Topics for Test #3, ISyE 6644 Online Masters in Analytics, Summer 2019

1. I've tried my best to make this list as complete as possible, but I may have missed a topic or two. That being said, you are responsible for everything that we did in class or homework.
2. As GT students, you are expected to formulate problems and solution strategies which are more than mere rote regurgitation of material you learned in class. Thus, you shouldn't be surprised if some questions cover natural extensions of material from class.
3. I'll supply all necessary tables, e.g., N(0,1), *t*, and χ^2, but you can feel free to use your own.
4. Summary: You'll be responsible for **EVERTHING IN THE ENTIRE COURSE!** Woo Woo!!
5. Study Hint: The material from Tests 1 and 2 will receive less emphasis than the subsequent stuff; see Practice Test 3.

OK, so off we go! Here's the list of all of the stuff we've done so far (don't be scared!)…

1. **Intro Material**
   a. Definition of simulation
   b. Advantages and disadvantages of simulation
   c. History of simulation
   d. Typical questions and applications

2. **Calculus, Probability, and Statistics Review**
   a. **Calculus [not really responsible for this material, except I might make you search for a zero]**
      i. Basic definitions
      ii. Derivatives
      iii. Solving for zeros
      iv. Integration
      v. Numerical integration
   b. Probability Preliminaries
      i. Conditional probability
      ii. Independent events
      iii. Definition of random variable
      iv. Discrete RV's and probability mass function
      v. Continuous RV's and probability density function
      vi. Cumulative distribution function
   c. Simulating RV's (first pass)
      i. Discrete uniform distribution
      ii. General discrete distribution
      iii. Inverse Transform Theorem for continuous RV's

iv. Exponential (and other) continuous distributions via IVT.
v. Generating U(0,1)'s via desert island algorithm, including walk-through of pseudo-code.

d. Expected Values
   i. Definition
   ii. Discrete and continuous examples of expected value
   iii. Law of the Unconscious Statistician
   iv. Moments, central moments, variance, standard deviation
   v. Discrete and continuous examples of LOTUS
   vi. Moment generating function
   vii. Examples and properties of mgf's
e. Functions of a RV
   i. Discrete examples
   ii. Continuous examples
   iii. IVT methods (again) with examples
   iv. Relationship with LOTUS
f. Jointly distributed RV's
   i. Definition of joint cdf
   ii. Marginal cdf's
   iii. Joint and marginal pmf's
   iv. Joint and marginal pdf's
   v. Examples for discrete and continuous cases
   vi. Independent RV's
   vii. Conditional pmf's and pdf's
   viii. **Conditional expectation [this won't be on the test]**
   ix. **Double expectation E(E(Y|X)) = EY, including examples [this won't be on the test]**
g. Covariance and correlation
   i. Definitions
   ii. Relationship between independence and correlation
   iii. Examples
   iv. Miscellaneous properties (e.g., Var(X+Y), bounds on correlation, etc.)
h. Probability distributions
   i. Discrete distributions
      1. Bernoulli
      2. Binomial
      3. Geometric
      4. Poisson (including discussion on Poisson processes)
   ii. Continuous distributions
      1. Uniform
      2. Exponential (including memoryless property)

3. Erlang, Gamma distributions
4. Triangular
5. Normal (including Standard Normal)
6. Other sampling distributions (including chi-square, t, F, and various relationships with each other)

i. Limit theorems
   i. Linear combinations of independent normal (including distribution of sample mean)
   ii. Convergence in distribution
   iii. Law of Large Numbers
   iv. Central Limit Theorem for independent and identically distributed data.
   v. Examples

j. Statistics Tidbits **[this material was eventually covered in modules 8, 9, 10, so it's in fair territory to have it on the final! ☺]**
   i. Properties of sample mean and sample variance
   ii. Confidence intervals for the mean and variance

3. Hand Simulations

a. Monte Carlo integration
b. Determining π via simulation (dart tossing on a circle and sphere)
c. Single-server queue (including FIFO and LIFO service disciplines)
d. (s,S) inventory system
e. Simulating RV's (repeats some material from the Prob/Stats review)
f. Spreadsheet simulation (e.g., stock portfolio in Excel)

4. General Simulation Principles

a. Steps in a simulation study
b. List of various simulation definitions (e.g., event, system state, simulation clock, etc.)
c. Event-Scheduling vs. Process Interaction modeling approaches
d. How are events processed?
e. Future events list + extended example
f. Simulation languages – what to look for

5. Arena

a. Layout of Arena screen (panels, modules, etc.)
b. Basic Process template: CREATE-PROCESS-DISPOSE modules
c. SEIZE-DELAY-RELEASE inside of the PROCESS module.
d. Resource, Schedule, Queue, Entity, and other spreadsheets
e. ASSIGN module
f. DECIDE module – probabilistic and conditional routing
g. Simple examples, e.g. (partial list),

i. Single-server queue
ii. Parallel servers
iii. Schedules for servers
iv. Multiple arrival streams

h. Displays, graphics, etc.
i. BATCH and SEPARATE modules
j. Run set-up and control
k. More-sophisticated queueing networks (e.g., two-channel manufacturing example, call center example)
    i. Advanced Process modules (e.g., SEIZE, DELAY, RELEASE modules)
    ii. Some primitive blocks (e.g., QUEUE)
    iii. Use of "pretend" customers
    iv. Nonhomogeneous Poisson arrivals
    v. Use of resource sets, including how to prioritize servers
    vi. Use of submodels
l. Inventory processes
m. **Crazy examples such as re-entrant queues [this won't be on the test]**
n. SMARTS files and other Rockwell examples
o. Manufacturing systems **(you're only responsible for the basics)**
    i. Advanced Transfer modules
    ii. Sequences of customer visitation locations
    iii. **Advanced sets of sequences [this won't be on the test]**

6. Uniform Random Number Generation

a. Overview – desirable properties of a pseudo-random number generator
b. Some generators we won't use, e.g.,
    i. PRN's from tables
    ii. Midsquare
c. Linear congruential generators
    i. Cycling
    ii. 16807 desert island generator (again)
    iii. RANDU (a bad generator)
d. Tausworthe generator
e. Combined generators
    i. L'Ecuyer's generator of cycle length 2^{191}
    ii. Mersenne Twister
f. Some theoretical considerations, e.g., from Knuth's book
g. Statistical tests for randomness
    i. Goodness-of-fit test – Chi-squared
    ii. Runs tests for independence
        1. Runs up and down

2. Runs above and below the mean
3. **Autocorrelation test [this won't be on the test]**

7. Random Variate Generation

a. Inverse Transform Theorem (yet again)
   i. Proof
   ii. Discrete example adaptations
   iii. Continuous examples
      1. East ones such as Exponential, Weibull, etc.
      2. Slightly harder examples such as Triangle distribution
      3. Normal distribution, both exact and approximate methods
   iv. Special case methods, e.g., Geometric
   v. **Empirical distributions [this won't be on the test]**

b. Convolution method
   i. Binomial
   ii. Triangle
   iii. Erlang
   iv. CLT
      1. Desert island sum of Uniforms to generate Normal
      2. Normal approximation to Poisson (including continuity correction)
   v. Cauchy
      1. Cauchy's add up to another Cauchy
      2. IVT method
      3. Ratio of two Normals

c. Acceptance-Rejection methods
   i. Trivial Uniform example
   ii. Some discussion on general method
   iii. **Proof of the general method [don't expect to see this on the final, unless I'm in a really bad mood!]**
   iv. Examples involving polynomial and half-normal p.d.f.'s
   v. Poisson distribution

d. **Composition [this won't be on the test]**

e. Special-case techniques
   i. Box-Muller method for Normal distribution
   ii. Extensions of B-M, e.g., Cauchy, Chi-squared.
   iii. Generating min's and max's of iid RV's, e.g., min of iid Exponentials.

f. Multivariate Normal
   i. Definition in 2 and then >2 dimensions
   ii. Cholesky decomposition method for generating realizations (exact expression in 2 dimensions, algorithm for >2 dimensions)

g. Stochastic processes

i. **Markov chains [this won't be on the test]**
ii. Poisson processes
iii. Nonhomogeneous Poisson processes (via thinning method)
iv. **Time series [this won't be on the test]**
    1. MA(1)
    2. AR(1)
    3. EAR(1)
    4. ARTOP
v. **M/M/1 queue waiting times [this won't be on the test]**
vi. Brownian motion
    1. Definition and history
    2. Elementary properties, including covariance structure
    3. General CLT
    4. How to generate
    5. Geometric BM and financial applications

8. Input Analysis

a. General discussion
    i. Careful about GIGO with respect to simulation input
    ii. What makes a good distribution
    iii. Identification of obvious distributions
b. Estimation review
    i. Unbiased estimators
        1. Definition
        2. Sample mean
        3. Sample variance
        4. Other examples such as Unif(0,θ)
    ii. Mean squared error
    iii. Maximum likelihood estimators
        1. Definition
        2. Examples such as Exponential
        3. Two-dimensional examples such as Normal with unknown mean and variance
        4. **Other tougher examples such as Unif(0,θ) and Gamma [this won't be on the test]**
        5. Invariance Property + examples
    iv. **Method of Moments [this won't be on the test]**
c. Goodness-of-fit tests for input distributions
    i. Chi-squared for Exponential
    ii. Chi-squared for Weibull, including search techniques such as bisection and Newton

iii. Kolmogorov-Smirnov
iv. More goodness-of-fit tests

d. Problem Children
   i. Little or no data
   ii. Data from an unusual distribution
   iii. Nonstationary data
   iv. Multivariate / correlated data

e. Arena Input Analyzer demo

9. Output Analysis

a. Introduction
   i. The need for output analysis in a proper statistical study
   ii. Simulation data isn't iid normal, and this is a problem
   iii. Types of output analysis – finite-horizon (terminating) and steady-state

b. A mathematical interlude related to the fact that the variance of the sample mean isn't $Var(X_i)/n$, and its consequences **[I might ask you to calculate the variance of the sample mean for a specific process, but nothing else.]**

c. Finite-horizon (terminating) simulations
   i. Examples
   ii. Confidence intervals for mean performance via the method of Independent Replications

d. Initialization problems

e. Steady-state analysis for a single system
   i. Examples
   ii. Confidence intervals for the steady-state mean via the method of Batch Means
   iii. **Properties of Batch Means [this won't be on the test]**
   iv. Overlapping Batch Means
   v. Other methods

10. Comparing Systems

a. Classical confidence interval for the mean of one normal population

b. Classical confidence interval to compare the means of two normal systems
   i. Variance completely unknown
   ii. Paired-t CI
   iii. Use in simulation scenarios

c. Variance reduction techniques
   i. Common random numbers
   ii. Antithetic random numbers **[no theory related to this will be on the test]**
   iii. **Control variates [this won't be on the test]**

d. Ranking and selection methods to compare means of >2 systems

i. Definition of problem
ii. Relevance to simulation
iii. Indifference-zone approach
iv. Normal means selection problem
    - Bechhofer's single-stage procedure
    - **Extensions [this won't be on the test]**
v. Bernoulli parameter selection problem
    - Sobel and Huyett single-stage procedure
    - **Extensions [this won't be on the test]**
vi. Multinomial cell selection problem
    - Multinomial review and motivation
    - Bechhofer, Elmaghraby, and Morse single-stage procedure
    - **Extensions [this won't be on the test]**