第八次作业_参考答案

郑盼盼

2024-11-22

目录

例模拟投掷一枚均匀骰子 1000 次的结果，并计算结果的算术平均值

```
x <- sample(1:6,100,T)
bar_x <- mean(x)
bar_x
```

```
## [1] 3.48
```

2.43 设 $\xi \sim B(1, 0.3)$，对于 $i = 1, 2, ..., 10000$，分别使用 R 语言模拟 ξ 的重复观测结果 $10i$ 次，并计算相应的重复观测结果的算术平均值 $\bar{x}_i$；绘制 $(i, \bar{x}_i)$ 的折线图，分析该折线随着 i 增大变化的趋势及原因

```
# 2.43 程序
i <- 1:10000
xi <- rbinom(10*max(i),1,0.3)
x_bar <- c()
for (j in 10*i){
  x_bar <- c(x_bar, mean(xi[1:j]))
}
plot(i, x_bar, type="l")
abline(h=0.3, col="red", lty=2, lwd=2)
```

```
# 2.43 程序
i <- 1:10000
x_bar <- c()
```

```
for (j in i){
  xi <- rbinom(10*j,1,0.3)
  x_bar <- c(x_bar, mean(xi))
}
plot(i, x_bar, type="l")
abline(h=0.3, col="red", lty=2, lwd=2)
```

- **答**：随着观测次数的提升，样本的均值愈发趋于实际的总体期望 ($np = 0.3$)；原因，根据大数定律可知，随着样本量的增大，样本均值会愈发接近总体的期望。

2.44 设 ξ 服从以 1000, 10 和 50 为参数的超几何分布，对于 $i = 1, 2, \dots, 10000$，分别使用 R 语言模拟 ξ 的重复观测结果 $10i$ 次，并计算相应的重复观测结果的算术平均值 $\bar{x}_i$；绘制 $(i, \bar{x}_i)$ 的折线图，分析该折线随着 i 增大变化的趋势及原因

```
# 2.44 程序
i = 1:10000
xi <- rhyper(max(i) * 10, 10, 990, 50)
x_bar <- c()
for (j in i){
  x_bar <- c(x_bar, mean(xi[1:j*10]))
}
plot(i, x_bar, type="l")
abline(h=0.5, lty=2, lwd=2, col="red")
```

```
# 2.44 程序
i = 1:10000
x_bar <- c()
for (j in i){
  xi <- rhyper(j * 10, 10, 990, 50)
  x_bar <- c(x_bar, mean(xi))
}
plot(i, x_bar, type="l")
abline(h=0.5, lty=2, lwd=2, col="red")
```

- **答**：随着观测次数的提升，样本的均值愈发趋于实际的总体期望 $\frac{nK}{N} = 0.5$；原因，根据大数定律可知，随着样本量的增大，样本均值会愈发接近总体的期望。

2.49 设 $\xi \sim B(10000, 0.3)$，写出蒙特卡洛方法近似计算 $F_\xi(5555)$ 的 R 语言程序代码，并将计算结果与 pbinom(5555,10000,0.3) 的计算结果相比较，分析近似计算的精度和重复观测次数之间的关系。(这道题出的不是很好。

```
# 2.49 程序
sim_freq <- c()
true_prob <- pbinom(5555, 10000,0.3)
for (n in c(10,100,1000)) {
  sim_B <- rbinom(n, 10000,0.3)
  tmp_freq <- mean(sim_B <= 5555)
  sim_freq <- c(sim_freq,tmp_freq)
  abs_diff <- abs(true_prob - tmp_freq)

  cat(" 重复观测数为 ",n," 的估计结果为 ", tmp_freq, " 其和真实值之间的差距 ",
  ↪ abs_diff,"\n")
}
```

```
## 重复观测数为  10  的估计结果为  1  其和真实值之间的差距  0
## 重复观测数为  100  的估计结果为  1  其和真实值之间的差距  0
## 重复观测数为  1000  的估计结果为  1  其和真实值之间的差距  0
```

- **答**:

2.52 设 $\xi \sim U(-10, 10)$，写出蒙特卡洛方法近似计算 $F_\xi(0.5)$ 的 R 语言程序代码，并将计算结果与 punif(0.5,-10,10) 的计算结果相比较，分析近似计算的精度和重复观测次数之间的关系。

```
# 2.52 程序
sim_freq <- c()
true_prob <- punif(0.5, -10, 10)
for (n in c(10,100,1000,10000)) {
  sim_U <- runif(n, -10,10)
  tmp_freq <- mean(sim_U <= 0.5)
  sim_freq <- c(sim_freq,tmp_freq)
  abs_diff <- abs(true_prob - tmp_freq)

  cat(" 重复观测数为",n," 的估计结果为", tmp_freq, " 其和真实值之间的差距",
  ↪ abs_diff,"\n")
}
```

```
## 重复观测数为 10 的估计结果为 0.7 其和真实值之间的差距 0.175
## 重复观测数为 100 的估计结果为 0.5 其和真实值之间的差距 0.025
## 重复观测数为 1000 的估计结果为 0.49 其和真实值之间的差距 0.035
## 重复观测数为 10000 的估计结果为 0.5256 其和真实值之间的差距 6e-04
```

- **答**:

2.54 试用蒙特卡洛方法估算定积分 $\int_0^1 x^2 e^{x^2}\, dx$

```r
# 2.54 程序
Y <- runif(10000, 0,1)
f_Y = Y^2 * exp(Y^2)
estim_int <- mean(f_Y) * (1-0)
true_int <- integrate(function(x) x^2 * exp(x^2), 0,1)
cat(" 蒙特卡洛模拟的结果为", estim_int, ", 实际的计算结果为", true_int$value)
```

```
## 蒙特卡洛模拟的结果为 0.6253453 , 实际的计算结果为 0.627815
```

2.60 若大学生中男生的身高均值为 173.61 cm，标准差为 4.96 cm。随机选取 16 名男大学生。在 RStudio 中，写出应用中心极限定理近似计算这 16 名大学生的身高均值落在区间 $(168.65, 178.57)$（单位：cm）内的概率的程序代码，并给出近似计算的结果。

- **答** 根据中心极限定理，对于 16 名男大学生的身高均值 $\bar{X}$，其近似于正态分布

$$N(\mathbb{E}(X), D(X)/16)$$

  其中

  - $\mathbb{E}(X) = 173.61$
  - $D(X) = 4.96^2$

```r
# 2.54 程序
mu = 173.61
sigma = 4.96^2
mean_sigma = sigma / 16
up = pnorm(178.57, mu, sqrt(mean_sigma))
low = pnorm(168.65, mu, sqrt(mean_sigma))
up - low
```

```
## [1] 0.9999367
```