Midterm
CS 661: System Simulation — Spring 2015
Prof. Marvin K. Nakayama

Print family (or last) name: _____

Print given (or first) name: _____

I have read and understand all of the instructions below, and I will obey the University Code on Academic Integrity.

_____

Signature and Date

Instructions

- This exam has 17 pages in total, numbered 1 to 17. Make sure your exam has all the pages.
- This is a closed book, closed-note exam. You are not allowed to use any electronic device (e.g., calculators or cellphones).
- For all problems, follow these instructions:
  - Show your work and give reasons.
  - For any problem that asks you to provide an algorithm, be sure to give a step-by-step algorithm. Do not explain how to implement your algorithm in Arena or Excel, but rather, you should provide an algorithm (i.e., pseudo-code) that can be implemented in any language. Also, for any random variates needed by your algorithm, be sure to explicitly provide the steps needed to generate the random variates. Also, be sure to explicitly specify all constants or parameter values used in your algorithm. You may assume that you have at your disposal a good unif[0,1] random number generator.
  - In any problem for which you use an approximation based on the central limit theorem, you can use a critical point from the normal distribution. Any confidence interval you provide should be at a 95% confidence level. Any hypothesis test you perform should be at significance level $\alpha = 0.05$.
  - If a problem requires you to solve for the root of a polynomial $p(x)$, then you must give the explicit root when the polynomial has order $k = 1$ or 2. For polynomials of order $k \geq 3$, just give the polynomial in the form $p(x) = c_k x^k + c_{k-1} x^{k-1} + \cdots + c_1 x + c_0$, explicitly giving each coefficient c_i, $0 \leq i \leq k$, and state that you need to solve for the roots of $p(x)$, without giving the roots.
  - You can use the backs of the sheets if you need more space.
  - The last two pages of the exam have some possibly useful reference material.

Problem	1	2	3	4	5	Total
Points						

1. **[20 points]** Suppose that we have a coin, which, on a single toss, has probability p of coming up heads, and probability $1-p$ of coming up tails. We flip the coin infinitely many times, where the coin flips are mutually independent. For each $k \geq 1$, let $N(k)$ be the number of heads observed up to and including the kth toss, and we define a stochastic process $N = [N(k) : k = 1, 2, 3, \ldots]$.

   (a) **[5 points]** For the coin-flipping experiment, define specific outcomes

   $$\omega_1 = \mathbf{HTTHHTTH}\ldots,$$
   $$\omega_2 = \mathbf{TTHTHTHT}\ldots,$$
   $$\omega_3 = \mathbf{THTTTTHT}\ldots,$$

   so in outcome ω_1, the first flip is heads, the second and third are tails, the fourth and fifth are heads, etc. There are infinitely many such outcomes, each of infinite length. For each of the specific outcomes $\omega_1, \omega_2, \omega_3$ given above, write out the values of $N(1), N(2), N(3), N(4), N(5)$.

(b) **[5 points]** What is the probability distribution of $N(5)$? What is the probability distribution of $N(16)$? Does the distribution of $N(k)$ depend on k?

(c) **[5 points]** Show that for each $k \geq 1$, $N(k)$ and $N(k+2)$ are dependent. Thus, conclude that the (marginal) distributions of $N(k)$ for each $k = 1, 2, 3, \ldots$ do not completely specify the probabilistic structure of the process N.

(d) **[5 points]** For any $k > j$, what is the distribution of $N(k) - N(j)$? For any $m \geq 1$, what is the distribution of $N(m+k) - N(m+j)$? Does the distribution of $N(m+k) - N(m+j)$ depend on m?

2. **[15 points]** Suppose that $X_1, X_2, \ldots, X_n$ is a sample of n independent and identically distributed observations from some CDF F.

   (a) **[8 points]** Suppose that we have determined that F is the geometric(p) distribution, which is a discrete distribution with parameter p, $0 < p < 1$, such that if X is a random variable with this distribution, then

$$\mathcal{P}(X = x) = \begin{cases} (1-p)^{x-1}p, & \text{for } x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases} .$$

   Derive the maximum likelihood estimator (MLE) of the parameter p.

   Answer: _____

(b) **[7 points]** Instead suppose now that F is unif$[a, b]$, where a and b are unknown. Derive the method-of-moments estimator of the parameters a and b.

Answer: _____

3. **[25 points]** Let (X, Y, Z) be a random vector with some unknown joint CDF F, and let

$$\theta = \mathcal{P}(-1 \leq 3X^5 \ln(Y)/\sin(X\sqrt{Z}) \leq 5).$$

Assume that $0 < \theta < 1$. Let $(X_1, Y_1, Z_1), (X_2, Y_2, Z_2), \ldots, (X_n, Y_n, Z_n)$ be a sample of size n of IID triples from F.

(a) **[5 points]** Give an estimator $\hat{\theta}_n$ of θ from the sample.

(b) **[5 points]** Show that your estimator $\hat{\theta}_n$ is an *unbiased* estimator of θ.

(c) **[4 points]** Argue that your estimator $\hat{\theta}_n$ satisfies a law of large numbers.

(d) **[5 points]** Argue that your estimator $\hat{\theta}_n$ satisfies the following central limit theorem (CLT):

$$\frac{\sqrt{n}}{\beta}(\hat{\theta}_n - \alpha) \overset{\mathcal{D}}{\approx} \mathcal{N}(0,1)$$

for large n, for some some constants α and β. Explicitly specify formulas for α and β.

(e) **[3 points]** Give an estimator $\hat{\beta}_n$ of β from the sample. Argue that $\hat{\beta}_n \approx \beta$ for large n.

(f) **[3 points]** Use the previous parts to derive an approximate 95% **upper-one-sided** confidence interval for θ.

4. **[30 points]** Our goal in this problem is to generate a random variate having an exponentially tilted Weibull distribution. Let W be a random variable having a Weibull distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$; i.e., the CDF F of W is given by

$$F(w) = \begin{cases} 1 - e^{-(w/\beta)^\alpha}, & w \geq 0 \\ 0, & w < 0 \end{cases},$$

and the density function f is given by

$$f(w) = \begin{cases} \alpha\, \beta^{-\alpha}\, w^{\alpha-1}\, e^{-(w/\beta)^\alpha}, & w \geq 0, \\ 0, & w < 0. \end{cases}$$

The moment generating function $m(\theta)$ of W is

$$m(\theta) = \int_0^\infty \alpha\, \beta^{-\alpha}\, w^{\alpha-1}\, \exp\{-(w/\beta)^\alpha + \theta w\}\, dw,$$

which in general cannot be evaluated in closed form and must be evaluated numerically; i.e., there is no simple formula for $m(\theta)$ in terms of θ, and we need to use numerical methods (such as the trapezoid method) to evaluate the above integral. The density function $h_\theta(x)$ of the exponentially tilted Weibull distribution is

$$h_\theta(x) = \frac{\alpha\, \beta^{-\alpha}\, x^{\alpha-1}\, \exp\{-(x/\beta)^\alpha + \theta x\}}{m(\theta)}$$

for $x \geq 0$, and $h_\theta(x) = 0$ for $x < 0$. Our goal now is to develop an acceptance-rejection method to generate an observation of X having the density h_θ when $\theta < 0$.

(a) **[3 points]** Define the function $t(x)$ as

$$t(x) = \frac{\alpha\, \beta^{-\alpha}\, x^{\alpha-1}\, e^{\theta x}}{m(\theta)}, \quad x \geq 0,$$

and $t(x) = 0$ for $x < 0$. Show that $h_\theta(x) \leq t(x)$ for all x.

(b) **[6 points]** Define $c = \int_0^\infty t(x)\,dx$. Calculate the value of c in terms of α, β, θ, and $m(\theta)$. [Hint: Recall that $\theta < 0$, and note that $t(x)$ resembles the density function of a gamma distribution.]

(c) **[2 points]** Define $g(x) = t(x)/c$ for all $x \geq 0$, and $g(x) = 0$ for $x < 0$. Show that $g(x)$ is the density function of a gamma distribution. Be sure to specify the shape and scale parameters of the gamma density.

(d) **[5 points]** Show that $\max\{h_\theta(x)/g(x) : x \geq 0\} = c$, where c is defined as in the previous parts.

For all of the following parts, we will now assume that α is an **integer**. Also, you may assume that you have a good unif$[0, 1]$ random number generator, but you must explicitly show how to generate observations from any other distribution.

(e) **[6 points]** Recall that $\theta < 0$. Give an algorithm to generate Y having density function g.

(f) **[6 points]** Use the results in the previous parts to develop an acceptance-rejection algorithm for generating a random variate X having the density h_θ when $\theta < 0$.

(g) **[2 points]** What is the expected number of unif$[0, 1]$ random numbers required by your algorithm of the previous part to generate X?

5. **[10 points]** Suppose that you download a program from the web. The program's author claims that when it is run, it will generate independent samples from an exponential distribution with mean 1. To test the program, you generate 100 outputs from the program, and you get that the average of the 100 values is 1.14. Would you be worried about the claim that the program generates independent exponentials with mean 1?

Some possibly useful facts

- The CDF F of random variable (RV) X is defined as $F(z) = P(X \le z)$. If X is a continuous random variable with density function f, then $F(z) = \int_{-\infty}^{z} f(x)dx$. If X is a discrete random variable with probability mass function p and support $\{x_1, x_2, x_3, \ldots\}$, then $F(z) = \sum_{x_i \le z} p(x_i)$.
- The expectation or mean μ of a RV X is $\mu = \mathcal{E}[X] = \sum_{x_i} x_i p(x_i)$ when X is a discrete random variable with probability mass function $p(\cdot)$ and support $\{x_1, x_2, \ldots\}$, and $\mu = \int_{-\infty}^{\infty} x f(x)\, dx$ when X is a continuous random variable with density function f.
- The variance of a RV X is $\mathcal{V}(X) = \mathcal{E}[(X-\mu)^2] = \mathcal{E}[X^2] - \mu^2$.
- The covariance of RVs X, Y with means μ_X, μ_Y is $\mathcal{C}(X,Y) = \mathcal{E}[(X-\mu_X)(Y-\mu_Y)] = \mathcal{E}[XY] - \mu_X\mu_Y$.
- Sample mean $\bar{X}(n) = \frac{1}{n}\sum_{k=1}^{n} X_i$, and sample variance $S^2(n) = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X}(n))^2$.
- The normal distribution with mean μ and variance $\sigma^2 > 0$ is denoted by $\mathcal{N}(\mu, \sigma^2)$ and has density function $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$. Its CDF cannot be evaluated in closed form.
- If $Z_1, Z_2, \ldots, Z_n$, $n \ge 1$, are IID $\mathcal{N}(0,1)$ random variables, then $X = Z_1^2 + Z_2^2 + \cdots + Z_n^2$ has a chi-squared distribution with n degrees of freedom, which is denoted by χ^2_n.
- If $Z \sim \mathcal{N}(0,1)$ and $V \sim \chi^2_n$ with Z and V independent, then $T = Z/\sqrt{V/n}$ has a Student-t distribution with n degrees of freedom, which is denoted by t_n.
- If $X_1, X_2, \ldots, X_n$ are IID $\mathcal{N}(\mu, \sigma^2)$, then $(n-1)S^2(n)/\sigma^2 \sim \chi^2_{n-1}$ and $(\bar{X}(n) - \mu)/\sqrt{S^2(n)/n} \sim t_{n-1}$.
- The uniform distribution over the interval $[a,b]$ has density function $f(x) = [1/(b-a)]I_{[a,b]}(x)$, where $I_A(x) = 1$ if $x \in A$ and 0 otherwise. Its CDF is $F(x) = (x-a)/(b-a)$ for $a \le x \le b$, and $F(x) = 0$ for $x < a$, and $F(x) = 1$ for $x > b$. Its mean is $(a+b)/2$ and its variance is $(b-a)^2/12$.
- The exponential distribution with (rate) parameter $\lambda > 0$ has density function $f(x) = \lambda e^{-\lambda x} I_{[0,\infty)}(x)$. Its CDF is $F(x) = [1 - e^{-\lambda x}]I_{[0,\infty)}(x)$. Its mean is $1/\lambda$ and its variance is $1/\lambda^2$.
- The gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ has density function $f(x) = [\beta^{-\alpha} x^{\alpha-1} \exp\{-x/\beta\}/\Gamma(\alpha)]I_{[0,\infty)}(x)$, where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$ and $\Gamma(\alpha) = (\alpha-1)!$ when α is a positive integer. In general, its CDF cannot be evaluated in closed form. Its mean is $\alpha\beta$ and its variance is $\alpha\beta^2$. In the special case that α is an integer k, then the distribution is known as an Erlang, and a random variable X having this distribution can be expressed as the sum of k IID exponentials, each with mean β.
- The Weibull distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ has density function $f(x) = [\alpha\beta^{-\alpha} x^{\alpha-1} \exp\{-(x/\beta)^\alpha\}]I_{[0,\infty)}(x)$. Its CDF is $F(x) = [1 - \exp\{-(x/\beta)^\alpha\}]I_{[0,\infty)}(x)$. Its mean is $(\beta/\alpha)\Gamma(1/\alpha)$ and its variance is $(\beta^2/\alpha)[2\Gamma(2/\alpha) - (1/\alpha)(\Gamma(1/\alpha))^2]$.
- If $Z \sim \mathcal{N}(\mu, \sigma^2)$, then $X = e^Z$ has a lognormal(μ, σ^2) distribution. The lognormal(μ, σ^2) distribution has density function $f(x) = [1/(x\sqrt{2\pi\sigma^2}) \exp(-(\ln x - \mu)^2/(2\sigma^2))]I_{[0,\infty)}(x)$. Its CDF cannot be evaluated in closed form. Its mean is $\exp(\mu + (\sigma^2/2))$, and its variance is $e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$.
- A Poisson distribution has a given parameter $\lambda > 0$ and has support $\{0, 1, 2, \ldots\}$. The probability mass function is $p(k) = e^{-\lambda}\lambda^k/k!$ for $k = 0, 1, 2, \ldots$. The mean and variance of a Poisson RV are both λ.
- Law of large numbers: if $X_1, X_2, \ldots, X_n$ are IID RVs with finite mean μ, then $\bar{X}(n) \approx \mu$ for large n.
- Central limit theorem: if $X_1, X_2, \ldots, X_n$ are IID RVs with finite mean μ and positive, finite variance σ^2, then $\frac{\sqrt{n}}{\sigma}(\bar{X}(n) - \mu) \stackrel{\mathcal{D}}{\approx} \mathcal{N}(0,1)$ for large n, where $\stackrel{\mathcal{D}}{\approx}$ means approximately equal to in distribution. Also, $\frac{\sqrt{n}}{S(n)}(\bar{X}(n) - \mu) \stackrel{\mathcal{D}}{\approx} \mathcal{N}(0,1)$ for large n.

- Confidence intervals: if $X_1, X_2, \ldots, X_n$ are IID RVs with finite mean μ and positive, finite variance σ^2, then an approximate $100(1-\alpha)\%$ confidence interval for μ is $[\bar{X}(n) \pm z_{\alpha/2} S(n)/\sqrt{n}]$, where $z_{\alpha/2}$ is the constant such that $\mathcal{P}(\mathcal{N}(0,1) > z_{\alpha/2}) = \alpha/2 = 1 - \mathcal{P}(\mathcal{N}(0,1) \leq z_{\alpha/2})$. For example, $\mathcal{P}(\mathcal{N}(0,1) > 1.65) = 0.05$, $\mathcal{P}(\mathcal{N}(0,1) > 1.96) = 0.025$, and $\mathcal{P}(\mathcal{N}(0,1) > 2.60) = 0.005$.

- $e = 2.718\ldots$, $e^{-1} = 0.368\ldots$, $\ln e = 1$, $\ln(ab) = \ln(a) + \ln(b)$, $\ln(a^b) = b \ln(a)$, and $e^{a \ln(b)} = b^a$.

- Parameter estimation: Suppose $X_1, \ldots, X_n$ are IID with distribution $F_{\theta_1,\ldots,\theta_d}$, where $\theta_1, \ldots, \theta_d$ are unknown and to be estimated.

  - MLE: If $F_{\theta_1,\ldots,\theta_d}$ is a continuous distribution with density function $f_{\theta_1,\ldots,\theta_d}$, then compute the likelihood $L(X_1, \ldots, X_n) = \prod_{i=1}^n f_{\theta_1,\ldots,\theta_d}(X_i)$. If $F_{\theta_1,\ldots,\theta_d}$ is a discrete distribution with probability mass function $p_{\theta_1,\ldots,\theta_d}$, then compute the likelihood $L(X_1, \ldots, X_n) = \prod_{i=1}^n p_{\theta_1,\ldots,\theta_d}(X_i)$. Then choose the values of $\theta_1, \ldots, \theta_d$ that maximize $L(X_1, \ldots, X_n)$, or equivalently that maximize $\ln L(X_1, \ldots, X_n)$.
  - Method-of-moments: Let $\mu_j = \mathcal{E}[X^j]$ be the jth (theoretical) moment of $F_{\theta_1,\ldots,\theta_d}$, and let $M_j = \frac{1}{n}\sum_{i=1}^n X_i^j$ be the jth sample moment of the data. Then set $\mu_j = M_j$ for $j = 1, 2, \ldots, d$, and solve for $\theta_1, \ldots, \theta_d$.

- Goodness-of-fit tests: Suppose $X_1, \ldots, X_n$ are IID from some hypothesized distribution F.

  - Chi-squared test: Divide samples into k bins. Let O_i be the number of observations in the ith bin, and let E_i be the expected number of observations in the ith bin. Compute test statistic $\chi^2 = \sum_{i=1}^k (O_i - E_i)^2/E_i$, which approximately has a chi-squared distribution with ν degrees of freedom, where $\nu =$ number of bins $-$ (1 + number of estimated parameters).
  - Kolmogorov-Smirnov (K-S) test: Compute empirical CDF $\hat{F}_n(x) = (\#\text{ observations} \leq x)/n$. Then compute $D_n = \max_{x \in \Re} |\hat{F}_n(x) - F(x)|$. We obtain the test statistic by centering and scaling D_n, where the centering and scaling depend on the hypothesized F.

- LCG: $Z_i = (aZ_{i-1} + c) (\text{mod } m)$ and $U_i = Z_i/m$, for constants a, c and m.

- Let $U_1, U_2, \ldots, U_n$ be n successive samples from random number generator. Null hypothesis is H_0 : samples are IID unif$[0,1]$. Fix level α of test, and define $z_{\alpha/2}$ such that $\mathcal{P}(\mathcal{N}(0,1) \leq z_{\alpha/2}) = 1 - \alpha/2$.

  - Runs-up/down test: Let R_n is the total number of runs, $\mu_n = \mathcal{E}[R_n] = (2n-1)/3$ and $\sigma_n^2 = \mathcal{V}[R_n] = (16n - 29)/90$. Reject H_0 at level α iff $R_n \notin [\mu_n - z_{\alpha/2}\sigma_n,\ \mu_n + z_{\alpha/2}\sigma_n]$.

- Inverse-transform method to generate sample X having CDF F. Generate $U \sim$ unif$[0,1]$, and return $X = F^{-1}(U)$.

- Convolution method: Suppose $X = Y_1 + Y_2 + \cdots + Y_n$, where $Y_1, Y_2, \ldots, Y_n$ are independent, and $Y_i \sim F_i$. We generate a sample of X by generating $Y_1 \sim F_1$, $Y_2 \sim F_2$, ..., $Y_n \sim F_n$, where $Y_1, \ldots, Y_n$ are generated independently, and returning $X = Y_1 + \cdots + Y_n$.

- Composition method: Suppose we can write CDF F as $F(x) = p_1 F_1(x) + p_2 F_2(x) + \cdots + p_d F_d(x)$ for all x, where $p_i > 0$ for each i and $p_1 + \cdots + p_d = 1$. We then generate a sample $X \sim F$ as follows. First generate I having PMF $(p_1, p_2, \ldots, p_d)$. If $I = j$, then independently generate $Y_j \sim F_j$, and return $X = Y_j$.

- Acceptance-rejection (AR) method: To generate X having density f, suppose Y is a RV with density g such that there exists a constant c for which $f(x) \leq cg(x)$ for all x. Then we can generate a sample of X using the AR method as follows. (i) Generate Y with density g. (ii) Generate $U \sim$ unif$[0,1]$, independent of Y. (iii) If $cg(Y)U \leq f(Y)$, then return $X = Y$ and stop; else, reject (Y, U) and return to step (i).