One page of handwritten notes, front and back. Closed book. 50 minutes.

Score ___________________________
1. True or false. (two points each)

(a) T F ← If truly random numbers are used on a 32-bit computer, then their period length is no more than $2^{31} - 1$.

(b) T F ← Mean squared error of a point estimator is unaffected by Monte Carlo sample size.

(c) T ← F If $U_1$ is U(0, 1), then the fractional part of $U = U_1 + 2.1475$ is U(0, 1).

(d) T F ← A performance measure, $\theta$, depends upon the number of micro replications.

(e) T ← F Generalized mean squared error is sometimes used to measure the quality of a simulation experiment.

(f) T ← F The inverse cdf transformation is defined for every univariate distribution.

(g) T ← F In micro-macro replications and a fixed sample size $n$, the best point estimator is obtained using a single macro replication of $n$ micro replications.

(h) T F ← If the new probability measure $h$ is monotonically increasing, then importance sampling is guaranteed to reduce the standard error of the point estimator.

(i) T F ← If $U$ is a random number and $X = \lfloor 6U \rfloor$ is used to toss a die, then $F_X^{-1}(U) = \lfloor 6U \rfloor$ is U(0, 1).

2. (six points) Increasing the sample size $n$ usually decreases the variance of the point estimator. Discuss (briefly) why this is not usually thought of as variance reduction.

For every Monte Carlo simulation experiment, $\text{var}(\hat{\theta}) = O(1/n)$, so there is little to discuss about increasing sample size to obtain a better point estimator.

Interesting discussion of variance-reduction methods centers on replacing the original experiment with a more-efficient experiment.

Efficiency is often measured with generalized mean-squared error, the product of computing effort $n$ and $\text{mse}(\hat{\theta}, \theta)$. Simply increasing $n$ to infinity does not decrease generalized mean-squared error to zero.
3. Consider the following tolerancing problem. The joint distribution of $X = (X_1, \ldots, X_d)$ and the vector of constants $a = (a_1, \ldots, a_d)$ is known. Twenty macro replications, each composed of $50000$ micro replications, are used to estimate $P(Y > 10)$ and $E(Y^3)$, where $Y = \sum_{l=1}^{d} a_l X_l$. The 16807 linear-congruential generator is used with the initial seed value 111,111,111.

(a) (three points) Which part(s) of the problem correspond(s) to the "input model"?

The joint distribution of $X = (X_1, \ldots, X_d)$.

(b) (three points) Which part(s) of the problem correspond(s) to the "logical model"?

$Y = \sum_{l=1}^{d} a_l X_l$

(c) (three points) Which part(s) of the problem correspond(s) to the "input data"?

$X_{ij}; i = 1, \ldots, 50000; j = 1, \ldots, 20$

(d) (three points) Which part(s) of the problem correspond(s) to the "output data"?

$Y_{ij}; i = 1, \ldots, 50000; j = 1, \ldots, 20$

(e) (three points) What part(s) of the problem correspond(s) to the "performance measure"?

$P(Y > 10)$ and $E(Y^3)$

(f) (three points) What part(s) of the problem correspond(s) to the "experiment parameters"?

initial seed 111,111,111. $k = 20$ macro replications of $m = 50000$ each.

(g) (five points) Suggest a point estimator for $E(Y^3)$.

$\hat{\theta} = \frac{\sum_{j=1}^{20} \sum_{i=1}^{50000} Y_{ij}^3}{(20)(50000)}$

(h) (six points) Discuss how you would recommend estimating the standard error of your point estimator from Part (g).

Let $\hat{\theta}_j = \frac{\sum_{i=1}^{50000} Y_{ij}^3}{50000}$. Then $\text{var}(\hat{\theta}) = \frac{(\sum_{j=1}^{20} \hat{\theta}_j^2) - 20 \hat{\theta}^2}{(19)(20)}$. Taking the sqrt yields $\text{ste}(\hat{\theta})$. 

4. (three points each) Based on the notation of Problem 3, circle one of "constant", "random variable", "event" or "meaningless".

(a) \( Y > 10 \) constant random variable event \( \leftarrow \) meaningless

(b) \( E(X_d) \) constant \( \leftarrow \) random variable event meaningless

(c) \( \sum_{i=1}^{d} a_i X_i \) constant random variable \( \leftarrow \) event meaningless

5. Suppose that the random variable \( Y \) has cumulative distribution function

\[
F_Y(x) = \begin{cases} 
1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right] & \text{if } x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

(five points) What do you know about \( F_Y(Y) \)?

The cdf of every continuous random variable is \( U(0, 1) \).

6. (six points) For any sample size \( n \), the left-most digits of a point estimator are meaningful and the right-most digits are not meaningful. Discuss the meaning of meaningful.

Many answers are correct. For example...

(a) A non-meaningful digits is the result of sampling \{0, 1,..., 9\} equally likely.

(b) A meaningful digit is likely to repeat if the experiment is replicated.

(c) A meaningful digit is likely to match its corresponding digit of \( \theta \).

(d) A digit is meaningful if \( \hat{\theta} \) rounded to that digit’s position is likely to match \( \theta \) rounded to that position.

(e) Digits of \( \hat{\theta} \) are meaningful if they match both the lower and upper confidence-interval bounds.
7. Suppose that common random numbers are used to obtain independent macro-replication estimators \( \tilde{\theta}_j^{(A)}, \tilde{\theta}_j^{(B)} \), \( j = 1, 2, \ldots, k \), for two systems, \( A \) and \( B \). We wish to estimate the difference \( \theta^{(A)} - \theta^{(B)} \).

(a) (five points) Suggest a point estimator.

\[
\tilde{\theta}^{(A)} - \tilde{\theta}^{(B)}, \text{ where } \tilde{\theta}^{(A)} = \sum_{j=1}^{k} \frac{\tilde{\theta}_j^{(A)}}{k} \text{ and } \tilde{\theta}^{(B)} = \sum_{j=1}^{k} \frac{\tilde{\theta}_j^{(B)}}{k}.
\]

or, equivalently,

\[
\bar{D} = \sum_{j=1}^{k} \frac{D_j}{k}, \text{ where } D_j = \tilde{\theta}_j^{(A)} - \tilde{\theta}_j^{(B)}.
\]

(b) (six points) Suggest a method to estimate the standard error of your point estimator.

Let \( D_j = \tilde{\theta}_j^{(A)} - \tilde{\theta}_j^{(B)} \) for \( j = 1, 2, \ldots, k \).

Then \( \bar{D} = \sum_{j=1}^{k} \frac{D_j}{k} \), the average of \( k \) independent differences.

Then \( \text{var}(\bar{D}) = \frac{\left(\sum_{j=1}^{k} D_j^2\right) - k \bar{D}^2}{(k-1)(k)} \). Taking the sqrt yields \( \text{sfe}(\bar{D}) \).

(Comment: This is the analysis of the paired-T test.)

(c) (six points) Do you have enough information to know whether common random numbers will reduce the standard error (compared to independent sampling)? Discuss.

No.

The nature of \( \theta_j \) is unknown. The method of random-variate generation is unknown.

(d) (six points) Which answers of (a), (b), and (c) above change if the macro replications are not independent? That is, suppose that \( (\tilde{\theta}_1^{(A)}, \tilde{\theta}_1^{(B)}) \) is not independent of \( (\tilde{\theta}_2^{(A)}, \tilde{\theta}_2^{(B)}) \).

Only the answer to Part (b) changes.

(Comment: Dependence could be because of antithetic variates being used with crn.)