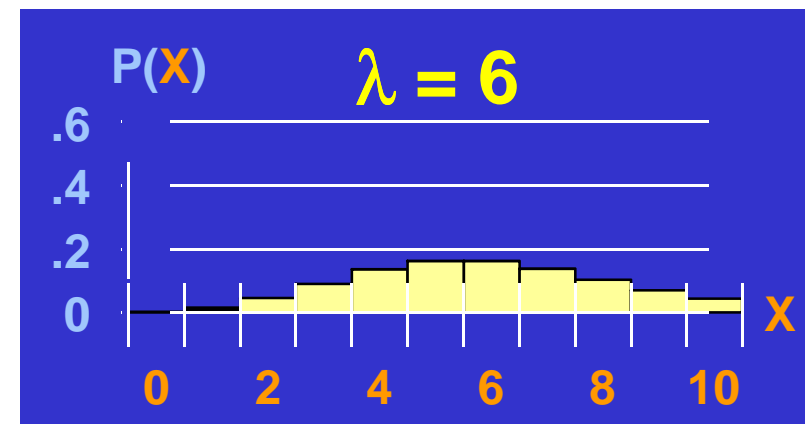
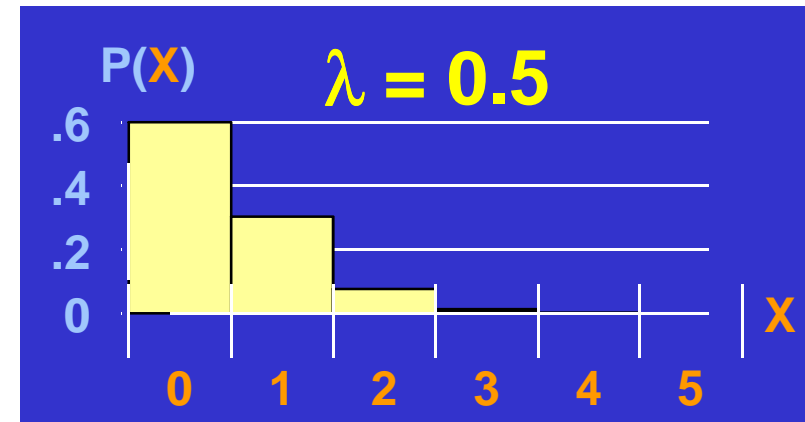
- Mean

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

$$= \sum_{i=1}^N X_i P(X_i)$$

- Standard Deviation and Variance

$$\sigma^2 = \lambda \quad \sigma = \sqrt{\lambda}$$



Continuos Random Variables

- ❑ *Continuous random variable:*
 - ❑ Values from interval of numbers (absence of gaps)
 - ❑ This implies, that for every single value, the probability that the variable takes on this value is 0
- ❑ *Continuous probability distribution:*
 - ❑ Distribution of continuous random variable
 - ❑ Transition from a histogram to a continuous function
 - ❑ The probability that the random variable takes on a value in the interval $[a, b]$ is the area under the function between the values a and b
- ❑ *Probability density function:*
 - ❑ The derivation of the probability distribution
- ❑ Important continuous probability distributions:
 - ❑ The *uniform distribution*
 - ❑ The *exponential distribution*
 - ❑ The *normal distribution*

The Uniform Distribution

- Properties:
 - The probability of occurrence of a value is equally likely to occur anywhere in the range between the smallest value a and the largest value b
 - Also called the rectangular distribution

- Mean:
$$\mu = \frac{(a + b)}{2}$$

- Variance:
$$\sigma^2 = \frac{(b - a)^2}{12}$$

The Uniform Distribution

- The probability density function of the uniform distribution

$$f(X) = \frac{1}{(b-a)} \quad \text{if } a \leq X \leq b$$

- Application: Selection of random numbers
 - E.g., A wooden wheel is spun on a horizontal surface and allowed to come to rest. What is the probability that a mark on the wheel will point to somewhere between the North and the East?

$$P(0^\circ < X < 90^\circ) = \frac{90^\circ}{360^\circ} = 0.25$$

Exponential Distributions

$$P(\text{arrival time} < X) = 1 - e^{-\lambda X}$$

X : any value of continuous random variable

λ : the population average number of arrivals per unit of time

$1/\lambda$: average time between arrivals

$$e = 2.71828$$

E.g., Drivers arriving at a toll bridge;
customers arriving at an ATM machine

Exponential Distributions

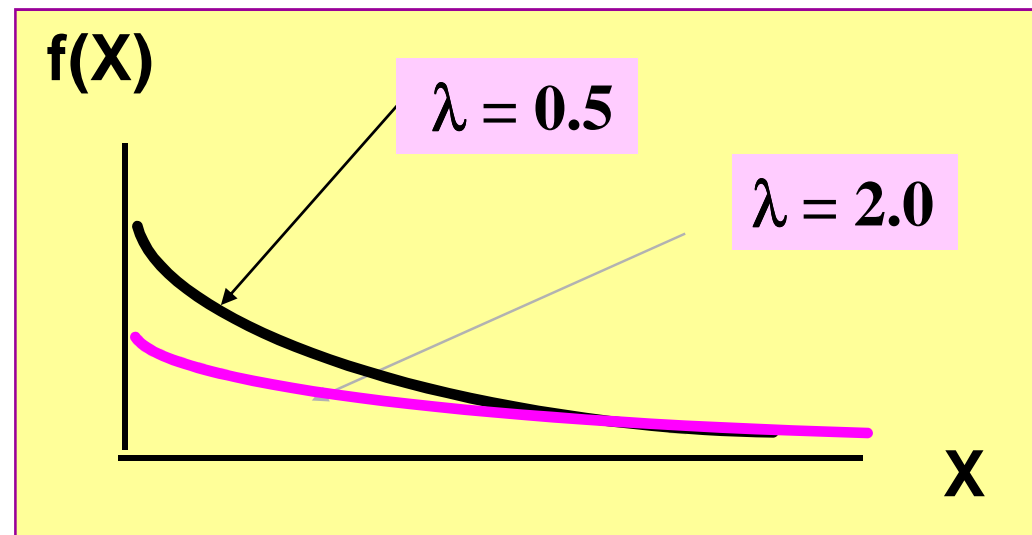
- Describe Time or Distance between Events
 - Used for queues

- Density Function

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

- Parameters

- $\mu = \lambda$ $\sigma = \lambda$



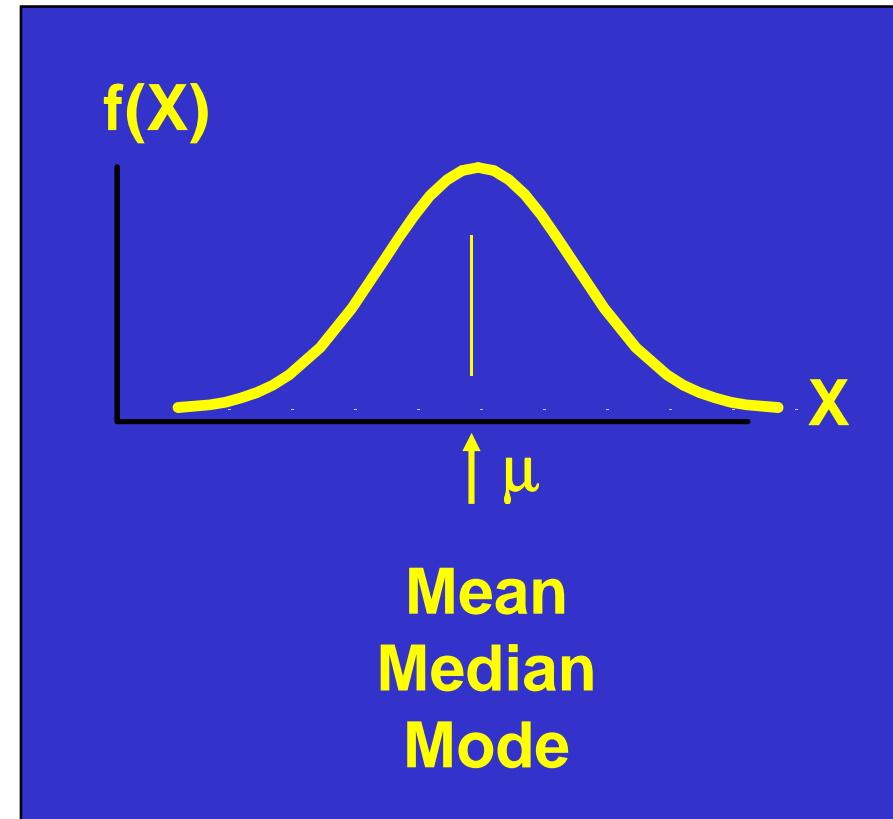
Exponential Distributions: Example

- ❑ Customers arrive at the checkout line of a supermarket at the rate of 30 per hour.
- ❑ What is the probability that the arrival time between consecutive customers will be greater than 5 minutes?

$$\lambda = 30 \quad X = 5 / 60 \text{ hours}$$

$$\begin{aligned}
 P(\text{arrival time} > X) &= 1 - P(\text{arrival time} \leq X) \\
 &= 1 - \left(1 - e^{-30(5/60)}\right) \\
 &= .0821
 \end{aligned}$$

- ❑ Properties:
 - ❑ “Bell Shaped”
 - ❑ Symmetrical
 - ❑ Mean, Median and Mode are Equal
 - ❑ Random Variable Has Infinite Range



$$f(X) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left[\frac{(X-\mu)}{\sigma}\right]^2}$$

$f(X)$: density of random variable X

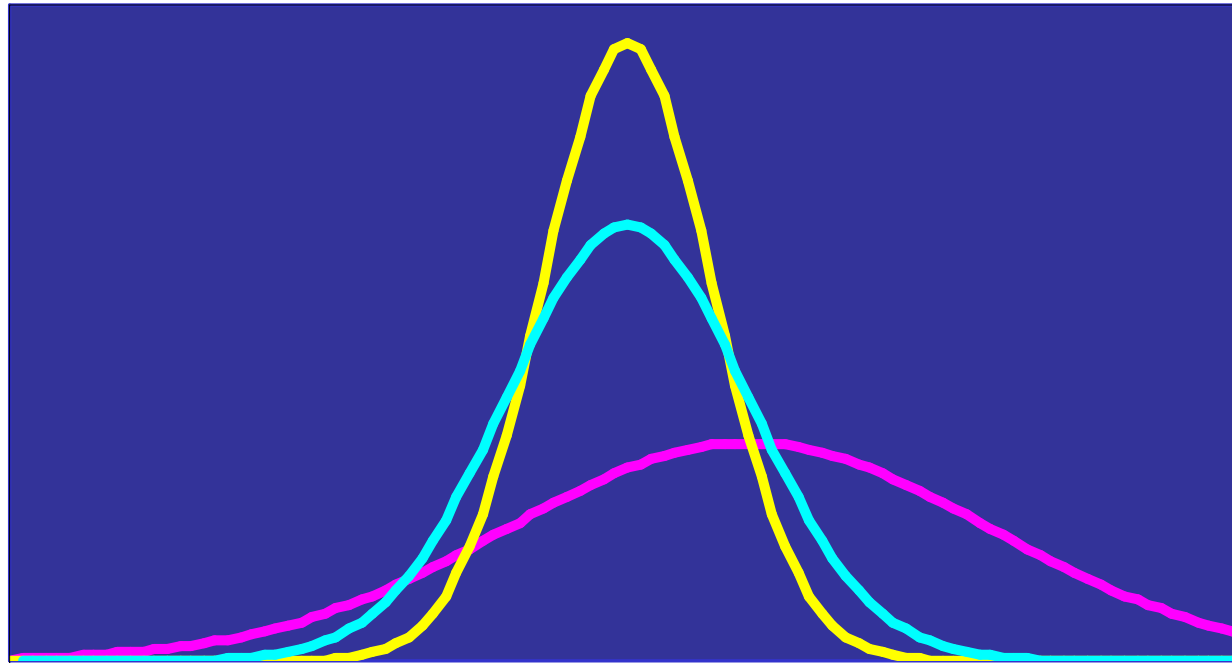
$\pi \approx 3.14159$; $e \approx 2.71828$

μ : population mean

σ : population standard deviation

X : value of random variable ($-\infty < X < \infty$)

There are an Infinite Number of Normal Distributions



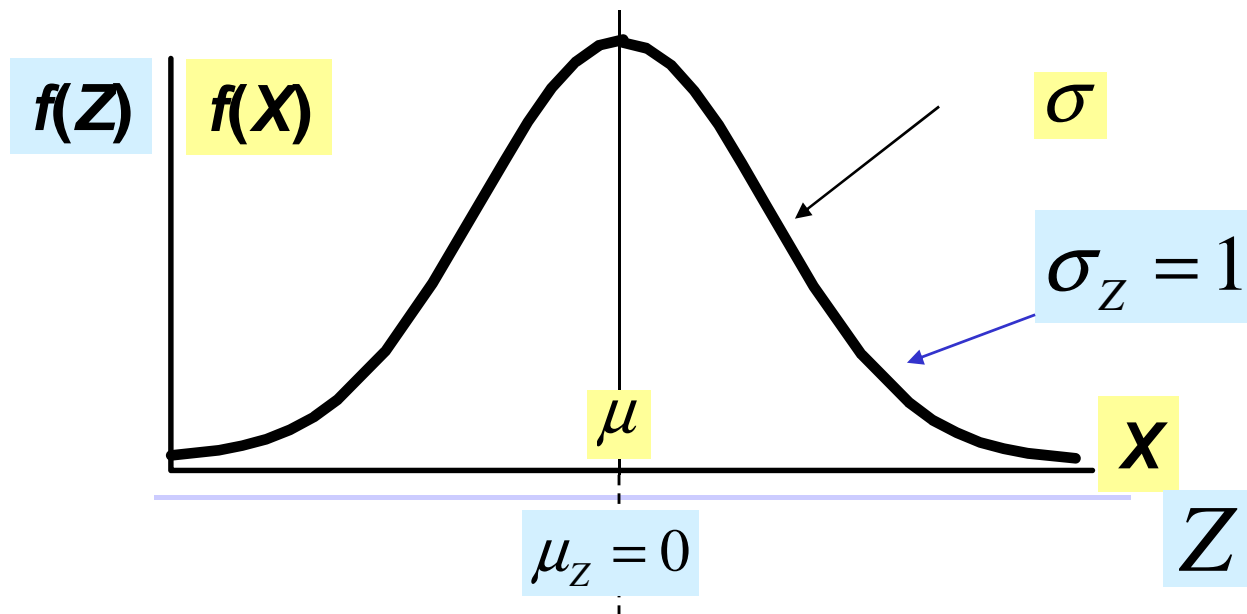
Varying the Parameters σ and μ , We Obtain Different Normal Distributions

The Standardized Normal Distribution

When X is normally distributed with a mean μ and a standard deviation σ ,

$$Z = \frac{X - \mu}{\sigma}$$

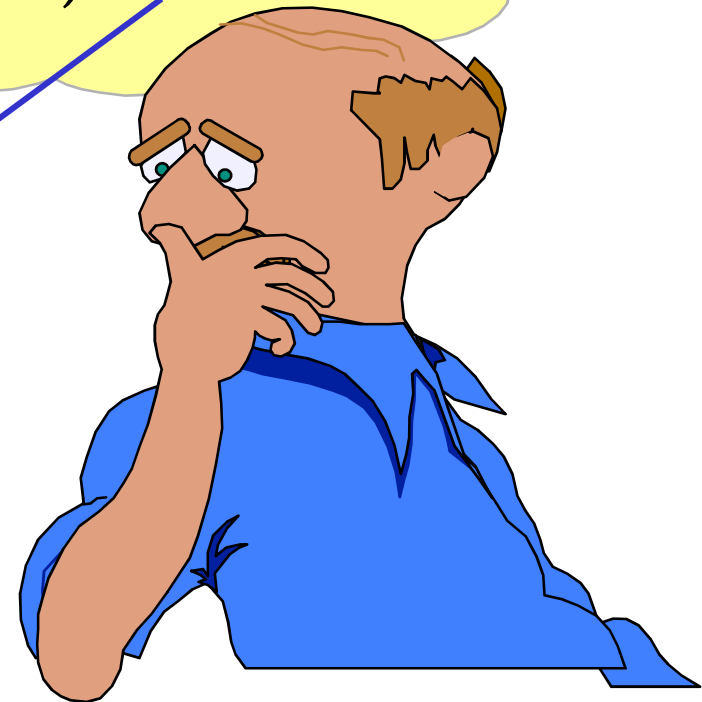
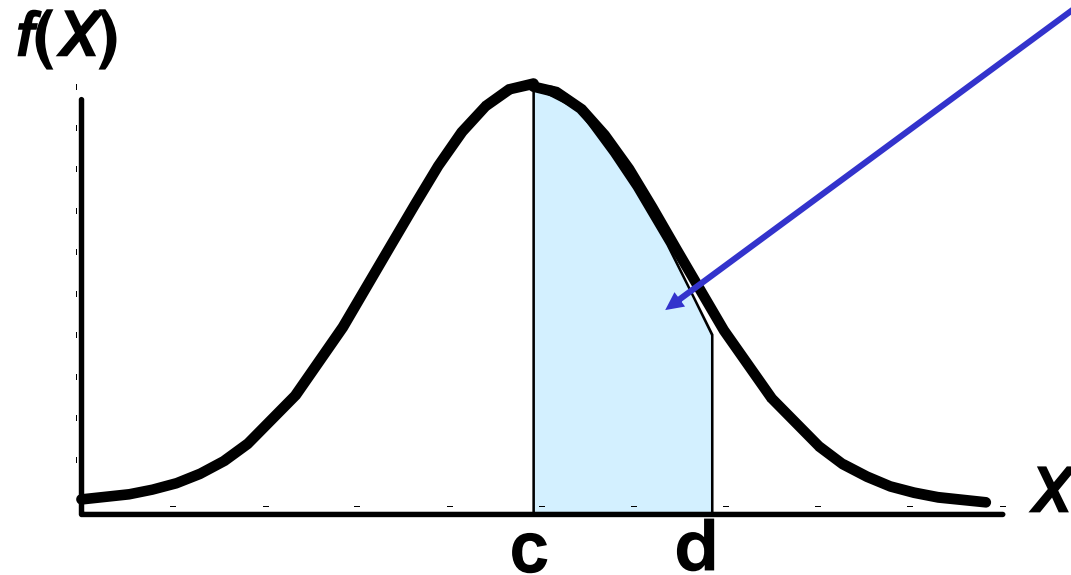
follows a standardized (normalized) normal distribution with a mean 0 and a standard deviation 1.



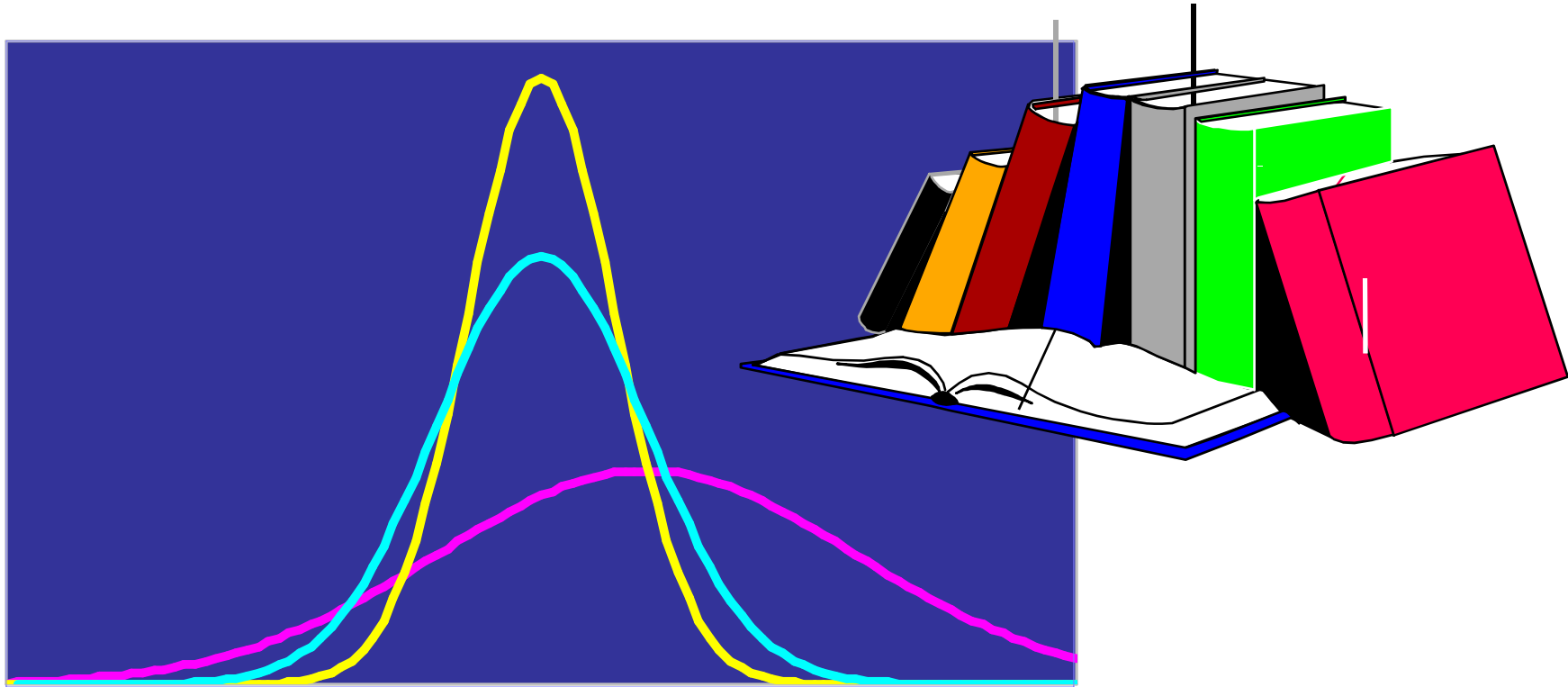
Finding Probabilities

Probability is
the area under
the curve!

$$P(c \leq X \leq d) = ?$$



Which Table to Use?



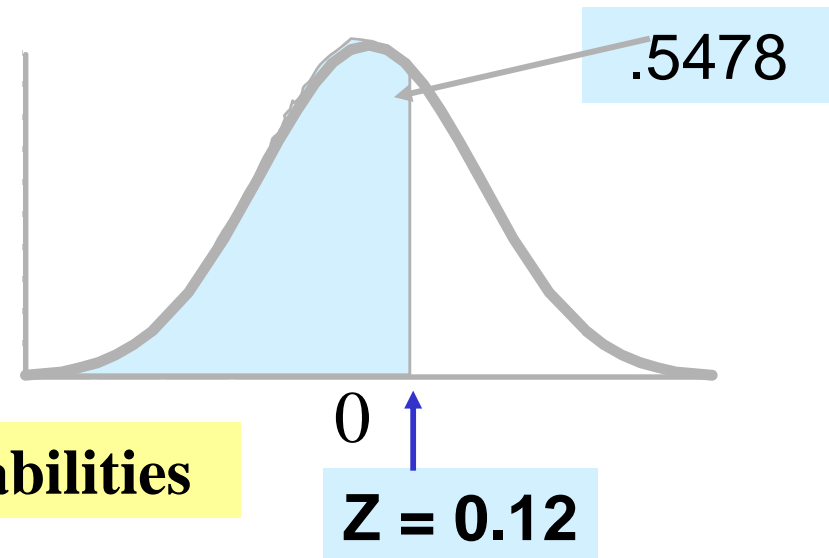
**Infinitely Many Normal Distributions
Means Infinitely Many Tables to Look Up!**

Solution: The Cumulative Standardized Normal Distribution

Cumulative Standardized Normal Distribution Table (Portion)

Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

$$\mu_z = 0 \quad \sigma_z = 1$$



Probabilities

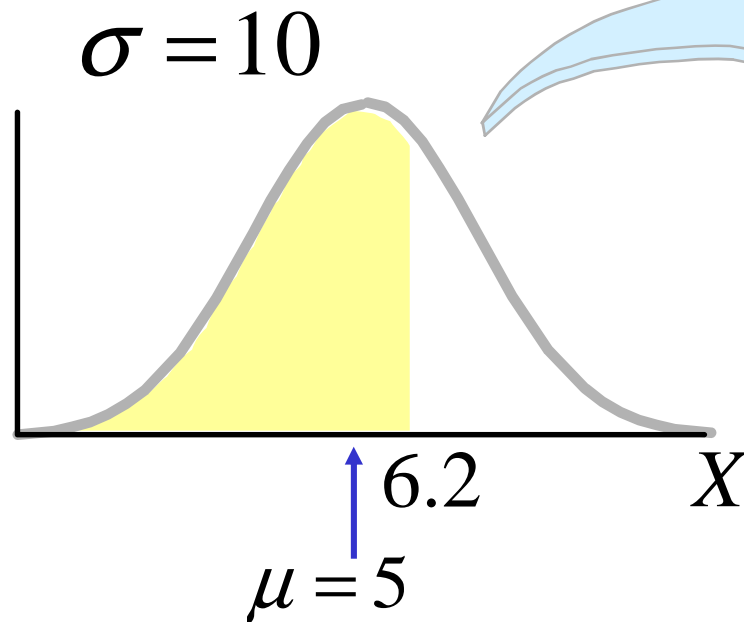
$$Z = 0.12$$

Only One Table is Needed

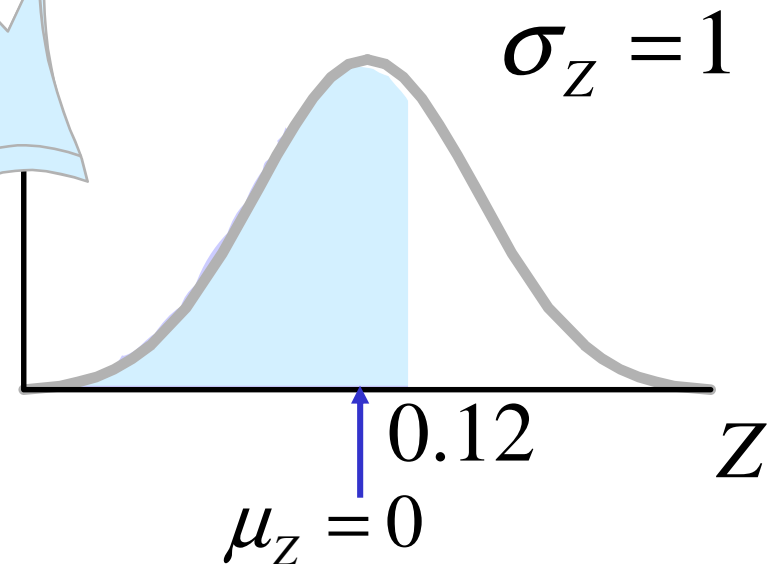
Standardizing Example

$$Z = \frac{X - \mu}{\sigma} = \frac{6.2 - 5}{10} = 0.12$$

Normal Distribution



Standardized Normal Distribution

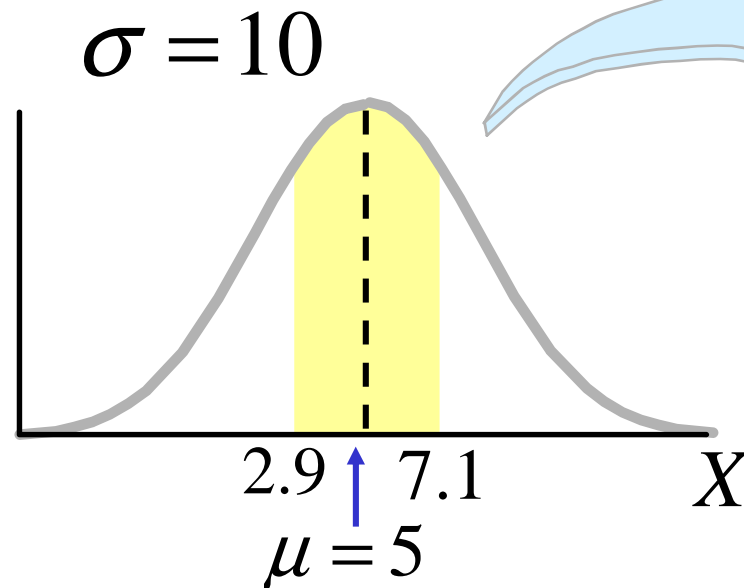


Example: $P(2.9 \leq X \leq 7.1) = .1664$

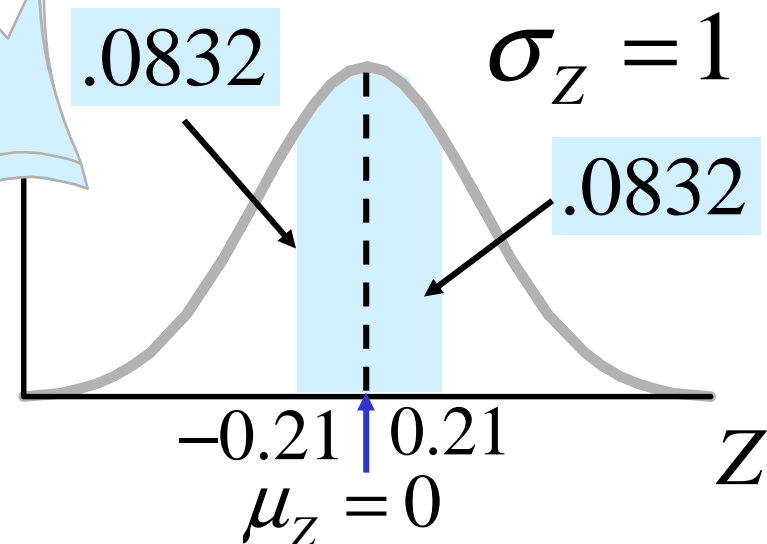
$$Z = \frac{X - \mu}{\sigma} = \frac{2.9 - 5}{10} = -.21$$

$$Z = \frac{X - \mu}{\sigma} = \frac{7.1 - 5}{10} = .21$$

Normal Distribution



Standardized Normal Distribution

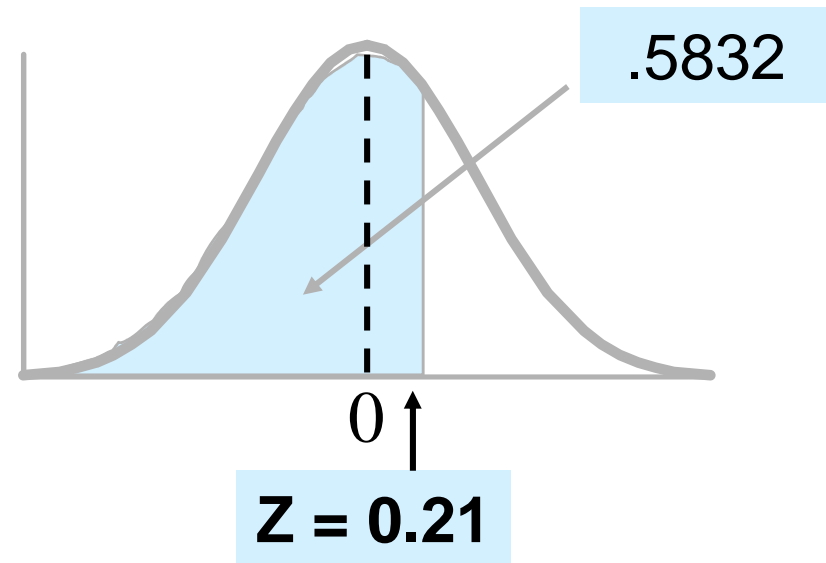


Example: $P(2.9 \leq X \leq 7.1) = .1664$

Cumulative Standardized Normal Distribution Table (Portion)

Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

$$\mu_Z = 0 \quad \sigma_Z = 1$$

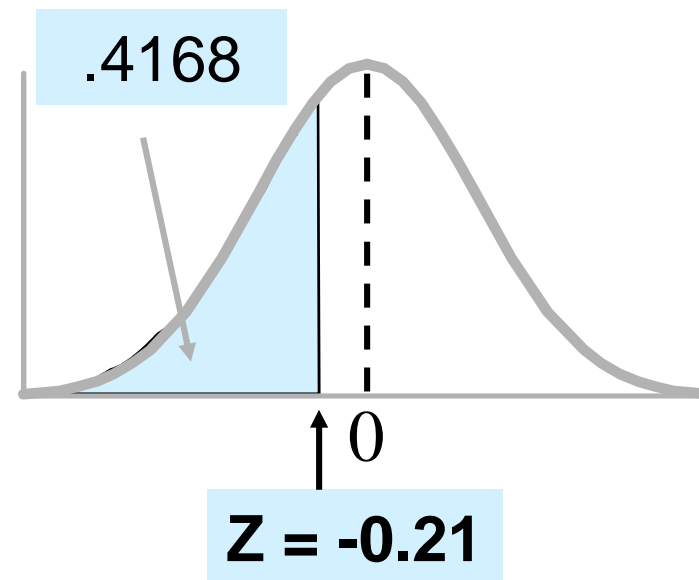


Example: $P(2.9 \leq X \leq 7.1) = .1664$

Cumulative Standardized Normal Distribution Table (Portion)

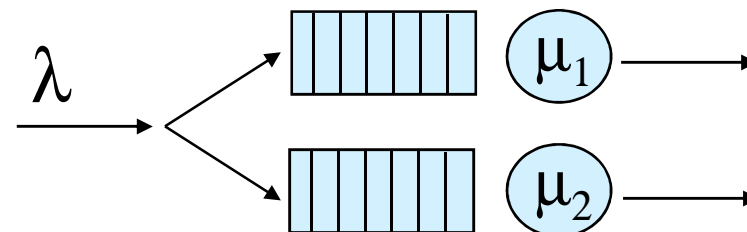
Z	.00	.01	.02
-0.3	.3821	.3783	.3745
-0.2	.4207	.4168	.4129
-0.1	.4602	.4562	.4522
0.0	.5000	.4960	.4920

$$\mu_Z = 0 \quad \sigma_Z = 1$$



Back to Simulation...

- ❑ When performing a simulation study, we are usually interested in obtaining answers to some quantitative questions
- ❑ One very common question in this respect is:
 - ❑ What is the mean value of some defined metric of the system under study?
- ❑ Example:
 - ❑ Packets arrive with an average inter-arrival time of $1/\lambda$ at a router
 - ❑ The router has two outgoing links and arriving packet joins link i with probability ϕ_i
 - ❑ The service time on link i is μ_i



(mostly compiled from [Tow04])

Gathering Performance Statistics

- Average delay at queue i:
 - Record D_{ij} : delay of customer j at queue i
 - Let N_i be # customers passing through queue i

$$\bar{T}_i = \frac{\sum_{j=1}^{N_i} D_{ij}}{N_i}$$

- Throughput at queue i, $\gamma_i = \frac{N_i}{\text{total simulated time}}$

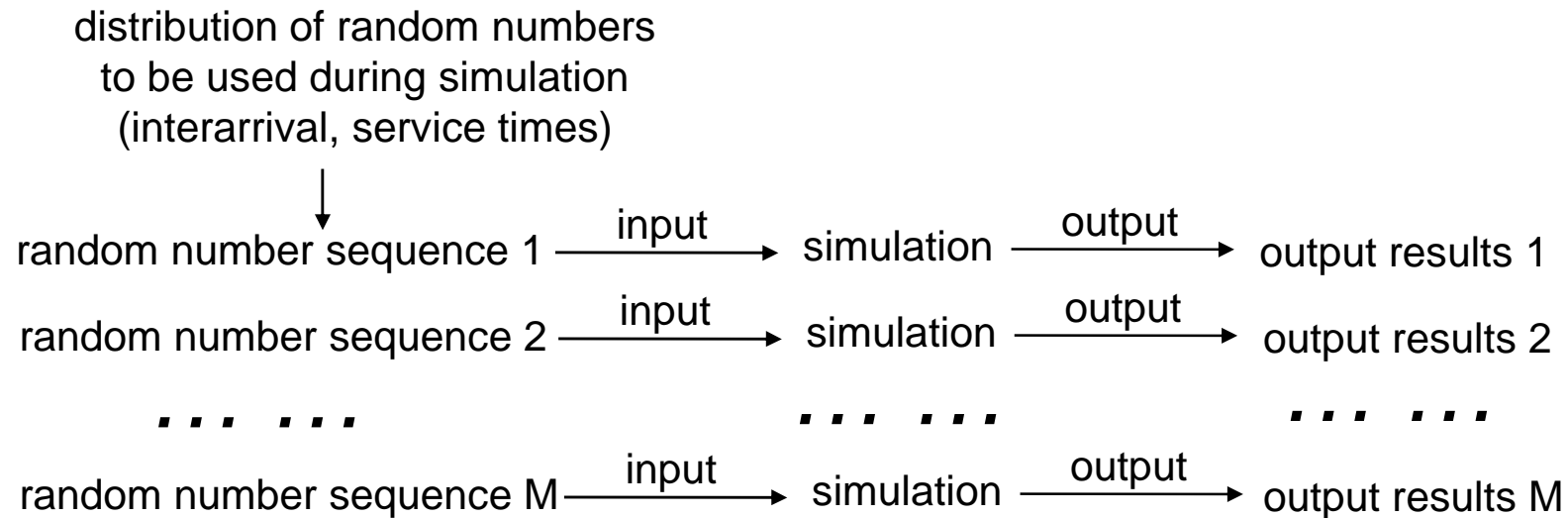
- Average queue length at i:

$$\bar{N}_i = \gamma_i \bar{T}_i$$

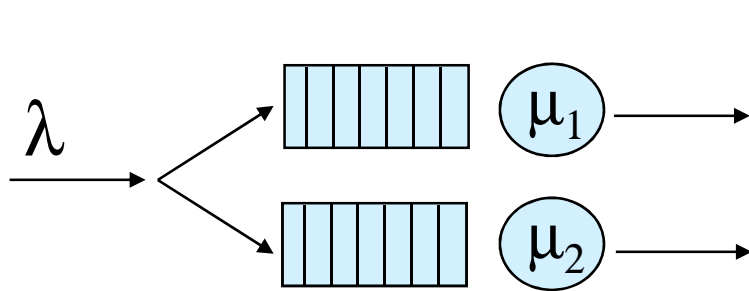
↑
Little's Law
(not treated here)

Analyzing Output Results (1)

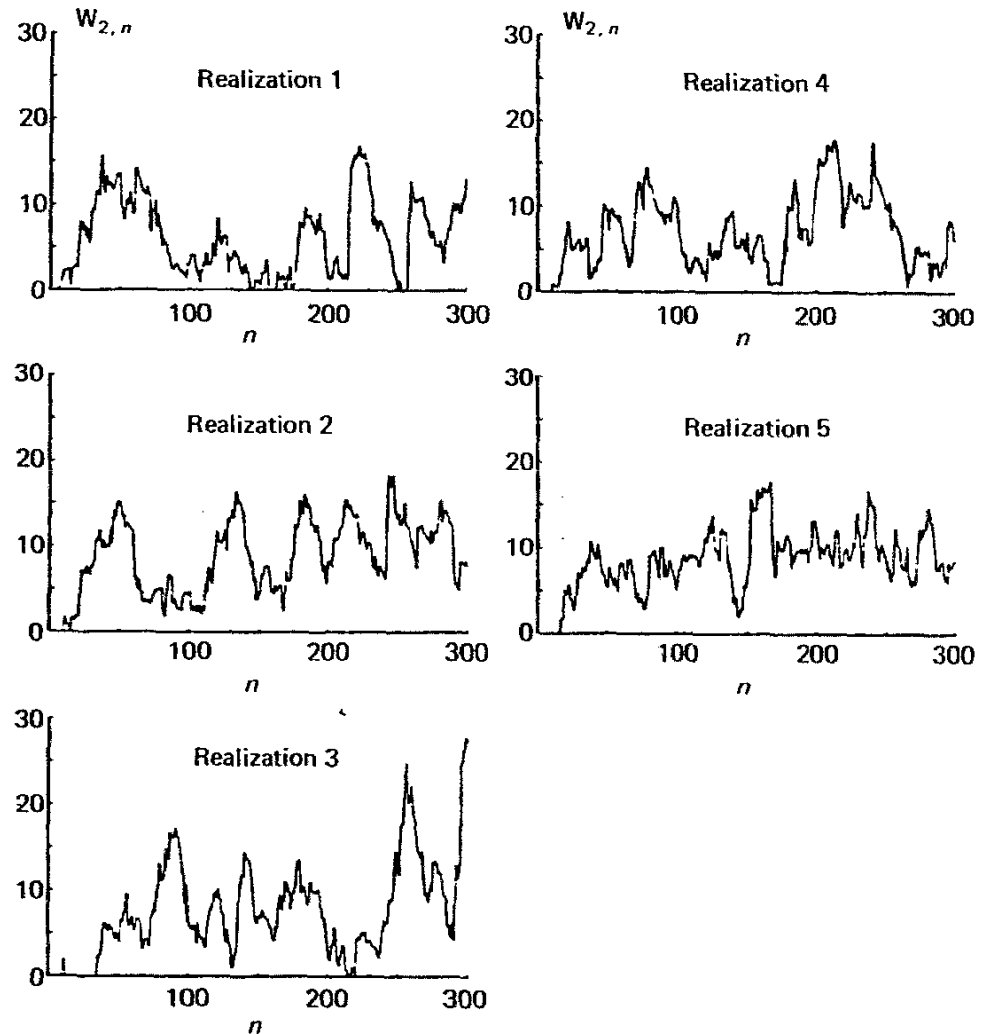
- Each time we run a simulation (using different random number streams), we will get different output results!



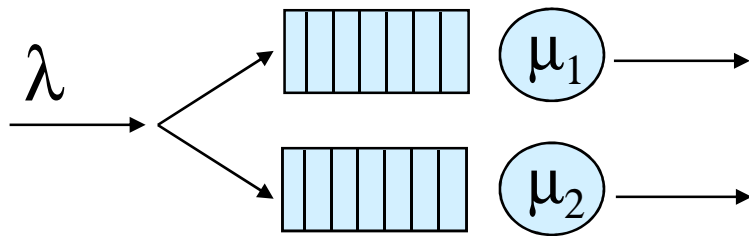
Analyzing Output Results (2)



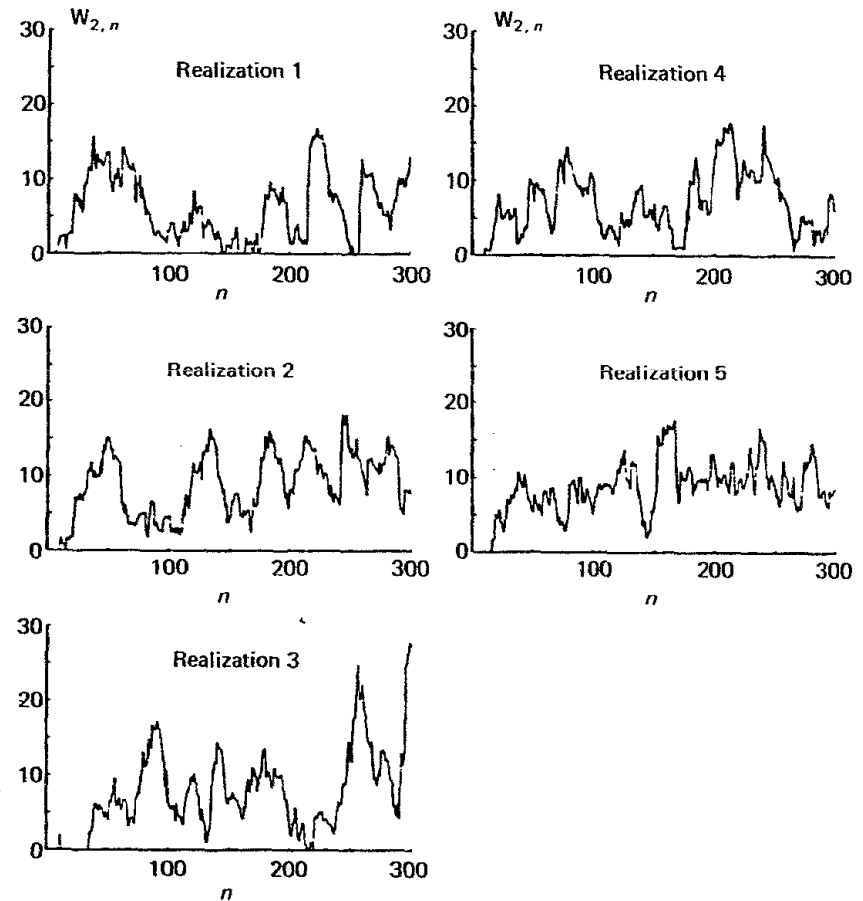
- Example: Delay experienced by each customer in queue 2
- $W_{2,n}$: delay of n th departing customer from queue 2



Analyzing Output Results (3)

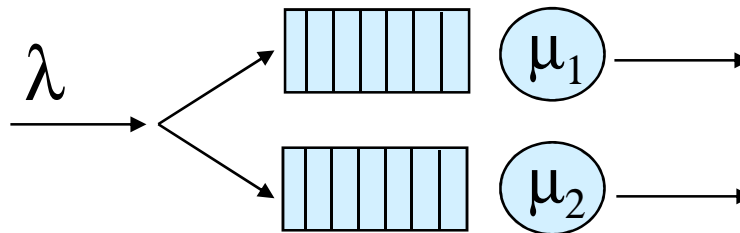


- ❑ Each run shows variation in customer delay
- ❑ One run different from next
- ❑ Statistical characterization of delay must be made
 - ❑ Expected delay of n-th customer
 - ❑ Behavior as n approaches infinity
 - ❑ Average of n customers



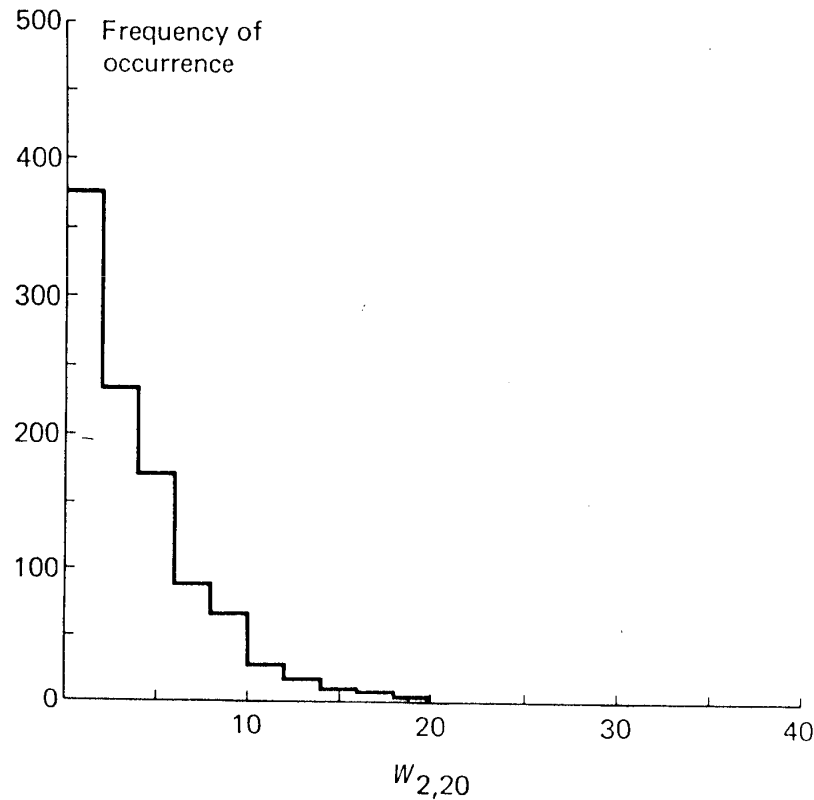
Transient Behavior

- ❑ Simulation outputs that depend on initial condition (i.e., output value changes when initial conditions change) are called *transient characteristics*
 - ❑ “Early” part of simulation
 - ❑ Example: The first customer entering a supermarket in the morning will always find an empty queue at the counter
 - ❑ Later part of simulation less dependent on initial conditions

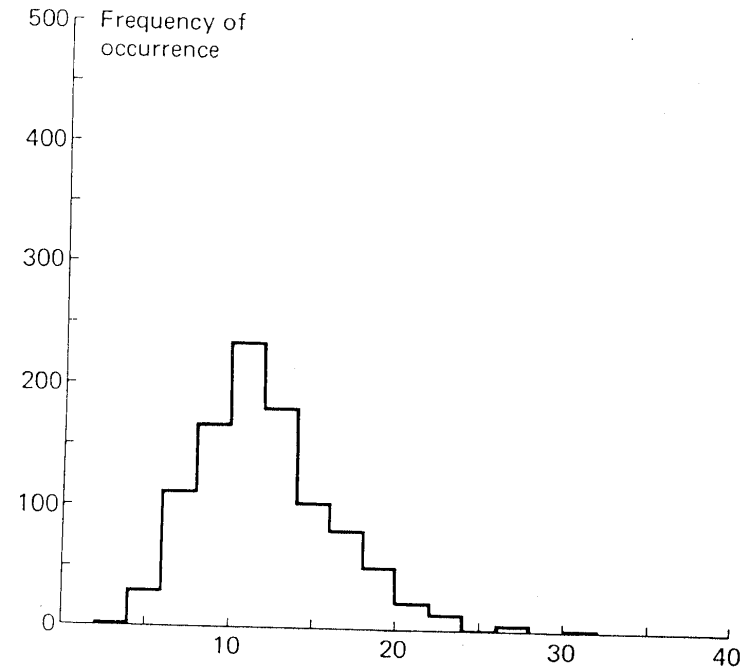


Effect of Initial Conditions

- Histogram of delay of 20th customer, given initially empty (1000 runs)

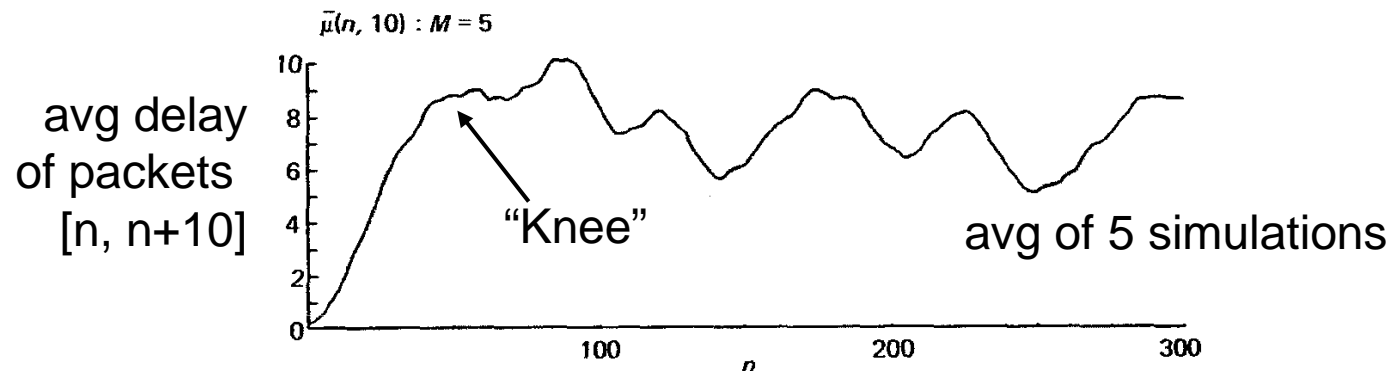


- Histogram of delay of 20th customer, given non-empty conditions



Steady state behavior

- Output results may converge to limiting “steady state” value if simulation run “long enough”



- Discard statistics gathered during transient phase, e.g., ignore first k measurements of delay at queue 2

$$\bar{T}_i = \frac{\sum_{j=k}^{N_i} D_{ij}}{N_i - k}$$

- Pick k so statistic is “approximately the same” for different random number streams and remains same as n increases

Estimating the Transient Phase (1)

- Perform m independent replications containing n observations each:

$$\{X_{i,j}, j = 1, \dots, n\}_{i=1}^m$$

- Calculate mean across all replications:

$$\bar{X}_j = \frac{1}{m} \sum_{i=1}^m X_{i,j}, \quad j = 1, \dots, n$$

- Calculate overall mean across all replications:

$$\bar{\bar{X}} = \frac{1}{n} \sum_{j=1}^n \bar{X}_j$$

- Remove first k observations when calculating overall mean:

$$\bar{\bar{X}}_k = \frac{1}{n-k} \sum_{j=k+1}^n \bar{X}_j, \quad k = 1, \dots, n-1$$

Estimating the Transient Phase (2)

- ❑ Compute and plot as a function of k : $\frac{\overline{\overline{X}}_k - \overline{\overline{X}}}{\overline{\overline{X}}}$
- ❑ Identify the position k_0 of the “knee” and discard the first k_0 observations in subsequent runs
- ❑ In order to facilitate computation of $\overline{\overline{X}}_k$, it is better to record the running sum of all prior values instead of recording the values themselves: $S_{i,k} = \sum_{j=1}^k X_{i,j}$
- ❑ Individual values can still be easily obtained by simple subtraction
- ❑ With this we get:

$$\overline{\overline{X}} = \frac{1}{n} \frac{1}{m} \sum_{i=1}^m S_{i,n} \quad \text{and} \quad \overline{\overline{X}}_k = \frac{1}{n-k} \frac{1}{m} \sum_{i=1}^m (S_{i,n} - S_{i,k})$$

Confidence Intervals (1)

- ❑ Run simulation: get estimate X_1 as estimate of performance metrics of interest
- ❑ Repeat simulation M times (each with new set of random numbers), get X_2, \dots, X_M – *all different!*
- ❑ Which of X_1, \dots, X_M is “right”?

- ❑ Intuitively, average of M samples should be “better” than choosing any one of M samples

$$\bar{X} = \frac{\sum_{j=1}^M X_j}{M}$$

How “confident” are we in \bar{X} ?

Confidence Intervals (2)

- ❑ We can not get perfect estimate of true mean, μ , with finite # samples
- ❑ Instead, we look for bounds: find c_1 and c_2 such that:

$$\text{Probability}(c_1 < \mu < c_2) = 1 - \alpha$$

$[c_1, c_2]$: *confidence interval*

$100(1-\alpha)\%$: *confidence level*

- ❑ One approach for finding c_1 , c_2 (suppose $\alpha=.1$)
 - ❑ Take k samples (e.g., k independent simulation runs)
 - ❑ Sort
 - ❑ Find largest value is smallest 5% $\rightarrow c_1$
 - ❑ Find smallest value in largest 5% $\rightarrow c_2$

Confidence Intervals – The Central Limit Theorem

- Central Limit Theorem: If samples X_1, \dots, X_M are independent and from same population (independent identically distributed, i.i.d.) with population mean μ and standard deviation σ , then

the sample mean:
$$\bar{X} = \frac{\sum_{j=1}^M X_j}{M}$$

is approximately normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{M}}$

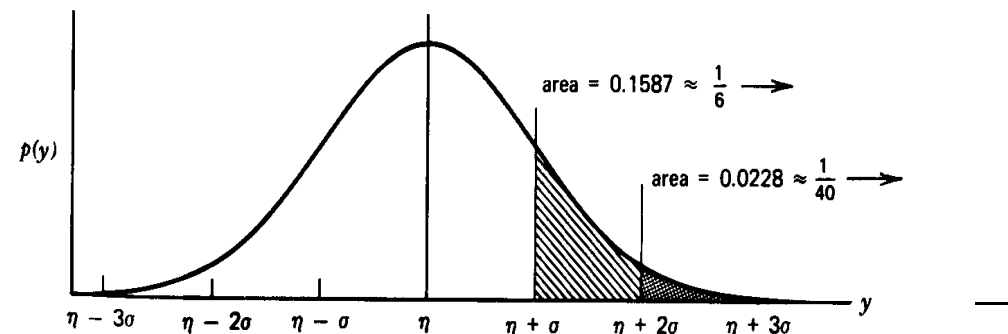


FIGURE 2.12. Tail areas of the normal distribution.

Confidence Intervals (3)

- ❑ However, we usually do not know the population's standard deviation
- ❑ So, we estimate it using the sample's (observed) standard deviation:

$$\sigma_X^2 = \frac{1}{M-1} \sum_{m=1}^M (X_m - \bar{X})^2$$

- ❑ Given \bar{X}, σ_X^2 we can now find upper and lower tails of normal distributions containing $\alpha 100\%$ of the mass

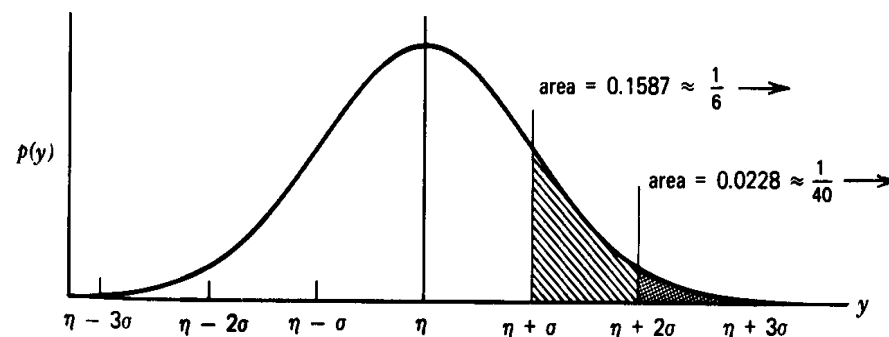


FIGURE 2.12. Tail areas of the normal distribution.

Interpretation of Confidence Intervals

- If we calculate confidence intervals as in the above recipe for $\alpha = 0.5$, 95% of the confidence intervals thus computed will contain the true (unknown) population mean

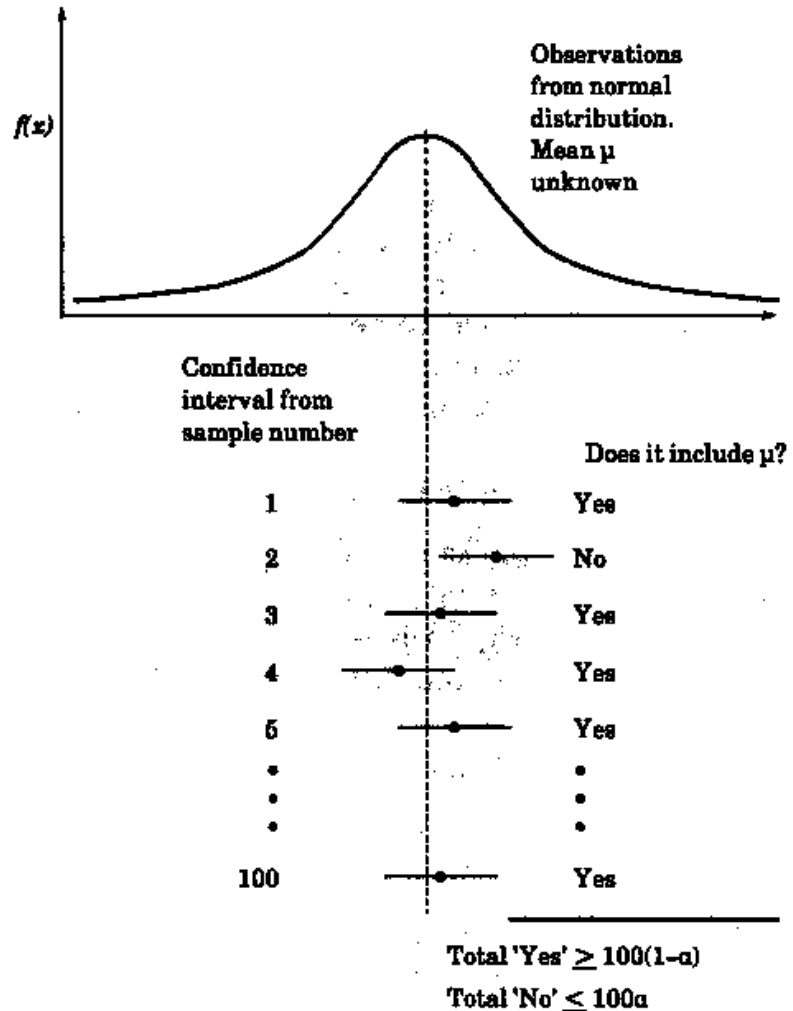


FIGURE 13.1 Meaning of a confidence interval.

Computing a Sample's Variance Simplified

- Computing the sum of squared samples leads to simpler computation:

$$\begin{aligned}
 \sigma_X^2 &= \frac{1}{M-1} \sum_{i=1}^M (X_i - \bar{X})^2 \\
 &= \frac{1}{M-1} \sum_{i=1}^M (X_i^2 - 2X_i\bar{X} + \bar{X}^2) \\
 &= \frac{1}{M-1} \left(\sum_{i=1}^M X_i^2 - 2\bar{X} \sum_{i=1}^M X_i + \bar{X}^2 \sum_{i=1}^M 1 \right) \\
 &= \frac{1}{M-1} \left(\sum_{i=1}^M X_i^2 - 2\bar{X}M\bar{X} + M\bar{X}^2 \right) \\
 &= \frac{\sum_{i=1}^M X_i^2 - M\bar{X}^2}{M-1}
 \end{aligned}$$

Confidence Intervals: The Recipe (1)

- Given samples X_1, \dots, X_M , (e.g., having repeated simulation M times), compute

$$\bar{X} = \frac{\sum_{j=1}^M X_j}{M}$$

$$\sigma_X^2 = \frac{1}{M-1} \sum_{m=1}^M (X_m - \bar{X})^2$$

95% confidence interval: $\bar{X} \pm \frac{1.96\sigma_X}{\sqrt{M}}$

Confidence Intervals: The Recipe (2)

- Why does this work?
- If the X_i have been observed running the same simulation program with different seed values for the random number generation (with enough “distance” between the seed values), then they are independent and identically distributed (i.i.d.) with some mean μ and variance σ^2

□ With
$$\bar{X}_M = \frac{1}{M} \sum_{i=1}^M X_i$$

we have:
$$E[\bar{X}_M] = E\left[\frac{1}{M} \sum_{i=1}^M X_i\right] = \frac{1}{M} E\left[\sum_{i=1}^M X_i\right]$$

$$= \frac{1}{M} \sum_{i=1}^M E[X_i] = \frac{1}{M} M\mu = \mu$$

Confidence Intervals: The Recipe (3)

□ Furthermore,
$$VAR\left[\sum_{i=1}^M X_i\right] = M \cdot VAR[X_i] = M\sigma^2$$

and
$$VAR\left[\frac{1}{M} \sum_{i=1}^M X_i\right] = \frac{1}{M^2} M\sigma^2 = \frac{\sigma^2}{M} = VAR[\bar{X}_M]$$

- Because of the central limit theorem, for large M the \bar{X}_M are normal distributed with expected value μ and standard deviation $\frac{\sigma}{\sqrt{M}}$
- In practice, this gives a good estimation already for $M > 30$

Confidence Intervals: The Recipe (4)

- ❑ Consider now the random variable $Z = \frac{\bar{X}_M - \mu}{\sigma / \sqrt{M}}$
- ❑ Z is normal distributed with mean 0 and variance 1
- ❑ We now look for $-Z_{\alpha/2}$ and $Z_{\alpha/2}$ so that

$$P\left[-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\right] = 1 - \alpha$$

- ❑ From the table of the N(0,1) normal distribution we obtain the value $Z_{\alpha/2}$ for a given α (~“the amount of confidence we are aiming at”)
- ❑ For $\alpha = 0.05$, we obtain $Z_{\alpha/2} = 1.96$

Confidence Intervals: The Recipe (5)

$$P\left[-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\right] = 1 - \alpha$$

$$\Leftrightarrow P\left[-Z_{\alpha/2} \leq \frac{\bar{X}_M - \mu}{\sigma / \sqrt{M}} \leq Z_{\alpha/2}\right] = 1 - \alpha$$

$$\Leftrightarrow P\left[-\bar{X}_M - Z_{\alpha/2} \frac{\sigma}{\sqrt{M}} \leq -\mu \leq -\bar{X}_M + Z_{\alpha/2} \frac{\sigma}{\sqrt{M}}\right] = 1 - \alpha$$

$$\Leftrightarrow P\left[\bar{X}_M - Z_{\alpha/2} \frac{\sigma}{\sqrt{M}} \leq \mu \leq \bar{X}_M + Z_{\alpha/2} \frac{\sigma}{\sqrt{M}}\right] = 1 - \alpha$$

- This leads to our formula $\bar{X} \pm \frac{1.96\sigma_X}{\sqrt{M}}$ for the confidence interval

Generating Confidence Intervals for Steady State Measures

- In order to obtain enough i.i.d. observations, we have two simple alternatives (others exist):
 - *Independent replications*: perform M independent runs with different seeds (also remember to delete k observations from transient phase):
 - However, if we want to use the central limit theorem, we should at least perform 30 independent simulation runs...
 - Alternatively, we can use a *student-t distribution* instead of the normal distribution if $M < 30$ (we do not explain why here)
 - *Batch means*: take a *single* run, delete first k observations, divide remainder into n groups and obtain \bar{X}_i for i -th group, $i = 1, \dots, n$
 - Follow procedure for independent replications
 - Complication due to non-independence of X_s
 - Potential efficiency gain due to deletion of only k observations

Additional References

- [BLK02] M. L. Berenson, D. M. Levine, T. K. Krehbiel. *Basic Business Statistics*. course slides to the ninth edition of the book. Prentice-Hall, 2002.
<http://myphlip.pearsoncmg.com/cw/mpbookhome.cfm?vbookid=462>
- [Jain91] R. Jain. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991.
- [Karl04] H. Karl. *Praxis der Simulation*. course slides, Technische Universität Berlin.
- [Rob01] T. Robinson. *Probability*. course slides, University of Bath,
<http://www.bath.ac.uk/~masar/UNIV0037/probability.ppt>
- [Tow04] Don Towsley. *Network Simulation*. slides for course “Advanced Foundations of Computer Networks”, University of Massachusetts, USA, 2004.
<http://www-net.cs.umass.edu/cs653/>
- [Var04] A. Varga. *OMNeT++: Object-Oriented Discrete Event Simulator*.
<http://www.omnetpp.org/>