

统计计算第二次作业

Code ▾

苏锦华 2017201620

题目要求

使用近似抽样法（butler）生成5000个标准正态分布随机变量绘图检验（直方图以及标准正态分布密度函数）生成的随机数是否合理。

解答

首先，设置m=20，即找出20个分点将密度曲线按面积等分，通过变量和擂台算法得到。

Hide

```
f = function(x)
  1 / sqrt(2* pi) *exp(-x*x/2)
find_split = function(m,x0,xn,f){
  target = seq(0,1,1/m)
  dist = rep(1,m+1)
  x = rep(0,m+1)
  for(i in 1:2000){
    for(j in 1:m+1){
      tmp = i/2000
      res = integrate(f,x0,tmp*(xn-x0)+x0)["value"][[1]]
      if(abs(res - target[j]) < dist[j]){
        dist[j] = abs(res - target[j])
        x[j] = tmp *(xn-x0)+x0
      }
    }
  }
  x[m+1] = 10.0
  x[1]= -10.0
  return(x)
}
x = find_split(20,-10,10,f)
x
```

[1]	-10.00	-1.64	-1.28	-1.04	-0.84	-0.67	-0.52	-0.39	-0.25	-0.13	0.00
0.13	0.25	0.39	0.52	0.67	0.84						
[18]	1.04	1.28	1.64	10.00							

上述结果对称工整，接下来将分两步抽样：

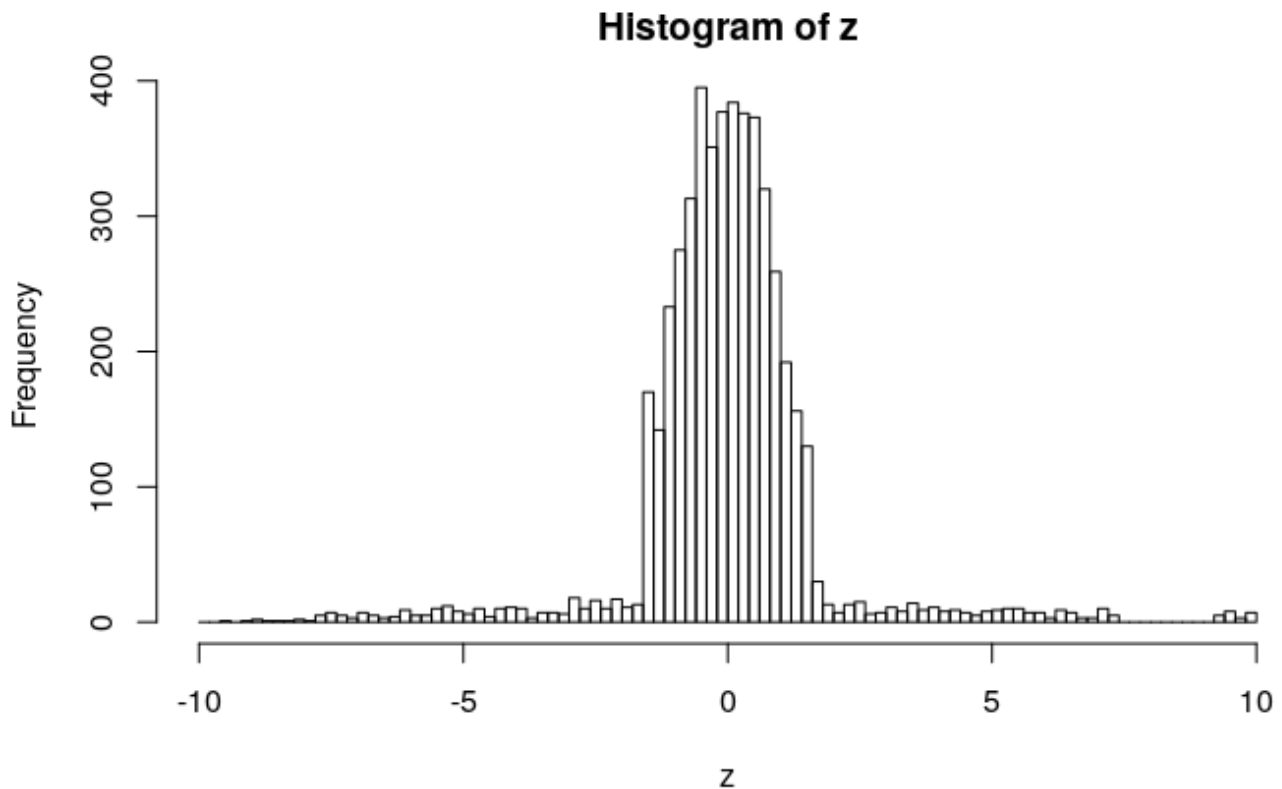
- 产生U(0,1)的r1,令idx=[mr+1]
- 根据idx计算三角形与梯形面积之比d
- 产生U(0,1)的r2,比较r2与d的大小关系和密度曲线上点的大小关系得到z

Hide

```

z = rep(0,5000)
r1 = runif(5000)
r2 = runif(5000)
m = 20
for (i in 1:5000) {
  idx = floor(r1[i]*m+1)
  left = f(x[idx])
  right = f(x[idx+1])
  d = abs(right - left+0.01)/(left+right)
  if (r2[i] < d){
    if(right>left){
      z[i] = x[idx] + (x[idx+1] - x[idx])*sqrt(r2[i])
    }else{
      z[i] = x[idx+1] - (x[idx+1] - x[idx])*sqrt(1 - r2[i])
    }
  }else{
    z[i] = x[idx] + (x[idx+1] - x[idx])*r2[i]
  }
}
hist(z,breaks = seq(-10,10,0.2))

```



从直方图观察非常像标准正态分布，下面再做出真正的标准正态密度曲线，发现两者差异不大。

Hide

```

x_ = seq(-10,10,0.1)
y = f(x_)
plot(x_,y)

```

