

Генерисање случајних величина (2.део)

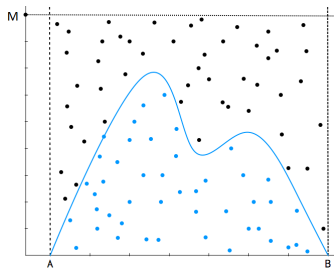
3.час

15. март 2016.

Моделовање непрекидних случајних величина

Метода одбацивања функције (Нојманова метода)

Претпоставимо случајна величина X има густину расподеле $f(x)$ која је ограничена на неком коначном интервалу.



Случајну променљиву X симулирамо на следећи начин:

- 1 Генеришемо два случајна броја γ_1 и γ_2 .
- 2 Рачунамо $x = A + (B - A)\gamma_1$ и $y = M\gamma_2$.

Ако је $y < f(x)$, тада је x реализована вредност случајне величине X .

Моделовање бета расподеле

Симулирајте $B(2, 4)$ расподелу користећи методу одбацивања.

Функција густине је

$$f(x) = B(2, 4)x(1 - x)^3 = \frac{5!}{1!3!}x(1 - x)^3 = 20x(1 - x)^3, \quad 0 < x < 1.$$

Моделовање нормалне расподеле

Постоје различите методе за моделовање нормалне расподеле, јер је овим методама посвећена посебна пажња, због значаја и честе примене нормалне расподеле.

Тврђење: Ако случајна величина X има нормалну расподелу $N(m, \sigma^2)$, тада случајна величина $Y = \frac{X-m}{\sigma}$ има стандардну нормалну расподелу $N(0, 1)$.

С обзиром на тврђење, довољно је навести поступке моделирања случајне величине која има стандардну нормалну расподелу $N(0, 1)$.

Моделовање нормалне расподеле на основу централне граничне теореме

Нека су дате независне случајне величине $Y_1, Y_2, \dots, Y_n$ са $U(0, 1)$ расподелом, тада случајна величина $S_n = \sum_{j=1}^n Y_j$ има математичко очекивање $E(S_n) = \frac{n}{2}$ и дисперзију $D(S_n) = \frac{n}{12}$.

Према централној граничној теорему: $S_n^* = \frac{S_n - E(S_n)}{\sqrt{D(S_n)}} \sim N(0, 1)$.

Конвергенција је брза и већ за $n = 12$ добијају се врло мала одступања, па ако су $\gamma_1, \dots, \gamma_{12}$ (псеудо)случајни бројеви, може се сматрати да је $\sum_{i=1}^{12} \gamma_i - 6$ реализована вредност случајне величине са $N(0, 1)$ расподелом.

На овај начин за моделовање једне вредности случајне величине са стандардном нормалном расподелом $N(0, 1)$ је потребно 12 (псеудо) случајних бројева.

Моделовање нормалне расподеле на основу централне граничне теореме

```
normalna_cgt_1 <- function (){
  gama<-c()
  for(i in 1:12)
    gama[i]<-runif(1)
  x<-sum(gama)-6
  return(x)
}

normalna_cgt_2 <- function(n){
  Y<-c()
  for(i in 1:n)
    Y[i]<-runif(1)
  x<-sqrt(3/n)*sum(2*Y-1)
  return(x)
}
```

```
> normalna_cgt_1()
[1] -1.755184
> normalna_cgt_2(6)
[1] 1.15483
```

Моделовање нормалне расподеле Вох-Muller-ов методом

Нормална нормирана расподела се може моделирати кориштењем Вох-Muller-ове трансформације.

Тврђење: Ако су Y_1 и Y_2 независне случајне величине са униформном расподелом $U(0, 1)$, тада су случајне величине:

$$Z_1 = \sqrt{-2 \ln Y_1} \cos(2\pi Y_2), \quad Z_2 = \sqrt{-2 \ln Y_1} \sin(2\pi Y_2)$$

независне са $N(0, 1)$ расподелама.

Дакле, ако су γ_1 и γ_2 два (псеудо)случајна броја, помоћу њих се добијају реализације x_1 и x_2 две независне случајне величине са нормалном нормираном расподелом, по формулама:

$$x_1 = \sqrt{-2 \ln \gamma_1} \cos(2\pi \gamma_2), \quad x_2 = \sqrt{-2 \ln \gamma_1} \sin(2\pi \gamma_2)$$

Моделовање нормалне расподеле Box-Muller-ов методом

```
normalna_bm <- function (){  
  x <- c()  
  gama1<-runif(1)  
  gama2<-runif(1)  
  x1<-sqrt(-2*log(gama1))*cos(2*pi*gama2)  
  x2<-sqrt(-2*log(gama1))*sin(2*pi*gama2)  
  x[1] <- x1  
  x[2] <- x2  
  return(x)  
}
```

```
> normalna_bm()  
[1] 0.06352467 0.20559947  
> normalna_bm()  
[1] 0.2426644 -0.6883668  
> normalna_bm()  
[1] -0.6225873 0.205279
```


Моделовање нормалне расподеле коришћењем поларних координата

Нормална нормирана расподела се може моделирати и преласком на поларне координате.

- 1 Генеришемо два случајна броја u и v у интервалу $(-1, 1)$ као $u = 2\gamma_1 - 1$ и $v = 2\gamma_2 - 1$
- 2 Рачунамо $\omega = u^2 + v^2$ и ако је $\omega > 1$ вратимо се на корак 1.
- 3 Рачунамо $x = uz$ и $y = vz$, где је $z = \sqrt{-2\frac{\ln \omega}{\omega}}$. Дакле, x и y су два независна случајна броја из стандардне нормалне расподеле.

Овај метод је бржи од претходног јер смо избегли рачунање $\sin$ и $\cos$. Изрази $\frac{u}{\sqrt{\omega}}$ и $\frac{v}{\omega}$ имају улогу $\sin$ и $\cos$ из претходне методе.

Моделовање нормалне расподеле коришћењем једне експоненцијалне расподеле

Нека су $Y \sim U(0, 1)$ и $V \sim \epsilon(1)$ независне случајне величине. Ако је $Y \leq e^{-\frac{(V-1)^2}{2}}$, тада се за реализацију случајне величине X са стандардном нормалном расподелом узима вредност случајне величине V .

Помоћу γ_2 се моделира вредност $v = -\ln \gamma_2$ случајне величине V . Ако је $\gamma_1 \leq e^{-\frac{(v-1)^2}{2}}$, тада се сматра да је v реализована вредност случајне величине са стандардном нормалном расподелом. Ако неједнакост не важи, поступак се понавља са наредним паром (псеудо)случајних бројева.

Овај поступак за моделовање једне вредности стандардне нормалне расподеле захтева коришћење бар два (псеудо)случајна броја γ_1 и γ_2 .

Моделовање нормалне расподеле коришћењем једне експоненцијалне расподеле

```
normalna_1_eksp <- function(){  
  gama1<-runif(1)  
  gama2<-runif(1)  
  v=-log(gama2)  
  while (gama1 > exp(-((v-1)^2)/2))  
  {  
    gama1<-runif(1)  
    gama2<-runif(1)  
    v=-log(gama2)  
  }  
  return(v)  
}
```

```
> normalna_1_eksp()  
[1] 0.2989716  
> normalna_1_eksp()  
[1] 0.8782449  
> normalna_1_eksp()  
[1] 0.03335193
```

Моделовање нормалне расподеле коришћењем две експоненцијалне расподеле

Нека су случајне величине Y_1 и Y_2 независне са истом експоненцијалном расподелом $\epsilon(1)$. Ако је $Y_2 \geq \frac{(Y_1-1)^2}{2}$, тада се за реализацију случајне величине X са стандардном нормалном расподелом узима вредност случајне величине Y_1 .

Овај поступак за моделовање једне вредности стандардне нормалне расподеле захтева коришћење бар два (псеудо)случајна броја γ_1 и γ_2 .

Помоћу γ_1 и γ_2 се моделирају вредности $y_1 = -\ln \gamma_1$ и $y_1 = -\ln \gamma_1$ случајне величине Y_i . Ако је $y_2 \geq \frac{(y_1-1)^2}{2}$, тада се сматра да је y_1 реализована вредност случајне величине X са стандардном нормалном расподелом. Ако неједнакост не важи, поступак се понавља са наредним паром (псеудо)случајних бројева.

Моделовање нормалне расподеле коришћењем две експоненцијалне расподеле

```
normalna_2_eksp <- function(){  
  gama1<-runif(1)  
  gama2<-runif(1)  
  v1<--log(gama1)  
  v2<--log(gama2)  
  while(v2<((v1-1)^2)/2){  
    gama1<-runif(1)  
    gama2<-runif(1)  
    v1<--log(gama1)  
    v2<--log(gama2)  
  }  
  return(gama1)  
}
```

```
> normalna_2_eksp()  
[1] 0.136578  
> normalna_2_eksp()  
[1] 0.2632056  
> normalna_2_eksp()  
[1] 0.7774272
```

Моделовање *log*-нормалне расподеле

Случајна променљива X има $LN(m, \sigma^2)$ расподелу ако случајна променљива $Y = \ln X$ има нормалну расподелу $N(m, \sigma^2)$.

Моделовање вредности случајне променљиве X добијамо по формули $x = e^y$, где је y моделована вредност случајне променљиве Y .

Моделовање *log*-нормалне расподеле

```
log_normalna <- function(){  
y<-normalna_1_eksp() # z je modelovana vrednost normalne raspodele  
pomocu 1 eksponencijalne funkcije  
x<-exp(y)  
return(x)  
}
```

```
> log_normalna()  
[1] 1.783015  
> log_normalna()  
[1] 1.335986  
> log_normalna()  
[1] 3.409166
```

Функције у R-у

Distribution	Functions			
Beta	pbeta	qbeta	dbeta	rbeta
Binomial	dbinom	qbinom	dbinom	rbinom
Cauchy	pcauchy	qcauchy	dcauchy	rcauchy
Chi-Square	pchisq	qchisq	dchisq	rchisq
Exponential	pexp	qexp	dexp	rexp
F	pf	qf	df	rf
Gamma	pgamma	qgamma	dgamma	rgamma
Geometric	pgeom	qgeom	dgeom	rgeom
Hypergeometric	phyper	qhyper	dhyper	rhyper
Logistic	plogis	qlogis	dlogis	rlogis
Log Normal	plnorm	qlnorm	dlnorm	rlnorm
Negative Binomial	pnbinom	qnbinom	dnbinom	rnbinom
Normal	pnorm	qnorm	dnorm	rnorm
Poisson	ppois	qpois	dpois	rpois
Student t	pt	qt	dt	rt
Studentized Range	ptukey	qtukey	dtukey	rtukey
Uniform	punif	qunif	dunif	runif
Weibull	pweibull	qweibull	dweibull	rweibull
Wilcoxon Rank Sum Statistic	pwilcox	qwilcox	dwilcox	rwilcox
Wilcoxon Signed Rank Statistic	psignrank	qsignrank	dsignrank	rsignrank

Задаци

- 1.(3 поена) Моделовати вредност случајне величине X која има негативну биномну расподелу са параметрима $n = 20$ и $p = 0.2$. Генерисати 1000 вредности и израчунати очекивање и дисперзију и нацртати хистограм. Упоредити добијене резултате са резултатима добијеним генератором `rnbinom()`.
- 2.(2 поена) Моделовати вредност случајне величине X која има двопараметарску Паретову расподелу, $F(x) = 1 - \left(\frac{b}{x}\right)^a$, $x \geq b$.
- 3.(2 поена) Написати функцију која генерише вредности случајне величине са χ^2 расподелом.
- 4.(3 поена) Симулирајте $B(2, 4)$ расподелу користећи методу одбацавања.