

花旗松原始森林特征研究项目代码

1. 分析样本区域的地市特征，绘制出等高线图，并分析还原三维树种分布图

#代码的解释放于相应代码的上方

#加载 MASS 包

```
library(MASS)
```

#读入研究数据并将数据集存入为d

```
d<-read.csv("forest.csv",header=T)
```

#查看前六列数据

```
head(d)
```

#链接到d

```
attach(d)
```

#估计数据的密度函数

```
a<-kde2d(X,Y,n=80,lims = c(range(X),range(Y)))
```

#等高线图高度相等的各点连成的闭合曲线，等高线上标注的数字为该等高线的海拔高度，等高线密集的地方代表山势比较陡，越密集越陡

```
contour(a,main="地势的等高线图")
```

#从山坡三维地势可以看出山坡连绵不断

```
persp(a,main="山坡三维地势",col="chartreuse4")
```

#安装 rgl 安装包

```
install.packages("rgl")
```

#加载 rgl 安装包

```
library(rgl)
```

#在数据d中加多一列col，代表CD对的颜色为红色，DF对的颜色为绿色,HL对的颜色为蓝色

```
d$col<-ifelse(d$SPECIES=="CD",2,  
             ifelse(d$SPECIES=="DF",3,4))
```

#红雪松的颜色为红色，花旗松的颜色为绿色，西部铁杉的颜色为蓝色

```
plot3d(X,Y,Z,col=d$col,main="树木分布大概状况")
```

2. 分析验证 HTC（树冠的高度，即光合作用发生的地方）与 HTT（树高）的关系，以及三个树种间这种关系的差异

#对d添加HTC列， $HTC=HTT-HTB$

```
d$HTC<-d$HTT-d$HTB
```

#处理缺失值后新数据并赋予dd，去除那些死去的树木的数据，dd为成活下来树木的数据

```
dd<-na.omit(d)
```

#查看dd数据前六列

```
head(dd)
```

#树冠散点图，呈线性递增趋势

```
plot(dd$HTT,dd$HTC,main="HTC和HTT的散点图")
```

#对HTC和HTT做一元线性回归模型

```
lm<-lm(HTC~HTT,data=dd)
```

#对HTC和HTT做幂函数回归模型

```
lm.p=lm(log(HTC)~log(HTT),data=dd)
#一元线性回归模型决定系数
summary(lm)$r.sq
#由两模型的决定系数可知，幂函数回归模型优化度、拟合度都比较好
summary(lm.p)$r.sq
#将窗口分成4分
par(mfrow=c(2,2))
#幂函数回归模型残差诊断图
plot(lm.p)
#查看dd$SPECIES的水平，结果："CD" "DF" "HL"
levels(dd$SPECIES)
#把dd$SPECIES=CD那些行的数据赋予cd、df、hl
cd<-dd[dd$SPECIES=="CD",]
df<-dd[dd$SPECIES=="DF",]
hl<-dd[dd$SPECIES=="HL",]

#红雪松树冠和树高的相关系数：0.9769148
cor(cd$HTC,cd$HTT)
#花旗松树冠和树高的相关系数：0.5109884
cor(df$HTC,df$HTT)
#西部铁杉树冠和树高的相关系数：0.9333573
cor(hl$HTC,hl$HTT)
#根据三个模型的树种各自的树冠和树高的相关系数，可知该关系与树种有差异，红雪松和西部铁杉树的树冠受树高影响比较大。
```

3. 分析死树对近邻的其它树种的 DBH（直径），HTT（树高），HTC（树冠高度）的相关性？检验辩证三个树种间这种关系的差异

```
#只抽取树木死亡的数据，并把数据赋予dead
dead<-d[(d$STATUS=="dead"),]
#查看数据dead的前六列
head(dead)
#图形区域只显示一个图
par(mfrow=c(1,1))
#红色代表红雪松死亡的位置，绿色代表花旗松死亡的位置，蓝色代表西部铁杉死亡的位置
plot(dead$X,dead$Y,main="死树所处的平面位置",xlab="x",ylab="y",col=dead$col)
#为成活下来树木进行分组，按x，y轴分组
dd$group<-ifelse(dd$X<=25&dd$Y<=40,1,
                 ifelse(dd$X<=25&dd$Y>40,2,
                 ifelse(dd$X<=50&dd$Y<=40,3,
                 ifelse(dd$X<=50&dd$Y>40,4,
                 ifelse(dd$X<=75&dd$Y<=40,5,
                 ifelse(dd$X<=75&dd$Y>40,6,
                 ifelse(dd$X<=max(dd$X)&dd$Y>40,7,8)))))))))
```

#从结果可以得到p 值=0.593，大于0.05，不能拒绝原假设，说明死树的存在与近邻的DBH影响不显著

```
anova(lm(dd$DBH~group,data=dd))
```

#从结果可以得到p 值=0.5521，大于0.05，不能拒绝原假设，说明死树的存在与近邻的HTT影响不显著

```
anova(lm(dd$HTT~group,data=dd))
```

#从结果可以得到p 值=0.1368，大于0.05，不能拒绝原假设，说明死树的存在与近邻的HTC影响不显著

```
anova(lm(dd$HTC~group,data=dd))
```

4. 验证树种的分布是否为均匀分布，各自集中分布的海拔的差异性

#用Kolmogorov-Smirnov 检验数据是否服从均匀分布。

#原假设: dd\$Z 具有均匀分布(dd\$Z 是存活的树木的山坡海拔高度)

#Ks 检验结果的p-value < 2.2e-16，其小于显著水平0.05，因此拒绝原假设，说明树种在山坡上的分布不是均匀

```
ks.test(dd$Z,"punif")
```

#三个树种的分布海拔情况

#由红雪松随海拔分布的直方图可知：红雪松分布特征比较均匀，但是在海拔地势偏中高的分布较多。

```
hist(cd$Z,xlim = range(cd$Z),main = "红雪松随海拔分布",xlab = "海拔高度",ylab = "红雪松的密度",nclass=30,freq=F)
```

#由花旗松随海拔分布的直方图可知：花旗松的分布特征在海拔地势比较高的地方。

```
hist(df$Z,xlim = range(df$Z),main = "花旗松随海拔分布",xlab = "海拔高度",ylab = "花旗松的密度",nclass=30,freq=F)
```

#由西部铁杉随海拔分布的直方图可知：西部铁杉的分布特征在海拔地5-15间。

```
hist(hl$Z,xlim = range(hl$Z),main = "西部铁杉随海拔分布",xlab = "海拔高度",ylab = "西部铁杉的密度",nclass=30,freq=F)
```