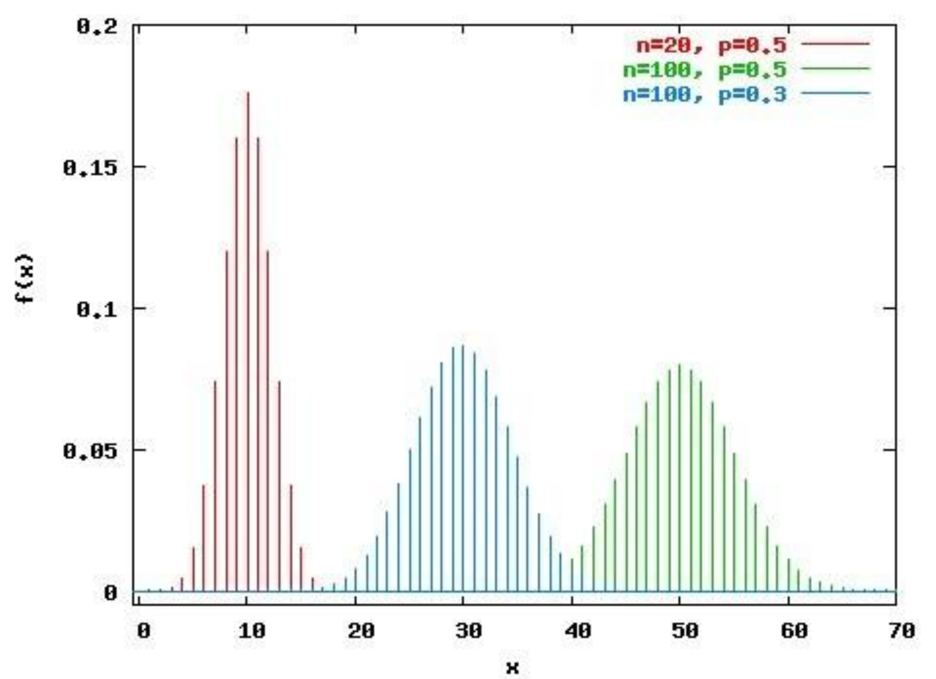
Study Program Statistics and Data Science
Departement of Statistics

IPB University — Bogor Indonesia —

STA1372 Metode Simulasi dan Resampling



Pembangkitan Peubah Acak Kontinu

IPB University
Bogor Indonesia

Metode Transformasi Langsung

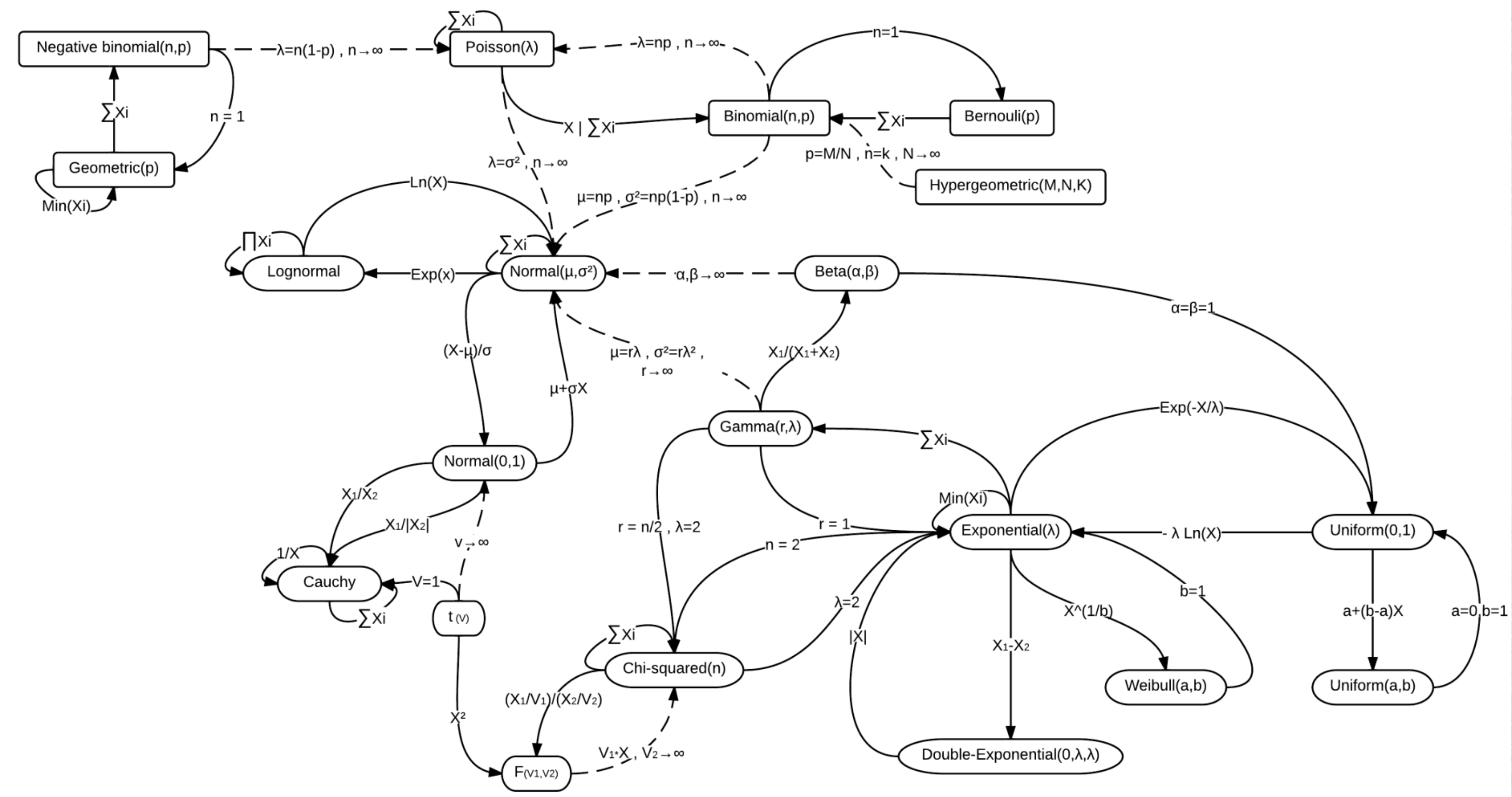
- Metode Transformasi Langsung (*Direct Transform Method*) , merupakan metode untuk menghasilkan peubah acak dengan sebaran target dari peubah acak lain yang sebarannya diketahui.
- Metode ini bergantung pada hubungan antar sebaran , misal Jika $Z \sim Normal(0,1)$ maka $V = Z^2 \sim \chi^2_{(1)}$
- Berarti jika kita ingin membangkitkan data $\chi^2_{(1)}$, kita dapat membangkitkan terlebih dahulu $Normal(0,1)$ kemudian kita kudratkan distribusi tersebut
- Secara umum Langkah-langkahnya adalah:
    1. Selidiki hubungan sebaran target dengan sebaran lain yang sudah tersedia pembangkitannya
    2. Bangkitkan data berdasarkan sebaran yang sudah tersedia tersebut
    3. Transformasi hasil pada Langkah 2 dengan menggunakan fungsi tertentu sedemikian sehingga menjadi sebaran target
    4. Lakukan Langkah 2 sampai 3 sampai banyaknya amatan yang diinginkan terpenuhi

Metode Transformasi Langsung

Beberapa contoh hubungan antar sebaran:

- Jika $Z \sim Normal(0,1)$ maka $V = Z^2 \sim \chi_1^2$
- Jika $U \sim \chi_m^2$ dan $V \sim \chi_n^2$ dimana U dan V saling bebas maka $F = \frac{U/m}{V/n} \sim F(m,n)$
- Jika $Z \sim Normal(0,1)$ dan $V \sim \chi_n^2$ Dimana Z dan V saling bebas maka $T = \frac{Z}{V/n} \sim t(n)$
- Jika $X_1, X_2, \dots, X_n \sim Eksponensial(\lambda)$ dan $X_1, X_2, \dots, X_n$ saling bebas maka $Y = \sum_{i=1}^{n} X_i \sim Gamma(n, \lambda)$
- Jika $U \sim Gamma(n, 1), \lambda > 0$ merupakan suatu konstanta maka $V = \lambda U \sim Gamma(n, \lambda)$
- Jika $X_1, X_2, \dots, X_n \sim Normal(0,1)$ dan $X_1, X_2, \dots, X_n$ saling bebas maka $Y = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2$

Metode Transformasi Langsung



Ilustrasi 1

Seorang Mahasiswa ingin melakukan pembangkitan data yang menyebar $Gamma(5,4)$. Mahasiswa tersebut diharuskan menggunakan software Julia untuk pembangkitan data tersebut. Sayangnya, Software Julia tersebut hanyalah untuk membangkitkan $Gamma(5,1)$. Bantulah mahasiswa tersebut untuk membangkitkan data tersebut sebanyak 100 amatan.

Ilustrasi 1

Hal pertama yang harus dilakukan adalah mengidentifikasi hubungan antara $Gamma(5,1)$ dan $Gamma(5,4)$. Identifikasi dilakukan menggunakan transformasi peubah acak. Diketahui pdf dari $Gamma(\alpha,\beta)$ adalah:

$$f(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} \quad ,x > 0$$

Misalkan $X \sim Gamma(5,1)$ maka pdf nya:

$$f(x;5,1) = \frac{1}{\Gamma(5)1^{5}} x^{5-1} e^{-x/1} \quad ,x > 0$$

Misal $Y = 4X$, menggunakan metode transformasi Jacobian diperoleh:

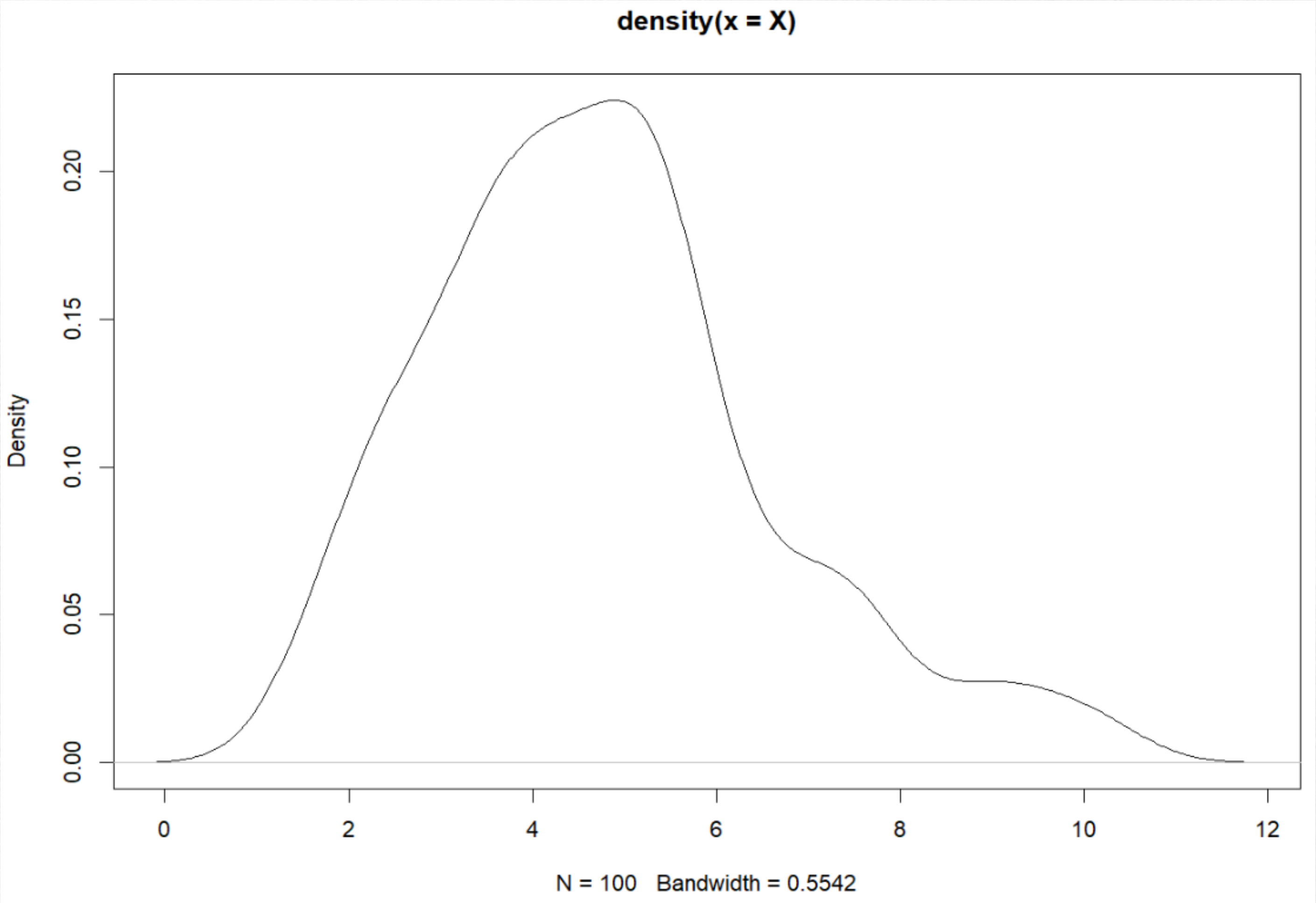
$$f(y) = \frac{1}{\Gamma(5)4^{5}} y^{5-1} e^{-y/4} \quad ,y > 0$$

Sehingga dapat disimpulkan hubungan antara $Gamma(5,1)$ dan $Gamma(5,4)$ adalah $Y = 4X$. Maka untuk membangkitkan data $Gamma(5,4)$ bisa menggunakan metode transformasi langsung, yaitu dengan membangkitkan data $Gamma(5,1)$ lalu dikalikan dengan konstanta 4 maka diperoleh data $Gamma(5,4)$

Ilustrasi 1

Syntax R:

```
#Membangkitkan Data Gamma(5,1)
> set.seed(123)
> X=rgamma(100,5,scale = 1)
> head(X)
[1] 3.389585 7.378957 1.629561
[4] 4.778440 8.873564 5.530862
> plot(density(X))
```

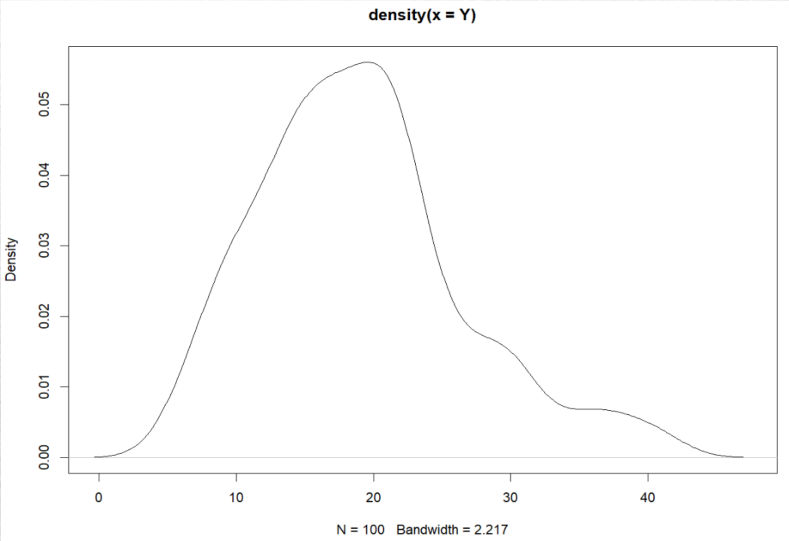


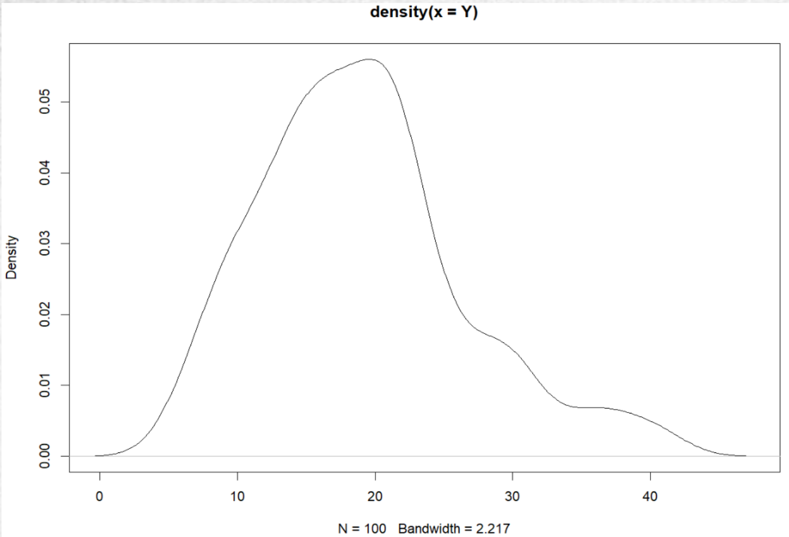
Ilustrasi 1

Syntax R:

```
#Transformasi Y=4X
> Y=4*X
> head(Y)
[1] 13.558339 29.515830 6.518243
[4] 19.113758 35.494257 22.123447
> plot(density(Y))

#Membandingkan hasil data yang dibangkitkan
#dengan fungsi build-in di R
> set.seed(123)
> Z=rgamma(100,5,scale=4)
> head(Z)
[1] 13.558339 29.515830 6.518243
[4] 19.113758 35.494257 22.123447
> plot(density(Z))
```





Metode Inverse Transform

- Metode inverse transform adalah metode yang digunakan untuk menghasilkan peubah acak dari sebaran uniform ke sebaran target.
- Metode ini bekerja dengan mentransformasikan peubah acak dari sebaran uniform ke sebaran target dengan menggunakan *inverse Cumulative Distribution Function* (CDF) atau sering disebut fungsi quantile.

CDF:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt = p$$

Fungsi Quantile:

$$Q(p) = F_X^{-1}(p)$$

Metode Inverse Transform

Langkah-Langkah metode inverse transform adalah sebagai berikut:

1. Tentukan bentuk CDF dari $X, F_X(x)$
2. Cari fungsi quantile $Q(p) = F_X^{-1}(p)$
3. Bangkitkan bilangan acak $Uniform(0,1)$, misalk dinotasikan dengan U
4. Hitung $X = Q(U)$
5. Nilai X akan memiliki sebaran target
6. Ulangi langkah 3-5 sebanyak amatan yang diinginkan

Ilustrasi 2

Bangkitkan suatu gugus data yang mengikuti sebaran $Eksponensial(3)$ sebanyak 1000 amatan.

Ilustrasi 2

Misal $X \sim Eksponensial\ (\lambda)$, maka:

$$f(x) = \lambda e^{-\lambda x}, x \geq 0$$

$$F(x) = 1 - e^{-\lambda x}, x \geq 0$$

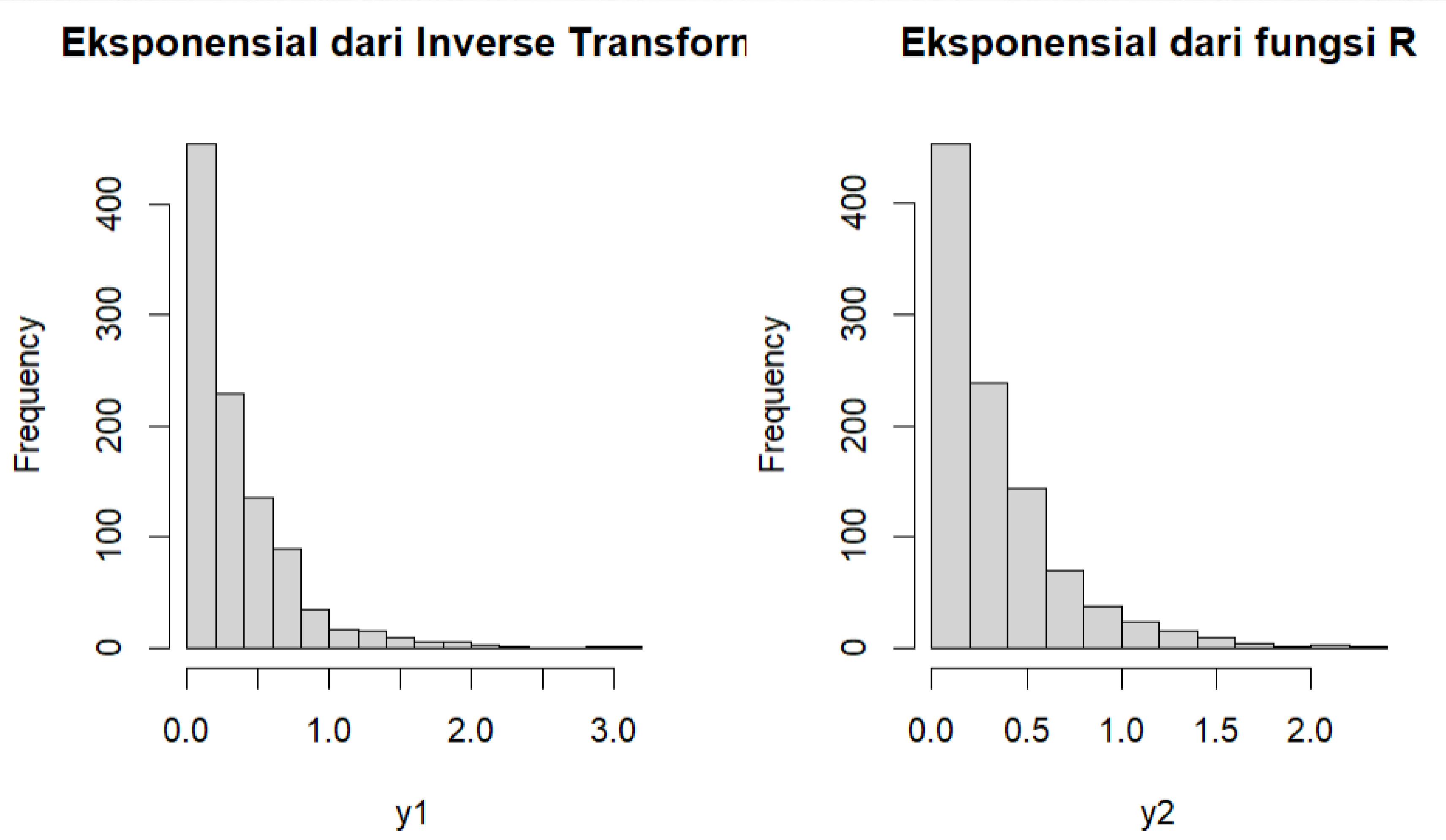
$$U = F(x) = 1 - e^{-\lambda x}, x \geq 0$$

$$X = -\frac{\ln(1-U)}{\lambda}$$

Syntax R:

```
set.seed(10)
eksponensial<-function(n, lambda){
   U<-runif(n)
   x<- -log(1-U)/lambda
   return(x)}
y1<-eksponensial(1000,3)
y2<-rexp(1000, rate = 3)
par(mfrow=c(1,2))
hist(y1, main = 'Eksponensial dari Inverse
Transform')
hist(y2, main = 'Eksponensial dari fungsi R')
```

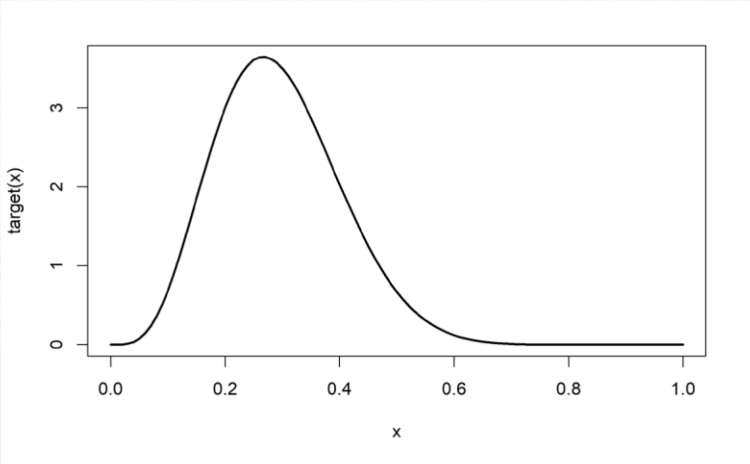
Ilustrasi 2

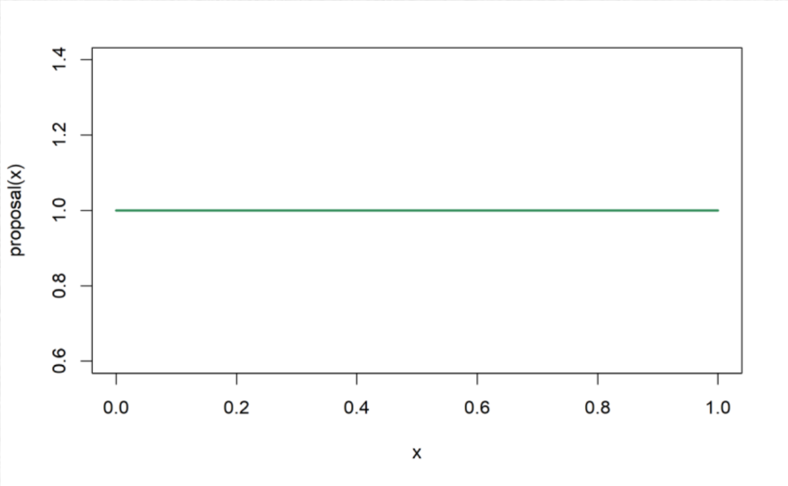


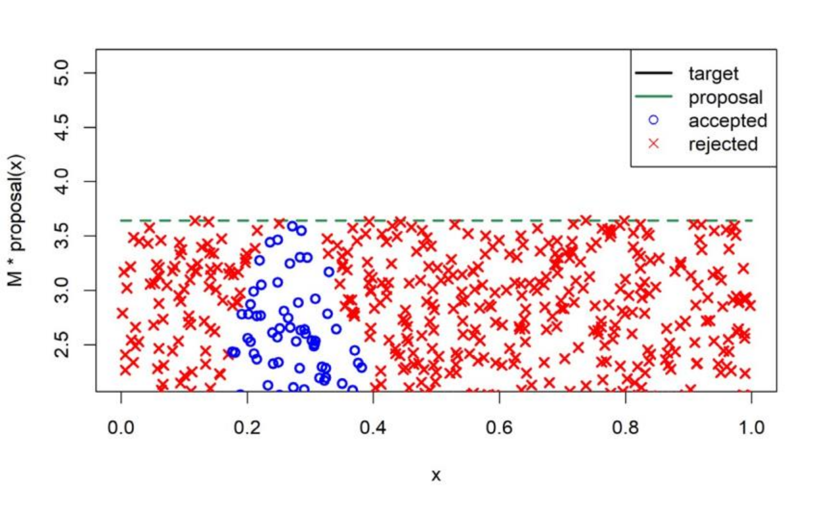
Eksponensial dari Inverse Transforn
Frequency
0
100
200
300
400
0.0
1.0
2.0
3.0
y1
Eksponensial dari fungsi R
Frequency
0
100
200
300
400
0.0
0.5
1.0
1.5
2.0
y2

Metode Acceptance-Rejection

Metode Acceptance-Rejection merupakan metode untuk membangkitkan peubah acak dari sebaran sederhana yang mudah dibangkitkan atau disebut **sebaran proposal**, dan kemudian menerima atau menolak nilai yang dihasilkan berdasarkan kemiripannya dengan sebaran target



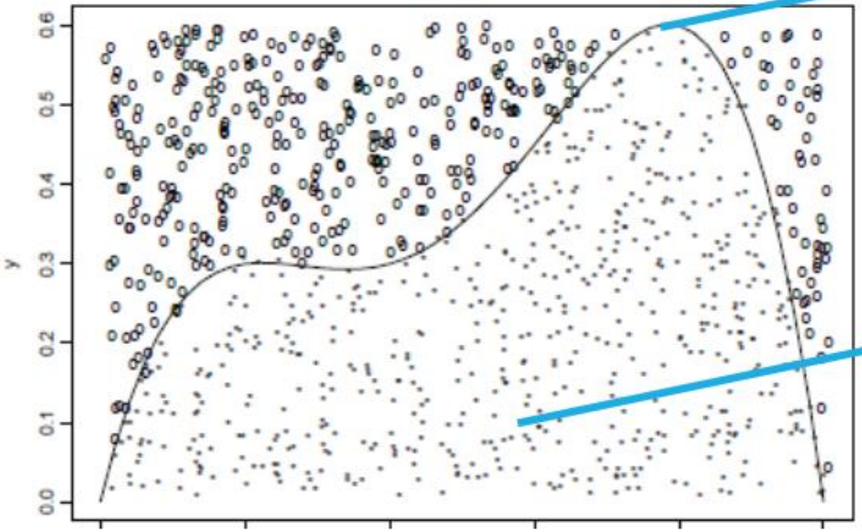




Metode Acceptance-Rejection

Ide dasar :

$y = f(x)$



kriteria x

Kita perlu membangkitkan (x, y) pada daerah tersebut

Terima x jika (x, y) di dalam daerah kurva

Tolak x jika (x, y) di luar daerah kurva

Metode Acceptance-Rejection

Langkah-Langkah Metode Acceptance-Rejection adalah sebagai berikut:

1. Pilih distribusi proposal: Pilih sebaran proposal, misal $g(x)$ yang memiliki domain sama dengan sebaran target, misal $f(x)$.
2. Tentukan konstanta c yang memenuhi $\frac{f(x)}{g(x)} \leq c$ untuk semua c atau $c = \max\left(\frac{f(x)}{g(x)}\right)$
3. Bangkitkan peubah acak X dari sebaran proposal $g(x)$
4. Bangkitkan peubah acak U dari sebaran $Uniform(0,1)$
5. Jika $U \leq \frac{f(x)}{cg(x)}$ maka terima X sebagai amatan dari peubah acak sebaran target Y, selain itu lanjutkan kembali ke langkah 3.
6. Ulangi langkah 3-5 sebanyak amatan yang diinginkan

Ilustrasi 3

Bangkitkan peubah acak $X \sim Beta(2,2)$ yang memiliki pdf:

$$f(x; 2,2) = 6x(1-x)\,, 0 < x < 1$$

dengan bantuan R sebanyak 1000 amatan menggunakan

a. Inverse transform method
b. Acceptance-Rejection Method

Ilustrasi 3

a. Metode Inverse – Transform

$X \sim Beta(2,2)$ dengan pdf:

$$f(x;2,2) = 6x(1-x)\ ,0 < x < 1$$

Maka CDF dari X adalah:

$$F(x;2,2) = \int_0^x f(t;2,2)\,dt = \frac{B(x;2,2)}{B(2,2)} = I(2,2)\ \ ,0 < x < 1$$

dengan $B(x;2,2)$ merupakan fungsi ***incomplete beta***, $I(2,2)$ fungsi ***regularized incomplete beta***, dan $B(2,2)$ adalah fungsi ***beta***

Ilustrasi 3

a. Metode Inverse – Transform

Selanjutnya fungsi quantile dari $Beta(2,2)$ adalah:

$$Q(p;\alpha,\beta) = I^{-1}(\alpha,\beta)$$

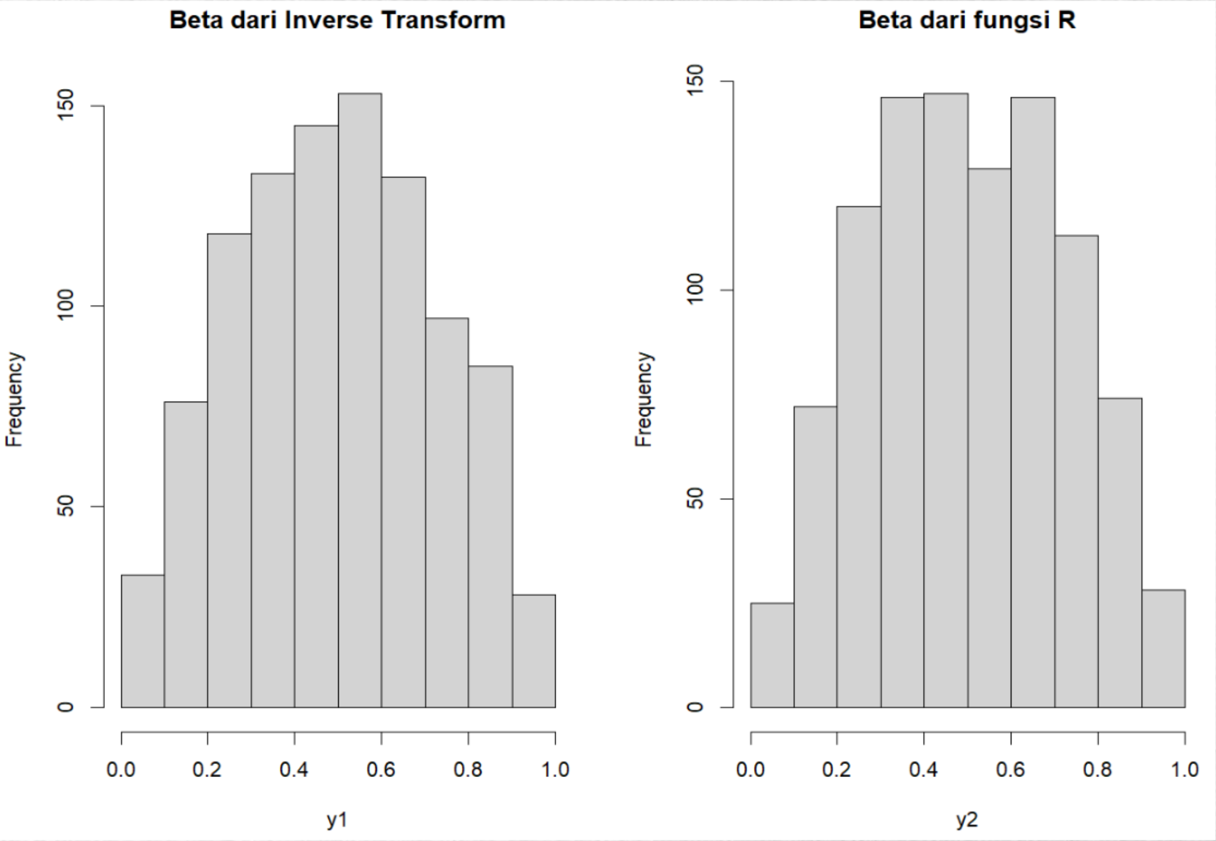
Fungsi quantile dapat dicari menggunakan fungsi di dalam R yaitu `qbeta()`.

Ilustrasi 3

a. Metode Inverse – Transform

Syntax R:

```
set.seed(55)
beta_22<-function(n=1000, alpha=2,beta=2){
        U<-runif(n)
        x<- qbeta(U,alpha,beta)
        return(x)}
y1<-beta_22(1000,2,2)
y2<-rbeta(1000, 2,2)
par(mfrow=c(1,2))
hist(y1, main = Beta dari Inverse Transform')
hist(y2, main = 'Beta dari fungsi R')
```



Ilustrasi 3

b. Metode Acceptance – Rejection

1. Tentukan sebaran proposal, misal $g(x)$, yang memiliki domain sama dengan sebaran target, misal $f(x)$. Dalam kasus ini dipilih $g(x) \sim Uniform(0,1)$ dengan $g(x) = 1 \ , 0 < x < 1$

2. Tentukan konstanta c yang memenuhi $\frac{f(x)}{g(x)} \leq c$ untuk semua c atau $c = \max\left(\frac{f(x)}{g(x)}\right)$.

   Misalkan $h(x) = \frac{f(x)}{g(x)}$.

$$h(x) = \frac{6x(1-x)}{1} = 6x(1-x) \ , 0 < x < 1$$

Ilustrasi 3

b. Metode Acceptance – Rejection

Nilai maksimum dari $h(x)$ dapat dicari dengan mencari turunan pertama.

$$\frac{d}{dx}h(x) = 6 - 12x \quad \rightarrow 0 = 6 - 12x \quad \rightarrow x = \frac{1}{2}$$

Cek turunan kedua: $\frac{d^2}{dx^2}h(x) = -12$. Sehingga $x = \frac{1}{2}$ memaksimumkan fungsi $h(x)$.

Substitusi $x = 1/2$ ke fungsi $h(x)$ → $h\left(\frac{1}{2}\right) = 6\left(\frac{1}{2}\right)\left(1 - \frac{1}{2}\right) = \frac{6}{4} = \frac{3}{2}$. Sehingga dipilih $c = \frac{3}{2}$

Nilai maksimum dari fungsi dapat juga dicari menggunakan R.

```
h <- function(x) 6*x*(1-x)

derivH <- optimize(f=h,interval=c(0,1),maximum=T)

derivH$objective
```

Ilustrasi 3

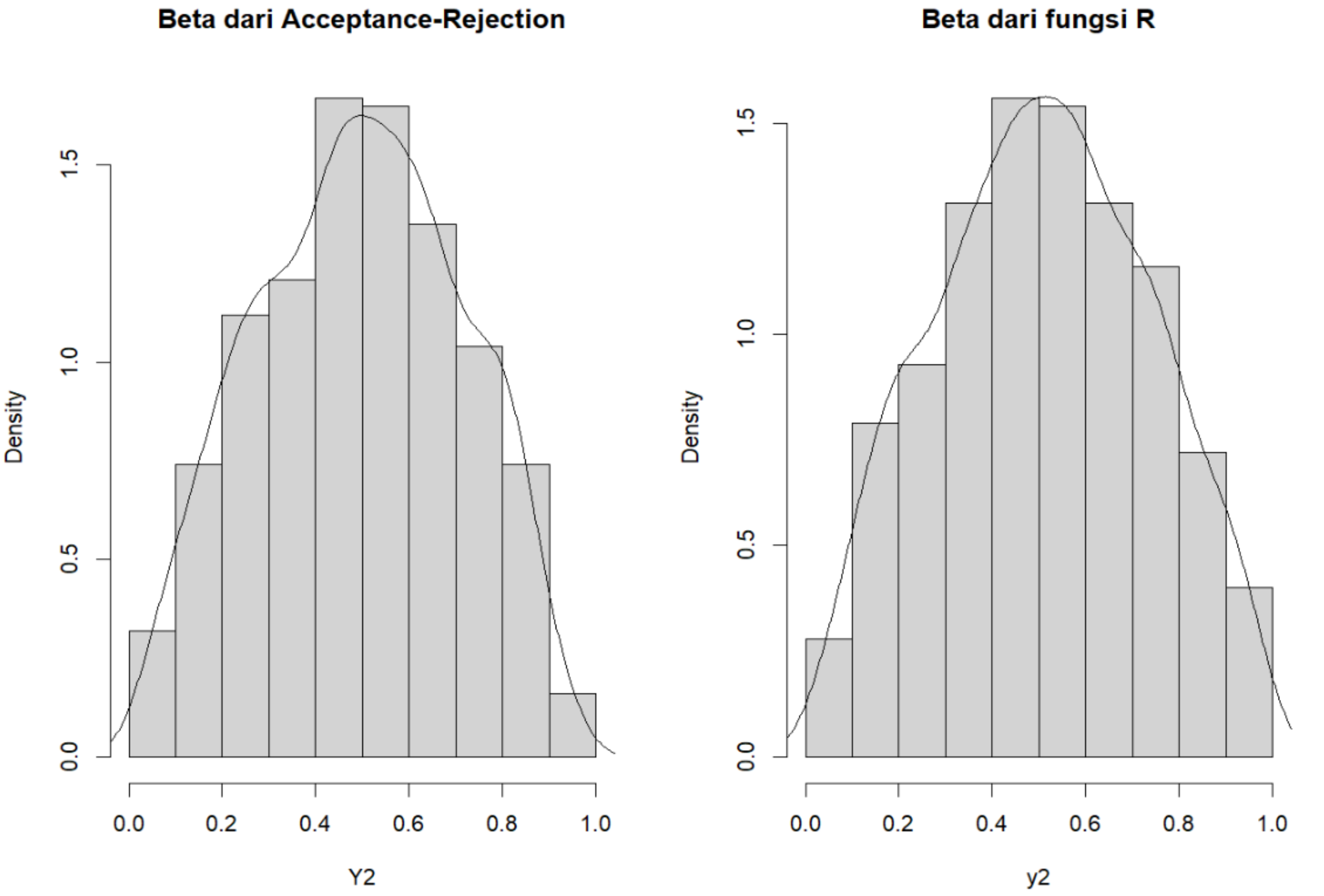
b. Metode Acceptance – Rejection

Syntax R:

```
set.seed(111)
X=runif(2000)
set.seed(333)
U=runif(2000)
f=function(x) 6*x*(1-x)
g=function(x) 1
h=function(x) 6*x*(1-x)
derivH=optimize(f=h,interval=c(0,1),
maximum=T)
nilaic=derivH$objective
```

```
#kriteria penerimaan
criteria=U < f(X)/(nilaic*g(X))
##banyaknya amatan yang sesuai kriteria
Y2=X[criteria][1:1000]
#plot
set.seed(123)
y2<-rbeta(1000, 2,2)
par(mfrow=c(1,2))
hist(Y2, prob = T, main='Beta dari Acceptance-
Rejection')
lines(density(Y2))
hist(y2, prob=T, main='Beta dari fungsi R')
lines(density(y2))
```

Ilustrasi 3


Beta dari Acceptance-Rejection
Density
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
Y2
Beta dari fungsi R
Density
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
y2