E3.5

```r
generate_random_sample <- function(pmf, n) {
  # 计算累积分布函数（CDF）
  cdf <- cumsum(pmf)

  # 将CDF转换为[0, 1]之间的均匀随机数
  u <- runif(n, min = 0, max = 1)

  # 使用逆变换方法生成随机样本
  sample <- sapply(u, function(x) min(which(cdf >= x)))

  return(sample)
}

# 定义概率质量函数
pmf <- c(0.1, 0.2, 0.2, 0.2, 0.3)

# 生成随机样本
set.seed(123)
sample <- generate_random_sample(pmf, 1000)

# 构造相对频率表
relative_freq_table <- table(factor(sample, levels=1:length(pmf))) / length(sample)

# 输出相对频率表
print(relative_freq_table)

# 比较理论概率与经验概率
theoretical_probs <- pmf
empirical_probs <- as.numeric(relative_freq_table)

# 输出结果
print(data.frame(Theoretical = theoretical_probs, Empirical = empirical_probs))
```

32:80 (Top Level) R Script

Console Terminal Background Jobs

R 4.4.1 · ~/

```
  Theoretical Empirical
1         0.1     0.099
2         0.2     0.197
3         0.2     0.211
4         0.2     0.198
5         0.3     0.295
>
```

E3.6

假设目标分布为 $f(x)$; $u \sim U(0,1)$, 且 $Mq(x) \geq f(x)$

$$p(x \mid accept) = \frac{p(x, accept)}{\int_x p(x, accept)\, dx}$$

$$p(x, accept) = p(accept \mid x)\, p(x)$$

$$= q(x) \int_U p(accept \mid x, u)\, p(u \mid x)\, dx$$

$$= q(x) \int_U p(accept \mid x, u)\, p(u)\, du$$

而其中 $p(accept \mid x, u) = \begin{cases} 1, & u \leq \frac{f(x)}{Mq(x)} \\ 0, & \text{otherwise} \end{cases}$

又∵ $\frac{f(x)}{Mq(x)} \in [0,1]$

$$\therefore p(x, accept) = q(x) \cdot \int_0^{\frac{f(x)}{Mq(x)}} p(u)\, du = q(x) \cdot \frac{f(x)}{Mq(x)} = \frac{f(x)}{M}$$

代入到 $p(x \mid accept)$

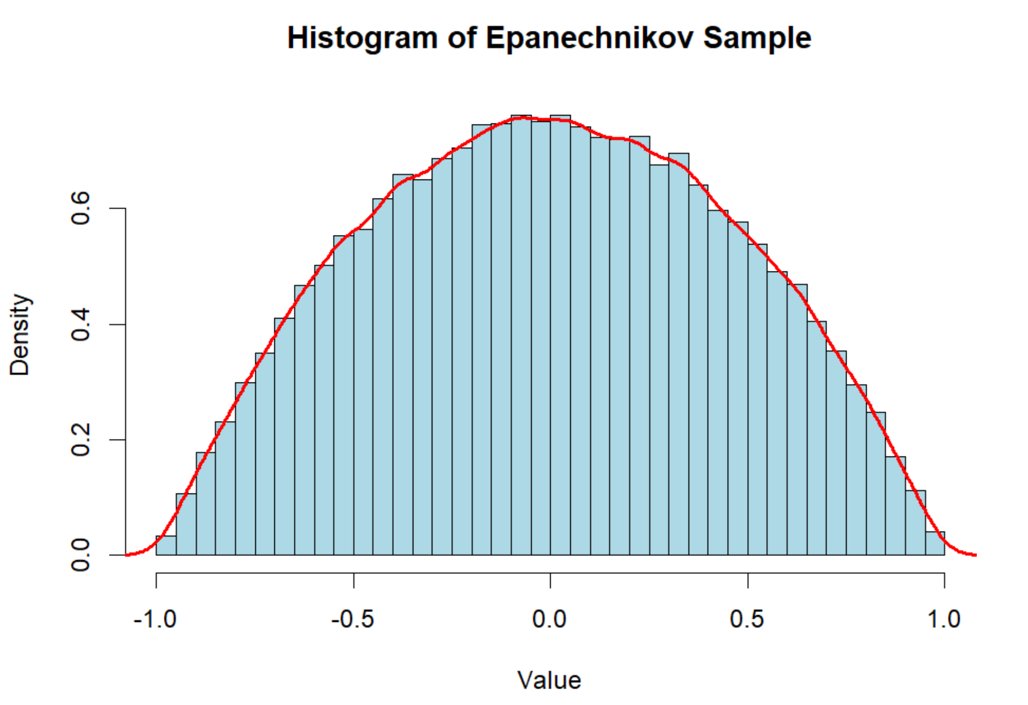
$$\therefore p(x \mid accept) = \frac{p(x, accept)}{\int_x p(x, accept)\, dx} = \frac{\frac{f(x)}{M}}{\int_x \frac{f(x)}{M}\, dx} = \frac{f(x)}{\int_x f(x)\, dx} = f(x)$$

E3.9

```r
# 定义 Epanechnikov 分布的密度函数
epanechnikov_density <- function(x) {
  ifelse(abs(x) <= 1, (3/4) * (1 - x^2), 0)
}

# 模拟 Epanechnikov 分布的随机变元
generate_epanechnikov <- function() {
  repeat {
    u1 <- runif(1, -1, 1)
    u2 <- runif(1, -1, 1)
    u3 <- runif(1, -1, 1)

    if (abs(u3) >= abs(u2) && abs(u3) >= abs(u1)) {
      return(u2)
    } else {
      return(u3)
    }
  }
}
# 生成大量 Epanechnikov 分布的随机变元
set.seed(100)
n <- 100000
epanechnikov_sample <- replicate(n, generate_epanechnikov())

# 构造直方图
hist(epanechnikov_sample, breaks = 50, freq = FALSE, main = "Histogram of Epanechnikov Sample",
     xlab = "Value", ylab = "Density", col = "lightblue")

# 添加密度曲线
epan_density <- density(epanechnikov_sample)
lines(epan_density, col = "red", lwd = 2)
```

Files Plots Packages Help Viewer Presentation

Zoom Export Publish




E3.10

$$f(x) = f_{U_3}(x \mid |U_3| \text{非最大}) \cdot P(|U_3| \text{非最大}) + f_{U_2}(x \mid |U_3| \text{最大}) \cdot P(|U_3| \text{最大})$$

$$= f_{U_3}(x) \cdot P(|U_3| \text{非最大} \mid U_3 = x) + f_{U_2}(x) P(|U_3| \text{最大} \mid U_2 = x)$$

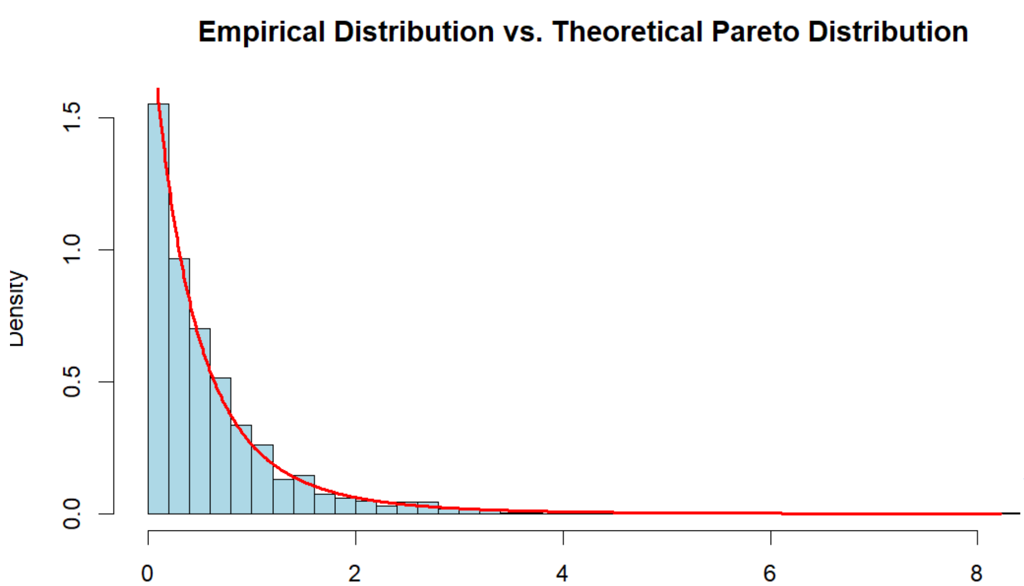
$$= \frac{1}{2}[1 - P(|U_3| \text{最大} \mid U_3 = x)] + \frac{1}{2} P(|U_3| \text{最大} \mid U_2 = x)$$

$$= \frac{1}{2}(1-x^2) + \frac{1}{2} \frac{\int_{|x|}^{1} t\,dt}{(1-|x|)} (1-|x|)$$

$$= \frac{3}{4}(1-x^2)$$

# E3.12 & E3.13.

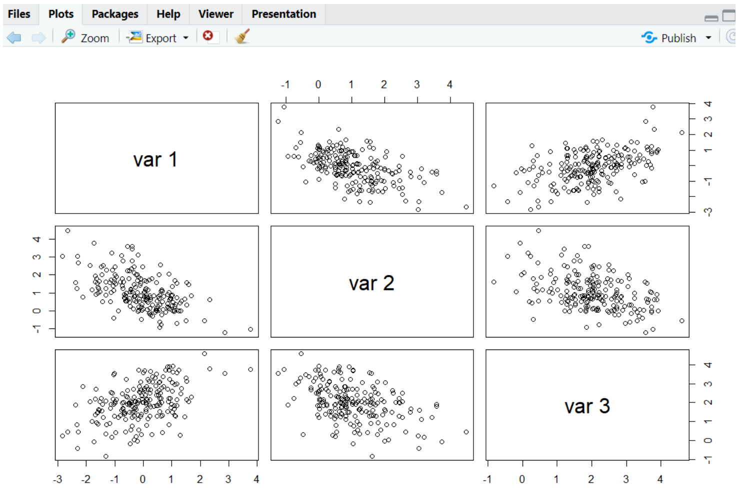
```r
#生成exp-gamma混合函数
Exp_Gamma_mixture <- function(r,beta,n){
  U1 = rgamma(n,r,beta)
  U2 = rexp(n,U1)
}

#生成目标函数
r = 4
beta = 2
results = Exp_Gamma_mixture(r,beta,1000)

#输出直方图以及曲线
hist(results,col='lightblue',freq = F,breaks=50,xlab='x',ylab = 'Density',
     main='Empirical Distribution vs. Theoretical Pareto Distribution')
xx = seq(min(results),max(results),length=1000)
yy = beta^r * r * (beta+xx)^(-r-1)
lines(xx,yy,col='red',lwd=2)
```




# E3.14

```r
#利用Cholesky分解生成变量
rmvn.Cholesky <-
  function(n, mu, Sigma) {
    # generate n random vectors from MVN(mu, Sigma)
    # dimension is inferred from mu and Sigma
    d <- length(mu)
    Q <- chol(Sigma) # Choleski factorization of Sigma
    Z <- matrix(rnorm(n*d), nrow=n, ncol=d)
    X <- Z %*% Q + matrix(mu, n, d, byrow=TRUE)
    X
  }

#代入协方差矩阵和均值向量
Sigma = c(1,-.5,.5,-.5,1,-.5,.5,-.5,1)
Sigma = matrix(Sigma,nrow=3)
mu = c(0,1,2)
samples = rmvn.Cholesky(200,mu,Sigma)

#输出散点图
X1 = samples[,1]
X2 = samples[,2]
X3 = samples[,3]
pairs(samples)
apply(samples,MARGIN = 2,FUN = mean)
cor(X1,X2)
cor(X1,X3)
cor(X2,X3)
```

```
> pairs(samples)
> apply(samples,MARGIN = 2,FUN = mean)
[1] -0.08972105  1.03274733  2.00825675
> cor(X1,X2)
[1] -0.5525612
> cor(X1,X3)
[1] 0.4652262
> cor(X2,X3)
[1] -0.4385157
```




E2.

先假定 $Z_1, Z_2$ 服从标准正态分布且独立，令 $p(Z_1)$ 和 $p(Z_2)$ 为其密度函数

$$p(z_1)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{z_1^2}{2}}, \quad p(z_2)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{z_2^2}{2}}$$

∵ $z_1, z_2$ 独立 ∴ $p(z_1,z_2)=\frac{1}{2\pi}\cdot e^{-\frac{z_1^2+z_2^2}{2}}$

将 $z_1, z_2$ 作坐标变换，使 $Z_1=R\cos\theta$，$Z_2=R\sin\theta$

则 $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}\cdot e^{-\frac{z_1^2+z_2^2}{2}}dz_1dz_2=\int_0^{2\pi}\int_0^{\infty}\frac{1}{2\pi}\cdot e^{-\frac{R^2}{2}}\cdot R\,d\theta\,dR=1$

∴ $P_R(R\le r)=\int_0^{2\pi}\int_0^{r}\frac{1}{2\pi}\cdot e^{-\frac{R^2}{2}}\cdot R\,d\theta\,dR=1-e^{-\frac{r^2}{2}}$

$P_\theta(\theta\le\phi)=\int_0^{\phi}\int_0^{\infty}\frac{1}{2\pi}\cdot e^{-\frac{R^2}{2}}R\,d\theta\,dR=\frac{\phi}{2\pi}$

∴ $\theta$ 服从 $[0, 2\pi]$ 上的均匀分布

再令 $F_R(r)=1-e^{-\frac{r^2}{2}}$ 则 $R=\sqrt{-2\ln(1-z)}$

∴ 当 $z$ 服从 $[0,1]$ 上的均匀分布时，$R$ 的分布满足 $F_R(r)$

因此可以选取两个服从 $[0,1]$ 均匀分布的 $U$ 和 $V$

其中 $U, V$ 满足：$\theta=2\pi V$，$1-z=U \to R=\sqrt{-2\ln U}$

再代入 $Z_1 = R\cos\theta$ , $Z_2 = R\sin\theta$

∴可得 $Z_1 = \sqrt{-2\ln U}\,\cos(2\pi V)$ $Z_2 = \sqrt{-2\ln U}\cdot\sin(2\pi V)$

满足 $Z_1$ 和 $Z_2$ 是独立的标准正态分布.

E3. $f(x_1, x_2) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}}\cdot\exp\left(-\frac{1}{2}\begin{pmatrix}x_1-\mu_1\\x_2-\mu_2\end{pmatrix}^T\cdot\Sigma^{-1}\begin{pmatrix}x_1-\mu_1\\x_2-\mu_2\end{pmatrix}\right)$

$$\Sigma^{-1} = \begin{pmatrix}\Sigma_{11}^{-1}+\Sigma_{11}^{-1}\cdot\Sigma_{12}M^{-1}\cdot\Sigma_{21}\Sigma_{11}^{-1} & -\Sigma_{11}^{-1}\Sigma_{12}M^{-1}\\ -M^{-1}\Sigma_{21}\Sigma_{11}^{-1} & M^{-1}\end{pmatrix}$$

其中 $M = \Sigma_{22}-\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$

∴ $\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2-\mu_2)$

$\Sigma_{1|2} = \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$