

# 统计计算 HW2

朱强强 17064001

2020 年 3 月 29 日

- (1) 在 R 中使用 `set.seed` 控制伪随机数初始值，并从  $N(3, 1)$  生成一个样本  $y_1, \dots, y_n$ ，样本量  $n = 10$ ，将样本数据记录在向量  $\mathbf{y}$  中。

```
> set.seed(123) # 控制伪随机数初始值
> y <- rnorm(10, 3, 1)
> print(y)
```

```
[1] 2.439524 2.769823 4.558708 3.070508 3.129288 4.715065 3.460916 1.734939
[9] 2.313147 2.554338
```

- (2) 假设已知样本服从正态分布且总体方差  $\sigma^2 = 1$ ，写出以上样本的对数似然函数。

当  $\sigma^2 = 1$  时

$$\begin{aligned} L(y) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_i - \mu)^2}{2}\right) \\ &= (2\pi)^{-n/2} \exp\left(-\sum_{i=1}^n \frac{(y_i - \mu)^2}{2}\right) \\ \log L(y) &= -\frac{n}{2} \ln(2\pi) - \sum_{i=1}^n \frac{(y_i - \mu)^2}{2} \end{aligned}$$

```
> # 对数似然函数
> llk <- function(mu, y) {
+   n <- length(y)
+   log.likelihood <- -(n/2)*log(2*pi)-sum((y-mu)^2)/2
+   return(log.likelihood)
+ }
> mu <- mean(y)
> llk(mu, y)
```

```
[1] -13.28305
```

- (3) 假设对于总体均值  $\mu$  有先验信息  $\mu \sim \text{Cauchy}(0, 1)$ 。设舍选法的目标抽样分布是  $\mu \sim p(\mu|\mathbf{x})$ ，试投分布是  $\mu \sim \text{Cauchy}(0, 1)$ 。请针对 (1) 中样本，自定义名为 `accp.rate` 函数，该函数返回舍选法步骤 2 中定义的接受率

$$\text{acceptance rate} = \frac{p(\mu|\mathbf{y})}{c_{\text{opt}}p(\mu)}$$

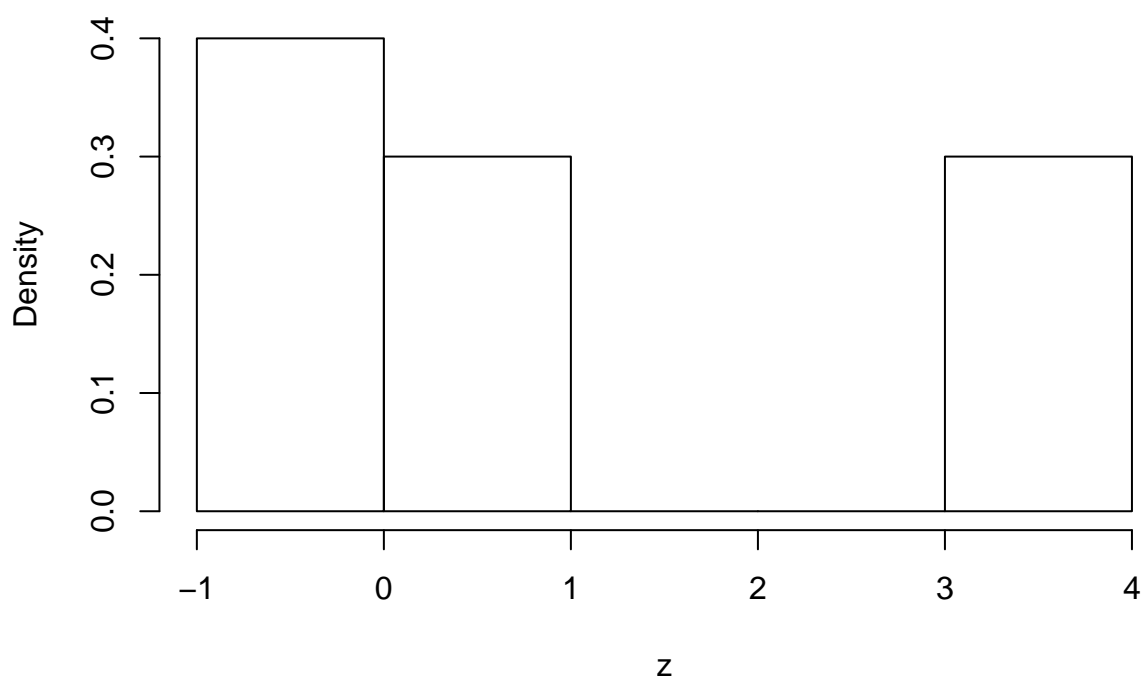
$$p(\mu|\mathbf{y}) = \frac{p(\mathbf{y}|\mu) \times p(\mu)}{\int_{-\infty}^{+\infty} p(\mathbf{y}|\mu) \times p(\mu) d\mu} \\ \propto p(\mu) \times L(\mu|\mathbf{y})$$

```
> accp.rate <- function(mu, y) {
+   p <- 1/(1+mu^2)*exp(-sum((y-mu)^2/2))
+   g <- 1/(1+mu^2)
+   c <- max(p/g)
+   accp_rate <- p/(c*g)
+   return(accp_rate)
+ }
```

- (4) 针对 (1) 中的样本，使用舍选法，从  $p(\mu|\mathbf{y})$  生成大量随机数。利用直方图，图示后验分布抽样结果。 $\mu$  后验分布期望的近似值是多少？请和  $\mu_0 = 3$  进行比较。

```
> sample <- function(n, mu, y) {
+   Z <- rep(NA, n)
+   i <- 1
+   while (i <= n) {
+     X <- rcauchy(n, 0, 1)
+     U <- runif(1)
+     if (U <= accp.rate(mu, y)) {
+       Z[i] <- X
+       i <- i + 1
+     }
+   }
+   return(Z)
+ }

>
> n <- 10
> mu <- 3
> z <- sample(n, mu, y)
> hist(z, prob=T, main="")
```

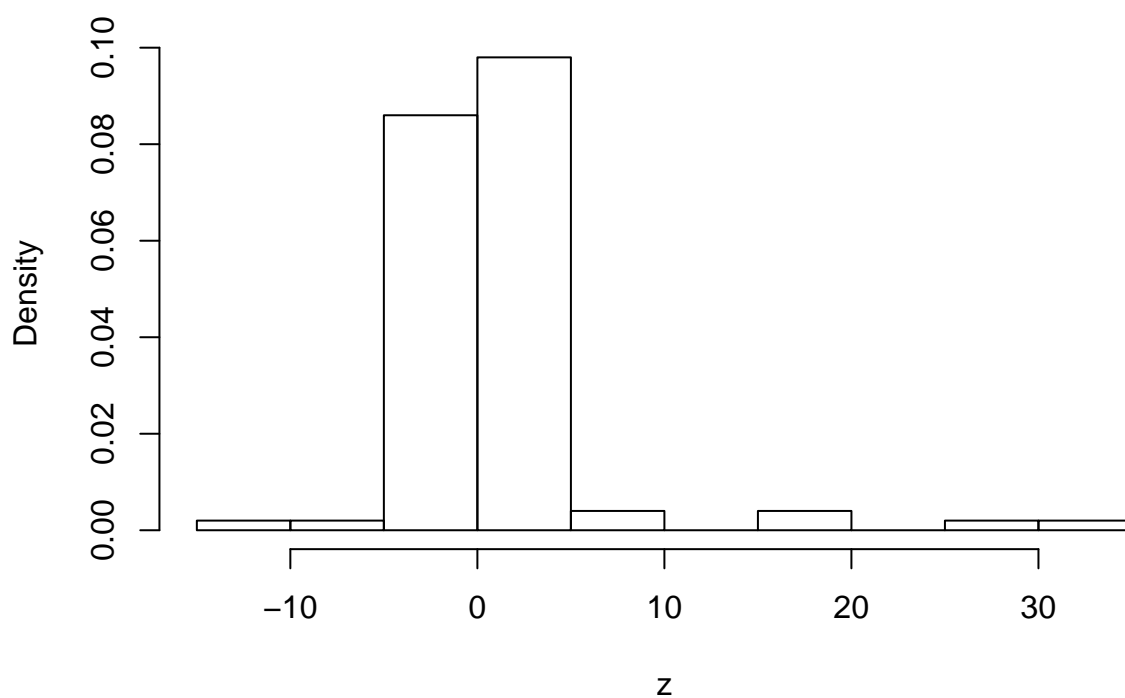


```
> cat("mu后验分布期望的近似值", mean(z), "\n")
```

mu后验分布期望的近似值 1.045074

(5) 令样本量  $n = 100$ 。重复步骤 (1) 和 (4)。

```
> set.seed(123)
> n <- 100
> y <- rnorm(n, 3, 1)
> z <- sample(n, mu, y)
> hist(z, prob=T, main="")
```



```
> cat("mu后验分布期望的近似值", mean(z), "\n")
```

mu后验分布期望的近似值 1.087303