**6.9 R 语言应用**

**# 陈文贤 着 《大话统计学》 清华大学出版社 2022年**

set.seed(100)

X <- c(76, 78, 79, 81, 86) # 例题6.2

sample(X, 10, replace = TRUE) # 重置抽样

sample(X, 3) # 不重置抽样

data(iris) ; str(data) # 鸢尾花数据, 数据框格式, 150个样本, 5个变量

select<-sample(150,60) ; iris[select,] # 抽样 60 个样本

# 3.4.2 节 事件概率的计算 – 相对次数

# 例题3.16 两个骰子 点数和是 7 的概率 = 6 / 36

mean(replicate(10000,{D2<-sample(x = 1:6,size=2,replace=TRUE);sum(D2)==7}))

# 两个骰子 都是红色 {1} 或 {4} 的概率 = 4 / 36

mean(replicate(10000,{D2<-sample(x = 1:6,size = 2,replace = TRUE);

((D2[1]==1)+(D2[1]==4))\*((D2[2]==1)+(D2[2]==4))})) #

# 两个骰子 第一个是红色{1,4} , 第二个是黑色{2,3,5,6}的概率 = 8/36

mean(replicate(10000,{D2<-sample(x = 1:6,size = 2,replace = TRUE);

((D2[1]==1)+(D2[1]==4))\*((D2[2]==2)+(D2[2]==3)+(D2[2]==5)+(D2[2]==6))}))

# 抽样25人, 当中没有两人相同生日的概率 (重复10000次抽样)

mean(replicate(n=10000,anyDuplicated(sample(x=1:365, size=25, replace=TRUE)))==0)

# 中心极限定理

if(!require(ConvergenceConcepts)){install.packages("ConvergenceConcepts")}

library(ConvergenceConcepts)

investigate()

# B(n, p) # 中心极限定理

p <- 0.3 ; n <- c(5, 10, 60, 100)

k <- 3000 ; par(mfrow=c(2,2))

for (i in n) { Zn <- numeric(k)

for (j in 1:k) {Xn = rbinom(1, i, p)

Zn[j] = (Xn/i - p) / sqrt(p\*(1-p)/i) }

hist(Zn, breaks =25, main=paste(" n = ", i), col =2 ) }

# χ2(n) # 中心极限定理

rand = function(n) { (cumsum(rchisq(n,1))-(1:n))/sqrt(2\*(1:n)) }

data =generate(randomgen=rand, nmax=2000,M=500)$data

rmin=-4 ; rmax=4 ; par(mfrow=c(2,2))

for (nn in c(1, 5, 10, 100)) {

plot.ecdf(ecdf(data[,nn]), do.points=FALSE, xlim=c(rmin, rmax), col.h="blue",

main = paste(" n = ", nn))

curve(pnorm, xlim=c(rmin, rmax), lty =2, col=2,

main = " n = ", add=TRUE) }

par(mfrow=c(1,1))

law.plot3d(data, pnorm)

# 分布的拟合 distribution fitting , fitting of distribution

if(!require(fitdistrplus)){install.packages("fitdistrplus")} ; library(fitdistrplus)

x <- iris$Sepal.Length

descdist(x, boot=1000)

fit.n <- fitdist(x, "norm") ; summary(fit.n)

fit.w <- fitdist(x, "weibull") ; summary(fit.w)

fit.g <- fitdist(x, "gamma") ; summary(fit.g)

par(mfrow=c(2 ,2))

plot.legend <- c("normal", "weibull", "gamma")

denscomp(list(fit.n, fit.w, fit.g), legendtext=plot.legend)

qqcomp(list(fit.n, fit.w, fit.g), legendtext=plot.legend)

cdfcomp(list(fit.n, fit.w, fit.g), legendtext=plot.legend)

ppcomp(list(fit.n, fit.w, fit.g), legendtext=plot.legend)