محاسبات آماری پیشرفته

ترم اول سال تحصیلی ۹۳

جلسه : تولید تحقق‌هایی از متغیرهای تصادفی

حسین باغیشنی

دانشگاه شاهرود

۲۴ آبان ۱۳۹۵

چند نکته دیگر در مورد روش رد

روش پذیرش ـ رد در هر بعدی قابل به کارگیری است، البته با این شرط که g تابع (چگالی) احتمال بر روی فضای یکسان f باشد.

چند نکته دیگر در مورد روش رد

روش پذیرش ـ رد در هر بعدی قابل به کارگیری است، البته با این شرط که g تابع (چگالی) احتمال بر روی فضای یکسان f باشد.

به ثابت c ، ثابت نرمال‌ساز هم می‌گویند.

چند نکته دیگر در مورد روش رد

روش پذیرش ـ رد در هر بعدی قابل به کارگیری است، البته با این شرط که g تابع (چگالی) احتمال بر روی فضای یکسان f باشد.

به ثابت c ، ثابت نرمال‌ساز هم می‌گویند.

در اجرای روش پذیرش ـ رد، تنها دانستن نسبت f/g کافی است و الگوریتم به ثابت نرمال‌ساز نیازی ندارد.

چند نکته دیگر در مورد روش رد

روش پذیرش ـ رد در هر بعدی قابل به کارگیری است، البته با این شرط که g تابع (چگالی) احتمال بر روی فضای یکسان f باشد.

به ثابت c، ثابت نرمال‌ساز هم می‌گویند.

در اجرای روش پذیرش ـ رد، تنها دانستن نسبت f/g کافی است و الگوریتم به ثابت نرمال‌ساز نیازی ندارد.

نیازی نیست کران $f \leq cg$ خیلی تند و تیز(!) باشد. با جایگزین کردن c با یک مقدار بزرگتر، الگوریتم باز هم به درستی اجرا می‌شود. هرچند کارایی آن کاهش می‌یابد.

چند نکته دیگر در مورد روش رد

روش پذیرش ـ رد در هر بعدی قابل به کارگیری است، البته با این شرط که g تابع (چگالی) احتمال بر روی فضای یکسان f باشد.

به ثابت c، ثابت نرمال‌ساز هم می‌گویند.

در اجرای روش پذیرش ـ رد، تنها دانستن نسبت f/g کافی است و الگوریتم به ثابت نرمال‌ساز نیازی ندارد.

نیازی نیست کران $f \leq cg$ خیلی تند و تیز(!) باشد. با جایگزین کردن c با یک مقدار بزرگتر، الگوریتم باز هم به درستی اجرا می‌شود. هرچند کارایی آن کاهش می‌یابد.

احتمال پذیرش برابر $1/c$ است. بنابراین تا جای ممکن c باید کوچک باشد تا کارایی بیشتری داشته باشد.

چند نکته دیگر در مورد روش رد

روش پذیرش ـ رد در هر بعدی قابل به کارگیری است، البته با این شرط که g تابع (چگالی) احتمال بر روی فضای یکسان f باشد.

به ثابت c، ثابت نرمال‌ساز هم می‌گویند.

در اجرای روش پذیرش ـ رد، تنها دانستن نسبت f/g کافی است و الگوریتم به ثابت نرمال‌ساز نیازی ندارد.

نیازی نیست کران $f \leq cg$ خیلی تند و تیز(!) باشد. با جایگزین کردن c با یک مقدار بزرگتر، الگوریتم باز هم به درستی اجرا می‌شود. هرچند کارایی آن کاهش می‌یابد.

احتمال پذیرش برابر $1/c$ است. بنابراین تا جای ممکن c باید کوچک باشد تا کارایی بیشتری داشته باشد.

یک ایراد روش رد و پذیرش، آن است که متغیرهای غیرمفید را، زمانی که مقدار پیشنهادی از g رد می‌شود، تولید می‌کند. برای مرتفع کردن این مشکل، بعدا روش نمونه‌گیری با اهمیت را معرفی می‌کنیم.

روش رد برای توزیع بتا

داریم از توزیع $Beta(۲/۷, ۶/۳)$ با روش رد و پذیرش، نمونه تولید کنیم. ابتدا دقت کنید که کران بالا، c، مقدار ماکسیمم تابع چگالی احتمال بتاست که با دستور زیر قابل دستیابی است:

```
c=optimize(f=function(x){dbeta(x,2.7,6.3)},
interval=c(0,1),maximum=T)$objective
```

تابع g را یکنواخت استاندارد در نظر بگیرید. بنابراین مقدار پیشنهادی y پذیرفته می‌شود هرگاه $c \times U < f(y)$.

روش رد برای توزیع بتا

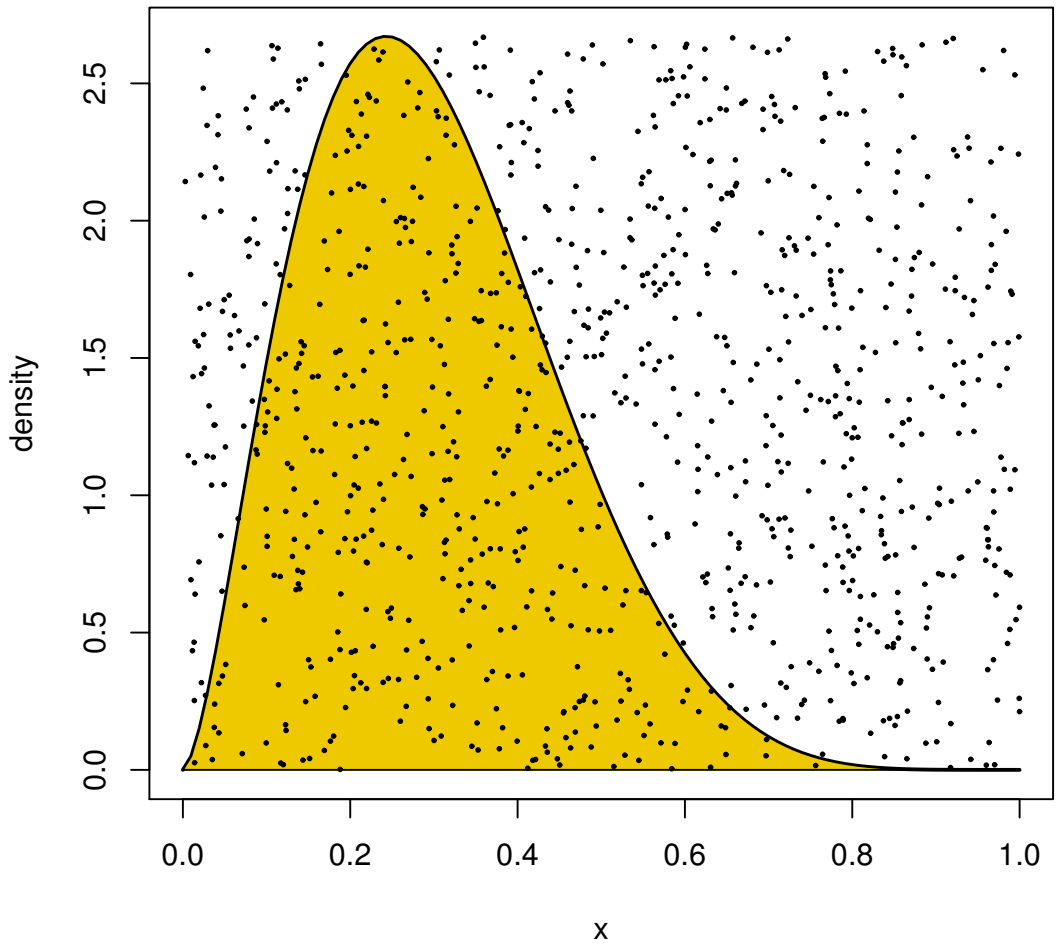
داریم از توزیع $Beta(۲/۷, ۶/۳)$ با روش رد و پذیرش، نمونه تولید کنیم. ابتدا دقت کنید که کران بالا، c، مقدار ماکسیمم تابع چگالی احتمال بتاست که با دستور زیر قابل دستیابی است:

```
c=optimize(f=function(x){dbeta(x,2.7,6.3)},
interval=c(0,1),maximum=T)$objective
```

تابع g را یکنواخت استاندارد در نظر بگیرید. بنابراین مقدار پیشنهادی y پذیرفته می‌شود هرگاه $c \times U < f(y)$.
دقت کنید که تولید $U \sim U(۰, ۱)$ و سپس ضرب آن در c، معادل است با تولید $U \sim U(۰, c)$.

```
ys=runif(1000)
us=c*runif(1000)
val=seq(0,1.,.01)
valf=dbeta(val,shape1=2.7,shape2=6.3)
plot(val,valf,type="l",xlab="x",ylab="density",lwd=2)
polygon(c(val,rev(val)),c(valf,0*valf),col="gold2")
points(ys,us,cex=.4,pch=20)
```

روش رد برای توزیع بتا



مقایسه کارایی روش‌ها

تولید تحقق‌هایی از توزیع کای‌دو با ۶ درجه آزادی، دو روش در نظر می‌گیریم:

```
test1=function(){
    U=runif(3*10^4)
    U=matrix(data=U,nrow=3) # matrix for sums
    X=-log(U) # uniform to exponential
    X=2*apply(X,2,sum)
return(X)
}
##
test2=function(){
    X=rchisq(10^4,df=6)
return(X)
}
##
> system.time(test1());system.time(test2())
   user  system elapsed
   0.11    0.00    0.08
   user  system elapsed
   0.00    0.00    0.01
```

تولید متغیر تصادفی کای‌دو غیرمرکزی

توزیع کای‌دو غیرمرکزی، فرم مشخصی برای تابع چگالی آن وجود ندارد.

تولید متغیر تصادفی کای‌دو غیرمرکزی

توزیع کای‌دو غیرمرکزی، فرم مشخصی برای تابع چگالی آن وجود ندارد.

برای تولید از این توزیع، چندین روش وجود دارد. یک روش کارا استفاده از روش تبدیل می‌باشد:

$$Z \sim \chi^{۲}_{p-۱},\ Y \sim N(\sqrt{\lambda}, ۱) \longrightarrow Z + Y^{۲} \sim \chi^{۲}_{p}(\lambda).$$

تولید متغیر تصادفی کای‌دو غیرمرکزی

توزیع کای‌دو غیرمرکزی، فرم مشخصی برای تابع چگالی آن وجود ندارد.

برای تولید از این توزیع، چندین روش وجود دارد. یک روش کارا استفاده از روش تبدیل می‌باشد:

$$Z \sim \chi^2_{p-1},\ Y \sim N(\sqrt{\lambda}, 1) \longrightarrow Z + Y^2 \sim \chi^2_p(\lambda).$$

```
rnchisq = function(n,p,lambda){
z = rchisq(n,df=p-1)
y = rnorm(n,sqrt(lambda),1)
x = z+y^2
return(x)
}
> xx = rnchisq(1000,3,3)
> mean(xx) # theoretical value is (p+lambda)=6
[1] 6.114251
> var(xx) # theoretical value is 2(p+2lambda)=18
[1] 18.25124
```

یک تمرین

تولید یک متغیر تصادفی $N(۰,۱)$ به وسیله روش پذیرش‌ـرد با توزیع نمایی دوگانه به عنوان منتخب با چگالی $g(x|\alpha) = \frac{\alpha}{۲}\exp\{-\alpha|x|\}$ را در نظر بگیرید.

۱. نشان دهید

$$\frac{f(x)}{g(x|\alpha)} \leq \sqrt{\frac{۲}{\pi}}\alpha^{-۱}e^{\frac{\alpha^۲}{۲}},$$

و نشان دهید می‌نیمم این کران بر حسب α در $\alpha = ۱$ به دست می‌آید

۲. نشان دهید احتمال پذیرش برابر $\sqrt{\frac{\pi}{۲e}} = ۰/۷۶$ است و برای تولید یک متغیر تصادفی نرمال به طور متوسط $\frac{۱}{۰/۷۶} = ۱/۳$ تولید متغیرهای یکنواخت لازم است

توزیع‌های آمیخته

متغیر تصادفی X دارای توزیع آمیخته گسسته است، اگر توزیع آن، برای دنباله‌ای از متغیرهای تصادفی $X_۱, X_۲, \cdots$، به صورت مجموع وزنی $F_X(x) = \sum \theta_i F_{X_i}(x)$ باشد به طوری که $\sum \theta_i = ۱$.

توزیع‌های آمیخته

متغیر تصادفی X دارای توزیع آمیخته گسسته است، اگر توزیع آن، برای دنباله‌ای از متغیرهای تصادفی $X_1, X_2, \cdots$، به صورت مجموع وزنی $F_X(x) = \sum \theta_i F_{X_i}(x)$ باشد به طوری که $\sum \theta_i = 1$.

به ثابت‌های θ_i، وزن‌های آمیختگی گویند.

توزیع‌های آمیخته

متغیر تصادفی X دارای توزیع آمیخته گسسته است، اگر توزیع آن، برای دنباله‌ای از متغیرهای تصادفی $X_1, X_2, \cdots$، به صورت مجموع وزنی $F_X(x) = \sum \theta_i F_{X_i}(x)$ باشد به طوری که $\sum \theta_i = 1$.

به ثابت‌های θ_i، وزن‌های آمیختگی گویند.

متغیر تصادفی X دارای توزیع آمیخته پیوسته است، اگر برای یک خانواده $X|\, Y = y$ که با مقادیر حقیقی y اندیس‌گذاری شده است، توزیع آن به صورت $F_X(x) = \int_{-\infty}^{\infty} F_{X|\, Y=y}(x) f_Y(y)\, dy$ باشد، به طوری که در آن f_y تابع وزن‌دهنده است و $\int_{-\infty}^{\infty} f_Y(y)\, dy = 1$.

توزیع‌های آمیخته

متغیر تصادفی X دارای توزیع آمیخته گسسته است، اگر توزیع آن، برای دنباله‌ای از متغیرهای تصادفی $X_1, X_2, \cdots$، به صورت مجموع وزنی $F_X(x) = \sum \theta_i F_{X_i}(x)$ باشد به طوری که $\sum \theta_i = 1$.

به ثابت‌های θ_i، وزن‌های آمیختگی گویند.

متغیر تصادفی X دارای توزیع آمیخته پیوسته است، اگر برای یک خانواده $X|\, Y = y$ که با مقادیر حقیقی y اندیس‌گذاری شده است، توزیع آن به صورت $F_X(x) = \int_{-\infty}^{\infty} F_{X|\, Y=y}(x) f_Y(y)\, dy$ باشد، به طوری که در آن f_y تابع وزن‌دهنده است و $\int_{-\infty}^{\infty} f_Y(y)\, dy = 1$.

دقت کنید توزیع‌های آمیخته با توزیع‌های حاصل از پیچش با هم فرق دارند. به مثال زیر توجه کنید.

آمیخته یا پیچش؟

فرض کنید $X_1 \sim N(0,1)$ و $X_2 \sim N(3,1)$ و مستقل از هم باشند. $S = X_1 + X_2$ پیچش دو متغیر را نشان می‌دهد. توزیع S نرمال با میانگین $\mu_1 + \mu_2 = 3$ و واریانس $\sigma_1^2 + \sigma_2^2 = 2$ است. برای تولید این پیچش داریم:

1. x_1 را از توزیع $N(0,1)$ تولید کن
2. x_2 از $N(3,1)$ تولید کن
3. قرار بده $s = x_1 + x_2$

آمیخته یا پیچش؟

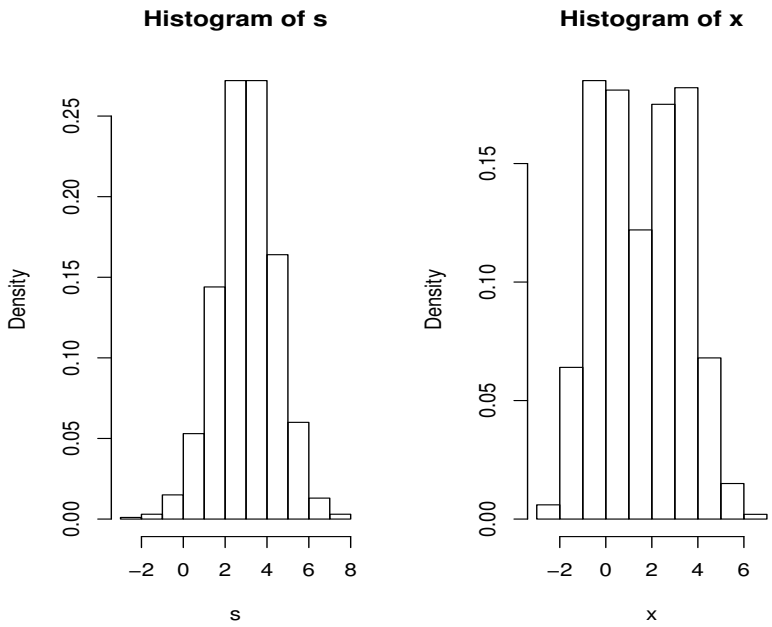
فرض کنید $X_1 \sim N(0,1)$ و $X_2 \sim N(3,1)$ و مستقل از هم باشند. $S = X_1 + X_2$ پیچش دو متغیر را نشان می‌دهد. توزیع S نرمال با میانگین $\mu_1 + \mu_2 = 3$ و واریانس $\sigma_1^2 + \sigma_2^2 = 2$ است. برای تولید این پیچش داریم:

1. x_1 را از توزیع $N(0,1)$ تولید کن
2. x_2 از $N(3,1)$ تولید کن
3. قرار بده $s = x_1 + x_2$

برای ساخت یک آمیخته ۵۰٪، یعنی $F_X(x) = 0.5F_{X_1}(x) + 0.5F_{X_2}(x)$، داریم:

1. عدد صحیح k را از مجموعه $\{1,2\}$ تولید کن به طوری که $P(1) = P(2) = 0.5$
2. اگر $k = 1$، x را از $N(0,1)$ تولید کن و اگر $k = 2$، x را از $N(3,1)$ تولید کن

```
n <- 1000
x1 <- rnorm(n, 0, 1)
x2 <- rnorm(n, 3, 1)
s <- x1 + x2                # the convolution
u <- runif(n)
k <- as.integer(u > 0.5)  # vector of 0's and 1's
x <- k * x1 + (1-k) * x2  # the mixture
par(mfcol=c(1,2))           # two graphs per page
hist(s, prob=TRUE)
hist(x, prob=TRUE)
par(mfcol=c(1,1))           # restore display
```



آمیخته‌ای از چند توزیع

مشابه بالا می‌توان برای بیشتر از دو توزیع هم، شکل آمیخته‌ای را نوشت. مثلاً فرض کنید آمیخته‌ای از ۵ توزیع گاما مورد نظر باشد. بنابراین $F_X = \sum_{i=1}^{5} \theta_i F_{X_i}$، به طوری که $X_i \sim Gamma(s = 3, \lambda_i = 1/i)$ و مستقل از هم هستند. همچنین وزن‌های آمیختگی برای $i = 1, \ldots, 5$ عبارتند از $\theta_i = i/15$.

آمیخته‌ای از چند توزیع

مشابه بالا می‌توان برای بیشتر از دو توزیع هم، شکل آمیخته‌ای را نوشت. مثلاً فرض کنید آمیخته‌ای از ۵ توزیع گاما مورد نظر باشد. بنابراین $F_X = \sum_{i=۱}^{۵} \theta_i F_{X_i}$، به طوری که $X_i \sim Gamma(s = ۳, \lambda_i = ۱/i)$ و مستقل از هم هستند. همچنین وزن‌های آمیختگی برای $i = ۱, \ldots, ۵$ عبارتند از $\theta_i = i/۱۵$.

برای تولید یک متغیر تصادفی با این توزیع، به صورت زیر عمل می‌کنیم:

۱. عدد صحیح $k \in \{۱, ۲, ۳, ۴, ۵\}$ را با احتمال‌های $P(k) = \theta_k$، برای $k = ۱, \ldots, ۵$ تولید می‌کنیم
۲. متغیر $Gamma(s, \lambda_k)$ را به عنوان مقدار تولید شده برمی‌گردانیم

برای تولید n نمونه باید دو مرحله بالا را n بار تکرار کنیم. برای این کار باید از یک حلقه for استفاده شود که می‌تواند برای حجم نمونه بزرگ ناکارا باشد.

بالا را می‌توان به صورت زیر با بردارسازی کاراتر کرد:

۱. نمونه تصادفی $k = (k_1, \ldots, k_n)$ را از اعداد صحیح تولید کن به طوری که $P(k) = \theta_k$. بنابراین $k[i]$ مشخص می‌کند که برای تولید عضو iام نمونه از کدام توزیع گاما استفاده شده است.

۲. بردار $rate$ را با طول n تنظیم کن که حاوی مقادیر λ_k است.

۳. با دستور $rgamma$ با پارامتر شکل ۳ و بردار $rate$ نمونه‌ای به حجم n تولید کن

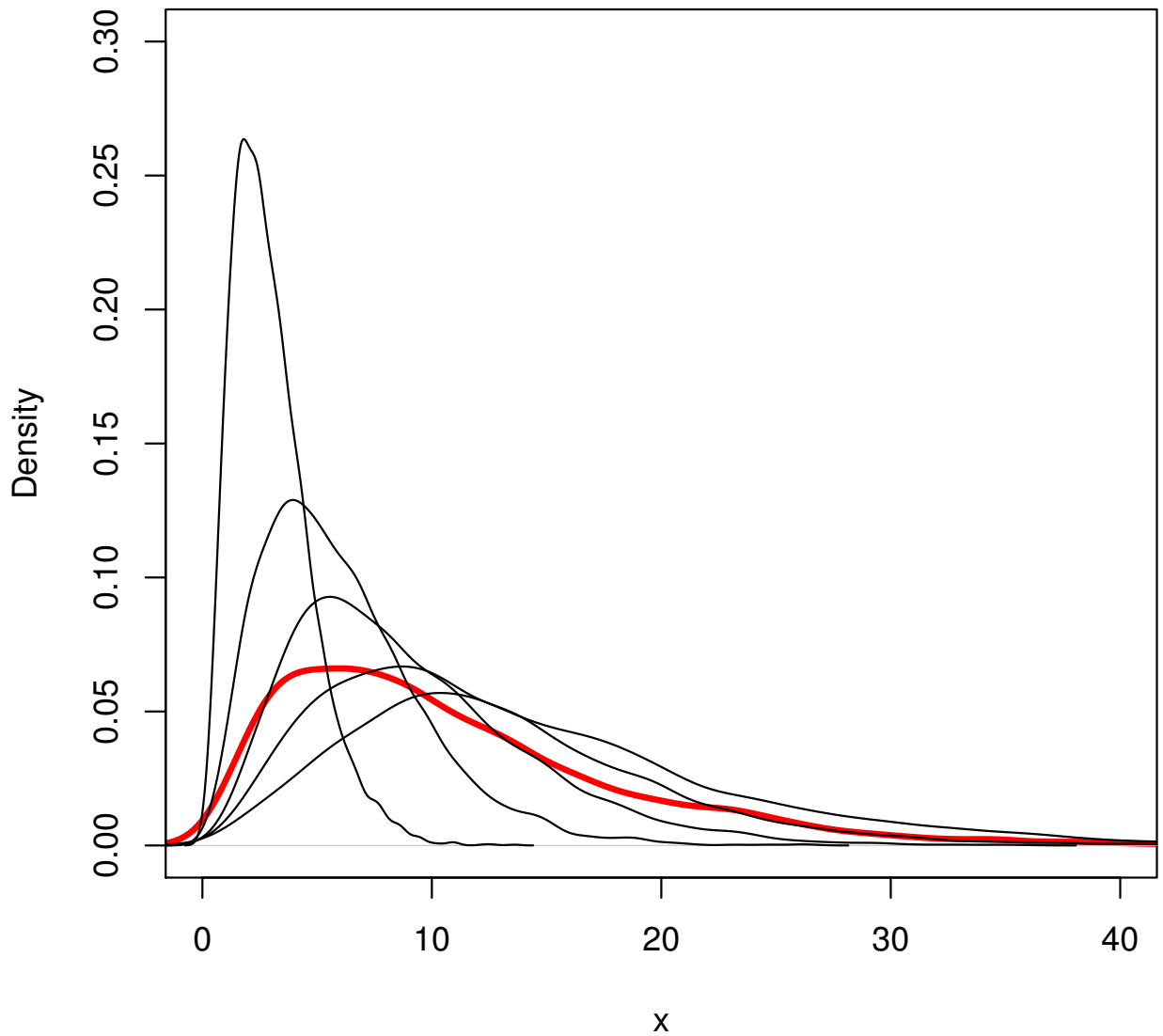
بالا را می‌توان به صورت زیر با بردارسازی کاراتر کرد:

۱. نمونه تصادفی $k = (k_1, \ldots, k_n)$ را از اعداد صحیح تولید کن به طوری که $P(k) = \theta_k$. بنابراین $k[i]$ مشخص می‌کند که برای تولید عضو iام نمونه از کدام توزیع گاما استفاده شده است.

۲. بردار $rate$ را با طول n تنظیم کن که حاوی مقادیر λ_k است.

۳. با دستور $rgamma$ با پارامتر شکل ۳ و بردار $rate$ نمونه‌ای به حجم n تولید کن

```
n <- 5000
k <- sample(1:5, size=n, replace=TRUE, prob=(1:5)/15)
rate <- 1/k
x <- rgamma(n, shape=3, rate=rate)
# plot the densities of the mixture alongside the components
plot(density(x), xlim=c(0,40), ylim=c(0,.3),
lwd=3, xlab="x", col="red", main="")
for (i in 1:5)
    lines(density(rgamma(n, 3, 1/i)))
```

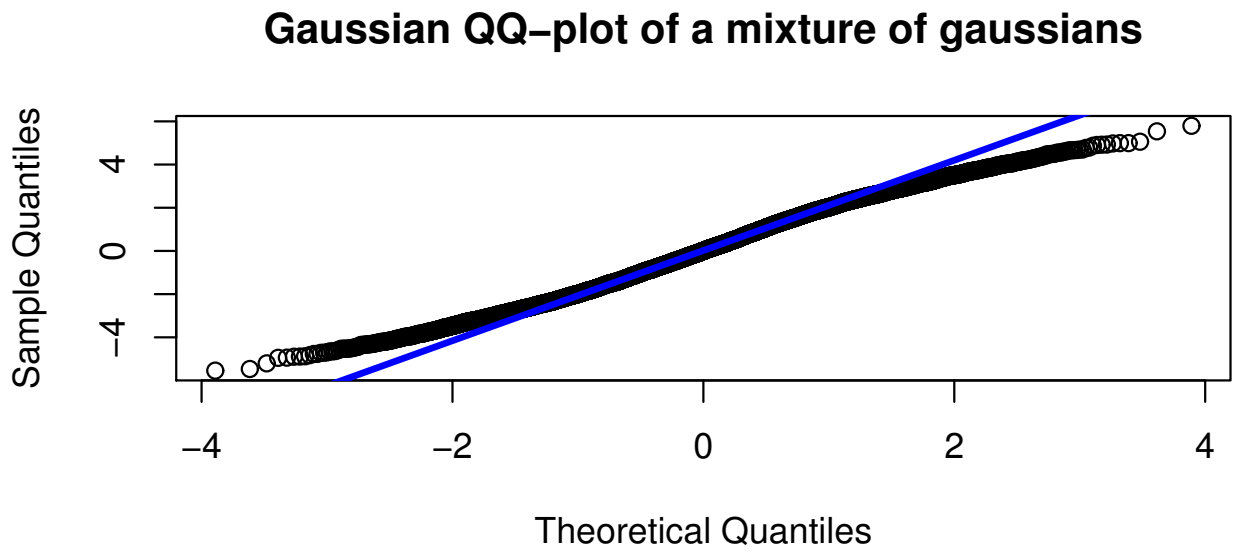

Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
10
20
30
40
x

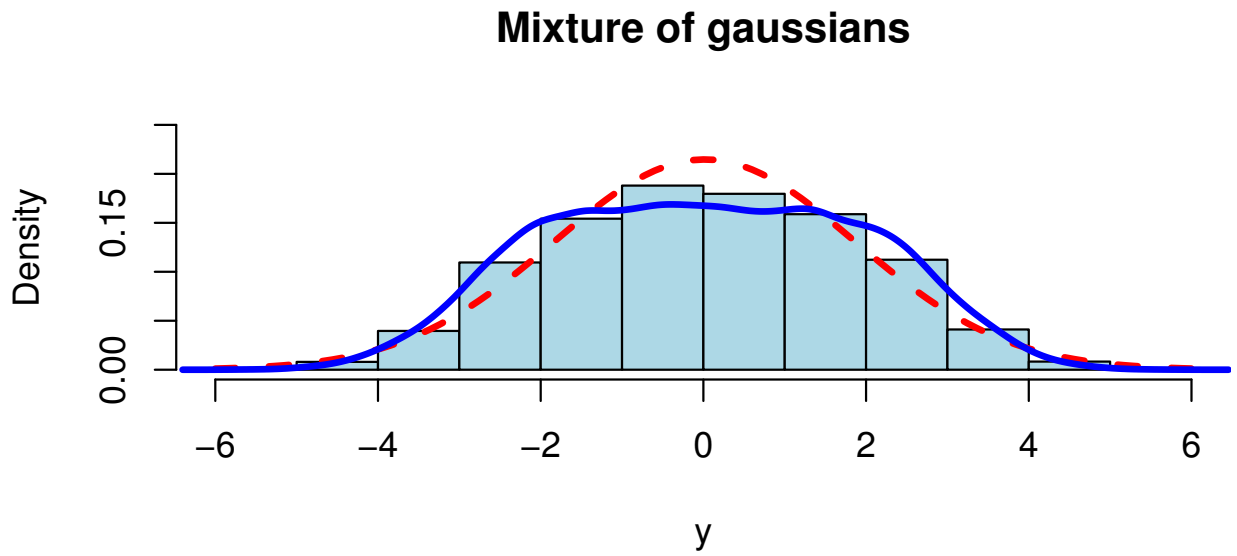
توزیع‌های آمیخته نرمال (متناهی)

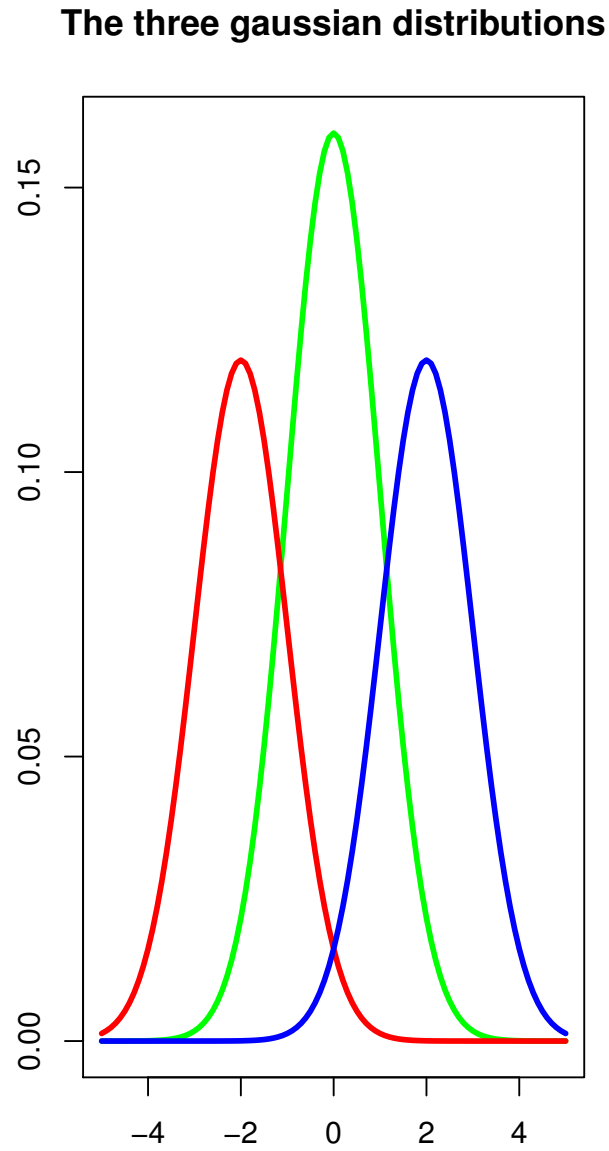
از توزیع‌ها را شامل توزیع‌های چوله، با دم‌های پهن و غیره، می‌توان با توزیع‌های آمیخته نرمال تولید کرد:

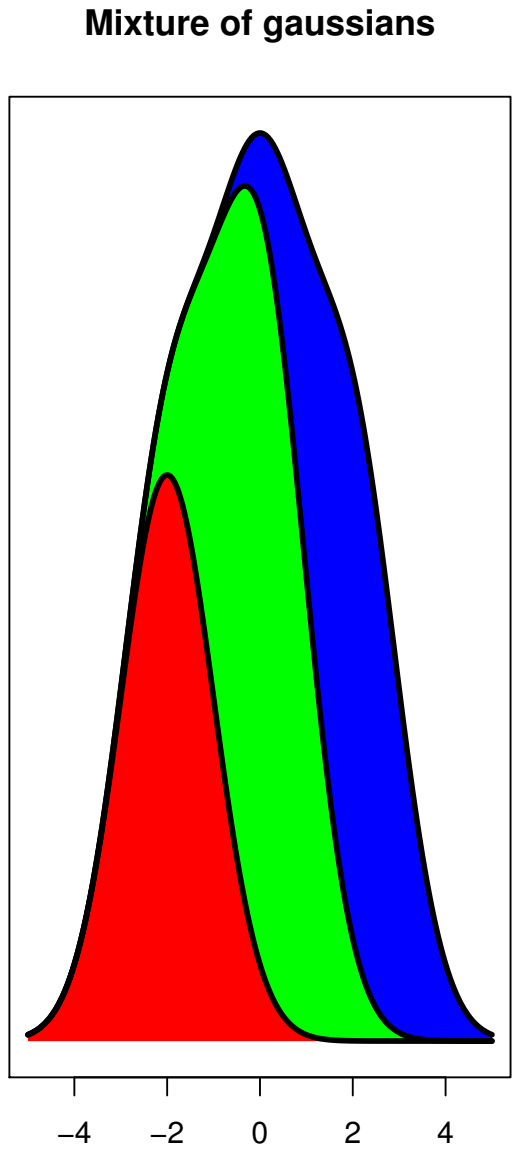
$$F_X(x) = ۰/۳N(-۲,۱) + ۰/۴N(۰,۱) + ۰/۳N(۲,۱)$$

```
n <- 10000
m <- c(-2,0,2)    # Means
p <- c(.3,.4,.3)  # Probabilities
s <- c(1, 1, 1)   # Standard deviations
x <- cbind( rnorm(n, m[1], s[1]),
            rnorm(n, m[2], s[2]),
            rnorm(n, m[3], s[3]) )
a <- sample(1:3, size=n, prob=p, replace=TRUE)
y <- x[ 1:n + n*(a-1) ]
par(mfrow=c(2,1))
qqnorm(y,
       main="Gaussian QQ-plot of a mixture of gaussians")
qqline(y, col="blue", lwd=3)
hist(y, col="light blue", probability=TRUE,
     ylim=c(0,.25),
     main="Mixture of gaussians")
curve(dnorm(x, mean=mean(y), sd=sd(y)),
      add=TRUE, col="red", lwd=3, lty=2)
lines(density(x), col="blue", lwd=3)
```


Gaussian QQ–plot of a mixture of gaussians
Sample Quantiles
−4
0
4
−4
−2
0
2
4
Theoretical Quantiles


Mixture of gaussians
Density
0.00
0.15
−6
−4
−2
0
2
4
6
y


The three gaussian distributions
0.15
0.10
0.05
0.00
−4
−2
0
2
4


Mixture of gaussians
−4
−2
0
2
4

رسم نمودار چگالی در آمیخته‌ها

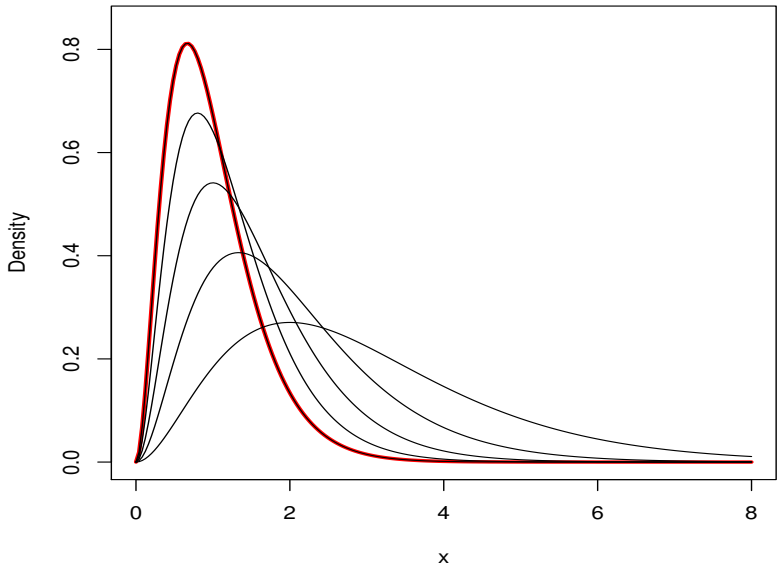
رسم تابع چگالی احتمال به عنوان مثال آمیخته گاما، باید ابتدا شکل $f(x) = \sum_{i=1}^{5} \theta_i f_i(x)$ را به عنوان یک تابع در R تعریف کنیم:

```
f <- function(x, lambda, theta) {
        # density of the mixture at the point x
        sum(dgamma(x, 3, lambda) * theta)
    }
```

این تابع مقدار چگالی آمیخته را به ازای هر x محاسبه می‌کند. دقت کنید در تابع $dgamma$، اگر x یک مقدار حقیقی باشد (نه بردار)، نتیجه برداری با طولی برابر طول $lambda$ خواهد بود، یعنی $(f_1(x), \ldots, f_5(x))$. همچنین برای همین تابع ضرب در θ.

```
x <- seq(0, 8, length=200)
dim(x) <- length(x)  # need for apply
# compute density of the mixture f(x) along x
y <- apply(x, 1, f, lambda=lambda, theta=p)
```

```
#plot the density of the mixture
plot(x, y, type="l", ylim=c(0,.85), lwd=3, col="red",
     ylab="Density")
for (j in 1:5) {
        # add the j-th gamma density to the plot
        y = apply(x, 1, dgamma, shape=3, rate=lambda[j])
        lines(x, y)
    }
```



تولید آمیخته پیوسته: آمیخته پواسون ـ گاما

دوجمله‌ای منفی، یک آمیخته از توزیع‌های پواسون $Poisson(\Lambda)$ است که در آن Λ دارای توزیع گاما است.

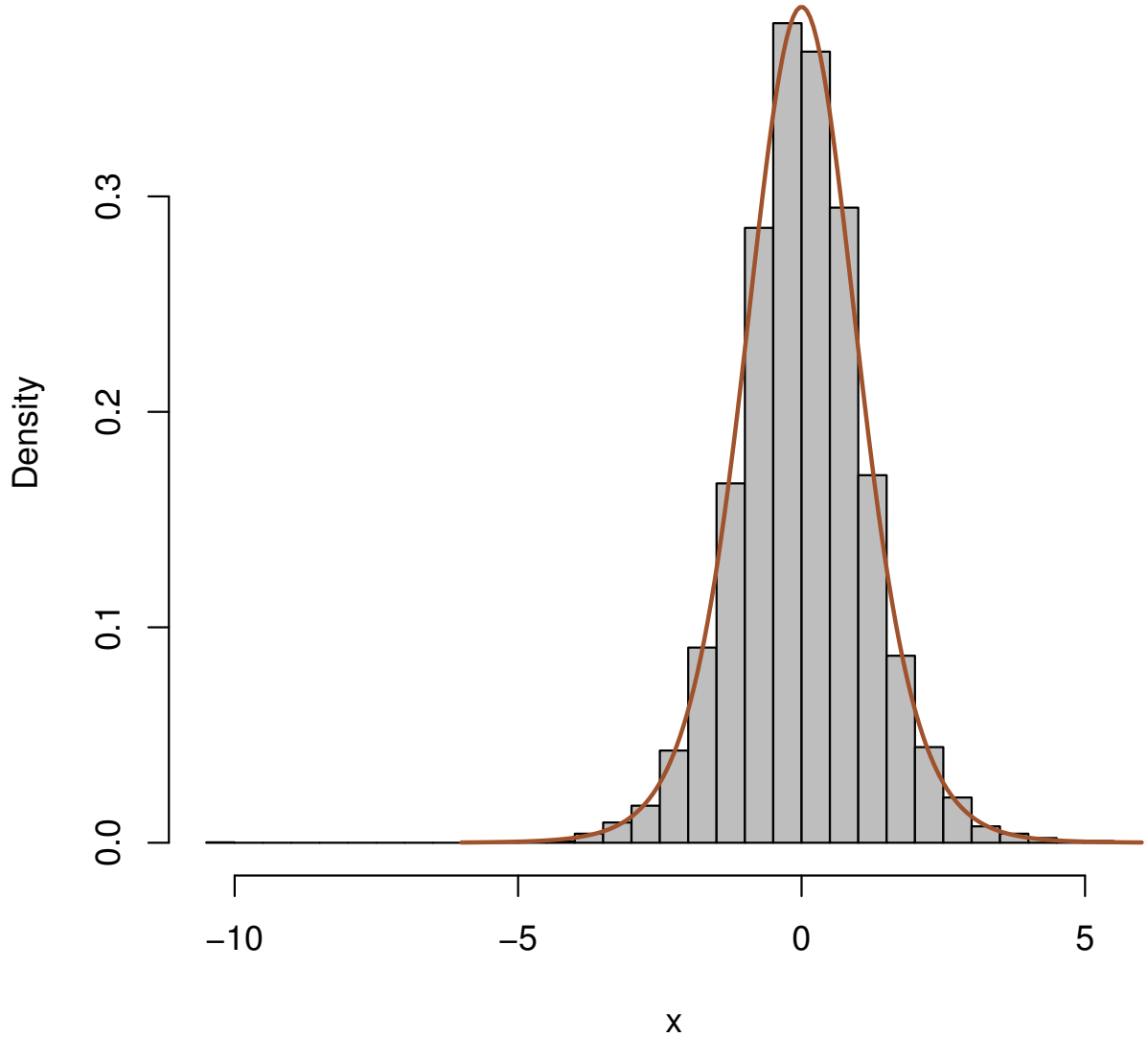
$$X|\Lambda = \lambda \sim Piosson(\lambda);\ \Lambda \sim Gamma(r, \beta),$$

$$X \sim NB(r, p = \frac{\beta}{۱ + \beta}).$$

```
# generate a Poisson-Gamma mixture
n <- 1000
r <- 4
beta <- 3
lambda <- rgamma(n, r, beta) #lambda is random
# now supply the sample of lambda's as the Poisson mean
x <- rpois(n, lambda)        #the mixture
# compare with negative binomial
mix <- tabulate(x+1) / n
negbin <- round(dnbinom(0:max(x), r, beta/(1+beta)), 3)
se <- sqrt(negbin * (1 - negbin) / n)
> round(rbind(mix, negbin, se), 3)
        [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]
mix    0.316 0.328 0.196 0.087 0.048 0.014 0.008 0.003
negbin 0.316 0.316 0.198 0.099 0.043 0.017 0.006 0.002
se     0.015 0.015 0.013 0.009 0.006 0.004 0.002 0.001
```

$$X|y \sim N(۰, \nu/y) \ \& \ Y \sim \chi^{۲}_{\nu} \longrightarrow X \sim T_{\nu}.$$

```
Nsim=10^4
nu=9
y=rchisq(Nsim,df=nu)
x=rnorm(Nsim,0,sqrt(nu/y))
hist(x,main="",freq=F,col="grey",breaks=40)
z = seq(-6,6,length.out=200)
lines(z,dt(z,nu),lwd=2,col="sienna")
> mean(x) # the mean of a t variable is 0
[1] 0.007723547
> var(x) # the variance of a t-variable is nu/nu-2=1.2857
[1] 1.294373
```


Density
0.0
0.1
0.2
0.3
−10
−5
0
5
x

توزیع‌های چندمتغیره: نرمال چندمتغیره

بردار $\mathbf{X} = (X_1, \ldots, X_d)$ دارای توزیع نرمال d متغیره است و با نماد $N_d(\mu, \Sigma)$ نمایش می‌دهند، هرگاه تابع چگالی احتمال آن به صورت زیر باشد:

$$f(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\},\ x \in \Re^d$$

به طوری که $\mu = (\mu_1, \ldots, \mu_d)^T$ بردار میانگین و $\Sigma = (\sigma_{ij})$ ماتریس $d \times d$ متقارن و معین مثبت با درایه‌های $\sigma_{ij} = Cov(X_i, X_j)$ است.

توزیع‌های چندمتغیره: نرمال چندمتغیره

بردار $\mathbf{X} = (X_1, \ldots, X_d)$ دارای توزیع نرمال d متغیره است و با نماد $N_d(\mu, \Sigma)$ نمایش می‌دهند، هرگاه تابع چگالی احتمال آن به صورت زیر باشد:

$$f(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\},\ x \in \Re^d$$

به طوری که $\mu = (\mu_1, \ldots, \mu_d)^T$ بردار میانگین و $\Sigma = (\sigma_{ij})$ ماتریس $d \times d$ متقارن و معین مثبت با درایه‌های $\sigma_{ij} = Cov(X_i, X_j)$ است.

ماتریس Σ^{-1} که معکوس ماتریس کوواریانس است، معروف به ماتریس دقت می‌باشد.

توزیع‌های چندمتغیره: نرمال چندمتغیره

بردار $\mathbf{X} = (X_1, \ldots, X_d)$ دارای توزیع نرمال d متغیره است و با نماد $N_d(\mu, \Sigma)$ نمایش می‌دهند، هرگاه تابع چگالی احتمال آن به صورت زیر باشد:

$$f(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\},\ x \in \Re^d$$

به طوری که $\mu = (\mu_1, \ldots, \mu_d)^T$ بردار میانگین و $\Sigma = (\sigma_{ij})$ ماتریس $d \times d$ متقارن و معین مثبت با درایه‌های $\sigma_{ij} = Cov(X_i, X_j)$ است.

ماتریس Σ^{-1} که معکوس ماتریس کوواریانس است، معروف به ماتریس دقت می‌باشد.

یک متغیر نرمال چندمتغیره را می‌توان در دو مرحله تولید کرد:

1. ابتدا تولید یک نمونه تصادفی ساده $\mathbf{Z} = (Z_1, \ldots, Z_d)$ از توزیع نرمال استاندارد
2. تبدیل بردار $\mathbf{Z}$ به طوری که دارای بردار میانگین و ماتریس کوواریانس مورد نظر باشد

توزیع‌های چندمتغیره: نرمال چندمتغیره

بردار $\mathbf{X} = (X_1, \ldots, X_d)$ دارای توزیع نرمال d متغیره است و با نماد $N_d(\mu, \Sigma)$ نمایش می‌دهند، هرگاه تابع چگالی احتمال آن به صورت زیر باشد:

$$f(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\},\ x \in \Re^d$$

به طوری که $\mu = (\mu_1, \ldots, \mu_d)^T$ بردار میانگین و $\Sigma = (\sigma_{ij})$ ماتریس $d \times d$ متقارن و معین مثبت با درایه‌های $\sigma_{ij} = Cov(X_i, X_j)$ است.

ماتریس Σ^{-1} که معکوس ماتریس کوواریانس است، معروف به ماتریس دقت می‌باشد.

یک متغیر نرمال چندمتغیره را می‌توان در دو مرحله تولید کرد:

1. ابتدا تولید یک نمونه تصادفی ساده $\mathbf{Z} = (Z_1, \ldots, Z_d)$ از توزیع نرمال استاندارد
2. تبدیل بردار $\mathbf{Z}$ به طوری که دارای بردار میانگین و ماتریس کوواریانس مورد نظر باشد

تبدیل مورد نظر نیازمند تجزیه ماتریس کوواریانس است

$$Z \sim N_d(\mu, \Sigma) \longrightarrow CZ + b \sim N_d(C\mu + b, C\Sigma C^T)$$
$$Z \sim N_d(0, I_d) \longrightarrow CZ + b \sim N_d(b, CC^T).$$

حال فرض کنید بتوانیم ماتریس کوواریانس را برای یک ماتریس C به صورت $\Sigma = CC^T$ تجزیه کنیم. پس:

$$CZ + \mu \sim N_d(\mu, \Sigma).$$

$$Z \sim N_d(\mu, \Sigma) \longrightarrow CZ + b \sim N_d(C\mu + b, C\Sigma C^T)$$
$$Z \sim N_d(0, I_d) \longrightarrow CZ + b \sim N_d(b, CC^T).$$

حال فرض کنید بتوانیم ماتریس کوواریانس را برای یک ماتریس C به صورت $\Sigma = CC^T$ تجزیه کنیم. پس:

$$CZ + \mu \sim N_d(\mu, \Sigma).$$

تجزیه مورد نیاز برای Σ را می‌توان به وسیله روش‌های تجزیه طیفی، تجزیه چولسکی ($Cholesky$)، یا تجزیه مقدار ویژه (SVD) انجام داد.

$$Z \sim N_d(\mu, \Sigma) \longrightarrow CZ + b \sim N_d(C\mu + b, C\Sigma C^T)$$
$$Z \sim N_d(0, I_d) \longrightarrow CZ + b \sim N_d(b, CC^T).$$

حال فرض کنید بتوانیم ماتریس کوواریانس را برای یک ماتریس C به صورت $\Sigma = CC^T$ تجزیه کنیم. پس:

$$CZ + \mu \sim N_d(\mu, \Sigma).$$

تجزیه مورد نیاز برای Σ را می‌توان به وسیله روش‌های تجزیه طیفی، تجزیه چولسکی ($Cholesky$)، یا تجزیه مقدار ویژه (SVD) انجام داد.

توابع متناظر آن‌ها در R عبارتند از $eigen$، $chol$، و svd.

$$
\begin{aligned}
Z &\sim N_d(\mu, \Sigma) \longrightarrow CZ + b \sim N_d(C\mu + b, C\Sigma C^T) \\
Z &\sim N_d(0, I_d) \longrightarrow CZ + b \sim N_d(b, CC^T).
\end{aligned}
$$

حال فرض کنید بتوانیم ماتریس کوواریانس را برای یک ماتریس C به صورت $\Sigma = CC^T$ تجزیه کنیم. پس:

$$CZ + \mu \sim N_d(\mu, \Sigma).$$

تجزیه مورد نیاز برای Σ را می‌توان به وسیله روش‌های تجزیه طیفی، تجزیه چولسکی $(Cholesky)$، یا تجزیه مقدار ویژه (SVD) انجام داد.

توابع متناظر آن‌ها در R عبارتند از $eigen$، $chol$، و svd.

در حالت ماتریسی، فرض کنید $Z = (Z_{ij})$ یک ماتریس $n \times d$ باشد که Z_{ij} ها نمونه تصادفی ساده از نرمال استاندارد هستند. در این حالت

$$X = ZQ + J\mu^T,$$

که در آن $Q^TQ = \Sigma$ و J یک بردار از ۱ ها می‌باشد.

تولید نمونه‌هایی از نرمال چندمتغیره

تولید یک نمونه n تایی از $N_d(\mu, \Sigma)$ باید:

1. یک ماتریس $n \times d$، Z، شامل nd متغیر نرمال استاندارد تولید کنیم
2. تجزیه $\Sigma = Q^T Q$ را انجام دهیم
3. تبدیل $X = ZQ + J\mu^T$ را اجرا کنیم
4. ماتریس X با بعد $n \times d$ را که سطرهای آن نمونه‌های یک نرمال چندمتغیره است را برگردانیم

تبدیل $X = ZQ + J\mu^T$ را در R می‌توان به صورت زیر اجرا کرد:

```
Z = matrix(rnorm(n*d), nrow = n, ncol = d)
X = Z%*%Q+matrix(mu, n, d, byrow=TRUE)
```

تجزیه ماتریس کوواریانس به روش طیفی

دوم ماتریس کوواریانس عبارتست از $\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}$ که در آن Λ ماتریس قطری با مقادیر ویژه ماتریس Σ است و P ماتریسی است که ستون‌های آن بردارهای ویژه Σ متناظر با مقادیر ویژه آن است.

تجزیه ماتریس کوواریانس به روش طیفی

دوم ماتریس کوواریانس عبارتست از $\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}$ که در آن Λ ماتریس قطری با مقادیر ویژه ماتریس Σ است و P ماتریسی است که ستون‌های آن بردارهای ویژه Σ متناظر با مقادیر ویژه آن است.

در این تجزیه داریم $P^{-1} = P^T$ و بنابراین $\Sigma^{1/2} = P\Lambda^{1/2}P^T$

تجزیه ماتریس کوواریانس به روش طیفی

دوم ماتریس کوواریانس عبارتست از $\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}$ که در آن Λ ماتریس قطری با مقادیر ویژه ماتریس Σ است و P ماتریسی است که ستون‌های آن بردارهای ویژه Σ متناظر با مقادیر ویژه آن است.

در این تجزیه داریم $P^{-1} = P^T$ و بنابراین $\Sigma^{1/2} = P\Lambda^{1/2}P^T$

بنابراین در این تجزیه، $Q = \Sigma^{1/2}$.

تجزیه ماتریس کوواریانس به روش طیفی

دوم ماتریس کوواریانس عبارتست از $\Sigma^{1/2} = P\Lambda^{1/2}P^{-1}$ که در آن Λ ماتریس قطری با مقادیر ویژه ماتریس Σ است و P ماتریسی است که ستون‌های آن بردارهای ویژه Σ متناظر با مقادیر ویژه آن است.

در این تجزیه داریم $P^{-1} = P^T$ و بنابراین $\Sigma^{1/2} = P\Lambda^{1/2}P^T$

بنابراین در این تجزیه، $Q = \Sigma^{1/2}$.

به عنوان مثال تولید نمونه از توزیع نرمال دومتغیره با بردار میانگین صفر و ماتریس کوواریانس

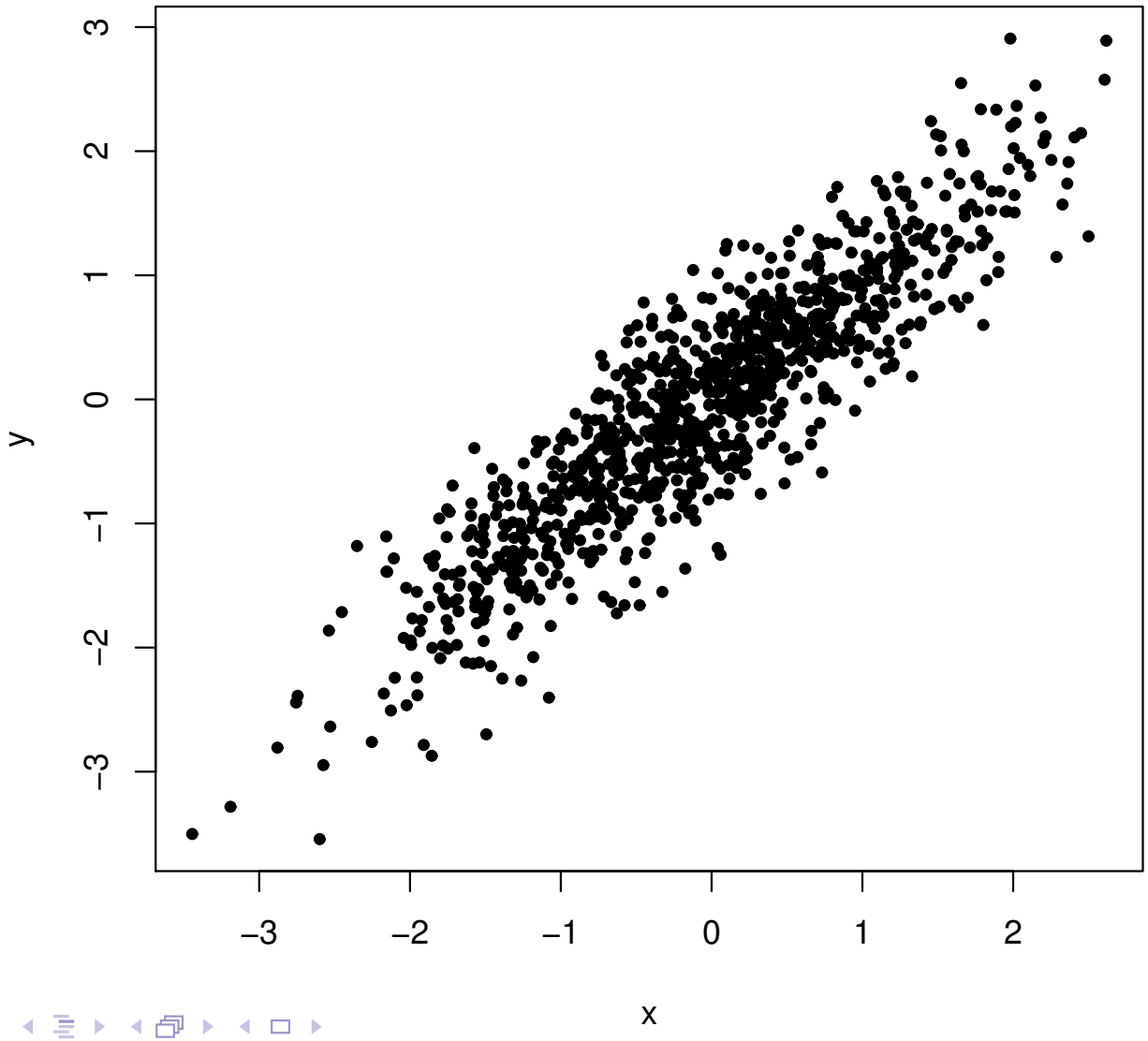
$$\Sigma = \begin{bmatrix} ۱/۰ & ۰/۹ \\ ۰/۹ & ۱/۰ \end{bmatrix}$$

را در نظر بگیرید

```
# mean and covariance parameters
mu <- c(0, 0)
Sigma <- matrix(c(1, .9, .9, 1), nrow = 2, ncol = 2)
```

```
rmvn.eigen = function(n, mu, Sigma) {
        # generate n random vectors from MVN(mu, Sigma)
        # dimension is inferred from mu and Sigma
        d = length(mu)
        ev = eigen(Sigma, symmetric = TRUE)
        lambda = ev$values
        V = ev$vectors
        R = V %*% diag(sqrt(lambda)) %*% t(V)
        Z = matrix(rnorm(n*d), nrow = n, ncol = d)
        X = Z %*% R + matrix(mu, n, d, byrow = TRUE)
        X
    }
##
X = rmvn.eigen(1000, mu, Sigma)
>     print(colMeans(X))
[1] -0.06803006 -0.06935001
>     print(cor(X))
          [,1]      [,2]
[1,] 1.0000000 0.8969038
[2,] 0.8969038 1.0000000
```

```
plot(X, xlab = "x", ylab = "y", pch = 20)
```



تجزیه ماتریس به روش SVD

تجزیه مقدار ویژه ($singular\ value$)، ایده بردارهای ویژه را به ماتریس‌های مستطیلی تعمیم می‌دهد. این تجزیه برای ماتریس X به صورت زیر است:

$$X = UDV^T,$$

که در آن D برداری شامل مقادیر ویژه X، U ماتریسی با ستون‌های شامل بردارهای ویژه چپ X و V ماتریسی با ستون‌های شامل بردارهای ویژه راست ماتریس X است.

تجزیه ماتریس به روش SVD

تجزیه مقدار ویژه ($singular\ value$)، ایده بردارهای ویژه را به ماتریس‌های مستطیلی تعمیم می‌دهد. این تجزیه برای ماتریس X به صورت زیر است:

$$X = UDV^T,$$

که در آن D برداری شامل مقادیر ویژه X، U ماتریسی با ستون‌های شامل بردارهای ویژه چپ X و V ماتریسی با ستون‌های شامل بردارهای ویژه راست ماتریس X است.

ماتریس مورد علاقه ما در اینجا Σ است که برای آن $UV^T = I$. تجزیه SVD یک ماتریس متقارن معین مثبت، رابطه $U = V = P$ را نتیجه می‌دهد. بنابراین $\Sigma^{1/2} = UD^{1/2}V^T$.

تجزیه ماتریس به روش SVD

تجزیه مقدار ویژه ($singular\ value$)، ایده بردارهای ویژه را به ماتریس‌های مستطیلی تعمیم می‌دهد. این تجزیه برای ماتریس X به صورت زیر است:

$$X = UDV^T,$$

که در آن D برداری شامل مقادیر ویژه X، U ماتریسی با ستون‌های شامل بردارهای ویژه چپ X و V ماتریسی با ستون‌های شامل بردارهای ویژه راست ماتریس X است.

ماتریس مورد علاقه ما در اینجا Σ است که برای آن $UV^T = I$. تجزیه SVD یک ماتریس متقارن معین مثبت، رابطه $U = V = P$ را نتیجه می‌دهد. بنابراین $\Sigma^{1/2} = UD^{1/2}V^T$.

برای این کاربرد، روش تجزیه SVD با تجزیه طیفی یکی است، اما از کارایی کمتری نسبت به تجزیه طیفی برخوردار است. زیرا تجزیه SVD از این اطلاع که ماتریس کوواریانس متقارن مربعی است، استفاده نمی‌کند.

```
rmvn.svd = function(n, mu, Sigma) {
        # generate n random vectors from MVN(mu, Sigma)
        # dimension is inferred from mu and Sigma
        d = length(mu)
        S = svd(Sigma)
        R = S$u %*% diag(sqrt(S$d)) %*% t(S$v)
        Z = matrix(rnorm(n*d), nrow=n, ncol=d)
        X = Z %*% R + matrix(mu, n, d, byrow=TRUE)
        X
    }
```

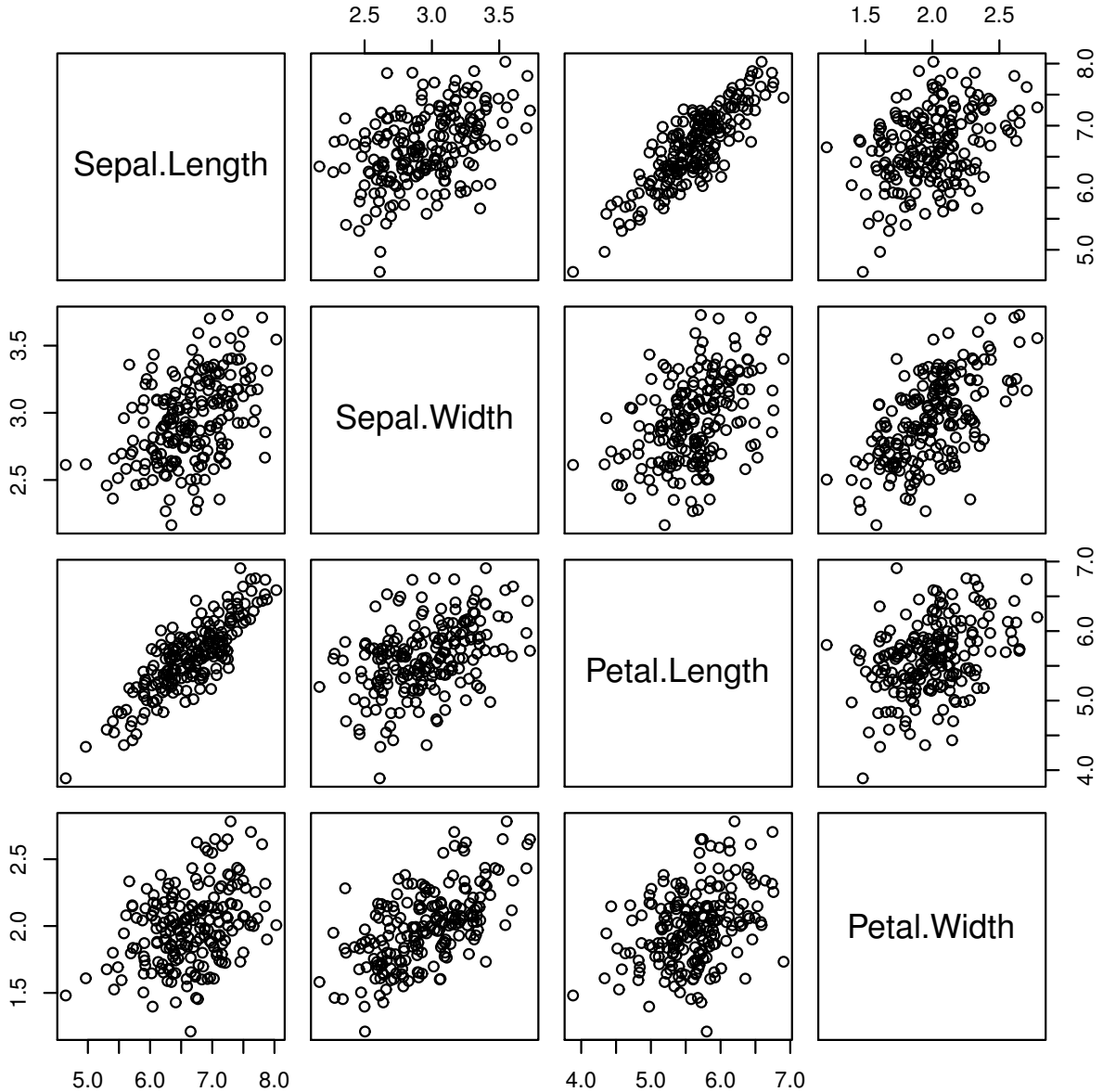
تجزیه به روش چولسکی

چولسکی یک ماتریس متقارن معین مثبت به صورت $X = Q^T Q$ است که در آن، Q یک ماتریس بالا مثلثی است. تابع $chol(X)$ در R این تجزیه را انجام می‌دهد و خروجی آن یک ماتریس بالا مثلثی مانند S است به طوری که $S^T S = \Sigma$.

```
rmvn.Choleski = 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
    }
```

تجزیه چولسکی برای داده‌های *iris*

```
y <- subset(x=iris, Species=="virginica")[, 1:4]
mu <- colMeans(y)
Sigma <- cov(y)
>      mu
Sepal.Length  Sepal.Width Petal.Length  Petal.Width
       6.588        2.974        5.552        2.026
>      Sigma
             Sepal.Length Sepal.Width Petal.Length Petal.Width
Sepal.Length   0.40434286  0.09376327   0.30328980  0.04909388
Sepal.Width    0.09376327  0.10400408   0.07137959  0.04762857
Petal.Length   0.30328980  0.07137959   0.30458776  0.04882449
Petal.Width    0.04909388  0.04762857   0.04882449  0.07543265
X <- rmvn.Choleski(200, mu, Sigma)
pairs(X)
```


Sepal.Length
Sepal.Width
Petal.Length
Petal.Width
2.5
3.0
3.5
1.5
2.0
2.5
5.0
6.0
7.0
8.0
2.5
3.0
3.5
4.0
5.0
6.0
7.0
1.5
2.0
2.5
5.0
6.0
7.0
8.0
4.0
5.0
6.0
7.0

مقایسه روش‌های تجزیه

سوال که در اینجا ممکن است مطرح شود، آن است که کدام یک از این روش‌ها ارجح است؟

مقایسه روش‌های تجزیه

سوال که در اینجا ممکن است مطرح شود، آن است که کدام یک از این روش‌ها ارجح است؟

یک معیار برای جواب دادن به این سوال مقایسه سرعت این چند روش است.

مقایسه روش‌های تجزیه

سوال که در اینجا ممکن است مطرح شود، آن است که کدام یک از این روش‌ها ارجح است؟

یک معیار برای جواب دادن به این سوال مقایسه سرعت این چند روش است.

این سه روش را با دو تابع $mvrnorm$ در بسته $MASS$ و $rmvnorm$ در بسته $mvtnorm$ مقایسه می‌کنیم.

مقایسه روش‌های تجزیه

سوال که در اینجا ممکن است مطرح شود، آن است که کدام یک از این روش‌ها ارجح است؟

یک معیار برای جواب دادن به این سوال مقایسه سرعت این چند روش است.

این سه روش را با دو تابع $mvrnorm$ در بسته $MASS$ و $rmvnorm$ در بسته $mvtnorm$ مقایسه می‌کنیم.

تابع $mvrnorm$ بر اساس تجزیه طیفی عمل می‌کند.

مقایسه روش‌های تجزیه

سوال که در اینجا ممکن است مطرح شود، آن است که کدام یک از این روش‌ها ارجح است؟

یک معیار برای جواب دادن به این سوال مقایسه سرعت این چند روش است.

این سه روش را با دو تابع $mvrnorm$ در بسته $MASS$ و $rmvnorm$ در بسته $mvtnorm$ مقایسه می‌کنیم.

تابع $mvrnorm$ بر اساس تجزیه طیفی عمل می‌کند.

عنوان شده است که: اگرچه تجزیه چولسکی سریعتر است، تجزیه طیفی پایدارتر است.

```
library(MASS)
library(mvtnorm)
n = 100 # sample size
d <- 30 # dimension
N <- 2000 # iterations
mu <- numeric(d)
##
> set.seed(100)
>     system.time(for (i in 1:N)
+         rmvn.eigen(n, mu, cov(matrix(rnorm(n*d), n, d))))
   user  system elapsed
   4.02    0.00    4.07
>     set.seed(100)
>     system.time(for (i in 1:N)
+         rmvn.svd(n, mu, cov(matrix(rnorm(n*d), n, d))))
   user  system elapsed
   5.05    0.00    5.12
>     set.seed(100)
>     system.time(for (i in 1:N)
+         rmvn.Choleski(n, mu, cov(matrix(rnorm(n*d), n, d))))
   user  system elapsed
   3.74    0.02    3.76
>     set.seed(100)
>     system.time(for (i in 1:N)
+         mvrnorm(n, mu, cov(matrix(rnorm(n*d), n, d))))
   user  system elapsed
   3.88    0.02    3.93
>     set.seed(100)
>     system.time(for (i in 1:N)
+         rmvnorm(n, mu, cov(matrix(rnorm(n*d), n, d))))
   user  system elapsed
   4.82    0.00    4.87
```

توزیع‌های آمیخته نرمال چندمتغیره

نرمال آمیخته با دو مولفه:

$$pN_d(\mu_1, \Sigma_1) + (1-p)N_d(\mu_2, \Sigma_2),$$

با تغییر p، به عنوان پارامتر آمیختگی، توزیع‌های متفاوت با شکل‌ها و خواص متفاوتی خواهیم داشت.

توزیع‌های آمیخته نرمال چندمتغیره

نرمال آمیخته با دو مولفه:

$$pN_d(\mu_1, \Sigma_1) + (1-p)N_d(\mu_2, \Sigma_2),$$

با تغییر p، به عنوان پارامتر آمیختگی، توزیع‌های متفاوت با شکل‌ها و خواص متفاوتی خواهیم داشت.

به عنوان مثال یک آمیخته ٪۵۰ توزیعی متقارن با دم‌های سبک است، در حالی که یک آمیخته ٪۹۰ چوله با دم‌های سنگین است.

توزیع‌های آمیخته نرمال چندمتغیره

نرمال آمیخته با دو مولفه:

$$pN_d(\mu_1, \Sigma_1) + (1-p)N_d(\mu_2, \Sigma_2),$$

با تغییر p، به عنوان پارامتر آمیختگی، توزیع‌های متفاوت با شکل‌ها و خواص متفاوتی خواهیم داشت.

به عنوان مثال یک آمیخته ٪۵۰ توزیعی متقارن با دم‌های سبک است، در حالی که یک آمیخته ٪۹۰ چوله با دم‌های سنگین است.

با توجه به این ویژگی توزیع‌های آمیخته نرمال، از آن‌ها در استنباط‌های تنومند زیاد استفاده می‌شود.

توزیع‌های آمیخته نرمال چندمتغیره

نرمال آمیخته با دو مولفه:

$$pN_d(\mu_1, \Sigma_1) + (1-p)N_d(\mu_2, \Sigma_2),$$

با تغییر p، به عنوان پارامتر آمیختگی، توزیع‌های متفاوت با شکل‌ها و خواص متفاوتی خواهیم داشت.

به عنوان مثال یک آمیخته ٪۵۰ توزیعی متقارن با دم‌های سبک است، در حالی که یک آمیخته ٪۹۰ چوله با دم‌های سنگین است.

با توجه به این ویژگی توزیع‌های آمیخته نرمال، از آن‌ها در استنباط‌های تنومند زیاد استفاده می‌شود.

از چنین توزیعی می‌توان به طریق زیر، نمونه تولید کرد:

1. تولید $U \sim Uniform(0,1)$
2. اگر $U \leq p$، X را از $N_d(\mu_1, \Sigma_1)$ تولید کن، در غیر این صورت از $N_d(\mu_2, \Sigma_2)$ تولید کن

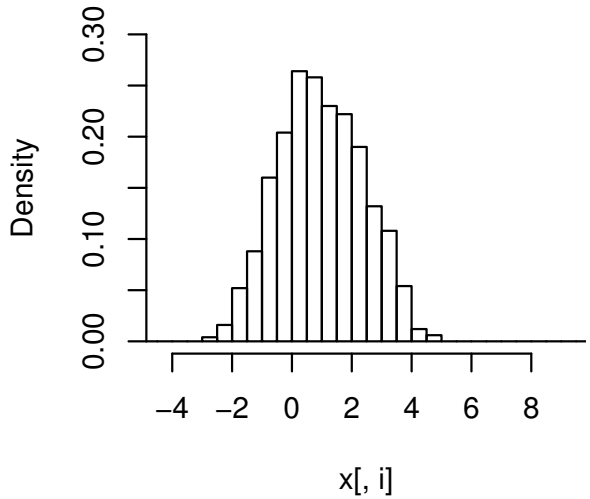
روش زیر هم می‌توان عمل کرد:

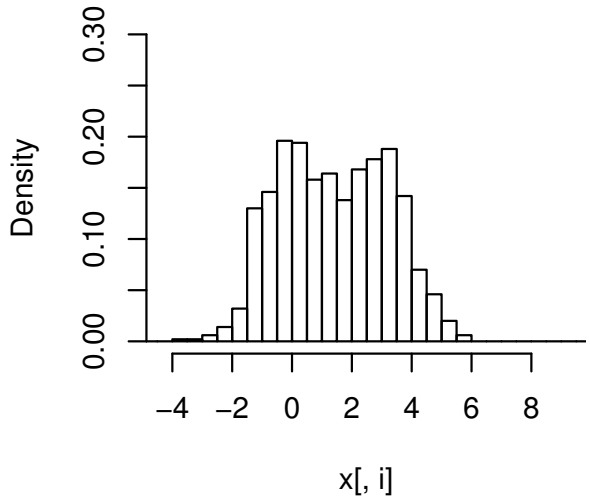
۱. N را از توزیع $Ber(p)$ تولید کن

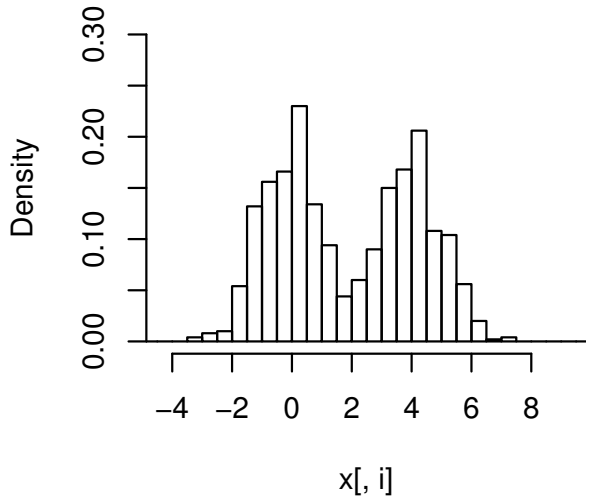
۲. اگر $N = ۱$، X را از $N_d(\mu_۱, \Sigma_۱)$ تولید کن در غیر این صورت از $N_d(\mu_۲, \Sigma_۲)$

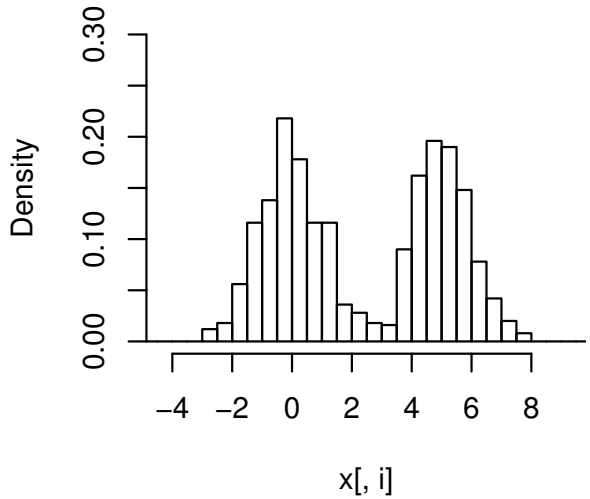
```
library(MASS)  #for mvrnorm
# ineffecient version loc.mix.0 with loops
loc.mix.0 <- function(n, p, mu1, mu2, Sigma) {
        # generate sample from BVN location mixture
        X <- matrix(0, n, 2)
        for (i in 1:n) {
            k <- rbinom(1, size = 1, prob = p)
            if (k)
                X[i,] <- mvrnorm(1, mu = mu1, Sigma) else
                X[i,] <- mvrnorm(1, mu = mu2, Sigma)
            }
        return(X)
    }
```

```
#more efficient version
    loc.mix <- function(n, p, mu1, mu2, Sigma) {
        #generate sample from BVN location mixture
        n1 <- rbinom(1, size = n, prob = p)
        n2 <- n - n1
        x1 <- mvrnorm(n1, mu = mu1, Sigma)
        x2 <- mvrnorm(n2, mu = mu2, Sigma)
        X <- rbind(x1, x2)              #combine the samples
        return(X[sample(1:n), ])        #mix them
    }
x <- loc.mix(1000, .5, rep(0, 4), 2:5, Sigma = diag(4))
r <- range(x) * 1.2
par(mfrow = c(2, 2))
for (i in 1:4)
        hist(x[ , i], xlim = r, ylim = c(0, .3), freq = FALSE,
        main = "", breaks = seq(-5, 10, .5))
```


Density
0.00
0.10
0.20
0.30
−4
−2
0
2
4
6
8
x[, i]


Density
0.00
0.10
0.20
0.30
−4
−2
0
2
4
6
8
x[, i]


Density
0.00
0.10
0.20
0.30
−4
−2
0
2
4
6
8
x[, i]


Density
0.00
0.10
0.20
0.30
−4
−2
0
2
4
6
8
x[, i]

توزیع ویشارت

اگر سطرهای ماتریس X $n \times d$ یک نمونه تصادفی ساده از توزیع $N_d(\cdot, \Sigma)$ باشد، آن‌گاه $M = X^T X \sim Wishart_d(\Sigma, n)$.

توزیع ویشارت

اگر سطرهای ماتریس X $n \times d$ یک نمونه تصادفی ساده از توزیع $N_d(۰, \Sigma)$ باشد، آن‌گاه $M = X^T X \sim Wishart_d(\Sigma, n)$.

به عنوان مثال ماتریس کوواریانس نمونه (S) دارای توزیع ویشارت است: $S \sim Wishart_d(\Sigma/n - ۱, n - ۱)$.

توزیع ویشارت

اگر سطرهای ماتریس X $n \times d$ یک نمونه تصادفی ساده از توزیع $N_d(۰, \Sigma)$ باشد، آن‌گاه $M = X^T X \sim Wishart_d(\Sigma, n)$.

به عنوان مثال ماتریس کوواریانس نمونه (S) دارای توزیع ویشارت است: $S \sim Wishart_d(\Sigma/n - ۱, n - ۱)$.

اگر $S \sim Wishart_d(\Sigma, n)$، و A یک ماتریس $q \times d$ با رتبه کامل q باشد، $ASA^T \sim Wishart_q(A\Sigma A^T, n)$.

توزیع ویشارت

اگر سطرهای ماتریس X $n \times d$ یک نمونه تصادفی ساده از توزیع $N_d(۰, \Sigma)$ باشد، آن‌گاه $M = X^T X \sim Wishart_d(\Sigma, n)$.

به عنوان مثال ماتریس کوواریانس نمونه (S) دارای توزیع ویشارت است: $S \sim Wishart_d(\Sigma/n - ۱, n - ۱)$.

اگر $S \sim Wishart_d(\Sigma, n)$، و A یک ماتریس $q \times d$ با رتبه کامل q باشد، $ASA^T \sim Wishart_q(A\Sigma A^T, n)$.

یکی از کاربردهای آن در مدل‌بندی بیزی مدل‌های رگرسیونی زمانی که برای ماتریس کوواریانس توزیع پیشین ویشارت انتخاب می‌شود.

توزیع ویشارت

اگر سطرهای ماتریس X $n \times d$ یک نمونه تصادفی ساده از توزیع $N_d(۰, \Sigma)$ باشد، آن‌گاه $M = X^T X \sim Wishart_d(\Sigma, n)$.

به عنوان مثال ماتریس کوواریانس نمونه (S) دارای توزیع ویشارت است: $S \sim Wishart_d(\Sigma/n - ۱, n - ۱)$.

اگر $S \sim Wishart_d(\Sigma, n)$، و A یک ماتریس $q \times d$ با رتبه کامل q باشد، $ASA^T \sim Wishart_q(A\Sigma A^T, n)$.

یکی از کاربردهای آن در مدل‌بندی بیزی مدل‌های رگرسیونی زمانی که برای ماتریس کوواریانس توزیع پیشین ویشارت انتخاب می‌شود.

نمونه‌گیری از توزیع ویشارت با تولید نمونه از نرمال چندمتغیره و ضرب ماتریسی آن‌ها انجام می‌شود. اما دو عیب اساسی دارد:

- برای nهای بزرگ کارا نیست
- برای nهای ناصحیح قابل کاربرد نیست

تجزیه بارتلت

یک روش کاراتر استفاده از تجزیه بارتلت می‌باشد.

تجزیه بارتلت

یک روش کاراتر استفاده از تجزیه بارتلت می‌باشد.

اگر $S \sim Wishart_d(I, n)$ دارای تجزیه چولسکی $S = LL^T$ باشد، آنگاه $p(p+۱)/۲$ عضو ناصفر L مستقل از هم و دارای توزیع‌های

- $t_{ii} \sim \sqrt{\chi^۲_{n-i+۱}}$
- $t_{ij} \sim N(۰, ۱)\ (i \neq j)$

هستند. برای تولید $Wishart_d(\Sigma, n)$ ، می‌توان تجزیه $\Sigma = AA^T$ را انتخاب کرد و مقدار ASA^T را برگرداند.

تجزیه بارتلت

یک روش کاراتر استفاده از تجزیه بارتلت می‌باشد.

اگر $S \sim Wishart_d(I, n)$ دارای تجزیه چولسکی $S = LL^T$ باشد، آنگاه $p(p+۱)/۲$ عضو ناصفر L مستقل از هم و دارای توزیع‌های

- $t_{ii} \sim \sqrt{\chi^۲_{n-i+۱}}$
- $t_{ij} \sim N(۰, ۱)\ (i \neq j)$

هستند. برای تولید $Wishart_d(\Sigma, n)$ ، می‌توان تجزیه $\Sigma = AA^T$ را انتخاب کرد و مقدار ASA^T را برگرداند.

در بسته $mixAK$، تابع $rWishart(n, df, S)$ و در بسته $bayesSurv$، تابع $rWishart(n, df, S)$ این کار را انجام می‌دهند. در این دو تابع، n تعداد ماتریس‌های ویشارت و df درجه آزادی است.