上 海 交 通 大 学 试 卷 （A卷）

（ 2019 至 2020 学年 第 一 学期 ）

班级号 ____________________ 学号 ____________________ 姓名 __________

课程名称 MG26018 仿真建模与分析 ____________________ 成绩 __________

注意事项

1 考试时间：6:00 PM – 8:20 PM。

2 共十题，满分 100 分。题中的数学符号（记号）、概念均遵从课件上的定义。

3 用中文或英文，在题目下方的空白处作答。

4 本次考试为开卷笔试，允许带的材料为：课件、作业、手写笔记、计算器、文具。

5 使用未经允许的材料及电子设备，或未经允许走动、交谈、传递物品，即视为作弊。作弊行为一律上报至学校处理。

6 考试开始 30 分钟内，以及考试结束前 30 分钟内，不得交卷。

我承诺，我将严格遵守考试纪律。承诺人： ______________

– 此页为空白页 –

第一题 [4 分]

为解决实际问题，我们经常会为研究的系统建立数学模型。如果引入较多的假设，那么该模型往往可以通过数学工具求得解析解；如果引入较少的假设，那么该模型往往只能通过一些数值方法（如，运行仿真）来求得数值解。请简述这两类模型各自的优点和缺点。

第二题 [5 分]

Weibull (α, β) 分布的概率密度函数（pdf）为 $f(x) = \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^{\alpha}}$，$x > 0, \alpha > 0, \beta > 0$。请使用反变换技术（inverse-transform technique）生成服从 Weibull $(1/2, 1/4)$ 分布的随机变量。使用随机数 0.8147，0.9058，0.1270，0.9134 来产生 4 个观测值 (random variates)。（保留 4 位小数。）

第三题 [12 分]

我们来为一个 $G/G/2$ 排队模型进行手动的仿真。下面表格中包含了我们已经生成好的 5 名顾客的到达间隔（interarrival times），服务台（server）1 和 2 中的服务时间（单位是一致的）。

到达间隔	2	1.5	2	0.5	1
服务台 1 的服务时间	3	5	2.5	2	$\cdots$
服务台 2 的服务时间	4	2	3	8	$\cdots$

这里服务台的服务时间，与顾客无关。比如说服务台 1，第一个进入服务台 1 的顾客（无论是谁）的服务时间为 3，第二个进入服务台 1 的顾客（无论是谁）的服务时间为 5，依次类推。服务台 2 也是如此。如服务台 1 和 2 都空闲，优先去服务台 1。

（1）利用这些数据来完成仿真，即，填充下表。（注意：运行仿真直至 **5** 名顾客全部离开；请根据需要自行延长下面的表格。按顺序依次使用上表中的数据；生成好的服务时间可能有多余。下表中 System State 表示系统中顾客的数量。） [8 分]

Clock	System State	Event Calendar		
		Next Arrival	Next Departure	
			Server 1	Server 2
0	0	$0+2=2$	∞	∞
2	1	$2+1.5=3.5$	$2+3=5$	∞

（2）计算，从 0 时刻到最后一个顾客离开这段时间内，系统中的平均人数和人均逗留时长。[4 分]

第四题 [10 分]
有一个机器一次可以生产一个产品。每一个产品花费的时间服从指数分布 $\mathrm{Exp}(a)$。生产出来的产品存放在一个容量为 b 的仓库中。当仓库满的时候，机器会马上被关掉；当仓库有剩余空间的时候，机器会马上被启动。顾客以泊松过程（速率为 c）来到仓库取产品（每人只取一件，取货时间可忽略；若仓库为空，顾客即刻离开）。
（1）如果我关注的是仓库中产品的数量，这个问题可不可以用课件 2 中的某一个经典排队模型来刻画？[2 分]

（2）如果可以，请写出这个排队模型及其对应的参数，并解释为什么可以；如果不可以，请添加或修改一些假设，使其满足课件 2 中某一个经典排队模型，并写出这个模型及其对应的参数。[8 分]

第五题 [5 分]
我们观察一个系统，记录了时间 $[0,T]$ 之间所有顾客到达的时间点 $\{s_1, s_2, \ldots, s_n\}$。我们想要用统计检验（第一类错误的概率 ≤ 0.05）的方式来判断顾客的到达过程是不是一个（近似的）泊松过程，但不想估计任何的参数，又不想人为选取任何参数。请问该使用哪种统计检验？ H_0 和 H_1 分别是什么？请简述检验的步骤，并写出**具体**的算式。（假设检验统计量的分布及其分位数我们都知道。）

第六题 [18 分]

为以下三种排队模型分别求解 L, W, L_Q, W_Q:

（1）$M/M/1$，$\lambda = 0.6$，$\mu = 1$。[2 分]

（2）$M/M/2$，$\lambda = 0.6$，$\mu = 0.5$。（和（1）相比，把一个快的服务台拆为两个慢的服务台。）[3 分]

（3）$M/G/1$，$\lambda = 0.6$，服务时间服从 Uniform $[0, 2]$。（和（1）相比，服务时间的分布改变了，但均值没变。注：如果 $X \sim \text{Uniform}[a, b]$，$\text{Var}(X) = \frac{(b-a)^2}{12}$。）[3 分]

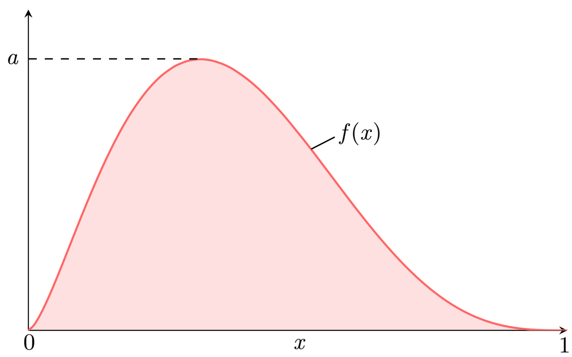
根据上面的计算结果，回答以下问题：

（4）（1）中的 $M/M/1$ 和（2）中的 $M/M/2$ 相比，哪一个系统总的运作效率更高？依据是什么？此外，再为这效率上的差别提供一个**直观的**解释，什么原因导致这样的差异？[5 分]

（5）（1）中的 $M/M/1$ 和（3）中的 $M/G/1$ 相比，哪一个系统总的运作效率更高？依据是什么？此外，再为这效率上的差别提供一个**直观的**解释，什么原因导致这样的差异？[5 分]

第七题 [8 分]

已知 $f(x)$，$0 \le x \le 1$，是随机变量 X 的概率密度函数（pdf），如下图所示。假设我们知道 $f(x)$ 的具体表达式。现在我们想在阴影所示的二维区域中均匀且随机地采点。假设我们运行函数 `Unif(a,b)` 即可生成 $[a,b]$ 之间均匀分布的随机样本。



（1）如果我们已经知道如何生成 X 的随机样本。比如，每次运行一次 `RANDX()` 函数，就可以得到 X 的一个随机样本。**利用**这个函数，如何均匀且随机地采点？[4 分]

（2）如果我们不知道如何生成 X 的随机样本，并且 X 的累积分布函数（cdf）$F(x)$ 无显式的反函数。在这种条件下，如何均匀且随机地采点？[4 分]

第八题 [12 分]

假设 $x_1, x_2, \ldots, x_n$ 是来自正态分布 $\mathcal{N}(\mu, \sigma^2)$ 的独立同分布样本点。用极大似然估计（MLE）方法来估计 μ 和 σ^2。（写出具体的推导步骤；注意严密性。）

第九题 [10 分]

我们有 $k > 2$ 个系统设计，其期望性能指标（mean performance）为 θ_i, $i = 1, 2, \ldots, k$。现在需要从中选出 θ_i 最大的那个。课件 8 中的 Bechhofer's Procedure（第 16 页），可以保证，当假设 1-4（见课件 8 第 15 页）都满足的时候，$\mathbb{P}\{\text{select the larget } \theta_i\} \geq 1-\alpha$。它的证明请见课件 8 第 17-18 页。现在我们取消假设 3。请给出严格的数学（概率）证明，当假设 1，2，4 满足的时候，

$$\mathbb{P}\Big\{\Big|\text{selected } \theta_i - \max_{1\leq i\leq k} \theta_i\Big| < \delta\Big\} \geq 1 - \alpha.$$

第十题 [16 分]

假设我们运行一个稳态的仿真，在一次运行中依次收集到离散输出 $Y_1, Y_2, \ldots, Y_n$。假设初始化偏差的影响可以忽略不计。我们用 $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ 作为稳态性能指标（steady-state performance measure，假设它为 ϕ）的点估计。一个容易犯的错误是，计算 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$，然后用 $\bar{Y} \pm t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}$ 作为该稳态性能指标的 $1-\alpha$ 置信区间。

（1）为什么这样是错的？请简述原因。 [4 分]

（2）假设 $Y_1, Y_2, \ldots, Y_n$ 是同分布但是正相关，请为如下的两个结果给出证明：$\mathbb{E}[S^2] < \mathrm{Var}(Y_1)$，$\mathbb{E}[S^2/n] < \mathrm{Var}(\bar{Y})$。[10 分]

（3）对于（2）中的情形，如果用 $\bar{Y} \pm t_{n-1,1-\alpha/2}\frac{S}{\sqrt{n}}$ 作置信区间，后果是什么？[2 分]