

云南大学数学与统计学院

实验报告

实验课名称： 随机过程实验

指导教师： 韩博 王晓波

专业（年级）： 统计学 2021 级

学生姓名： 枫叶 学号：

实验名称： 复合泊松过程的随机模拟

实验成绩：

《随机过程实验》实验报告_6

实验名称	复合泊松过程的随机模拟		实验成绩		
学号			姓名	枫叶	
实验时间	2024 年 4 月 22 日	实验地点	格物楼 3508	指导教师	韩博 王晓波

一、实验目的

学习使用 R 软件对复合泊松过程进行模拟和计算。

二、实验要求

1. 对所使用的方法与所得到的结果进行适当的文字描述。
2. 在实验结果的相应部分附上完整的代码与适当的注释。
3. 采用一定的可视化方法体现出对应计算结果。

三、实验内容

1、设游客以每小时 50 人的平均速率走进拉斯维加斯的一家赌场，这一过程可以用 Poisson 过程来描述。其中 10% 的游客不会参与赌博，且每位游客是否参与赌博互不影响，也与进入赌场的游客数无关。假设每位参与赌博的游客在赌场损失的金额服从均匀分布 $U[0, \$1500]$ ，且每位游客损失的金额是独立的。

- (1) 假定从营业开始，有 20 位游客进入赌场，请进行一次模拟实验计算截止到第 20 位游客到达，赌场的总收益是多少，并绘制该复合 Poisson 过程的样本路径（横轴为顾客到达时间 t ，纵轴为赌场在 $(0, t]$ 时间内的收益）；
- (2) 写出 12 小时内在该赌场赌博游客数的均值，并通过 1000 次模拟实验得到经验均值，随后进行比较；

(3) 写出赌场 12 小时内的平均收益以及收益的方差，并通过 1000 次模拟实验得到经验均值和经验方差，随后进行比较。

2、到达加油站的汽车数目可以用参数为 λ 的泊松过程来描述。设每个司机购买汽油的行为是独立的，购买汽油的金额价格服从 Gamma 分布 $\Gamma(\alpha, \beta)$ ，花费的金额和汽车的数量是独立的。汽车到达时刻（分钟）和花费金额的数据见下表：

Arrival Time	Amount Spent, in \$	Arrival Time	Amount Spent, in \$	Arrival Time	Amount Spent, in \$
0.15	23.67	28.81	69.67	47.94	20.38
3.81	25.55	32.36	25.39	49.73	34.95
5.67	38.54	32.76	30.86	50.72	29.23
6.61	31.31	32.92	50.53	50.86	36.51
13.14	74.20	33.22	24.93	51.99	37.77
13.57	32.78	33.51	27.49	52.36	34.41
15.68	29.70	34.40	22.56	52.89	23.35
22.83	35.83	35.76	21.38	53.64	32.95
23.35	22.17	39.08	45.53	55.03	21.27
23.77	34.96	41.03	39.14	55.29	37.32
23.77	24.20	42.05	26.02	56.82	20.30
24.69	26.01	42.38	21.35	63.02	32.59
26.94	24.07	45.66	33.88		

(1) 该加油站的营业额可以用复合 Poisson 过程来描述，请写出建模过程，并给出每个参数/符号的解释；

(2) 将表中数据带入模型，估计模型中的未知参数 λ, α, β ；

(3) 绘制时间间隔的直方图和拟合的分布曲线。

四、 实验软件

R 语言

五、 实验结果

加载包

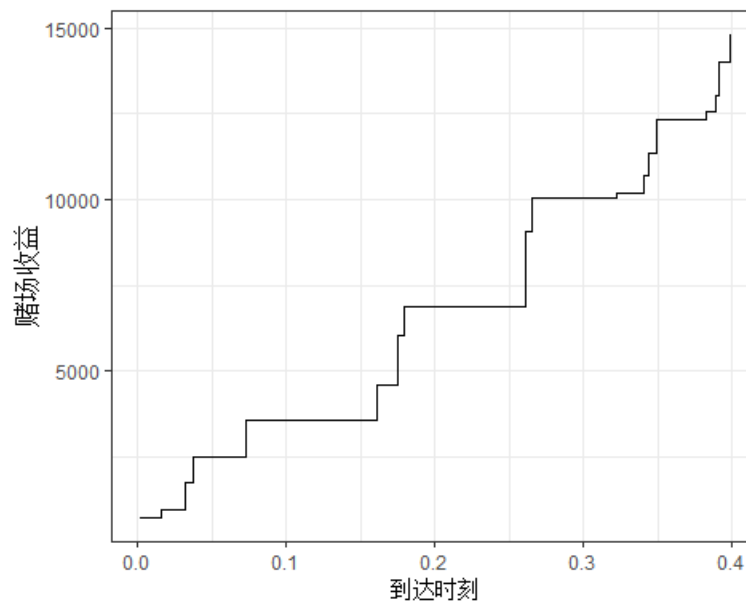
```
library(dplyr)
library(purrr)
library(ggplot2)
library(readxl)
```

第一题

第一问

#沿用上次实验的函数

```
poisson_gen <- function(n=1,lambda=10,maxtime=10,set_seed=NA){
  maxperson <- 2*lambda*maxtime
  if(!is.na(set_seed)) set.seed(set_seed)
  data.frame(id = n,
             Time = rexp(maxperson,lambda) %>% cumsum(),
             persons = 1:maxperson) %>%
    filter(Time <= maxtime)
}
data <- poisson_gen(n = 1,lambda = 50,maxtime = 1) %>%
  filter(persons<=20) %>%
  mutate(involve=sample(c(1,0),20,T,c(0.9,0.1)),
         loss=involve*runif(20,0,1500),
         casino_income=cumsum(loss))
ggplot(data) +
  geom_step(aes(x=Time,y=casino_income)) +
  xlab("到达时刻") +
  ylab("赌场收益") +
  theme_bw()
```



某次模拟中得到赌场收益为 14807.71，该复合 Poisson 过程的样本路径见上图

第二问

```
map(1:1000,poisson_gen,lambda = 50,maxtime = 12) %>%
  list_rbind() %>%
  group_by(id) %>%
  summarise(persons=max(persons)) %>%
  summarise(mean_persons=mean(persons)) %>%
  .$mean_persons -> mean_persons
```

理论均值为 600 人，一千次模拟实验的均值为 599.41，与理论均值相差 0.59，这里为了比较差异，保留了模拟均值的小数

第三问

```
final_data <- map(1:1000,poisson_gen,lambda = 50,maxtime = 12) %>%
  list_rbind() %>%
  group_by(id) %>%
  mutate(involve=sample(c(1,0),max(persons),T,c(0.9,0.1)),
         loss=involve*runif(max(persons),0,1500),
         casino_income=cumsum(loss)) %>%
  summarise(casino_income=max(casino_income)) %>%
  ungroup() %>%
  summarise(mean_income=mean(casino_income),
            var_income=var(casino_income))
```

理论平均收益为 405000，模拟得到的经验平均收益为 403813.27，二者相差 0.293%

理论方差为 405000000，模拟得到的经验方差为 401549790，二者相差 0.852%

第二题

第一问

t 时刻总共达到加油站的汽车数记为 $N(t)$ ，其为 Poisson 过程，再记第 i 个司机购买汽油的金额为 Y_i ，则营业额为 $X(t) = \sum_{i=1}^{N(t)} Y_i$

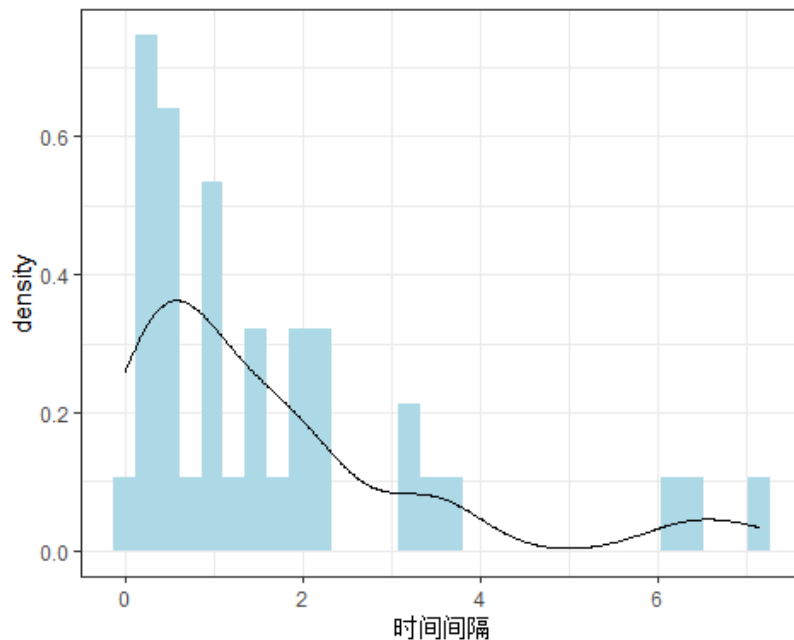
第二问

```
data <- read_xlsx("D:/预删除文件夹/大三下/随机过程/随机过程实验6数据.xlsx")
lambda <- mean(c(0.15,diff(data$AT)))
alpha <- mean(data$AS)^2/var(data$AS)
beta <- mean(data$AS)/var(data$AS)
```

矩估计结果为 $\lambda = 1.6584211$, $\alpha = 7.2928754$, $\beta = 0.2266443$, 其中为了充分利用样本信息, 基于时间间隔的指数分布进行 λ 的矩估计

第三问

```
diff_time <- data.frame(diff_time=c(0.15,diff(data$AT)))
ggplot(diff_time) +
  geom_histogram(aes(x=diff_time,y=after_stat(density)),fill="lightblue") +
  geom_density(aes(x=diff_time)) +
  xlab("时间间隔") +
  theme_bw()
```



样本的密度曲线大致与指数分布的理论密度曲线相近

六、 教师点评