George Mason University
CS 700

<div style="text-align: right">

Mid-term Exam
March 16, 2004

</div>

Important Notes:
1. This is an open book/notes exam.
2. You can take your time answering the questions. I estimate that you should be able to finish the exam in 90 minutes, but you have until 7:10 pm to complete the exam.
3. There are 5 questions in all. The points for each question are indicated in parentheses at the end of the question.
4. In answering the questions, make sure you show all the steps taken. That way even if your calculations end up being wrong you will get credit for using the correct approach.
5. Please answer the questions in the space provided below the question. If you need additional space use the blank pages at the end. You can use your own paper for rough work if necessary.

Q1. Two computer systems A and B were compared using several benchmark programs as shown in the table below:

Benchmark Program	System A (Runtime in Seconds)	System B (Runtime in Seconds)
Payroll	43	55
Linear programming (simplex)	201	195
Least Squares	192	204
Queuing Network Solver	52	40
Puzzle	10	12
Simulation	305	290
Statistical test	10	13
Synthetic 1	1	1
Synthetic 2	350	320
Synthetic 3	59	65

What metric would you use if you had to characterize the performance of systems A and B with a single number? Compute that metric for systems A and B, and use it compare the relative performance of the two systems. Justify the metric you selected and explain why you believe it is more suitable than other metrics for characterizing the performance of the systems. ***(10)***

A1

Q2. Consider the benchmark data reported in the table in Q1. Based on this data, can you state with 90% confidence that one system has better performance than the other? Show the steps used to derive your answer.   *(15)*

A2

Q3. Given the following twenty measurements of the mean queue length of a CPU queue, compute the best estimates of the mean queue length and variance of the queue length.

3.00	2.87	3.58	3.28	3.87
4.14	5.23	3.86	2.88	4.37
4.75	4.33	3.17	2.85	4.16
4.03	3.57	3.68	3.95	3.58

1. Find the 90% confidence intervals for the mean and variance of the mean queue length. Assume that the queue length distribution is approximately normal.
2. Suppose our goal is to estimate the mean queue length within 1% of the true value with 99% confidence. How many measurements of the mean queue length are needed to achieve this goal?

(25)

A3.

Q4. Elapsed times for a program were measured on two different computers. The program was run 15 times on both systems and the following statistics were gathered:

On system A, mean run time = 104 seconds, sample variance = 290 seconds^2

On system B, mean run time = 114 seconds, sample variance = 510 seconds^2

What can you say about the relative performance of the program on systems A and B for a confidence level of 95%? (***20***)

A4

Q5. Associated with a job are two random variables: CPU time required (Y) and the number of disk I/O operations (X). Consider the following datapoints:

i	Time (sec) Y_i	Number X_i
1	40	396
2	38	390
3	42	410
4	50	502
5	30	305
6	60	590
7	20	210
8	25	252
9	40	398
10	39	392

Assume we want to predict a CPU time requirement given an I/O request count. Performa linear regression $Y = aX + b$ for this purpose.

1. Compute point estimates of a and $b,$ as well as 90% confidence intervals.
2. Suppose we want to predict the CPU time if the number of I/O operations = 800. Find an estimate of the predicted CPU time and the associated 90% confidence intervals.

(30)

A5