Department of Industrial Engineering and Engineering Management
University of Sharjah

0405324: Stochastic System Simulation
0405325: Stochastic Simulation
Course Overview

Today’s Agenda

- **General Information**
  - Instructor’s information
  - Course information
    - Outcomes
    - Topics
    - Textbooks
    - Grading scale
    - Exams
  - Class format
- **A Refresher on Probability and Statistics**

Instructor's Information

- Instructor: **Dr. Ali Cheaitou**
- Office: **W9-123 (W9-110)**
- Phone: **Ext. 3921; 3966**
- Email: **acheaitou@sharjah.ac.ae**

- **Office Hours**
  - **Available most of the time (working hours) on MS Teams, but would be available in particular:**
    - **Immediately after the lectures.**
  - **You can also take an appointment by email.**

Instructor’s Information

✓**Education**

✓French ministry of higher education qualification (Industrial Engineering)

✓PhD in Industrial Engineering (Supply Chain Optimization), Ecole Centrale Paris - Paris Saclay University, France.

✓M.Sc., Industrial Systems Engineering, Ecole Centrale Paris - Paris Saclay University, France.

✓M. Eng., Mechanical Engineering, Lebanese University, Beirut, Lebanon.

✓**Professional experience**

✓2018 – : Chairman of the IEEM dept., University of Sharjah, UAE.

✓2016 – 2017: Coordinator of the PhD program in Engineering Management, University of Sharjah, UAE

✓2013 – 2016: Coordinator of the MSc and PhD programs in Engineering Management, University of Sharjah, UAE

✓2019 - : Associate Professor, Industrial Engineering, University of Sharjah, UAE

✓2011 – 2018: Assistant Professor, Industrial Engineering, University of Sharjah, UAE

✓2009 – 2011: Assistant Professor, SCM, Euromed Management, France

✓2009 – 2013: Part-time lecturer, ESAA (Ecole Supérieure Algérienne des Affaire), Algiers, Algeria.

✓2008 : Part time lecturer, Engineering School, Pôle Universitaire Léonard De Vinci, Paris, France

✓2007 – 2009 : Consultant in Supply Chain Management – Information Systems (ERP), L’Oréal SA, L’Oréal China

✓2004 – 2007 : Lecturer and researcher, Ecole Centrale Paris, France

Course Information

- Course materials
  - Check blackboard
    - Syllabus + course outline
    - Handouts
    - Formulas and tables
    - Tutorials
    - Grades
  - Check the course information folder/your email for announcements
  - Please see the PDF file (Course outline) for important notes

Course Information

- Course outcomes - Mapping

	Course Outcome	Link to Student outcomes
1	Recognize and formulate **engineering** problems where simulation is used.	1
2	Observe **waiting lines and** systems for data collection and fit distributions to data.	6
3	Build valid and credible simulation models using specialized software.	6
4	Design simulation experiments on production and service systems.	6, 2
5	Analyze and interpret simulation outputs and improve performance.	6
6	Communicate effectively and work in teams.	

Course Information

- ## Program outcomes – ABET Terminology

Upon successful completion of the B.Sc. program in IEEM, graduates will have:

1 An ability to apply knowledge of mathematics, science, management and engineering in identifying, formulating, and solving Industrial Engineering and Engineering Management problems.

2 An ability to apply engineering design of integrated systems of people, materials, information, facilities and technology, to solve Industrial Engineering and Engineering Management problems with consideration of health, environment, safety, globalization as well as cultural, social and economic factors.

3 An ability to communicate effectively with diverse audience.

4 An ability to recognize professional and ethical responsibilities in solving Industrial Engineering and Engineering Management problems using contemporary knowledge with consideration of the impact of engineering solutions in global, economic, environmental, and social context .

5 An ability to work in a multi-disciplinary team that establishes goals, plans tasks, meets objectives and provides leadership in a collaborative and inclusive environment.

6 An ability to design and conduct experiments, analyze and interpret data to draw conclusions using modern Industrial Engineering and Engineering Management tools.

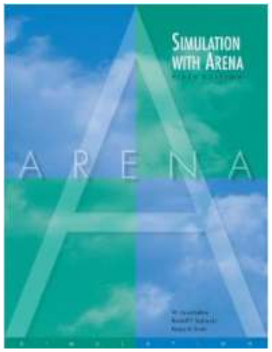
7 An ability to acquire and apply new knowledge as needed, using appropriate learning strategies.

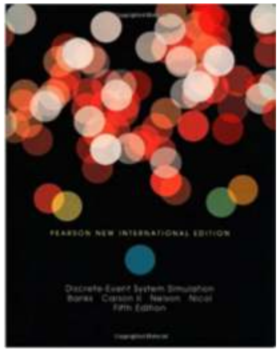
Course Information

- Course topics (Syllabus)
  1. Introduction to discrete event simulation **and queuing theory**.
  2. **Manual and computerized** simulation of the single server queue.
  3. Building conceptual models.
  4. Structuring simulation models.
  5. Preparing simulation input.
  6. Identifying probability distributions from data.
  7. Methods for random numbers and random variables generation.
  8. Verification and validation of simulation models.
  9. Design of experiments for simulation runs.
  10. Output analysis.
  11. Use of simulation language in the lab.

Course Information

- Textbook/Reference
  - Simulation with Arena, By W. Kelton, R. Sadowski, and N. Swets, McGraw Hill, 5th Edition, 2010.





  - **Discrete-event system simulation, by J. Banks, J.S. Carson, B.L. Nelson, D.M. Nicol, Pearson, fifth edition, 2010 (Available in Blackboard and in the new University online System)**

  - Lecture notes
  - Any book on Simulation

Grading Criteria

• Quiz (theory)	15%
• Lab (performance/quizzes/assignments)	20%
• Midterm	20%
• Final Exam	45%

Grading Scale

Letter Grade	Percentage of Total Points
A	90% or above
B+	85 - 89%
B	80 - 84%
C+	75 - 79%
C	70 - 74%
D+	65 - 69%
D	60 - 64%
F	< 60%

Quizzes/Exam

- **Quizzes/homework**

  Total 3 quizzes (best 2 will be considered). **Please note that there will be no makeup quiz.**

- **Exam**
  - Based on homework/lecture material/ lab material
  - If you somehow **miss any exam**, please apply with supporting document to the department within 3 days from the exam date. The department will decide whether your makeup exam will be considered. **No application will be accepted after three days from the date of the exam**. **Make-up exam may be complicated than the regular exam.**
  - Mid-term Exam : **Monday 28 June, 2021**
  - The **quiz dates** are related to the progress in the material and will be scheduled as follows (the dates will be on any teaching day and during the class time):

    - Quiz 1 : Chapter 5

    - Quiz 2: Chapter 6

    - Quiz 3: Chapters 7-8

Plagiarism

- Plagiarism

  "the act of stealing someone else's property and claiming it as one's own. This property might be a paper, a book, an idea, a computer program, an experiment, an exam paper, an answer, etc" (University Bylaws).

- Penalty for plagiarism
  - **Penalty for minor plagiarism (up to 25% copy)-** The student is issued a warning ***and*** 50% of the marks for the specific assessment item are deducted (as a minimum). The instructor, in consultation with the department chairman, has the right to raise the deduction up to 100% of the marks for the specific assessment item.
  - **Penalty for significant plagiarism (more than 50% copy)-** If a significant amount of plagiarism is demonstrated, the student is issued with a warning ***and*** the student is assigned zero marks for the specific assignment.

Lecture Format

- **Lecture material + <u>Recordings</u> will be delivered on Blackboard**
  - Material will be added to blackboard at the beginning of the semester
  - There may have periodic in-class problem solving sessions
  - Minor changes to the slides may be made later and you will be notified about it.
  - **<u>Copyrights</u>** are reserved to the instructor and the textbook writer.
- **All added (hand written) material is your responsibility**
  - Hand written material added by instructor will NOT be available on Blackboard

Attendance

- **If your absence exceeds 20% of the total number of hours without any excuse acceptable to the course instructor and approved by the Dean, you will be barred from the final exam and shall be given a grade of "FA" (University By-Laws, Section-six, Article(27))**

- **I will take the attendance automatically in Blackboard and by raising hands (Blackboard).**

Syllabus

Introduction to discrete event simulation, simulation of single server queue, systems simulation structure, conceptual models; generation of random numbers and random variables; system simulation languages, model verification and validation, design of experiments for simulation runs, output analysis; applications to industrial situations; **introduction to queueing theory**.

The course contains simulation lab.

Note. Detail is in the "Course syllabus"

Questions?

A Refresher on Probability and Statistics

Probability Basics

- ***Experiment*** **– activity with uncertain outcome**
  - Flip coins, throw dice, pick cards, draw balls from urn, …
  - Drive to work tomorrow – Time? Accident?
  - Operate a (*real*) call center – Number of calls? Average customer hold time? Number of customers getting busy signal?
  - *Simulate* a call center – same questions as above

Probability Basics

- **Population-** samples are collected from populations that are *collections of all individuals or individual items of a particular type.*

- **Sample-** Information is gathered in the form of *samples*, or collections of *observations*

Sample space – complete list of all possible individual outcomes of an experiment

- Could be easy or hard to characterize
- May not be necessary to characterize

Probability Basics

- ***Event*** **– a subset of the sample space**
  - Describe by either listing outcomes, “physical” description, or mathematical description
  - Usually denote by E, F, E_1, E_2, etc.
  - Union, intersection, complementation operations
- ***Probability*** **of an event is the relative likelihood that it will occur when you do the experiment**
  - If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes corresponds to event A, then the probability of event A is

$$P(A) = \frac{n}{N}$$

Probability Basics

- **Some properties of probabilities**

  If S is the sample space, then $P(S) = 1$

  Can have event $E \neq S$ with $P(E) = 1$?

  If $o_1, o_2, \ldots$ are the individual outcomes in the sample space, then

$$\sum_{\text{all } i} P(o_i) = 1$$

Random Variables

- **One way of quantifying, simplifying events and probabilities**
- **A *random variable* (RV) is a number whose value is determined by the outcome of an experiment**
  - Technically, a function or mapping from the sample space to the real numbers, but can usually define and work with a RV without going all the way back to the sample space
  - Think: RV is a number whose value cannot be known with certainty but one can usually know something about what it can be or is likely to be
  - Usually denoted as capital letters: X, Y, W_1, W_2, etc.
- **Probabilistic behavior described by *distribution function***

Discrete vs. Continuous RVs

- ***Discrete*** **– can take on only certain separated values**
  - Number of possible values could be finite or infinite
- ***Continuous*** **– can take on any real value in some range**
  - Number of possible values is always infinite
  - Range could be bounded on both sides, just one side, or neither

Discrete Distributions

- **Let X be a discrete RV with possible values (range) x_1, x_2, … (finite or infinite list)**
- ***Probability mass function*** **(PMF)**

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

  - The statement "$X = x_i$" is an event that may or may not happen, so it has a probability of happening, as measured by the PMF
  - Can express PMF as numerical list, table, graph, or formula
  - Since X must be equal to *some* x_i, and since the x_i's are all distinct, $\sum_{\text{all } i} p(x_i) = 1$

Discrete Expected Values

- **Data set has a "center" – the average (mean)** X
- **RVs have a "center" –** ***expected value***

$$E(X) = \sum_{\text{all } i} x_i p(x_i)$$

  - Also called the *mean* or *expectation* of the RV *X*
  - Other common notation: μ, μ_X

Discrete Variances and Standard Deviations

- **Data set has measures of "dispersion" –**
  - Sample variance $s^2 \quad \frac{1}{n-1}\sum_{i=1}^{n}(X_i - X)^2$
  - Sample standard deviation $s = \sqrt{s^2}$
- **RVs have corresponding measures**

$$Var(X) = \sum_{\text{all } i}(x_i - \mu)^2 p(x_i)$$

  - Other common notation: σ^2, σ_X^2

Continuous Distributions

- ***Probability density function*** **(PDF) is a function *f*(*x*) with the following three properties:**
  - $f(x) \geq 0$ for all real values x
  - The total area under $f(x)$ is 1: $\int_{-\infty}^{\infty} f(x)\,dx = 1$
  - For any fixed a and b with $a \leq b$, the probability that X will fall between a and b is the area under $f(x)$ between a and b:

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

Continuous Expected Values, Variances, and Standard Deviations

- **Expectation or mean of X is**

$$\mu = \mu_X = E(X) = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

  - Roughly, a weighted "continuous" average of possible values for X
  - Same interpretation as in discrete case: average of a large number (infinite) of observations on the RV X

- **Variance of X is**

$$\sigma^2 = \sigma_X^2 = Var(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\, dx$$

- **Standard deviation of X is**

$$\sigma = \sigma_X = +\sqrt{Var(X)}$$

Correlation Between RVs

- ***Covariance***

$$\sigma_{X_1X_2} = Cov(X_1, X_2) = E\left[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})\right]$$
$$= E[X_1X_2] - \mu_{X_1}\mu_{X_2}$$

  - Positive (negative) if the points $(X_1; X_2)$ with positive probability fall along a line with positive (negative) slope
  - Measures the linear relationship between X_1 and X_2.

- ***Correlation*** **(coefficient) between X_1 and X_2 is**

$$\rho_{X_1X_2} = Cor(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sigma_{X_1}\sigma_{X_2}}$$

Correlation Between RVs

- ***Covariance / Correlation*** **(coefficient)**

  - If $X_1 and\ X_2$ are independent random variables then:

$$\rho_{X_1 X_2} = \sigma_{X_1 X_2} = 0$$

Sampling

- ***Statistical analysis*** **– estimates or infers something about a *population* or *process* based on only a *sample* from it**
  - Think of a RV with a distribution governing the population
  - *Random sample* is a set of *independent and identically distributed* (IID) observations $X_1, X_2, \ldots, X_n$ on this RV
  - In simulation, sampling is making some runs of the model and collecting the output data
  - Don't know *parameters* of population (or distribution) and want to estimate them or infer something about them based on the sample

Sampling (cont'd.)

- **Population parameter**

  Population mean $\mu = E(X)$

  Population variance σ^2

  Population proportion
- **Parameter – need to know whole population**
- **Fixed (but unknown)**

- **Sample estimate**

  Sample mean

  Sample variance

  Sample proportion
- ***Sample statistic*** **– can be computed from a sample**
- **Varies from one sample to another – is a RV itself, and has a distribution, called the *sampling distribution***

Confidence Intervals

- **A point estimator is just a single number, with some uncertainty or variability associated with it**
- ***Confidence interval*** **quantifies the likely imprecision in a point estimator**
  - An interval that contains (*covers*) the unknown population parameter with specified (high) probability $1 - \alpha$
  - Called a $100\,(1 - \alpha)\%$ confidence interval for the parameter
- **Confidence interval for the population mean μ:**

$$\overline{X} \pm t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}$$

$t_{n-1,\alpha/2}$ is point below which is area $\alpha/2$ in Student's t distribution with $(n - 1)$ degrees of freedom

- **CIs for some other parameters – in text**

Confidence Intervals in Simulation

- Run simulations, get results
- View each replication of the simulation as *a* data point
- Random input $\Rightarrow$ random output
- Form a confidence interval
- Brackets (with probability $1 - \alpha$) the "true" expected output (what you'd get by averaging an infinite number of replications)

Hypothesis Tests

- **Test some assertion about the population or its parameters**
- **Can never determine truth or falsity for sure – only get evidence that points one way or another**
- ***Null hypothesis* (*H_0*) – what is to be tested**
- ***Alternative hypothesis* (*H_1* or *H_A*) – denial of *H_0***

  H_0: $\mu = 6$ vs. H_1: $\mu \neq 6$ (two-sided alternative hypothesis)

  H_0: $\sigma = 10$ vs. H_1: $\sigma > 10$ (one-sided alternative hypothesis)

  H_0: $\mu_1 = \mu_2$ vs. H_1: $\mu_1 \neq \mu_2$
- **Develop a decision rule to decide on *H_0* or *H_1* based on sample data**

Errors in Hypothesis Testing: Type I and type II errors

- **Type I Error: Rejecting the null hypothesis when it is actually true.**
  - Level of significance is the maximum probability of committing a type I error. We want to limit this Type I Error to some small number.
- **Type II Error: Accepting the null hypothesis when it is actually false.**

		Decision based on the sample statistics	
		Accept null	**Reject null**
State of nature	**Null true**	**Correct decision**	**Type I error** α
	Null false	**Type II error** β	**Correct decision**

- **Rejecting the null hypothesis is a strong conclusion, while failing in rejecting H_0 is a weak conclusion.**

p-Values for Hypothesis Tests

- **Alternate method– compute *p-value* of the test**
  - p-value: probability that a test statistic (for instance $\overline{X}$) would take a value at least as extreme as its observed value (based on the sample) when H_0 is true.
  - p-value: minimum level of significance (α) at which H_0 can be rejected.
  - Small *p-value* (< 0.01) is convincing evidence against H_0
  - Large *p-value* (> 0.20) indicates lack of evidence against H_0
- **Credibility of H_0 or the risk in rejecting H_0.**
- **Connection to traditional method**
  - If *p-value* < α, reject H_0
  - If *p-value* $\geq \alpha$, do not reject H_0

Hypothesis Testing in Simulation

- **Input side**
  - Specify input distributions to drive the simulation
  - Collect real-world data on corresponding processes
  - "Fit" a probability distribution to the observed real-world data
  - Test H_0: the data are well represented by the fitted distribution
- **Output side**
  - Have two or more "competing" designs modeled
  - Test H_0: all designs perform the same on output vs. H_1: at least one design is different from at least on other design (use analysis of variance (ANOVA))
  - Selection of a "best" model scenario

Continued in Lecture-1