ECE/CS 541
Computer System Analysis:
Confidence Intervals

Mohammad A. Noureddine
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

Fall 2018

Announcements and Reminders

- Project presentations on December 15
  - Will try to start at 5:00 pm to finish early

- Homework 4 is out and tomorrow!

- Submit papers on the 17th via EasyChair
  - Will send the link soon
  - You will get 3 *anonymous* reviews
    - I wonder who the reviewers are!

- **ICES forms!!!**
  - Please show and fill up the ICES forms
  - Get 1 point of the participation credits

Outline for the next week

- **Today**
  - Output Analysis
  - Course wrap up
  - ICES forms

Learning Objectives

- Or what is this course about?

- At the start of the semester, you should have
  – Basic programming skills (C++, Python, etc.)
  – Basic understanding of probability theory (ECE313 or equivalent)

- At the end of the semester, you should be able to
  – Understand different system modeling approaches
    - Combinatorial methods, state-space methods, etc.
  – Understand different model analysis methods
    - Analytic/numeric methods, simulation
  – Understand the basics of discrete event simulation
  – **Design simulation experiments and analyze their results**
  – **Gain hands-on experience with different modeling and analysis tools**

Today's Lecture

- Confidence Intervals

- Course Wrap-up

Recall: Types of Simulation

- We distinguished between two types of simulation approaches: terminating and non-terminating

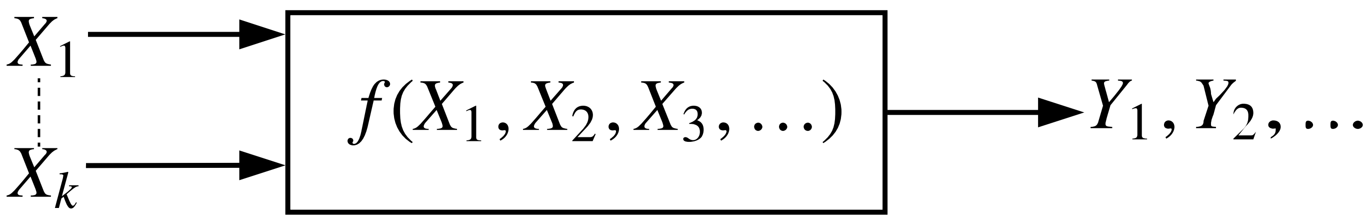
This class

- **Terminating simulation**
  - Run for a specific duration T_E or until a specific event E happens
  - Start at time 0 under well-specified initial conditions
  - Example: Bank opens at 8:30 a.m., closes at 4:30 p.m.
    - We are only interested in these 480 minutes

- **Non-terminating simulation**
  - Run for a very long period of time
  - Goal is to study the long run (steady-state) properties of the system
    - Properties that are not influenced by the initial conditions
  - Example: network packet switching, queueing systems, hospital emergency rooms

Output Data

- The outputs of a model consist of one or more random variables
  - Why?
  - We can look at a model as an input-output transformation of a given set of random variables (e.g. exponentially distributed inter-arrival times)



- Example: Consider an M/G/1 queueing example
  - Poisson arrival rate
  - Normal service time
- Suppose we are interested in the long-run mean queue length, i.e. L_Q
  - We run the simulation for 5000 time units, and we divide it into 5 equal sub-intervals of 1000 time units
  - Average the number of customers in the queue in each sub-interval

Current Setup

- We run k independent simulations and obtain k i.i.d. random variables

$$Y_1, Y_2, \ldots, Y_k$$

- where,

$$E[Y_i] = \theta, \text{ and, } Var(Y_i) = \sigma^2$$

Quantity we want to estimate

- Our purpose is to devise an estimator $\hat{\theta}$ of θ
  - Evaluate how good of an estimator that is
  - Obtain insight into when to stop running simulations, i.e., how large must k be

The Sample Mean

- **Definition**:
  - The **sample mean** is the arithmetic average of $Y_1, Y_2, \ldots, Y_k$ defined as

$$\hat{Y} = \frac{1}{k} \sum_{i=1}^{k} Y_i$$

- **Definition**:
  - An estimator $\hat{\theta}$ of a quantity θ is said to be **unbiased** if

$$E[\hat{\theta}] = \theta$$

- **Claim**:
  - The sample mean $\hat{\theta} = \hat{Y}$ is an **unbiased estimator** of $\theta = E[Y_i]$
  - Proof: What is $E[\hat{\theta}]$?

How Good is the Sample Mean?

- A common measure of the goodness of an estimator is to look at its **mean-squared error (MSE)**, i.e.

$$E\left[\left(\hat{\theta}-\theta\right)^2\right]$$

- So let's look at what the MSE for the sample mean is

$$\begin{aligned} E\left[\left(\hat{\theta}-\theta\right)^2\right] &= E\left[\hat{\theta}^2+\theta^2-2\hat{\theta}\theta\right] \\ &= E\left[\hat{\theta}^2\right]+\theta^2-2\theta E\left[\hat{\theta}\right] \\ &= E\left[\hat{\theta}^2\right]-\theta^2 = E\left[\hat{\theta}^2\right]-E\left[\hat{\theta}\right]^2 = Var\left(\hat{\theta}\right) \\ &= \frac{1}{k^2}\sum_{i=1}^{k} Var(Y_i) = \frac{\sigma^2}{k} = \mathrm{MSE}(\hat{\theta}) \end{aligned}$$

MSE of the Sample Mean

- We have shown that

$$\mathrm{MSE}(\hat{\theta}) = \frac{\sigma^2}{k}$$

- The error of our sample mean actually depends on the variance of our random variables
  – This means that when $\frac{\sigma}{\sqrt{k}}$, $\hat{\theta}$ is a good estimator of θ
- We know that a random variable is unlikely to be too many standard deviations from its mean
  – Why? Because of two main theorems:
    - Chebychev's inequality
    - The Central Limit Theorem
    - Proof in class

MSE and Variance

- We will show in class that, using some magic

$$P\left(|\hat{\theta} - \theta| > \frac{1.96\sigma}{\sqrt{k}}\right) \approx 0.05$$

- So what does this mean to us?
  - For large k, we have a **probabilistic approximate bound** on how far is $\hat{\theta}$ from θ

- That's good. But what seems to be missing?
  - We don't know what σ is, so how do we obtain it?
  - We can also estimate it from the data we have!

Sample Variance

- **Definition**:
  - The **sample variance** of $Y_1, Y_2, \ldots, Y_k$ is defined as

$$S^2 \equiv \sum_{i=1}^{n} \frac{\left(Y_i - \hat{Y}\right)^2}{k-1}$$

- **Claim**:
  - The sample variance is an unbiased estimator of σ^2
  - Proof: We will show in class that

$$E[S^2] = \sigma^2$$

- We will call S the sample standard deviation

Now What?

- So far, we have established that, using our sample mean $\hat{\theta}$ and our sample standard deviation S that we have

$$P\left(|\hat{\theta} - \theta| > \frac{1.96S}{\sqrt{k}}\right) \approx 0.05$$

- How does this help us figure out what k is?
  - Let $d = \dfrac{S}{\sqrt{k}}$
  - Then choose an acceptable value for d
    - Keep increasing k until we hit that value
  - Problems?

Variance Based Stopping Condition

- A problem might arise from the fact that we are also estimating σ
  - So we need k to be good enough for S a good estimate of σ

- Algorithm now looks like:
- Choose d and a lower bound $k_0 = 30$

Run k_0 simulation runs
$k \leftarrow k_0$
while $\frac{s}{\sqrt{k}} > d$ **do**
    Run simulation and generate output y_k
    $k \leftarrow k + 1$
    Compute new estimate s
end while

$\bar{x} \leftarrow \sum_{i=1}^{k} \frac{y_i}{k}$

Confidence Intervals

- From our previous discussion,
  - we know that $\hat{\theta}$ is close to θ
  - we have computed an approximate probabilistic upper bound on how far is $\hat{\theta}$ from θ

- However, it would be meaningful if
  - we could also assert, with a certain confidence, whether the real θ falls within a given interval

- Enter confidence intervals!
  - Notice that we used *confidence intervals* and not *probability intervals*
  - We will see later on why we used that term specifically!

Recall our Setup

- We run *k* simulation runs, and obtain *k* i.i.d random variables

$$Y_1, Y_2, \ldots, Y_k$$

- We know that

$$E[Y_i] = \theta, \text{ and, } Var(Y_i) = \sigma^2$$

- **Goal:** Assert, with a certain confidence, that θ falls within a certain interval,

$$\left(\hat{\theta} - \gamma, \hat{\theta} + \gamma\right)$$

Confidence Intervals

- We know from the Central Limit Theorem, that for large *k*,

$$\sqrt{k}\left(\frac{\hat{\theta}-\theta}{\sigma}\right) \dot{\sim} \mathrm{N}(0,1)$$

approximately

- Also, from a results know as Slutsky's theorem, we can write

$$\sqrt{k}\left(\frac{\hat{\theta}-\theta}{S}\right) \dot{\sim} \mathrm{N}(0,1)$$

Sample standard deviation

Derivations

- Let $Z \sim \mathrm{N}(0,1)$ and z_α be such that

$$P\left(Z > z_\alpha\right) = \alpha$$

- For example,

$$z_{0.025} = 1.96$$

- By symmetry of the normal distribution (specifically, the standard normal in our case)

$$z_{1-\alpha} = z_\alpha$$

- i.e.,

$$P(Z > z_{1-\alpha}) = 1 - \alpha, \quad P(Z > z_\alpha) = \alpha$$

- It follows that

$$P\left(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

Derivations Continued

- Therefore, for large *k*, we can write that

$$P\left(-z_{\frac{\alpha}{2}} < \sqrt{k}\,\frac{\hat{\theta}-\theta}{S} < z_{\frac{\alpha}{2}}\right) \approx 1-\alpha$$

- Rearranging the above equation,

$$\boxed{P\left(\hat{\theta} - z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}} < \theta < \hat{\theta} - z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}}\right) \approx 1-\alpha}$$

- What does this mean?
  - With approximate probability $1-\alpha$, the population mean (i.e., θ) will lie within the region

$$\hat{\theta} \pm z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}}$$

- But wait, we're still saying probability, where does confidence come into play?

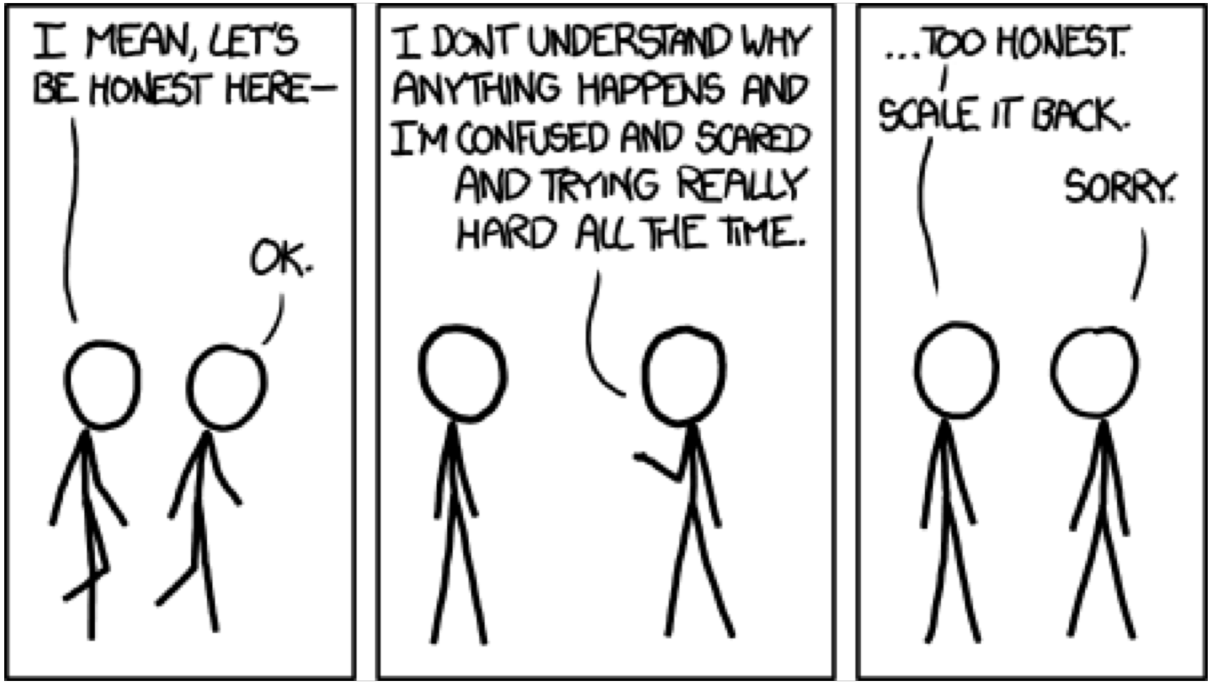
Interpretation

- Recall our previous finding,

$$P\left(\hat{\theta} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{k}} < \theta < \hat{\theta} - z_{\frac{\alpha}{2}} \frac{S}{\sqrt{k}}\right) \approx 1 - \alpha$$

- Let $\bar{x}$ and s be **observed** values of the sample mean $\hat{\theta}$ and sample standard deviation S
  - Then we call the interval $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{s}{\sqrt{k}}$
    - an (approximate) $100(1-\alpha)\%$ confidence interval estimate of θ
  - Note that for $\alpha = 0.05$
    - We get that $z_{\frac{\alpha}{2}} = 1.96$
      - seems familiar?

Confusion Alert!



https://xkcd.com/1146/

Correct Interpretation

- What we have derived is that for the sample mean $\hat{\theta}$ and the sample variance S^2 of our k output random variables $\{Y_1, Y_2, \ldots, Y_k\}$

$$P\left(\hat{\theta} - z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}} < \theta < \hat{\theta} - z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}}\right) \approx 1 - \alpha$$

  - Or, the true mean θ will fall in the interval $\hat{\theta} \pm z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}}$ with approximate probability $1 - \alpha$
  - Recall that $\hat{\theta}$ is also a random variable!

- However, after observing $\bar{x}$ and s, then the true mean θ
  - either falls in $\bar{x} \pm z_{\frac{\alpha}{2}}\frac{S}{\sqrt{k}}$ or it does not!
    - so talking about the probability here is irrelevant!
  - however, we are $100(1-\alpha)\%$ confident that $\hat{\theta}$ will fall into that interval

Correct Interpretation cont.

- However, after observing $\bar{x}$ and s, then the true mean θ
  - either falls in $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{S}{\sqrt{k}}$ or it does not!
    - so talking about the probability here is irrelevant!
  - however, we are $100(1-\alpha)\%$ confident that $\hat{\theta}$ will fall into that interval

- Or, alternatively, in prose
  - "I am *k* runs of the simulation and create a 95% confidence interval for θ. This particular interval (i.e., $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{S}{\sqrt{k}}$) may or may not contain θ. However, if I were to create many confidence intervals like this one many time over, then approximately 95% of them will contain the true mean θ"

Leemis, Lawrence M., and Stephen Keith Park. *Discrete-event simulation: A first course*. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

Stopping Condition

- Now we have two parameters we can control
  - The confidence level $100(1-\alpha)$
  - The length of the confidence interval l

- Choose α and l and a lower bound $k_0 = 30$

Choose α and interval length l
Run k_0 simulation runs
$k \leftarrow k_0$
while $2z_{\frac{\alpha}{2}}\frac{s}{\sqrt{k}} \geq l$ **do**
    Run simulation and generate output y_k
    $k \leftarrow k+1$
    Compute new estimate s
end while

$\bar{x} \leftarrow \sum_{i=1}^{k} \frac{y_i}{k}$

What did we do in this course?

- At the start of the semester, you should have
  - Basic programming skills (C++, Python, etc.)
  - Basic understanding of probability theory (ECE313 or equivalent)

- At the end of the semester, you should be able to
  - Understand different system modeling approaches
    - Combinatorial methods, state-space methods, etc.
  - Understand different model analysis methods
    - Analytic/numeric methods, simulation
  - Understand the basics of discrete event simulation
  - Design simulation experiments and analyze their results
  - **Gain hands-on experience with different modeling and analysis tools**