

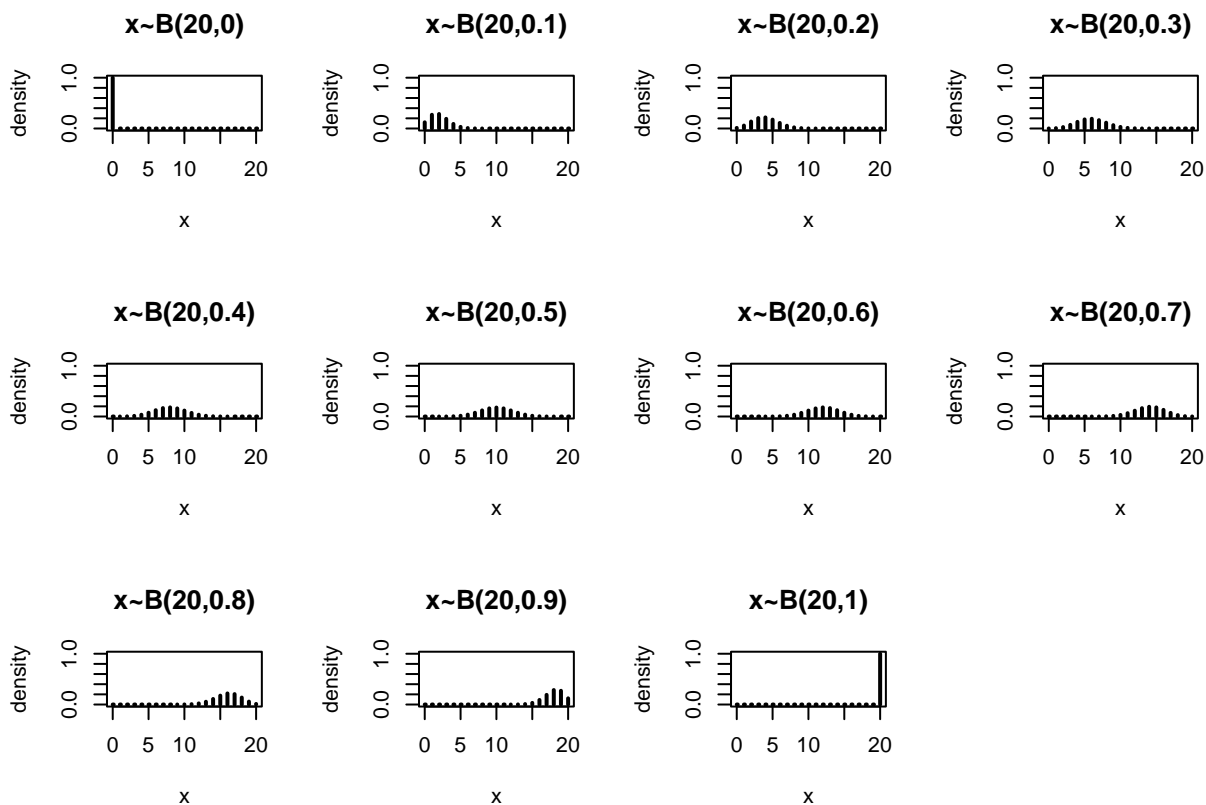
E04 上机实验 04

郑盼盼

2024-10-30

1. 对于 $p = 0, 0.1, \dots, 0.9, 1$ 绘制 $B(20, p)$ 的密度函数图像，考察随着成功概率由小到大的变化，密度图像的变化特征。

```
x <- 0:20
par(mfrow = c(3, 4))
for (p in seq(0,1, by=0.1)){
  y <- dbinom(x,20,p)
  plot(x, y,
        type="h",
        lwd=2,
        xlab = "x",
        ylab = "density",
        xlim = c(0,20),
        ylim = c(0,1),
        main=paste("x~B(20, ",p, ")", sep=""))
}
```



2. 已知 $X \sim B(4, \frac{1}{6})$ ，模拟 X 的 $m = 10$ 次重复观测值，用 f_i 表示 m 次观测中 $\{X = i\}$ 的频率，绘制密度矩阵

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ f_0 & f_1 & f_2 & f_3 & f_4 \end{pmatrix}$$

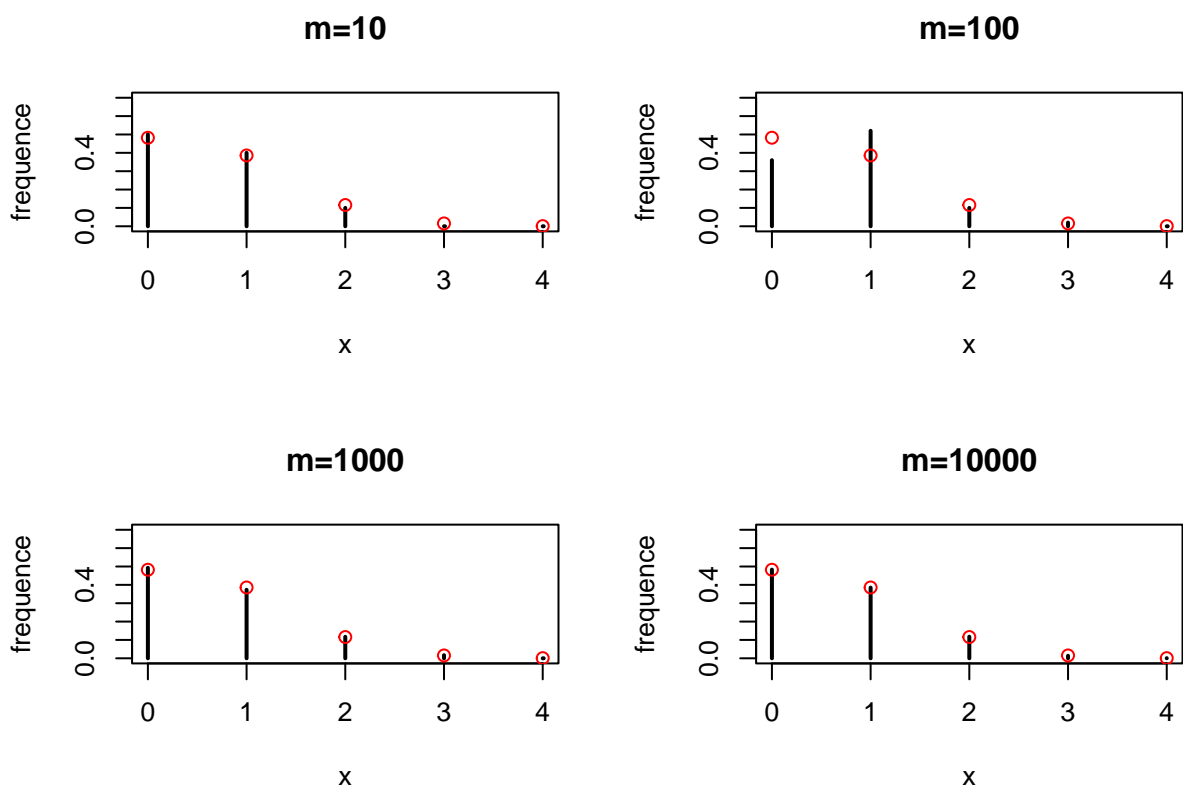
的密度图像，当 $m = 100, 1000, 10000$ 时，上述密度图像的变化规律是什么。

```
rm(list=ls())
all_df = data.frame("x" = 0:4)
par(mfrow=c(2,2))
for (i in 1:4){
  m <- 10^i
  x <- rbinom(m, 4, 1/6)
  tmp_freq = c()
  k <- 0:4
  for (j in k){
    tmp_freq <- c(tmp_freq, sum(x == j)/m)
  }
  plot(k, tmp_freq,
       type="h",
       lwd=2,
```

```

    xlab="x",
    ylab="frequence",
    xlim=c(0,4),
    ylim=c(0,0.7),
    main=paste("m=",m,sep="")
  )
  points(k, dbinom(k,4,1/6), col="red")
}

```



3. 袋中有 10 个红球和 10 个黑球。从袋中取后放回的方法依次任取 9 个球，用 X 表示取出红球的个数；从袋中用取后不放回的方法依次任取 9 个球，用 Y 表示取出红球的个数，将 X 和 Y 的分布函数曲线用不同颜色绘制在同一图中（绘制坐标位于区间 $[0, 9.5]$ 内），解释 X 和 Y 的分布函数为什么不同。由题意可知， $X \sim B(9, 0.5)$ ， $Y \sim H(20, 10, 9)$

```

rm(list=ls())
par(mfrow=c(1,1))
k <- 0:9
X_probs <- pbinom(k, 9, 0.5)
Y_probs <- phyper(k, 10, 10, 9)
plot(c(0,1),c(X_probs[1],X_probs[1]),

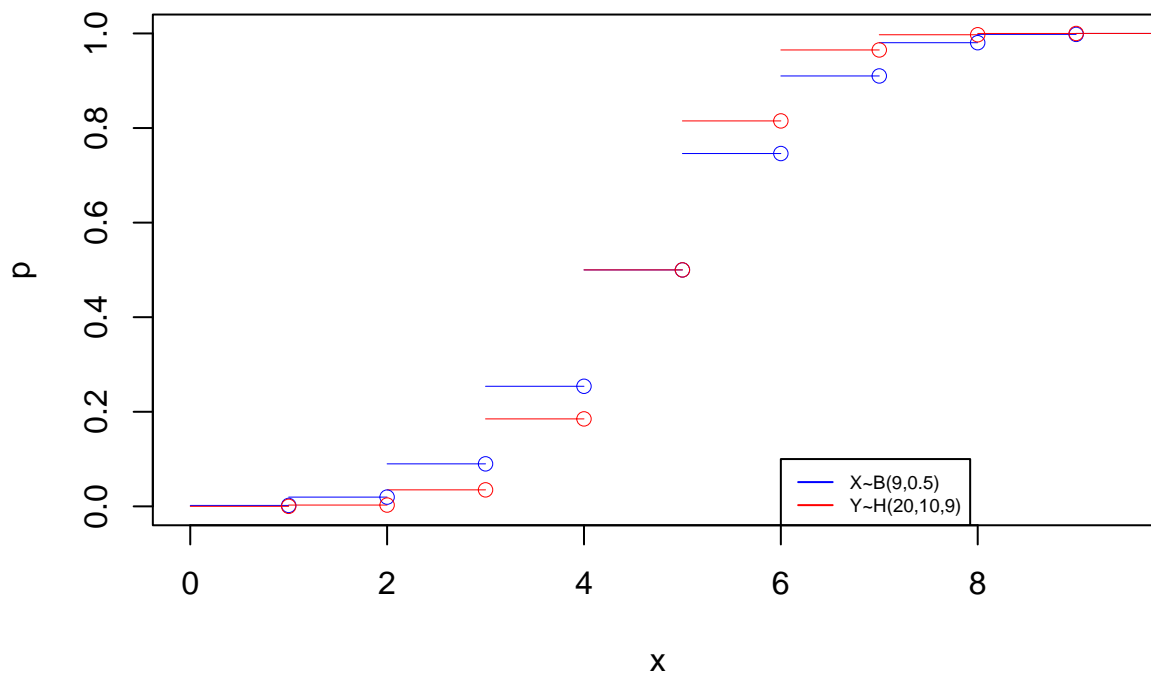
```

```

    type="l",
    lwd=.5,
    col="blue",
    xlim=c(0,9.5),
    ylim=c(0,1),
    xlab="x",
    ylab="p")
lines(c(0,1), c(Y_probs[1], Y_probs[1]),
      type="l",
      lwd=.5,
      col="red",
      )
for (i in 2:10){
  lines(c(i-1,i), c(X_probs[i], X_probs[i]),
        type = "l",
        lwd = .5,
        col = "blue")
  lines(c(i-1,i), c(Y_probs[i], Y_probs[i]),
        type = "l",
        lwd = .5,
        col = "red")
  points(i-1, X_probs[i-1],
         col="blue",
         lwd=.5)

  points(i-1, Y_probs[i-1],
         col="red",
         lwd=.5)
}
legend(6, 0.1, legend=c("X~B(9,0.5)", "Y~H(20,10,9)"),
      col=c("blue", "red"), lty=c(1,1), cex=0.6)

```



- 对于 X （有放回抽样），红球每次被取出的概率固定，因此分布函数较为平滑，概率集中在 $X = 4$ 附近。
- 对于 Y （无放回抽样），取出的红球数量影响后续抽样的概率，因此分布更加集中，分布曲线变化较为陡峭。