

第二次个人作业

齐柏珂 统计学院 2015201572

题目 1:

势函数 $g(\lambda) = \begin{cases} \alpha(\lambda) & 1 < \lambda < 2 \\ 1 - \beta(\lambda) & 0 < \lambda < 1 \end{cases}$ 由于多个泊松分布的线性组合仍服从泊松分布，所以检验统计量 T 服从参数为 (n, λ) 的泊松分布。

(1) 固定 $n=10$, 选择不同的 C :

【代码】

```
lambda <- seq(0.01,2.00,by=0.01) ##生成一个向量作为 X 轴

power1 <- ppois(5,10*lambda)## 生成 C=5 时的分布函数

power2 <- ppois(7,10*lambda)## 生成 C=7 时的分布函数

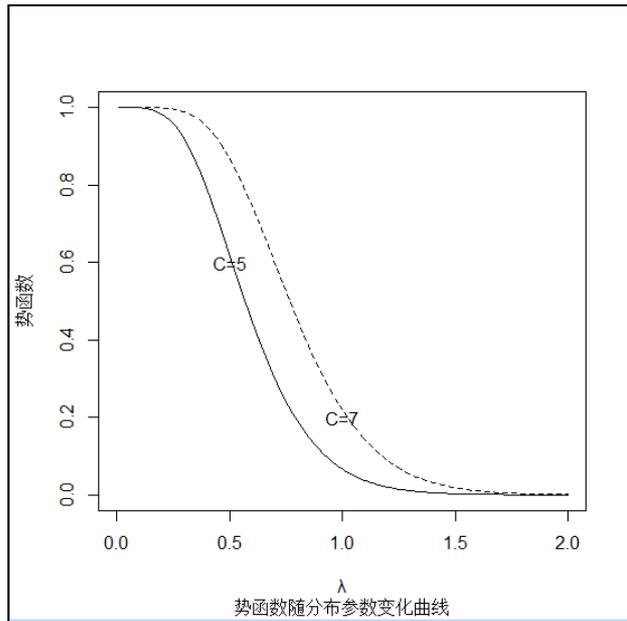
plot(lambda,power1,type="l",xlab="λ",ylab="势函数",sub="势函数随分布参数变化曲线",xlim=c(0,2)) ##用 plot 画图 在图形上添加坐标轴标签、副标题、并限定 X 轴 取值范围

lines(lambda,power2,type="l",lty=2)##使用 lines()添加 C=7 时的势函数曲线

text(0.5,0.6,"C=5")##添加文本标注

text(1.0,0.2,"C=7")
```

【运行结果】

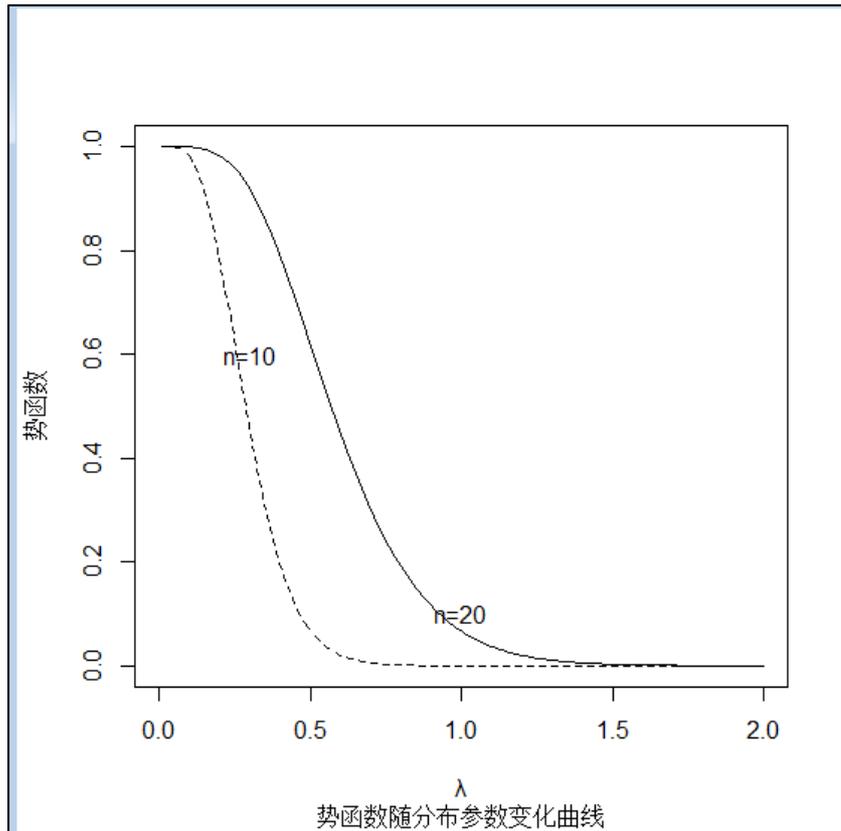


(2) 固定 $C=5$, 选择不同的 n , 只需修改上述代码的个别变量

【代码】

```
lambda <- seq(0.01,2.00,by=0.01)
power1 <- ppois(5,10*lambda)
power2 <- ppois(5,20*lambda)
plot(lambda,power1,type="l",xlab="λ",ylab="势函数",sub="势函数随分布参数变化曲线",xlim=c(0,2))
lines(lambda,power2,type="l",lty=2)
text(0.3,0.6,"n=10")
text(1.0,0.1,"n=20")
```

【运行结果】



T 1.7

(1) 在 $H_0: F=G$ 之下, X_i 和 Y_i 独立同分布, 则 $(X_i - Y_i)$ 关于 0 对称分布, $\theta = P$

$$P(X_1 + X_2 < Y_1 + Y_2) = P(X_1 - Y_1 + X_2 - Y_2 < 0) = P(X_1 - Y_1 + X_2 - Y_2 > 0) = 1/2$$

(2) 令 $h(X_1, X_2, Y_1, Y_2) = I(X_1 + X_2 < Y_1 + Y_2) = \begin{cases} 1, & x_1 + x_2 < y_1 + y_2 \\ 0, & \text{其他} \end{cases}$

易知: $E(h(X_1, X_2, Y_1, Y_2)) = \theta$, 由 $h(X_1, X_2, Y_1, Y_2)$ 张成的 U 统计量定义为

$$U_{mn} = \frac{1}{\binom{2}{m}\binom{2}{n}} \sum_{(1,2)} \sum_{(1,2)} h(X_1, X_2, Y_1, Y_2) = \frac{4}{n(n-1)m(m-1)} \sum_{i!=j} \sum_{h!=k} I(X_1 + X_2 < Y_1 + Y_2)$$

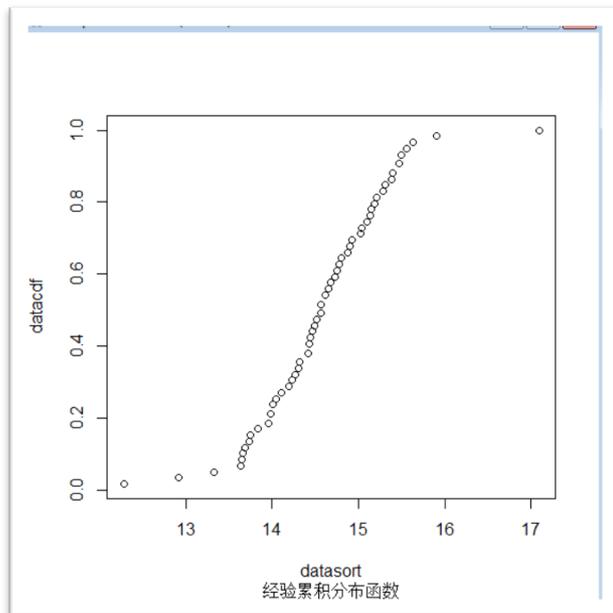
T1.11

画经验分布函数既可以通过编写程序, 也可以调用函数 `ecdf()`.

【经验累积分布函数 代码 1】

```
beenswax <- read.table(file.choose(),head=T) ##读取数据,并指明第一行包含变量名  
datasort <- sort(beenswax$Hydrocarbon) ##对碳氢化合物的数据排序  
datarank <- rank(datasort) ##返回数据的次序  
datacdf <- datarank/length(datarank) ##生成Y轴数据  
plot(datasort,datacdf,sub="经验累积分布函数") ##画经验累积分布函数
```

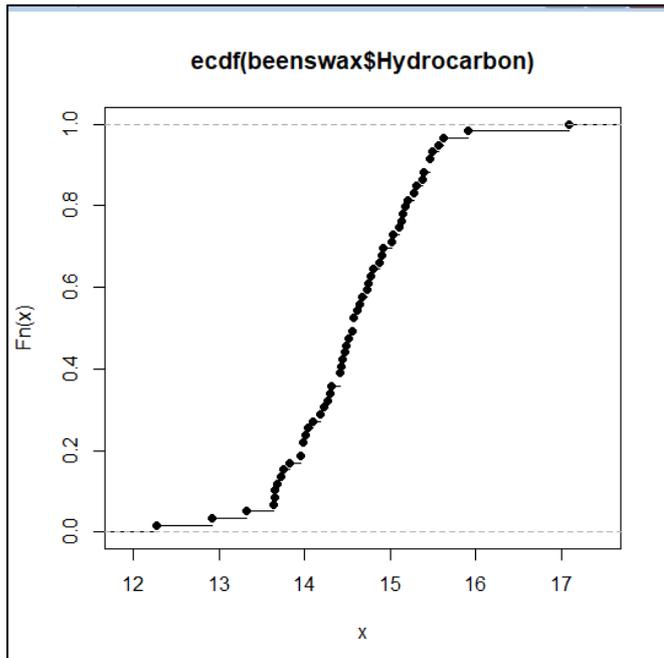
【运行结果】



【代码 2】

```
beenswax <- read.table(file.choose(),head=T)  
plot(ecdf(beenswax$Hydrocarbon))
```

【运行结果】

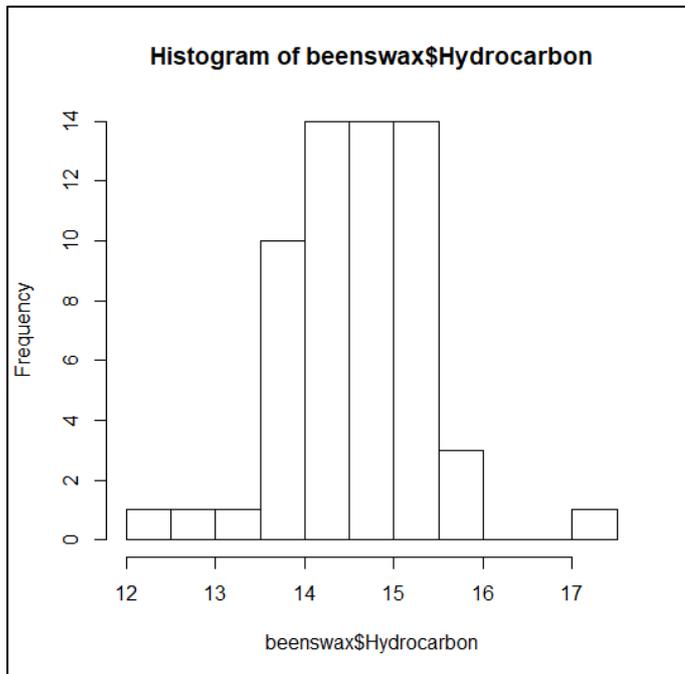


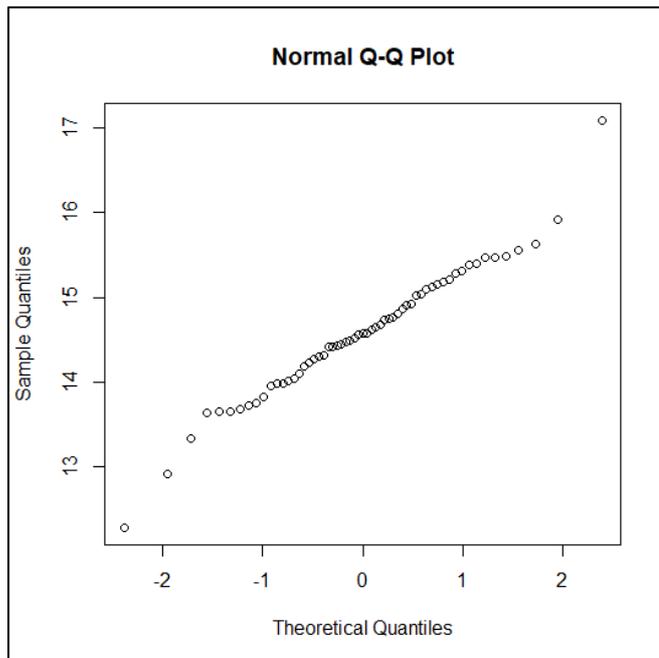
【直方图、QQ 图代码】

```
hist(beenswax$Hydrocarbon)
```

```
qqnorm(beenswax$Hydrocarbon)
```

【运行结果】





(2)

求分位数可以使用 R 中的 `quantile()` 函数

【代码】

```
quantile(beenswax$Hydrocarbon,c(0.90,0.75,0.50,0.25,0.10))
```

【运行结果】

```
 90%    75%    50%    25%    10%
15.470 15.115 14.570 14.070 13.676
```

(3) 通过对经验累积分布图，直方图，QQ 图的观察来判定数据是否服从高斯分布

【代码】

```
par(mfrow=c(3,1))

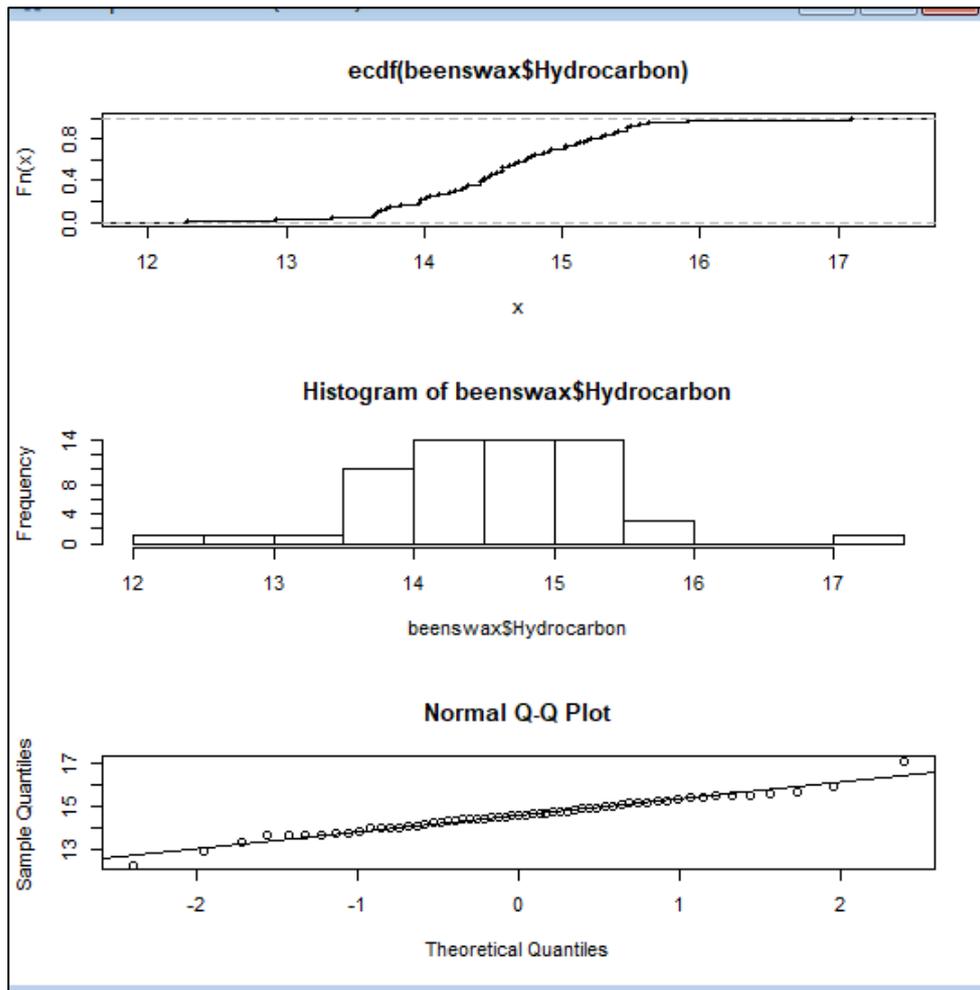
plot(ecdf(beenswax$Hydrocarbon))

hist(beenswax$Hydrocarbon)

qqnorm(beenswax$Hydrocarbon)

qqline(beenswax$Hydrocarbon)
```

【运行结果】



由 QQ 图可知，数据基本来自于高斯分布。

T1.12

由于 $S_n(t) = 1 - F_n(t)$ ，可用经验分布函数估计生存函数，表示寿命超过 t 的数据所占比例。

用 R 画出经验分布函数即可分析出不同治疗方法的作用效果是否有差异

【代码】

```

none <- c(174,224,260,255,165,237,191,100,115,189)

placebo <- c(263,213,231,291,168,121,137,102,89,433)

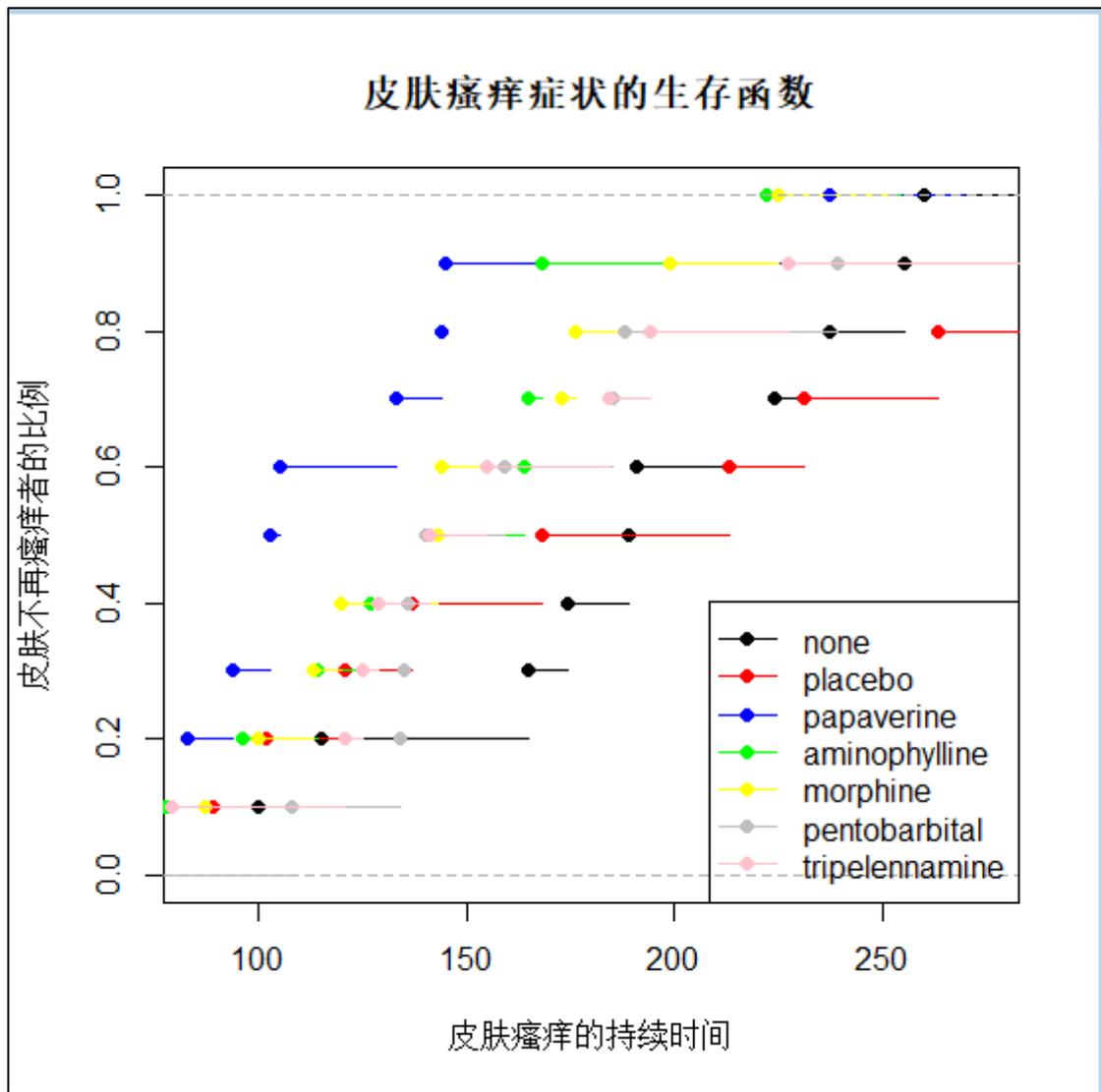
papaverine <- c(105,103,145,103,144,94,35,133,83,237)

aminophylline <- c(141,168,78,164,127,114,96,222,165,168)

```

```
morphine <-c(199,143,113,225,176,144,87,120,100,173)
pentobarbital <-c(108,341,159,135,239,136,140,134,185,188)
tripelennamine <-c(141,184,125,227,194,155,121,129,79,317)
sort <- sort(none)
rank <- rank(sort)
y <- rank/length(rank)
plot(sort,y, main="皮肤瘙痒症状的生存函数",xlab="皮肤瘙痒的持续时间",ylab="皮肤不再瘙痒者的比例")
lines(ecdf(placebo),col="red")
lines(ecdf(papaverine),col="blue")
lines(ecdf(aminophylline),col="green")
lines(ecdf(morphine),col="yellow")
lines(ecdf(pentobarbital),col="gray")
lines(ecdf(tripelennamine),col="pink")
legend("bottomright",c("none","placebo","papaverine","aminophylline","morphine",
,"pentobarbital","tripelennamine"),lty=rep(1,length=7),pch=rep(16,length=7),col=c
("black","red","blue","green","yellow","gray","pink"))
```

【运行结果】



由运行结果可知，药物 papaverine 的治疗效果最好，其余四种不相上下。