CPSC 531:
System Modeling and Simulation

Carey Williamson
Department of Computer Science
University of Calgary
Fall 2017

"A person with one watch knows what time it is.
A person with two watches is never quite sure."

-Segal's Law

Simulation Output Analysis

- Purpose: estimate system performance from simulation output

- Understand:
  - Terminating and non-terminating simulations
  - Transient and steady-state behavior

- Learn about statistical data analysis:
  - Computing confidence intervals
  - Determining the number of observations required to achieve a desired confidence interval

Outline

- Measure of performance and error
- Transient and steady state
  - Types of simulations
  - Steady-state analysis
  - Initial data deletion
  - Length of simulation run
- Confidence interval
  - Estimating mean and variance
  - Confidence interval for small and large samples
  - Width of confidence interval

Outline

- Measure of performance and error
- Transient and steady state
  - Types of simulations
  - Steady-state analysis
  - Initial data deletion
  - Length of simulation run
- Confidence interval
  - Estimating mean and variance
  - Confidence interval for small and large samples
  - Width of confidence interval

Stochastic Nature of Simulation Output

- Output data are random variables, because the input variables are stochastic, and model is basically an input-output transformation

- A queueing example: Banff park entry booth
  - Arrival rate ~ Poisson arrival process (λ per minute)
  - Service time $\sim Exponential(\mu = 1.5)$ minutes
  - System performance: long-run average queue length
  - Question 1: does simulation model result agree with M/M/1 model?
  - Question 2: is queueing better/same/worse for HyperExp() service?
  - Question 3: how much better would it be with two servers?
  - Suppose we run the simulation 3 times, i.e., 3 replications
    - Each replication is for a total of 5,000 minutes
    - Divide each replication into *5* equal subintervals (i.e., batches) of *1000* minutes
    - $Y_{i,j}$: Average number of cars in queue from time $(j-1) \times 1000$ to $j \times 1000$ in replication i
    - Y_i: Average number of cars in queue in replication i

Stochastic Nature of Simulation Output

- queueing example (cont'd):
  - Batched average queue length for 3 independent replications:

Batching Interval (minutes)	Batch, j	Replication 1, Y_{1j}	Replication 2, Y_{2j}	Replication 3, Y_{3j}
[0, 1000)	1	3.61	2.91	7.67
[1000, 2000)	2	3.21	9.00	19.53
[2000, 3000)	3	2.18	16.15	20.36
[3000, 4000)	4	6.92	24.53	8.11
[4000, 5000)	5	2.82	25.19	12.62
[0, 5000)		3.75	15.56	13.66

  - Inherent variability in stochastic simulation both within a single replication and across different replications
  - The average across 3 replications, i.e., $\mathrm{Y}_1, \mathrm{Y}_2, \mathrm{Y}_3$, can be regarded as independent observations, but averages within a replication, e.g., $Y_{11}, Y_{12}, Y_{13}, Y_{14}, Y_{15}$, are not.

Measure of Performance

- Consider estimating a performance parameter Θ
  - The true value of Θ is unknown
  - Can only observe simulation output
  - Estimate Θ using independent observations obtained from independent simulation runs (i.e., replications)

- $\widehat{\Theta}$: estimation of Θ
  - Is unbiased if: $\mathrm{E}[\widehat{\Theta}] = \Theta$ ⇐ **Desired**
  - Is biased if: $\mathrm{E}[\widehat{\Theta}] \neq \Theta$
  - Estimation bias = $\mathrm{E}[\widehat{\Theta}] - \Theta$

Measure of Error

- Confidence Interval (CI):
  - We cannot know for certain how far $\widehat{\Theta}$ is from Θ
  - CI attempts to bound the estimation error $|\Theta - \widehat{\Theta}|$
  - The more replications we make, the lower the error in $\widehat{\Theta}$

- Example: queueing system
  - Y: long-run average queue length
  - Y_i: average queue length in simulation run i
  - Define estimator $\widehat{\mathrm{Y}} = \overline{\mathrm{Y}}$, i.e.,

$$\widehat{\mathrm{Y}} = \frac{1}{8}(\mathrm{Y}_1 + \cdots + \mathrm{Y}_8) = 14.814$$

  - Can calculate a 95% confidence interval for Y such that:

$$\mathbb{P}\left(\left|\mathrm{Y} - \widehat{\mathrm{Y}}\right| \leq \epsilon\right) = 0.95$$

  - For instance: $11.541 \leq \mathrm{Y} \leq 18.087$

i	Y_i
1	15.028
2	13.385
3	18.891
4	10.559
5	8.866
6	15.883
7	18.598
8	17.302

Outline

- Measure of performance and error
- Transient and steady state
  - Types of simulations
  - Steady-state analysis
  - Initial data deletion
  - Length of simulation run
- Confidence interval
  - Estimating mean and variance
  - Confidence interval for small and large samples
  - Width of confidence interval

1. Terminating simulation:
   there is a natural event that specifies the length of the simulation:
   - Runs for some duration of time T_E, where E is a specified event that stops the simulation
   - Starts at time 0 under well-specified initial conditions
   - Ends at specified stopping time T_E

- Example: Simulating banking operations over "one day"
  - Opens at 8:30 am (time 0) with no customers present and 8 of the 11 tellers working (initial conditions), and closes at 4:30 pm (Time $T_E = 480$ minutes)

Types of Simulations

2. Non-terminating simulation:

No natural event specifying length of the simulation

- Runs continuously, for a very long period of time
- Initial conditions defined by the performance analyst
- Runs for some analyst-specified period of time T_E
- Of interest are *transient* and *steady state* behavior

- Example: Simulating banking operations to compute the “long-run” mean response time of customers

UNIVERSITY OF CALGARY

Types of Simulations

- Whether a simulation is considered to be terminating or non-terminating depends on both
  - The objectives of the simulation study, and
  - The nature of the system.

- Similar statistical techniques applied to both types of simulations to estimate performance and error

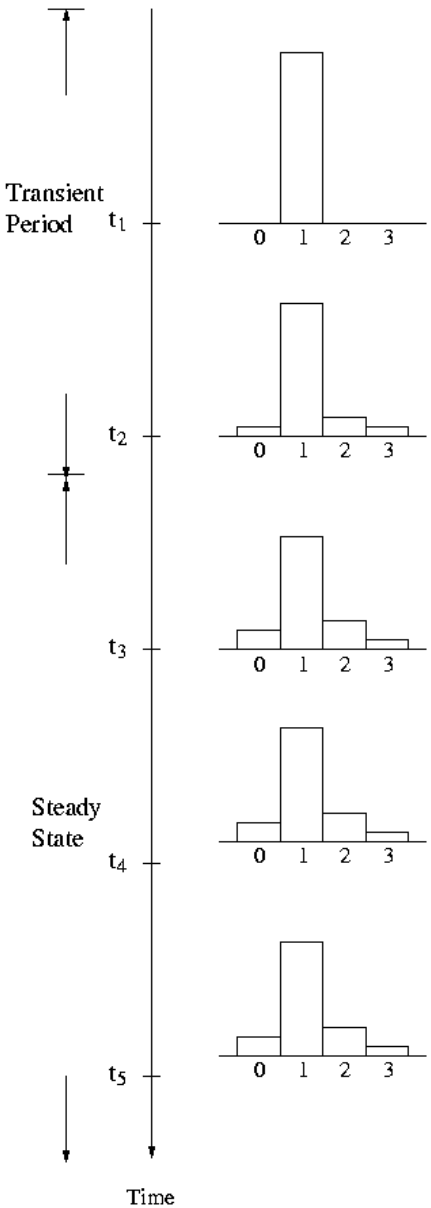
- For non-terminating simulations:
  - Transient and steady-state behavior are different
  - Generally, steady-state performance is of interest

Transient and Steady-State Behavior

Consider a queueing system

- Define $P(n,t) = \mathbb{P}(n \text{ in system at time } t)$
  - Depends on the initial conditions
  - Depends on time t
- Steady state behavior
  - System behavior over long-run: $P(n) = \lim_{t\to\infty} P(n,t)$
  - Independent of the initial conditions
  - Independent of time

Transient and Steady-State Behavior



Steady-State Output Analysis

- General approach based on independent replications:
  - Choose the initial conditions
  - Determine the length of simulation run
  - Run the simulation and collect data

- Problem: steady-state results are affected by using artificial and potentially unrealistic initial conditions

- Solutions:
  1. Intelligent initialization
  2. Simulation warmup (initial data deletion)

Intelligent Initialization

- Initialize the simulation in a state that is more representative of long-run conditions

- If the system exists, collect data on it and use these data to specify typical initial conditions

- If the system can be simplified enough to make it mathematically solvable (e.g. queueing models), then solve the simplified model to find long-run expected or most likely conditions, and use that to initialize the simulation

Simulation Warmup: Initial Data Deletion

- Divide each simulation into two phases:
  - Initialization phase, from time 0 to time T_0
  - Data-collection phase, from T_0 to stopping time $T_0 + T_E$

- Important to do a thorough job of investigating the initial-condition bias:
  - Bias is not affected by the number of replications, rather, it is affected only by deleting more data (i.e., increasing T_0) or extending the length of each run (i.e. increasing T_E)

- How to determine T_0 and T_E?

Simulation Warmup: Initial Data Deletion

- How to determine T_0?
  - After T_0, system should be more nearly representative of steady-state behavior
  - System has reached steady state: the probability distribution of the system state is close to the steady-state probability distribution
  - No widely accepted, objective and proven technique to guide how much data to delete to reduce initialization bias to a negligible level
  - Heuristics such as plotting the moving averages can be used

- How to implement data deletion in DES?
  - At initialization, schedule a `reset` event at `clock + ` T_0
  - `reset` event routine: reset all statistical counters (for data collection) to their initial values
  - At the end of simulation, statistical counters contain data collected after the transient period

Length of Simulation Run

- How to determine T_E?
  - Too short: results may not be reliable
  - Too long: wasteful of resources

- Method to determine length of run
  - Perform independent replications
  - For each replication, perform initial data deletion
  - Select length of run and number of replications such that the confidence intervals for the performance measures of interest narrow to the desired widths

Outline

- Measure of performance and error
- Transient and steady state
  - Types of simulations
  - Steady-state analysis
  - Initial data deletion
  - Length of simulation run
- Confidence interval
  - Estimating mean and variance
  - Confidence interval for small and large samples
  - Width of confidence interval

Terminology

- Observation: a single value of a performance measure from an experiment
  Example: mean response time of a web server

- Sample: the set of observations of a performance measure from an experiment

Sample Versus Population

- Generate several million random numbers with a given distribution and draw a sample of m observations
- Sample mean $\neq$ population mean
- In discrete-event simulation, population characteristics such as mean and variance are unknown
  - Need to estimate them using simulation output

Estimating Mean

- Consider a sample of m observations, denoted by $Y_1, Y_2, \ldots, Y_m$
- Example: Y_i is the mean response time of a web server in the i-th experiment
- Sample mean:

$$\bar{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i$$

  - Sample mean $\bar{Y}$ is an unbiased estimator for the unknown population mean

Estimating Variance

- Sample variance:

$$s^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \bar{Y})^2$$

  – s^2 is an unbiased estimator for the unknown population variance

- The divisor for s^2 is $m-1$ and not m
  – This is because only $m-1$ of the m differences are independent
  – Given $m-1$ differences, m-th difference can be computed since the sum of all m differences must be zero
  – The number of independent terms in a sum is also called its *degrees of freedom*

Confidence Interval for Mean

Consider a simulation study

- Y: random variable denoting the performance measure corresponding to the simulation output
  - Example: average wait time of customers in a bank

- Problem: Y varies across different simulation runs
  - Consider m simulation runs
  - Y_i: simulation output in simulation run i
  - Generally, $Y_1 \neq Y_2 \neq \cdots \neq Y_m$
- Solutions:
  1. Characterize distribution of Y (e.g., CDF)
  2. Characterize statistics of Y (e.g., mean and variance)

Confidence Interval for Mean

Consider a simulation study

- Y: random variable denoting the performance measure corresponding to the simulation output
  - Example: average wait time of customers in a bank

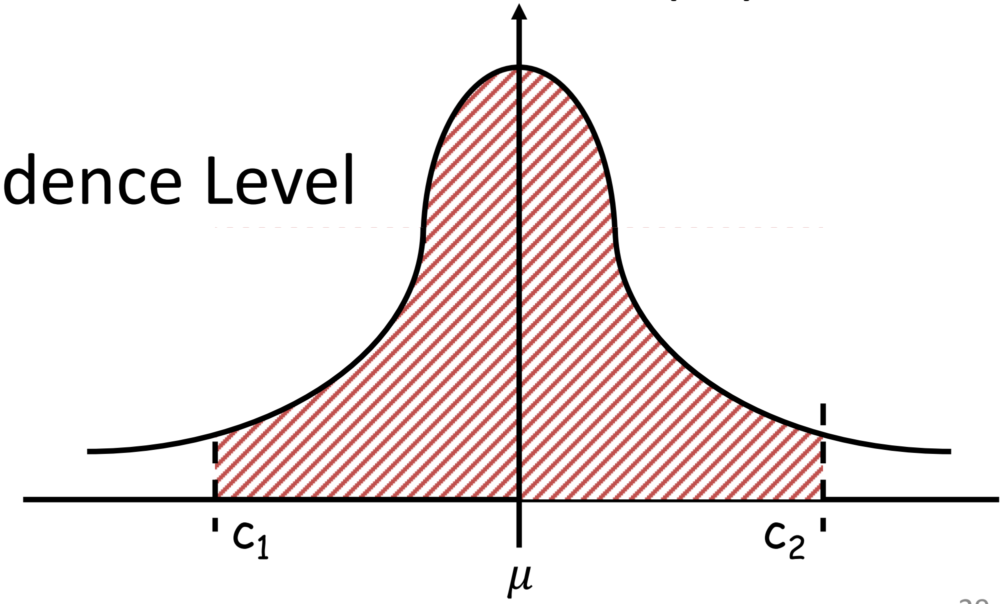
- Objective is to characterize the unknown mean $\mu = E[Y]$
  - Algorithm:
    - Make m independent simulation runs to obtain m observations $Y_1, Y_2, \dots, Y_m$
    - Sample mean $\bar{Y}$ is an unbiased estimator for μ
  - Question: How far is $\bar{Y}$ from μ?
    - Determine confidence interval for mean

- Determine bounds c_1 and c_2 such that:

  $\mathbb{P}(c_1 \leq \mu \leq c_2) = 1 - \alpha$

- $[c_1, c_2]$: $(1-\alpha)100\%$ Confidence Interval (CI)

- $(1-\alpha)100\%$: Confidence Level



Determining Confidence Interval

- $\bar{Y}$ is a random variable:

$$\bar{Y} = \frac{Y_1 + Y_2 + \cdots + Y_m}{m}$$

where the $Y_i's$ are IID with the same distribution as $Y \sim N(\mu, \sigma^2)$

- We have:

$$E[\bar{Y}] = \frac{E[Y_1] + \cdots + E[Y_m]}{m} = \frac{m\,\mu}{m} = \mu$$

$$V[\bar{Y}] = \frac{V[Y_1] + \cdots + V[Y_m]}{m^2} = \frac{m\,\sigma^2}{m^2} = \frac{\sigma^2}{m}$$

- μ and σ^2 unknown but can be estimated by $\bar{Y}$ and s^2

Determining Confidence Interval

- Define the normalized random variable X as

$$X = \frac{\bar{Y} - \mu}{s/\sqrt{m}}$$

- Theorem: The distribution of X is independent of unobservable parameter μ
  - For large m: X follows a standard normal distribution
  - For small m: X follows a Student's t-distribution

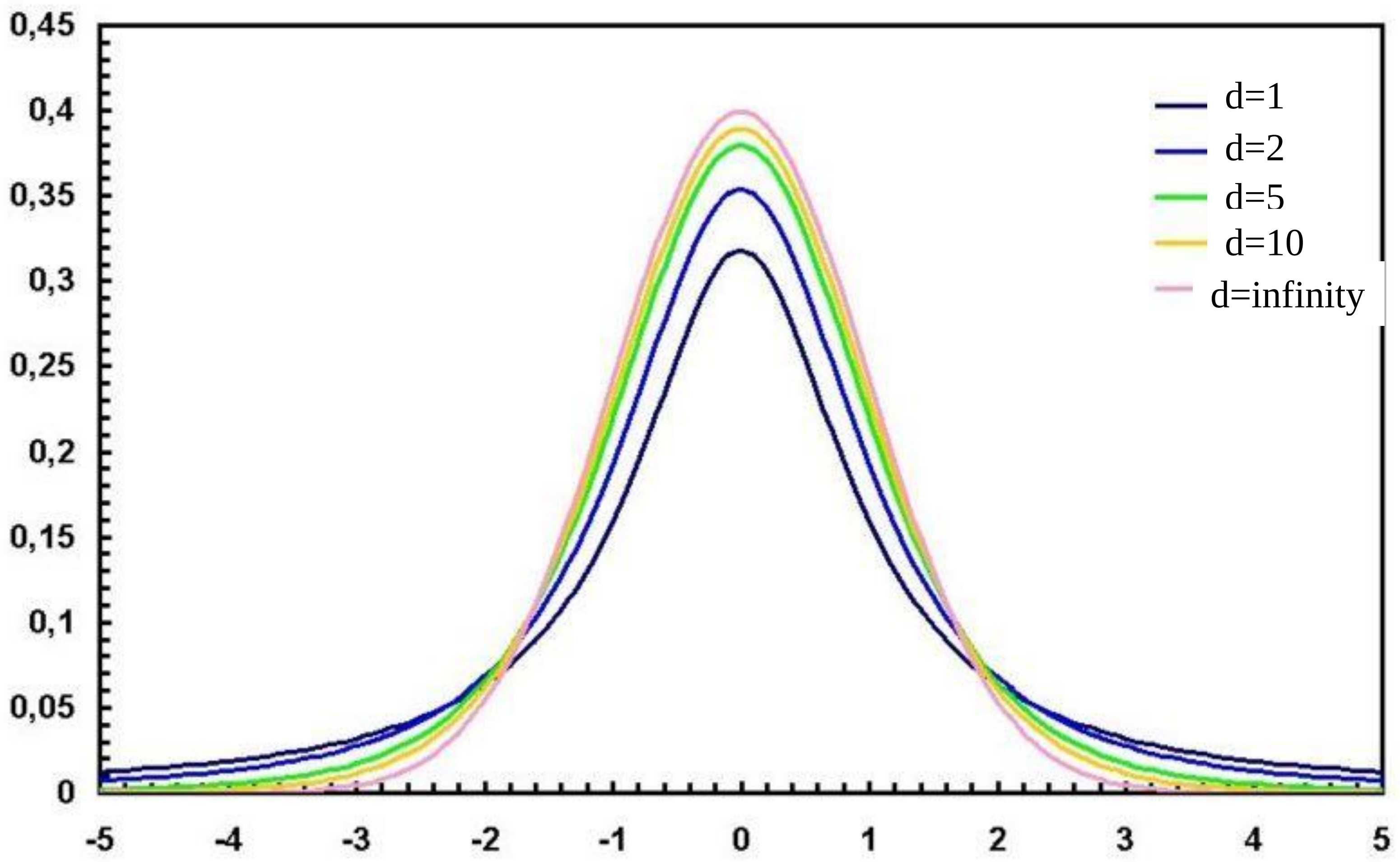
Confidence Interval for Small Samples

- Define the normalized random variable T as

$$T = \frac{\bar{Y} - \mu}{s/\sqrt{m}}$$

- T has a standard Student's t-distribution with $d = m - 1$ degrees of freedom
  - It describes the distribution of the mean of a sample of m observations
  - Symmetric distribution
  - $E[T] = 0$, and $V[T] = \frac{d}{d-2}$ for $d \geq 2$
  - As $m \to \infty$, we have $T \sim N(0, 1)$

Student's t-Distribution

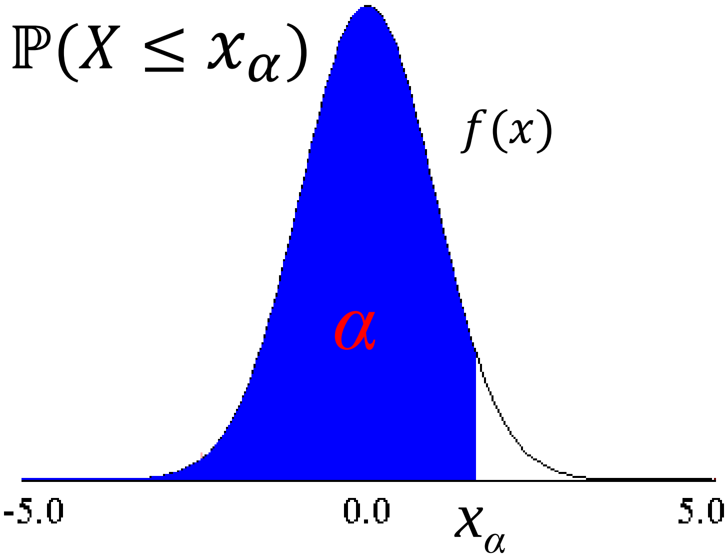


Probability Density Function (PDF)

Determining Confidence Interval

- **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:

$$\alpha = F(x_\alpha) = \mathbb{P}(X \le x_\alpha)$$

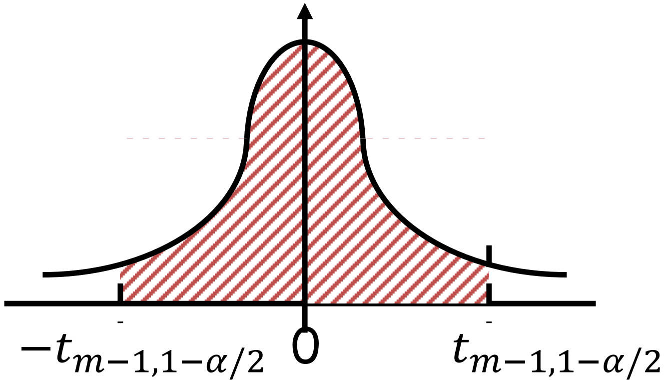


- Example: X has standard normal distribution
  - 95-percentile = 0.95-quantile = 1.6449
  - 25-percentile = 0.25-quantile = - 0.6745

Determining Confidence Interval

- Define $c = t_{m-1,1-\alpha/2}$ as:
  $(1-\alpha/2)$-quantile of T with $d = m-1$ degrees of freedom:

$$\mathbb{P}\left(T \leq t_{m-1,1-\alpha/2}\right) = 1 - \alpha/2$$



- Note that $t_{m-1,1-\alpha/2}$ does not depend on the value of the unobservable population mean μ

Determining Confidence Interval

- Therefore

$$\mathbb{P}(-c \leq T \leq c) = 1 - \alpha$$

- Which means

$$\mathbb{P}\left(-c \leq \frac{\bar{Y} - \mu}{s/\sqrt{m}} \leq c\right) = 1 - \alpha$$

$$\Rightarrow \mathbb{P}\left(\bar{Y} - c\frac{s}{\sqrt{m}} \leq \mu \leq \bar{Y} + c\frac{s}{\sqrt{m}}\right) = 1 - \alpha$$

- *(1-α)100%* confidence interval of μ is given by

$$\left[\bar{Y} - c\frac{s}{\sqrt{m}}, \qquad \bar{Y} + c\frac{s}{\sqrt{m}}\right]$$

UNIVERSITY OF CALGARY

Example

- Sample:
  -0.04, -0.19, 0.14, -0.09, -0.14, 0.19, 0.04, and 0.09

- Mean = 0, Sample standard deviation = 0.138

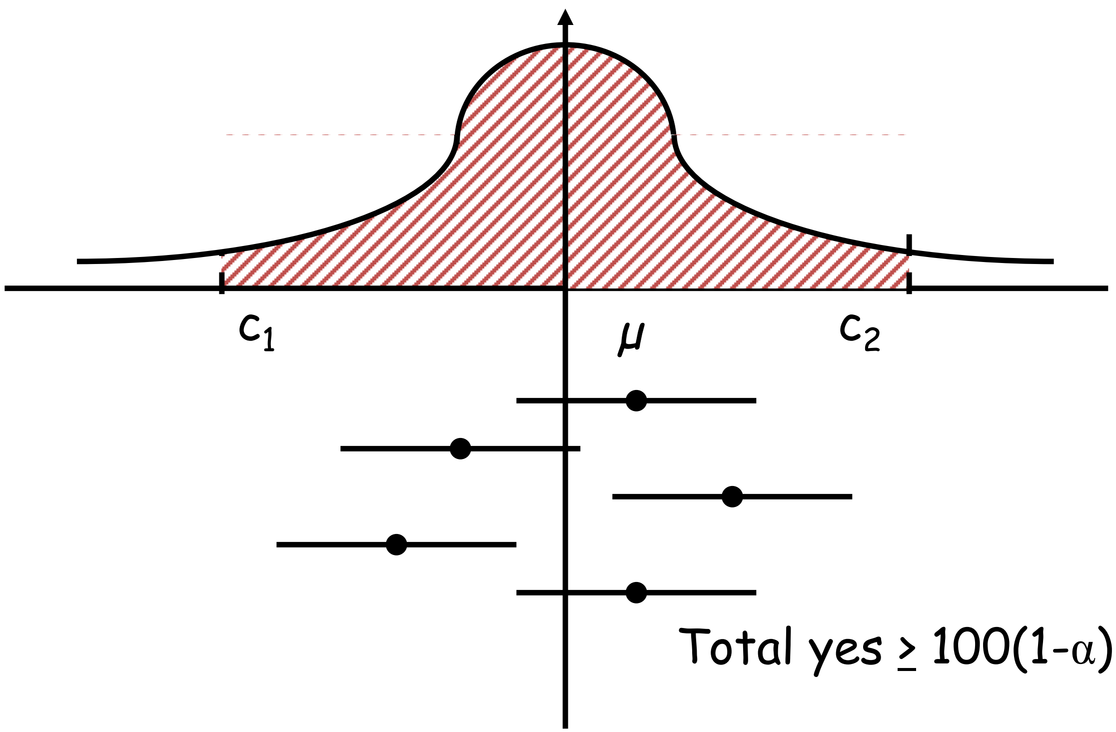
- For 90% confidence interval: $t_{7,\ 0.95}$ = 1.895

- Confidence interval for the mean

$$0 \mp 1.895 \times \frac{0.138}{\sqrt{8}} = 0 \mp 0.093 = (-0.093,\ 0.093)$$

Confidence Interval: Meaning

- If we take 100 samples and construct confidence interval for each sample, the interval would include the population mean in 90 cases.



Example: Assignment 2

- 10 replications of Banff park entry gate simulation
- Warmup: 10,000 minutes
- Number of cars: 60,000

λ	$1/\mu$	ρ	Mean Q	Std Dev
0.5	1.5	0.75	3.019	0.109
0.55	1.5	0.825	4.715	0.174
0.60	1.5	0.90	9.042	0.980
0.65	1.5	0.975	39.876	12.76

- See graph online for 90% confidence intervals

Example: Assignment 2 (cont'd)

- 10 batches from Banff park entry gate simulation
- Warmup: 0 minutes
- Number of cars: 500,000

λ	1/μ	ρ	Mean Q	Std Dev
0.5	1.5	0.75	2.997	0.088
0.55	1.5	0.825	4.813	0.206
0.60	1.5	0.90	9.033	0.608
0.65	1.5	0.975	42.22	14.72

Confidence Interval for Large Samples

- Define normalized random variable Z as

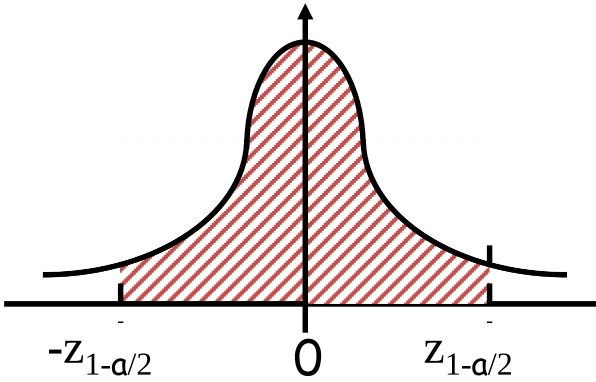
$$Z = \frac{\bar{Y} - \mu}{s/\sqrt{m}}$$

where s is the sample standard deviation

- From Central Limit Theorem: Z has standard normal distribution for large m

- (1-α)100% confidence interval for μ:
  $z_{1-\alpha/2}$ = (1-α/2)-quantile of $N(0,1)$

$$\left[\overline{Y} - z_{1-\alpha/2} \frac{s}{\sqrt{m}},\ \overline{Y} + z_{1-\alpha/2} \frac{s}{\sqrt{m}} \right]$$



Example

- $\bar{Y}$= 3.90, s = 0.95, and m = 32
- A 90% confidence interval for the mean
  = $3.90 \mp (1.645)(0.95)/\sqrt{32} = (3.62,\ 4.17)$

- We can state with 90% confidence that the population mean is between 3.62 and 4.17 The

$$\text{A 95\% confidence interval for the mean} = 3.90 \mp (1.960)(0.95)/\sqrt{32} = (3.57,\ 4.23)$$

$$\text{A 99\% confidence interval for the mean} = 3.90 \mp (2.576)(0.95)/\sqrt{32} = (3.46,\ 4.33)$$

Width of Confidence Interval

- Width of the confidence interval is

$$2 \cdot t_{m-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\sqrt{m}}$$

- Width can be reduced by
  - Using a larger m (i.e., more simulation runs)
  - Using a smaller s (i.e., longer simulation runs)

Number of Simulation Runs

- Suppose the desired width of the confidence interval is δ, and m replications have been made but the desired width is not met:
  - Total number of replications required can be estimated by

$$m^* = \left(2 \cdot t_{m-1,1-\frac{\alpha}{2}} \cdot \frac{s}{\delta}\right)^2$$

  - Number of additional replications required = $m^* - m$

Length of Simulation Runs

- An alternative to increasing m is to increase total run length $T_0 + T_E$ for each replication

- Approach: for $\beta \geq 1$
  - Increase run length from $(T_0 + T_E)$ to $\beta(T_0 + T_E)$, and
  - Delete additional amount of data, from time 0 to time βT_0