**数据统计与分析第5章作业**

**学号： 姓名： 班级： 总分：**

要求写出解题过程。**请将作业转换为PDF后再上传到教学平台**

第1题

设在10个产品中有2个不合格品，若从中随机取出4个，则其中不合格品数*X*是离散随机变量，它仅可取0,1,2三个值。

(1) *X*取这些值得概率为多少？

(2) 对于同样的问题，若用放回抽样，则从10个产品(其中2个不合格品)种随机取出4个，其中不合格品数*Y*是另一个随机变量，它可取0,1,2,3,4五个值。*Y*取这些值的概率为多少？

第2题

在一次制造过程中，不合格品率为0.1，如今从成品中随机取出6个，记*X*为6个成品中不合格品数，则*X*服从二项分布*b*(6,0.1)，简记为。现研究如下三个问题：

(1) 恰有1个不合格品的概率是多少？

(2) 不超过1个不合格品的概率是多少？

(3) 二项分布*b*(6,0.1)的均值、方差、标准差分别是多少？

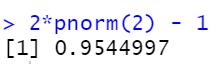
第3题

自动车床生产的零件长度*X*(毫米)服从，若零件的长度在毫米之间为合格品，求生产的零件是合格品的概率。

解：即是求概率



或者根据.注意这里X是总体。记从而



第4题

抽样调查表明，考生的外语成绩(总分为100分)近似服从正态分布，平均成绩为72分，96分以上占总数的2.3%。试求考生外语成绩在60分至84分之间的概率。

解：记考生成绩为score,则

根据

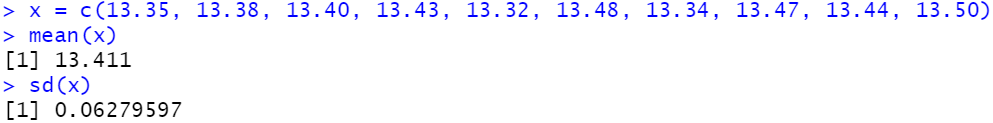
R 语言 2\*pnorm(1.995/2) – 1 = 0.6815

第5题

从某厂生产的一批铆钉中随机抽取10个，测得其直径(单位：毫米)分别为：13.35, 13.38, 13.40, 13.43, 13.32, 13.48, 13.34, 13.47, 13.44, 13.50。试求铆钉直径这一总体的均值的估计。

解：一、点估计

为的无偏估计,



二、区间估计

方差未知，，给定区间[a, b]使得，可以取

可以得到.

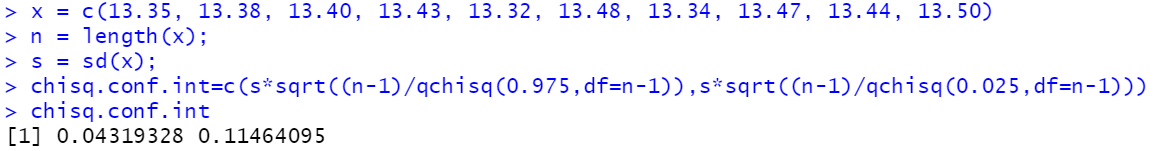


,即区间[13.36608,13.45592]包含总体均值的概率为95%。

同理，,给定区间[a,b]，使得，可取

*,*即区间[]包含的概率为95%

显然可得，区间



即区间[0.04319328, 0.11464095]包含的概率为95%。

第6题

续第3章习题3数据.

(1) 请用sample函数随机抽取月收入的100个数据并计算样本均值与样本方差；

(2) 请用sample函数对月收入分别进行样本容量为5,200的抽样，对每种样本容量各抽取1000次，画出各种样本容量下样本平均值的直方图(直方图如示例，作业中请去掉示例图中“抽样示例：”字样)，验证(说明理由)各样本容量下的抽样是否符合中心极限定理。

ch3\_3 = read.csv("ch3-3.csv")

income =ch3\_3$月收入;

sample200=seq(0,0,length=200);

meanSample200=seq(0,0,length=1000);

for (i in 1:1000){

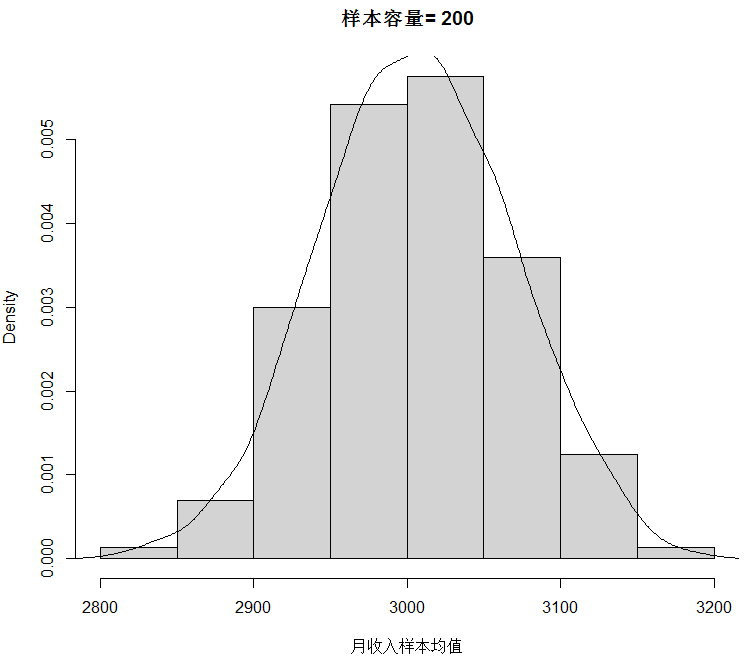
sample200 = sample(income,200,replace = T);

meanSample200[i] = mean(sample200);

}

hist(meanSample200,main=paste("样本容量=",200),xlab = "月收入样本均值",prob = T);

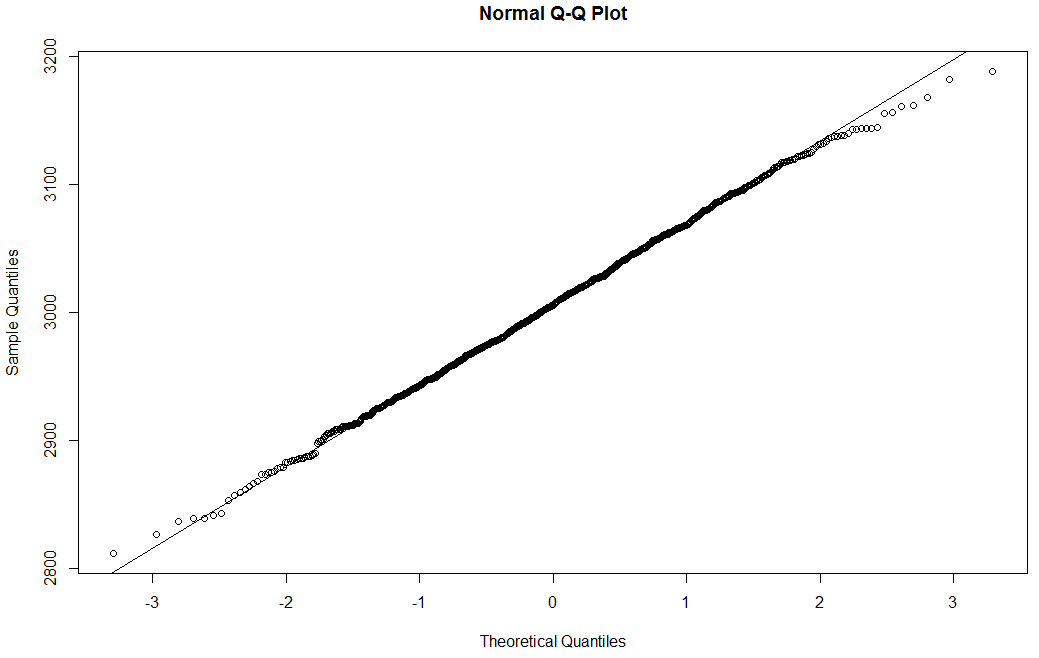
lines(density(meanSample200));



从直方图和密度曲线图看，样本均值的抽样分布大致是正态分布。下面在通过QQ图观察。

qqnorm(meanSample200)

qqline(meanSample200)

从QQ来看，样本均值的抽样分布大致符合正态分布。