Δειγματοληψία από τυχαίες μεταβλητές

Δειγματοληψία με την μέθοδο του αντίστροφου μετασχηματισμού (Inverse Transform Method)

Η συνεχής περίπτωση

Εάν $X \sim f_X$ και η αθροιστική συνάρτηση κατανομής F_X αντιστρέφεται αναλυτικά (γνωρίζουμε ότι η F_X είναι μονότονη και άρα έχει μοναδικό αντίστροφο, αλλά θέλουμε την F_X^{-1} σε κλειστή μορφή), τότε

$$U \sim \mathcal{U}(0,1) \Rightarrow F_X^{-1}(U) \sim f_X .$$

Το παραπάνω ισχύει διότι εάν θέσουμε $\tilde{X} = F^{-1}(U)$ παίρνουμε

$$F_{\tilde{X}}(x) = P\left\{\tilde{X} \le x\right\} = P\left\{F_X^{-1}(U) \le x\right\} = P\left\{U \le F_X(x)\right\}$$
$$= \int_{u=-\infty}^{F_X(x)} \mathcal{U}(x \mid 0,1)\, du = \int_{u=0}^{F_X(x)} du = F_X(x) .$$

εφόσον $0 \le F_X(x) \le 1$. Δηλαδή οι τ.μ. X και $\tilde{X}$ είναι ισόνομες, συμβολικά

$$F_{\tilde{X}}(x) = F_X(x) \Leftrightarrow \tilde{X} \overset{\mathcal{D}}{=} X ,$$

έτσι μπορούμε να μετασχηματίσουμε μια $\mathcal{U}(0,1)$ i.i.d. ακολουθία $\{U_k : 1 \le k \le N\}$ στην αντίστοιχη $X \sim f_X$ i.i.d. ακολουθία $\left\{X_k = F_X^{-1}(U_k) : 1 \le k \le N\right\}$.

Πρόταση

Ισχύει και το αντίστροφο. Δηλαδή όταν $U \sim \mathcal{U}(0,1)$ έχουμε και ότι

$$F_X^{-1}(U) \overset{\mathcal{D}}{=} X \Rightarrow U \overset{\mathcal{D}}{=} F_X(X).$$

Έστω ότι $T : y = F_X(x) \Leftrightarrow T^{-1} : x = F_X^{-1}(y)$ τότε

$$f_Y(y) = f_X\left(F_X^{-1}(y)\right)\left|\left(F_X^{-1}\right)'(y)\right|,\ 0 \le y \le 1.$$

Επειδή $F_X : \mathbb{R} \to [0,1]$ είναι γνησίως αύξουσα στη συνεχή περίπτωση, έχουμε:

$$1 = \left(F_X \circ F_X^{-1}\right)'(y) = F_X'\left(F_X^{-1}(y)\right)\cdot\left(F_X^{-1}\right)'(y) = f_X\left(F_X^{-1}(y)\right)\cdot\left(F_X^{-1}\right)'(y),$$

από όπου και $\left(F_X^{-1}\right)'(y) = \dfrac{1}{f_X\left(F_X^{-1}(y)\right)} > 0$, που δίνει

$f_Y(y) = 1,\ 0 \le y \le 1$ ή ότι $f_Y(y) = \mathcal{U}(y \mid 0,1)$.

<u>Δειγματοληψία από την εκθετική κατανομή</u>

$$f_X(x) = \mathrm{Exp}(x \mid \lambda) = \lambda e^{-\lambda x}\mathcal{I}(x > 0),\ \lambda > 0 \Leftrightarrow F_X(x) = \left(1 - e^{-\lambda x}\right)\mathcal{I}(x > 0)$$

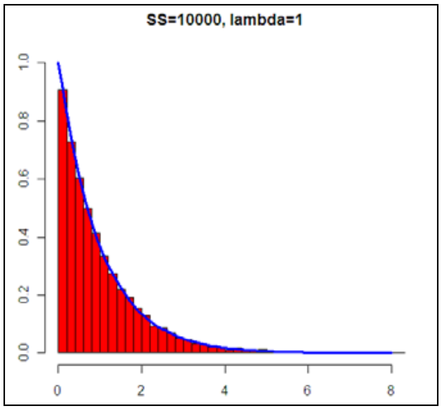
$$\Rightarrow F_X^{-1}(x) = -\frac{1}{\lambda}\log(1 - x),\ 0 < x < 1.$$

Επειδή $U_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1) \Leftrightarrow 1 - U_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1)$ θα έχουμε ότι

$$X_i = -\frac{1}{\lambda}\log(1 - U_i) \overset{\mathcal{D}}{=} -\frac{1}{\lambda}\log(U_i) \overset{i.i.d.}{\sim} \mathrm{Exp}(\lambda).$$

<u>Παράδειγμα</u>

```
# Sample a n-realization from the exponential distribution.

SampleExp <- function(n,lambda,realization=1){

    set.seed(realization); return((-1/lambda)*log(runif(n)))

}

# here n=10000, and lambda=1.

v<- SampleExp(n=10^4, lambda=1)

# plot the  probability histogram of v.

hist(v, breaks=50, freq=FALSE, ylim=c(0, 1), xlim=c(0, 8),

     main="n=10000, lambda=1", col="red")

curve(dexp(x, rate=1),col="blue",lwd=3,add=TRUE)
```



<u>Δειγματοληψία από την Weibull</u>

$$f_X(x) = \mathcal{W}ei(x \mid \lambda, \rho) = \begin{cases} \lambda\rho(\lambda x)^{\rho-1} e^{-(\lambda x)^{\rho}} & x > 0 \\ 0 & \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

1. Τότε α.σ.κ. $y = F_X(x) = 1 - e^{-(\lambda x)^{\rho}}$ για $x > 0$,

που αντιστρέφεται σε κλειστή μορφή

$x = F_X^{-1}(y) = \lambda^{-1}\{-\log(1-y)\}^{1/\rho}$, $0 < y < 1$

$U_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1) \Rightarrow X_i = \lambda^{-1}\{-\log(U_i)\}^{1/\rho} \overset{i.i.d.}{\sim} \mathcal{W}ei(\lambda,\rho)$.

2. $X \sim \mathrm{Exp}(1) \Leftrightarrow Y = \lambda^{-1} X^{1/\rho} \sim \mathcal{W}ei(\lambda,\rho)$.

$$f_Y(y) = \mathrm{Exp}\left((\lambda y)^{\rho} \mid 1\right)\left|\frac{d}{dy}(\lambda y)^{\rho}\right| = \lambda\rho(\lambda x)^{\rho-1} e^{-(\lambda x)^{\rho}} = \mathcal{W}ei(y \mid \lambda,\rho).$$

Δειγματοληψία από την Pareto

$$f_X(x) = \mathcal{P}a(x \mid \lambda, c) = \frac{\lambda c^{\lambda}}{x^{\lambda+1}} \mathcal{I}(x > c)$$

1. $y = F_X(x) = \left\{1 - \left(\frac{c}{x}\right)^{\lambda}\right\} \mathcal{I}(x > c) \Rightarrow x = F_X^{-1}(y) = c(1-y)^{-1/\lambda},\ 0 < y < 1 \Rightarrow$

$U_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1) \Rightarrow X_i = \lambda^{-1} U_i^{-1/\lambda} \overset{i.i.d.}{\sim} \mathcal{P}a(\lambda, c)$.

2. $X \sim \mathrm{Exp}(1) \Leftrightarrow Y = ce^{X/\lambda} \sim \mathcal{P}a(\lambda, c)$

$$f_Y(y) = \mathrm{Exp}\left(\lambda \log(y/c) \mid 1\right) \left| \frac{d}{dy} \left(\lambda \log(y/c)\right) \right| = \frac{\lambda}{y} e^{-\lambda \log(y/c)} \mathcal{I}\left(\lambda \log(y/c) > 0\right)$$

$$= \lambda c^{\lambda} y^{-(\lambda+1)} \mathcal{I}(y > c)$$

Δειγματοληψία από την Logistic

$$f_X(x) = \mathcal{L}o(x \mid \mu, s) = \frac{e^{-\frac{x-\mu}{s}}}{s\left\{1 + e^{-\frac{x-\mu}{s}}\right\}^2},\ x \in \mathbb{R}.$$

1. $F_X(x) = \left\{1 + e^{-\frac{x-\mu}{s}}\right\}^{-1} \Rightarrow F_X^{-1}(x) = \mu - s\log\left(x^{-1} - 1\right),\ 0 < x < 1$

$U_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1) \Rightarrow \mu - s\log\left(u^{-1} - 1\right) \overset{i.i.d.}{\sim} \mathcal{L}o(\mu, s)$.

2. Εάν $X \sim \mathrm{Exp}(1)$ και $Y \sim \mathrm{Exp}(1)$ ανεξάρτητες, τότε $Z = \mu - s\log\left(\frac{X}{Y}\right) \sim \mathcal{L}(\mu, s)$

Επειδή $u \sim \mathcal{U}(0,1) \Rightarrow \mu - s\log\left(u^{-1}-1\right) \sim \mathcal{L}(\mu,s)$, αρκεί να δείξουμε ότι $\frac{X}{Y} \overset{\mathcal{D}}{=} u^{-1}-1$.

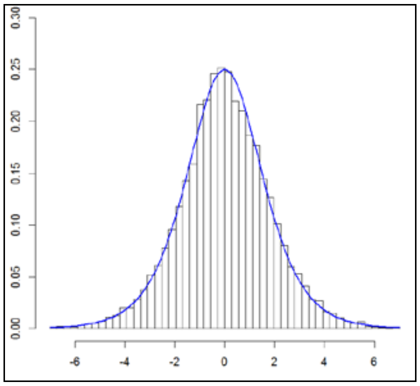
Θέτουμε $T:\left\{z=\frac{x}{y},\, u=x\right\} \Leftrightarrow T^{-1}:\left\{x=u,\, y=\frac{u}{z}\right\} \Rightarrow Jac\left(T^{-1}\right)=\frac{u}{z^2} \Rightarrow$

$$f_{Z,U}(z,u)=f_{X,Y}(x,y)\frac{u}{z^2}=\exp\left\{-u\left(1+\frac{1}{z}\right)\right\}\frac{u}{z^2}$$

$$\Rightarrow f_z(z)=\int_{u=0}^{\infty}\exp\left\{-u\left(1+\frac{1}{z}\right)\right\}\frac{u}{z^2}du=(z+1)^{-2},\ z>0$$

Ενώ εάν $u \sim \mathcal{U}(0,1)$ και $z=\frac{1}{u}-1$ τότε $f_z(z)=(z+1)^{-2}$ για $z>0$.

```
samplelogistic <- function(n, realization=1){
      set.seed(realization); return(-log(1/runif(n)-1)))
}
   v <- samplelogistic(n=10000);
   b<-7; inf <- -b; sup <- b
   mybreaks <- seq(from=inf,to=sup, by=(sup-inf)/50)
   v <- v[abs(v) < b]
   hist(v, breaks=mybreaks, freq=FALSE, ylim=c(0, 0.3), xlim=c(inf, sup))
   curve(dlogis(x, location = 0, scale = 1), col="blue", lwd=2, add=TRUE)
```



Δειγματοληψία με την μέθοδο του αντίστροφου μετασχηματισμού

Η διακριτή περίπτωση

Εάν $X \sim \Pi(\,\cdot\,) = \sum_{j=1}^{\infty} \pi_j \, \delta_j(\,\cdot\,)$ όπου $P(X = j) = \pi_j$, και $\sum_{j=1}^{\infty} \pi_j = 1$ και F_X η αντίστοιχη αθροιστική συνάρτηση κατανομής, ορίζουμε σαν την γενικευμένη αντίστροφη (generalized inverse) F_X^- την συνάρτηση

$$F_X^-(u) := \inf\left\{x \in X(\Omega) : u \le F_X(x)\right\}.$$

Παρατηρούμε ότι εάν η F_X αντιστρέφεται (συνεχής περίπτωση), τότε $F_X^{-1} = F_X^-$, πράγματι

$$F_X^-(u) = \inf\left\{x \in X(\Omega) : u \le F_X(x)\right\} = \inf\left\{x \in X(\Omega) : x \ge F_X^{-1}(u)\right\} = F_X^{-1}(u).$$

Ισχύει ότι

$$\boxed{u \sim \mathcal{U}(0,1) \;\Rightarrow\; F_X^-(u) = \inf\left\{x \in X(\Omega) : u \le F_X(x)\right\} \sim f_X(\,\cdot\,)}$$

Το παραπάνω ισχύει διότι θέτοντας $\tilde{X} := F^-(U)$ θα έχουμε:

$$P\left\{\tilde{X} = j\right\} = P\left\{F_X^-(U) = j\right\} = P\left\{F_X(j-1) < U \le F_X(j)\right\}$$
$$= \int_{F_X(j-1)}^{F_X(j)} \mathcal{U}(u \mid 0,1)\, du = F_X(j) - F_X(j-1) = \pi_j \; .$$

<u>Παράδειγμα</u>

Έστω διακριτή τυχαία μεταβλητή $X : \Omega \to \mathbb{R}$ με χώρο καταστάσεων

$$S_X = \left\{i \in \mathbb{Z} : p_X(i) > 0\right\} = \{1, \ldots, m\},$$

που ακολουθεί την διακριτή κατανομή $p_X(i) = P\{X = j\} = \pi_j$ για $1 \le j \le m$ με $\sum_{j=1}^{m} \pi_j = 1$,

συμβολικά $X \sim \mathcal{DU}\{1, \ldots, m\}$ είτε χρησιμοποιώντας το μέτρο Dirac $X \sim \sum_{k=1}^{m} \pi_k \delta_k(\,\cdot\,)$.

Θέλουμε να παράγουμε τυχαίο δείγμα από την X. Το σχήμα δειγματοληψίας θα είναι:

$$\boxed{\begin{array}{l} sample\ u \sim \mathcal{U}(0,1) \\ if\ \ u \le \underbrace{\pi_1}_{F_1} \ \Rightarrow\ x = 1 \\ \qquad else\ if\ \ u \le \underbrace{\pi_1 + \pi_2}_{F_2} \ \Rightarrow\ x = 2 \\ \qquad \quad \vdots \\ \qquad else\ if\ \ u \le \underbrace{\pi_1 + \pi_2 + \cdots + \pi_i}_{F_i} \ \Rightarrow\ x = i \\ \qquad \quad \vdots \\ \qquad else\ if\ \ u \le \underbrace{\pi_1 + \pi_2 + \cdots + \pi_{m-1}}_{F_{m-1}} \ \Rightarrow\ x = m-1 \\ \qquad else\ \Rightarrow\ x = m \\ end\ if \end{array}}$$

ή ότι

$$\begin{aligned} X &\overset{\mathcal{D}}{=} F_X^{-}(U) \\ &\overset{\mathcal{D}}{=} \mathcal{I}(F_0 \le U \le F_1) + 2\mathcal{I}(F_1 \le U \le F_2) + \cdots + m\mathcal{I}(F_{m-1} \le U \le F_m), \quad F_0 = 0 \end{aligned}$$

Πράγματι, σύμφωνα με το παραπάνω σχήμα δειγματοληψίας έχουμε

$$P\{X = i\} = P\{F_{i-1} < U \le F_i\} = \int_{u=F_{i-1}}^{F_i} \mathcal{U}(u \mid 0,1)\, d$$

$$= \int_{u=F_{i-1}}^{F_i} du = F_i - F_{i-1} = \pi_i, \quad i \in \{1, \ldots, m\}.$$

Έτσι μπορούμε να μετασχηματίσουμε μια $\mathcal{U}(0,1)$ i.i.d. ακολουθία $\{U_k : 1 \le k \le N\}$ στην i.i.d. ακολουθία $\{X_k : 1 \le k \le N\}$ από την τ.μ. X .

Επίσης έχουμε ότι

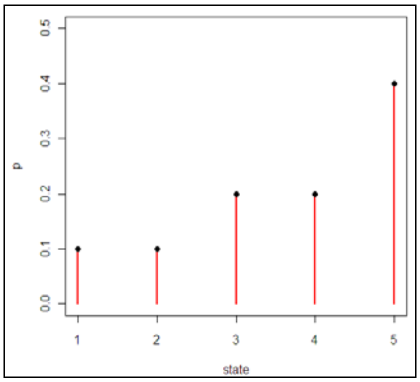
$$\hat{\pi}_{i,N} = \frac{1}{N}\sum_{k=1}^{N} \mathcal{I}(X_k = i) \xrightarrow[N \to \infty]{w.p.1} P\{X = i\} = \pi_i,$$

και επειδή $\mathcal{I}(X_k = i) \sim \mathrm{Bin}(1, \pi_i)$, παίρνουμε:

$$\sqrt{N}\left(\hat{\pi}_{i,N} - \pi_i\right) \overset{\mathcal{D}}{\approx} \mathcal{N}\left(0, \pi_i\left(1 - \pi_i\right)\right) \Leftrightarrow \hat{\pi}_{i,N} \overset{\mathcal{D}}{\approx} \mathcal{N}\left(\pi_i, \frac{1}{N}\pi_i\left(1 - \pi_i\right)\right)$$

Plot the specific discrete distribution (state, probs).

```
state <- c(1,2,3,4,5); probs <- c(0.1,0.1,0.2,0.2,0.4)
plot(state, probs, type="h", lwd=2, col="red", ylim=c(0,0.5))
points(state, probs, pch=16, cex=1, col="black")
```



#Sample from the General Discrete random Variable with finite state space.

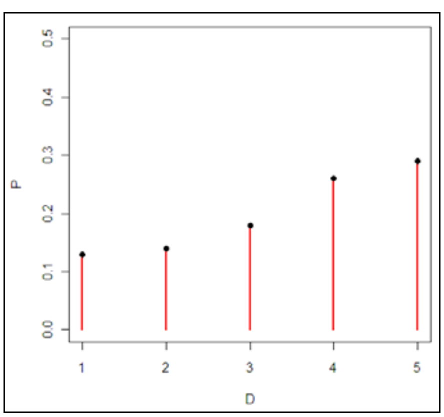
```
SampleVarX <- function(N=100, probs, state){
   sample <- c()
   for(i in 1:N) {
        F <- 0
        u <- runif(1)
        for(j in 1:length(probs)){
            F <- F + probs[j]
            if(u<F){
              sample <- c(sample, state[j])
              break
            }
        }
   }
   return(sample)
}
```

#Get the Relative Frequencies of the sample-vector.

```
GetRF <- function(sample){
   states <- c(); probs <- c(); L <- sample; N <- length(sample)
   while(length(L) != 0){ states <- c(states, L[1]); L <- L[L!=L[1]] }
   for(i in 1:length(states)){
       x <- states [i]; counter <- 0
       for(j in 1:N) if(x == sample[j]) counter <- counter+1
       probs <- c(probs, counter/N)
   }
    v<-c(states, probs)
    return(v)
}
```

ΕΚΤΕΛΕΣΗ

```
> set.seed(1)
> mysample <- SampleVarX(N=100, p=c(0.1, 0.1, 0.2, 0.2, 0.4), state=c(1,2,3,4,5))
> mysample
  [1] 3 5 5 1 5 3 2 3 4 5 2 5 4 3 4 2 5 1 3 5 2 4 3 5 4 4 5 4 5 5 5 1 3 4 5 5 3
 [38] 5 5 2 2 4 5 1 3 4 2 1 3 5 2 2 3 5 5 2 5 5 1 2 5 4 5 5 2 2 2 5 5 5 5 5 3 4
 [75] 4 2 3 5 5 1 5 5 2 3 4 2 3 5 2 5 3 2 1 1 5 5 2 5 5 1
> v <- GetRF(mysample); v
[1] 3.00 4.00 5.00 1.00 2.00
    0.25 0.19 0.43 0.07 0.06
```



```
#Ισοδύναμα με την built-in συνάρτηση της R
> table(v)
 1  2  3  4  5
 7  6 25 19 43
> table(v)/length(v)
      1       2      3      4      5
     0.07    0.06   0.25   0.19   0.43
```

Δειγματοληψία από την τ.μ. X με την μέθοδο της αντιστροφής στις εξής περιπτώσεις:

1. $X \sim \mathcal{G}eo(p)$
2. $X \sim \mathcal{N}b(n,p)$
3. $X \sim Bernoulli(p)$
4. $X \sim Bin(n,p)$

1. Θεωρούμε την γεωμετρική κατανομή με στήριγμα το $X(\Omega)=\mathbb{N}$. Εάν $X=k$ είναι ο αριθμός των ανεξάρτητων δοκιμών Bernoulli με πιθανότητα επιτυχίας p, έως την πρώτη επιτυχία, τότε

$$f_X(k)=P\{X=k\}=\mathcal{G}eo(k\mid p)=p(1-p)^{k-1}\,\mathcal{I}(k\geq 1),$$

με αθροιστική συνάρτηση κατανομής:

$$F_X(k)=P\{X\leq k\}=\sum_{j=1}^{k}P\{X=j\}=p\sum_{j=1}^{k}(1-p)^{j-1}$$

$$=p\sum_{l=0}^{k-1}(1-p)^{l}=p\left\{\frac{1-(1-p)^{k}}{1-(1-p)}\right\}=1-(1-p)^{k}\ .$$

Τότε

$$F_X^-(u) = \inf\left\{x \in X(\Omega) : u \le F_X(x)\right\} = \inf\left\{x \in X(\Omega) : u \le 1-(1-p)^x\right\}$$
$$= \inf\left\{x \in X(\Omega) : x \ge \frac{\log(1-u)}{\log(1-p)}\right\} = \left\lceil \frac{\log(1-u)}{\log(1-p)} \right\rceil,$$

όπου $\lceil x \rceil$ η συνάρτηση ceiling του x.

Έτσι μία παρατήρηση $k \in \mathbb{N}$ από την γεωμετρική με παράμετρο p θα είναι η

$k = \left\lceil \frac{\log u}{\log(1-p)} \right\rceil \sim \mathcal{G}eo(p)$, όπου $u \sim \mathcal{U}(0,1)$.

2. Γνωρίζουμε ότι

$$p_X(k) = P\{X = k\} = \mathcal{N}b(k \mid n, p) = \binom{k-1}{n-1} p^n (1-p)^{k-n}\, \mathcal{I}(k \ge n),$$

και ότι

$$(X_i \mid p) \overset{i.i.d.}{\sim} \mathcal{G}eo(p),\ 1 \le i \le n \ \Rightarrow\ \textstyle\sum_{i=1}^n (X_i \mid p) \sim \mathcal{N}b(n, p).$$

Έτσι μία παρατήρηση $k \in X(\Omega) = \mathbb{N}_0 + n$ από την αρνητική διωνυμική με παράμετρο (n, p), θα είναι η

$$u_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1),\ 1 \le i \le n \ \Rightarrow\ k = \sum_{i=1}^n \left\lceil \frac{\log(u_i)}{\log(1-p)} \right\rceil \sim \mathcal{N}b(n, p).$$

3.– 4. Γνωρίζουμε ότι:

$$(X_i \mid p) \overset{i.i.d.}{\sim} Beroulli(p),\ 1 \le i \le n \ \Rightarrow\ \textstyle\sum_{i=1}^n (X_i \mid p) \sim Bin(n, p).$$

Μία παρατήρηση $k_i \in \{0,1\}$ από την Bernoulli με παράμετρο p θα είναι η

$$u_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1) \ \Rightarrow\ k_i = \mathcal{I}(p \le u_i) \sim Bernoulli(p).$$

έτσι μία παρατήρηση $k \in \{0,1,\ldots,n\}$ από την διωνυμική με παράμετρο (n, p) θα είναι η

$$u_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1),\ 1 \le i \le n \ \Rightarrow\ k = \sum_{i=1}^n k_i = \sum_{i=1}^n \mathcal{I}(p \le u_i) \sim Bin(n, p).$$

Πρόταση

Ισχύει ότι εάν $X \sim \mathrm{Exp}\left(-\log(1-p)\right)$ τότε $\lceil X \rceil \sim Geo(p)$.

Δίνεται ότι $X \sim \mathrm{Exp}(\lambda)$, τότε $U \sim \mathcal{U}(0,1) \Leftrightarrow -\frac{1}{\lambda}\log(1-U) \sim \mathrm{Exp}(\lambda)$. Κάνοντας την αντικατάσταση $\lambda = -\log(1-p)$ έχουμε ότι

$$\frac{\log(1-U)}{\log(1-p)} \sim \mathrm{Exp}\left(-\log(1-p)\right)$$

$$\Leftrightarrow \left\lceil \frac{\log(1-U)}{\log(1-p)} \right\rceil \overset{\mathcal{D}}{=} \left\lceil \frac{\log(u)}{\log(1-p)} \right\rceil \overset{\mathcal{D}}{=} \left\lceil \mathrm{Exp}\left(-\log(1-p)\right) \right\rceil \overset{\mathcal{D}}{=} Geo(p).$$

Παρατήρηση

Εάν $X_i \overset{i.i.d.}{\sim} \mathrm{Exp}\left(-\log(1-p)\right) = \mathcal{G}a\left(1, -\log(1-p)\right)$ για $1 \le i \le n$ τότε και $\lceil X_i \rceil \sim Geo(p)$ για $1 \le i \le n$ από όπου και $\sum_{i=1}^{n} X_i \sim \mathcal{G}a\left(n, -\log(1-p)\right) \Rightarrow \sum_{i=1}^{n} \lceil X_i \rceil \sim \mathcal{N}b(n,p)$.

Παράδειγμα

Να γίνει δειγματοληψία από με την μέθοδο της αντιστροφής από τις εξής περικομμένες (truncated) κατανομές:

1. $p_{X|Y}(x|y) \propto Geo(x|p)\mathcal{I}(x \ge y)$ με $y \in \mathbb{N}$
2. $f_{X|Y}(x|y) \propto Exp(x|\lambda)\mathcal{I}(x > y)$ με $y \in (0,\infty)$
3. $f_{X|A}(x|y) \propto Exp(x|\lambda)\mathcal{I}(x \in A)$ με $A \in \mathcal{B}(0,\infty)$

1. $p_{X|Y}(x|y) \propto Geo(x|p)\mathcal{I}(x \ge y) \propto (1-p)^x \mathcal{I}(x \ge y)$,

έτσι εάν

$$p_{X|Y}(x|y) = c(1-p)^x \mathcal{I}(x \ge y) \Rightarrow 1 = \sum_{x \ge y} f(x|y) = c \sum_{x \ge y} (1-p)^x,$$

από όπου και

$$c^{-1}=\sum_{x\geq y}(1-p)^{x}=\sum_{x=y}^{\infty}(1-p)^{x}=\frac{(1-p)^{y}}{1-(1-p)}=p^{-1}(1-p)^{y}$$

οπότε

$$p_{X|Y}(x\mid y)=p(1-p)^{-y}(1-p)^{x}\,\mathcal{I}(x\geq y)=p(1-p)^{x-y}\,\mathcal{I}(x\geq y),$$

με αθροιστική συνάρτηση κατανομής:

$$F_{X|Y}(x\mid y)=\sum_{l=y}^{x}p_{X|Y}(l\mid y)=\sum_{l=y}^{x}p(1-p)^{l-y}$$
$$=p\sum_{r=0}^{x-y}(1-p)^{r}=p\left\{\frac{1-(1-p)^{x-y+1}}{1-(1-p)}\right\}=1-(1-p)^{x-y+1}\,.$$

Τότε για κάθε $y\geq 1$ θα έχουμε:

$$F_{X|Y}^{-}(u\mid y)=\inf\left\{x:F(x\mid y)\geq u\right\}=\inf\left\{x:1-(1-p)^{x-y+1}\geq u\right\}$$
$$=\inf\left\{x:x-y+1\geq\frac{\log(1-u)}{\log(1-p)}\right\}=\left\lceil\frac{\log(1-u)}{\log(1-p)}\right\rceil+y-1\,.$$

Έτσι μία παρατήρηση x από την περικομμένη γεωμετρική $p_{X|Y}(x\mid y)$, με παράμετρο p θα είναι η

$$x=\left\lceil\frac{\log(u)}{\log(1-p)}\right\rceil+y-1 \text{ όπου } u\sim\mathcal{U}(0,1).$$

2. $f_{X|Y}(x\mid y)\propto\mathrm{Exp}(x\mid\lambda)\,\mathcal{I}(x>y)\propto e^{-\lambda x}\mathcal{I}(x>y)$

$$f_{X|Y}(x\mid y)=ce^{-\lambda x}\mathcal{I}(x\geq y)\ \Rightarrow 1=\int_{x=y}^{\infty}ce^{-\lambda x}dx=\frac{c}{\lambda}e^{-\lambda y}\Leftrightarrow c=\lambda e^{\lambda y}$$

από όπου

$$f_{X|Y}(x\mid y)=\lambda e^{-\lambda(x-y)}\mathcal{I}(x\geq y),$$

και

$$F_{X|Y}(x \mid y) = \int_{t=y}^{x} f(t \mid y)\,dt = \int_{t=y}^{x} \lambda e^{-\lambda(t-y)}dt = 1 - e^{-\lambda(x-y)}.$$

Θέτοντας

$$u = F_{X|Y}(x \mid y) \Leftrightarrow x = -\frac{1}{\lambda}\log(1-u) + y$$

Έτσι μία παρατήρηση x από την περικομμένη εκθετική $f_{X|Y}(x \mid y)$, με παράμετρο λ θα είναι

$$x = -\frac{1}{\lambda}\log(1-u) + y \text{ όπου } u \sim \mathcal{U}(0,1).$$

3. $f(x \mid A) \propto \mathrm{Exp}(x \mid \lambda)\mathcal{I}(x \in A) \Rightarrow f(x \mid A) = C\,e^{-\lambda x}\mathcal{I}(x \in A) \Rightarrow$

$$\Rightarrow 1 = \int_{\mathbb{R}} f(x \mid A)\,dx = C\int_{A} e^{-\lambda x}dx \Rightarrow C^{-1} = \int_{A} e^{-\lambda x}dx .$$

Εάν $A = (a,b)$ έτσι ώστε $0 < a < b$ θα έχουμε $C^{-1} = \frac{1}{\lambda}\left(e^{-\lambda a} - e^{-\lambda b}\right)$

$$u = F(x \mid A) = C\int_{t=a}^{x} e^{-\lambda t}dt = \frac{C}{\lambda}\left(e^{-\lambda a} - e^{-\lambda x}\right) \Rightarrow x = -\frac{1}{\lambda}\log\left(e^{-\lambda a} - \lambda u C^{-1}\right),$$

οπότε

$$u \sim \mathcal{U}(0,1) \Rightarrow x = -\frac{1}{\lambda}\log\left(e^{-\lambda a} - u\left(e^{-\lambda a} - e^{-\lambda b}\right)\right) \sim \mathrm{Exp}(x \mid \lambda)\mathcal{I}(a < x < b)$$

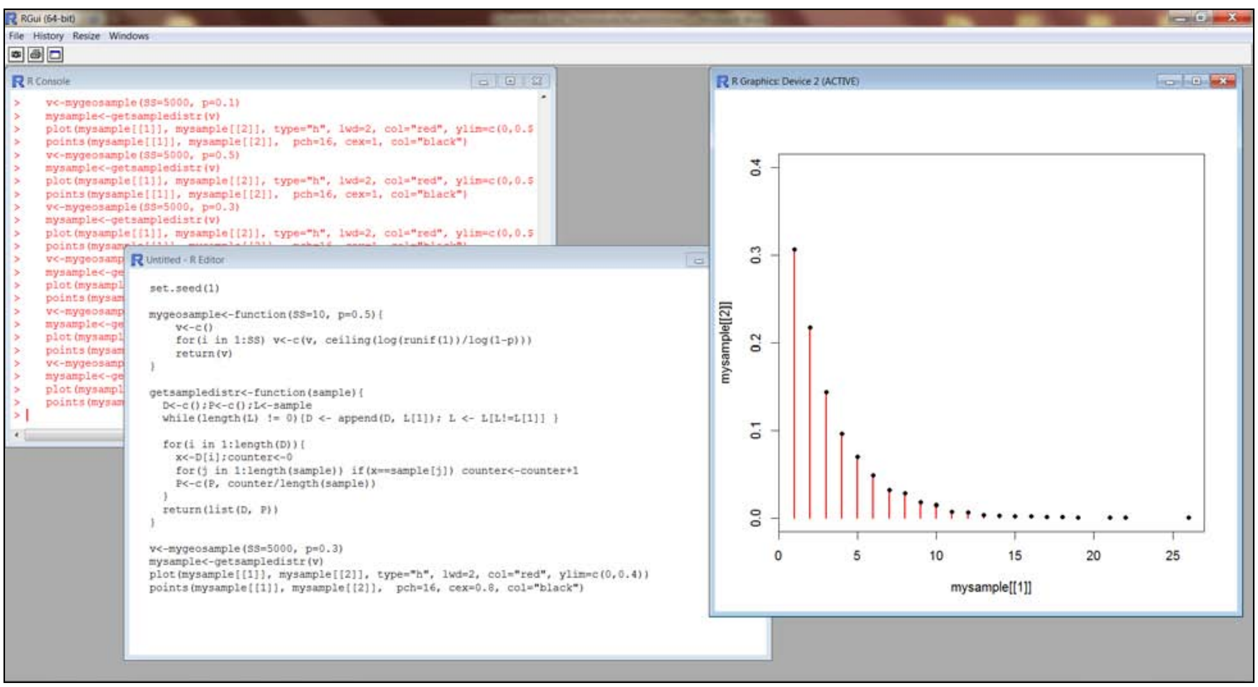
Παράδειγμα

Να γραφτεί R – script που να προσομοιώνει δείγμα μεγέθους N από την γεωμετρική και την διωνυμική κατανομή.

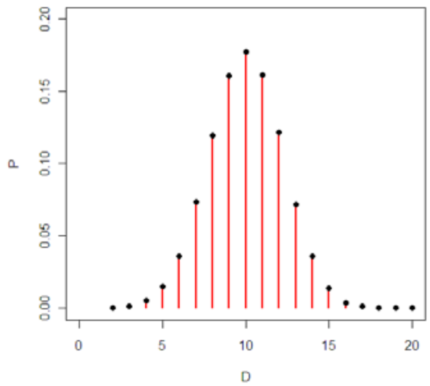
1.

```
set.seed(1)
mygeosample<-function(SS=10, p=0.5){
    v<-c()
```

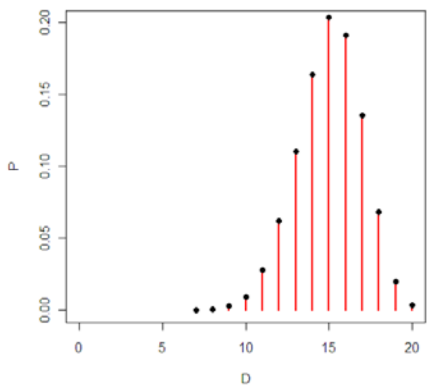
```
    for(i in 1:SS) v<-c(v, ceiling(log(runif(1))/log(1-p)))
    return(v)
}
getsampledistr<-function(sample){
  D<-c();P<-c();L<-sample
  while(length(L) != 0){D <- append(D, L[1]); L <- L[L!=L[1]] }
  for(i in 1:length(D)){
    x<-D[i];counter<-0
    for(j in 1:length(sample)) if(x==sample[j]) counter<-counter+1
    P<-c(P, counter/length(sample))
  }
  return(list(D, P))
}
v<-mygeosample(SS=5000, p=0.3); mysample<-getsampledistr(v)
plot(mysample[[1]], mysample[[2]], type="h", lwd=2, col="red", ylim=c(0,0.4))
points(mysample[[1]], mysample[[2]],  pch=16, cex=0.8, col="black")
```



```
# Sampling a vector of independent Bernoulli random deviates
BerSampler <- function(p=0.5, N=100){
   v<-c()
   for(i in 1:N) if(runif(1)<p) v <- c(v, 1) else v <- c(v, 0)
   return(v)
}

# Here we sample a vector of independent binomial random deviates.
BinSampler <- function(p=0.5, n=20, N=20000){
   v <- c()
   for(i in 1:N) {
       sum<-0
       for(j in 1:n) sum<-sum+BerSampler(p=p, N=1)
       v<-append(v, sum)
   }
   return(v)
}

getsampledistr<-function(sample){
   D<-c();P<-c();L<-sample
   while(length(L)!=0){D<-append(D, L[1]); L<-L[L!=L[1]]}

   for(i in 1:length(D)){
       x<-D[i];counter<-0
       for(j in 1:length(sample)) if(x==sample[j]) counter<-counter+1
       P<-append(P, counter/length(sample))
    }
    print(D);print(P)
    plot(D, P, type="h", lwd=2, col="red", ylim=c(0,0.20), xlim=c(0,20))
    points(D, P,  pch=16, cex=1, col="black")
}
> w<-BinSampler(p=0.5, n=20, N=40000); getsampledistr(w)
```




```
> w<-BinSampler(p=0.75, n=20, N=40000); getsampledistr(w)
```



Δειγματοληψία από διακριτές μίξεις κατανομών χρησιμοποιώντας την μέθοδο της αντιστροφής

Θέλουμε να κάνουμε δειγματοληψία από την τ.μ. X με πυκνότητα

$$X \sim \sum_{j=1}^{n} p_j f_{X_j}(\cdot), \quad p_j > 0, \quad \sum_{j=1}^{n} p_j = 1,$$

όπου f_{X_j} για $1 \le j \le n$ είναι πυκνότητες με αθροιστικές συναρτήσεις κατανομής F_{X_j} .

Πρώτα κάνουμε δειγματοληψία από την τ.μ. $J \sim \sum_{j=1}^{n} p_j \delta_j(\cdot)$ έτσι ώστε $P\{J = j\} = p_j$ (το μέτρο μίξης) για $1 \le j \le n$ και $\sum_{j=1}^{n} p_j = 1$, και στην συνέχεια από την τ.μ. $[X \mid J = j]$ έτσι ώστε:

$$[X \mid J = j] \sim f_{X|J}(\cdot \mid j) = f_{X_j}(\cdot).$$

Από όπου και

$$f_X(x)=\sum_{j=1}^{n} f_{X,J}(x,j)=\sum_{j=1}^{n} P\{J=j\} f_{X|J}(x\mid j)=\sum_{j=1}^{n} p_j f_{X_j}(x).$$

<u>Παράδειγμα</u>

Θέλουμε να κάνουμε δειγματοληψία από την πυκνότητα

$$f_X(x)=p\lambda_1 e^{\lambda_1 x}\mathcal{I}(x<0)+(1-p)\lambda_2 e^{-\lambda_2 x}\mathcal{I}(x\geq 0),$$

όπου λ_1 και λ_2 θετικοί παράμετροι.

Η πυκνότητα $f_X(x)=f(x)$ είναι μίξη των πυκνοτήτων

$$f_{X_1}(x)=f_1(x)=\lambda_1 e^{\lambda_1 x}\mathcal{I}(x<0) \text{ και } f_{X_2}(x)=f_2(x)=\lambda_2 e^{-\lambda_2 x}\mathcal{I}(x\geq 0),$$

με μέτρο μίξης $P\{J=1\}=p$, $P\{J=2\}=1-p$. Μπορούμε να θεωρήσουμε ότι έχουμε έχει τρείς παραμέτρους p,λ_1 και λ_2, συμβολικά $X\sim\mathcal{GDe}(p,\lambda_1,\lambda_2)$ (Generalized Double exponential).

Η αθροιστική συνάρτηση κατανομής της X θα είναι:

$$F(x)=p\int_{t=-\infty}^{x}\lambda_1 e^{\lambda_1 t}\mathcal{I}(t<0)\,dt+(1-p)\int_{t=-\infty}^{x}\lambda_2 e^{-\lambda_2 t}\mathcal{I}(t\geq 0)\,dt$$

$$=\begin{cases} p\int_{-\infty}^{x}\lambda_1 e^{\lambda_1 t}dt & x\leq 0\\ p\int_{-\infty}^{0}\lambda_1 e^{\lambda_1 t}dt+(1-p)\int_{0}^{x}\lambda_2 e^{-\lambda_2 t}dt & x\geq 0\end{cases}$$

$$=p\begin{Bmatrix} e^{\lambda_1 x} & x\leq 0\\ 1 & x\geq 0\end{Bmatrix}+(1-p)\begin{Bmatrix} 0 & x\leq 0\\ 1-e^{-\lambda_2 x} & x\geq 0\end{Bmatrix}.$$

Το σχήμα δειγματοληψίας λοιπόν θα είναι:

$$u_1, u_2 \overset{iid}{\sim} \mathcal{U}(0,1) \Rightarrow \frac{1}{\lambda_1}\log(u_2)\mathcal{I}(u_1 \le p) - \frac{1}{\lambda_2}\log(u_2)\mathcal{I}(u_1 > p) \sim f \ .$$

<u>Παρατήρηση</u>

Είδαμε ότι εάν $X \sim f_X$ με $f_X(x) = p\lambda_1 e^{\lambda_1 x}\mathcal{I}(x<0) + (1-p)\lambda_2 e^{-\lambda_2 x}\mathcal{I}(x \ge 0)$ τότε

$$X \overset{\mathcal{D}}{=} \frac{1}{\lambda_1}\log(U_2)\mathcal{I}(U_1 \le p) - \frac{1}{\lambda_2}\log(U_2)\mathcal{I}(U_1 > p).$$

Εμφανώς εάν $U_1, U_2 \overset{i.i.d.}{\sim} \mathcal{U}(0,1)$ τότε $\mathbb{E}(X) = \mathbb{E}(Y)$ ενώ εάν $P\{U_1 = U_2\} = 1,$ καταλήγουμε στο <u>άτοπο</u> $\mathbb{E}(Y) = \mathbb{E}(X) + \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right) p \log p$.

Πράγματι

$$\mathbb{E}(X) = \int_{\mathbb{R}} x\left\{p\lambda_1 e^{\lambda_1 x}\mathcal{I}(x<0) + (1-p)\lambda_2 e^{-\lambda_2 x}\mathcal{I}(x \ge 0)\right\}dx$$

$$= p\int_{\mathbb{R}} x\lambda_1 e^{\lambda_1 x}\mathcal{I}(x<0)\,dx + (1-p)\int_{\mathbb{R}} x\lambda_2 e^{-\lambda_2 x}\mathcal{I}(x \ge 0)\,dx$$

$$= p\int_{x=-\infty}^{0} x\lambda_1 e^{\lambda_1 x}dx + (1-p)\int_{x=0}^{\infty} x\lambda_2 e^{-\lambda_2 x}dx = p\left(-\frac{1}{\lambda_1}\right) + (1-p)\left(\frac{1}{\lambda_2}\right)$$

$$\mathbb{E}\left[\log(u_2)\mathcal{I}(u_1 < p)\right] = \mathbb{E}\left[\log(u_2)\right]\mathbb{E}\left[\mathcal{I}(u_1 < p)\right]$$

$$= \int_{u_2=0}^{1} \log(u_2)\,du_2 \int_{u_1=0}^{1} \mathcal{I}(u_1 < p)\,du_1 = (-1)(p)$$

$$\mathbb{E}\left[\log(u_2)\mathcal{I}(u_1 \ge p)\right] = \mathbb{E}\left[\log(u_2)\right]\mathbb{E}\left[\mathcal{I}(u_1 \ge p)\right]$$

$$= \int_{u_2=0}^{1} \log(u_2)\,du_2 \int_{u_1=0}^{1} \mathcal{I}(u_1 \ge p)\,du_1 = (-1)(1-p)$$

από όπου και $\mathbb{E}(Y) = -\frac{p}{\lambda_1} + \frac{1-p}{\lambda_2}$.

Ενώ εάν $P\{U_1 = U_2\} = 1$, θα έχουμε:

$$\mathbb{E}\left[\log(U)\mathcal{I}(U \le p)\right] = \int_{u=0}^{p} \log u \, du = p \log p - p$$

$$\mathbb{E}\left[\log(U)\mathcal{I}(U > p)\right] = \int_{u=p}^{1} \log u \, du = -p \log p + p - 1$$

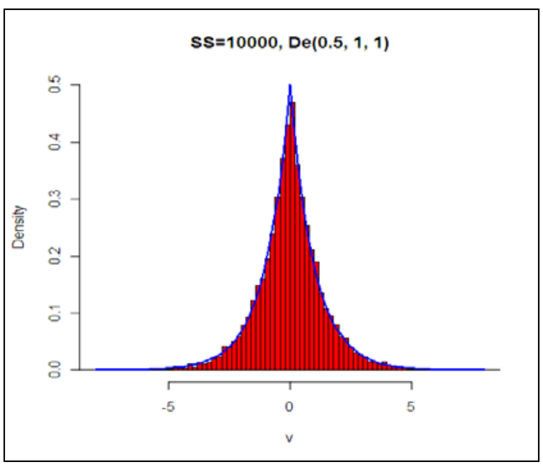
$$\mathbb{E}(Y) = \frac{1}{\lambda_1}(p \log p - p) - \frac{1}{\lambda_2}(-p \log p + p - 1) = \mathbb{E}(X) + \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right) p \log p \ .$$

<u>Παρατήρηση</u>

Όταν $p = 1/2$ και $\lambda_1 = \lambda_2 = \lambda$ η πυκνότητα της X παίρνει την μορφή $f_X(x) = \frac{\lambda}{2} e^{-\lambda|x|}$ για $x \in \mathbb{R}$ που είναι η κατανομή $X \sim \mathcal{D}e(\lambda) = \mathcal{GD}e(1/2, \lambda, \lambda)$ η κατανομή Laplace ή Double exponential.

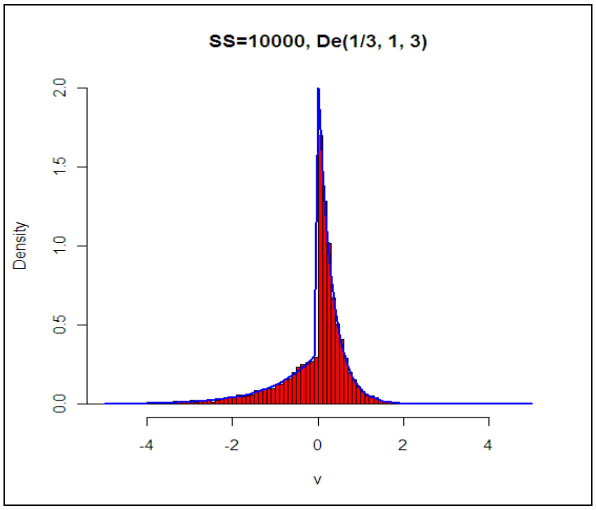
```
# sample from the generalized Double exponential distribution GDe(0.5, 1, 1)
Dexponential <- function(SS=10000, p=0.5, lam1=1, lam2=1) {
    v <- NULL
    for(i in 1:SS) {
         u1<-runif(1);u2<-runif(1)
         if(u1<p)
             v <- append(v, (1/lam1)*log(u2))
          else
             v <- append(v, (-1/lam2)*log(u2))
    }
    return(v)
}
# Plot the density function of the generalized double exponential distribution.
DDEXP<-function(x, p=0.5, lambda1=1, lambda2=1) {
```

```
    v <- c()
    for(i in 1:length(x)) {
        if(x[i] < 0)
            branch <- p*lambda1*exp(lambda1*x[i])
        else
            branch <- (1-p)*lambda2*exp(-lambda2*x[i])
        v <- append(v, branch)
    }
    return(v)
}

#EXPERIMENT 1.
# Take an independent sample of size SS=10000, from the Double Exponential distribution
# GDe(0.5,1,1).
v <- Dexponential(SS=10000, p=0.5, lambda1=1, lambda2=1)
# plot the probability histogram of v.
hist(v, breaks=100, freq=FALSE, ylim=c(0, 0.5), xlim=c(-8, 8),
       main="SS=10000, De(0.5, 1, 1)", col="red")
# Add the graph of the Double Exponential density to the probability histogram
curve(DDEXP(x), from=-8 ,to=8, col="blue", lwd=2, add=TRUE)
```




```
# EXPERIMENT 2.
v <- Dexponential(SS=10000, p=1/3,  lambda1=1, lambda2=3)
hist(v, breaks=100, freq=FALSE, ylim=c(0, 2), xlim=c(-5, 5),
       main="SS=10000, De(1/3, 1, 3)", col="red")
```

```
curve(DDEXP(x, p=1/3, lambda2=3), from=-5, to=5, col="blue", lwd=2, add=TRUE)
```



Δειγματοληψία από την πυκνότητα $g(x)$ με αποδοχή-απόρριψη (accept-reject), όταν το στήριγμα της $g(x)$ έχει πεπερασμένο μήκος.

Θέλουμε να κάνουμε δειγματοληψία από την πυκνότητα g που έχει στήριγμα $S_g = \{x \in \mathbb{R} : g(x) > 0\}$ πεπερασμένου μήκους, και γνωρίζουμε ότι για κάθε $x \in S_g$, $g(x) \le m$ είτε (ισχυρότερα), $m = \sup_{x \in S_g} g(x) < \infty$.

Έστω u και u' ανεξάρτητα δείγματα από τις ομοιόμορφες με στήριγμα τα σύνολα S_g και $(0, m)$ αντιστοίχως. Τότε το u' δοθέντος ότι $u' < g(u)$, προέρχεται από την g, δηλαδή:

$$\left[u \mid u' < g(u)\right] \sim g .$$

Στην ουσία, αυτό που κάνουμε όταν ισχύει ότι $u' < g(u)$, είναι δειγματοληψία από την ομοιόμορφη που ορίζεται στο δισδιάστατο χωρίο Γ_g μεταξύ του γραφήματος της g και του S_g. Δηλαδή έχουμε ότι:

$$\left(u' < g(u)\right) \Leftrightarrow \left((u, u') \sim \mathcal{U}(\Gamma_g)\right) \Rightarrow \left[u \mid u' < g(u)\right] \sim g .$$

Το accept – reject σχήμα δειγματοληψίας λοιπόν είναι:

Όταν $g(x) \le m, \ \ x \in S_g$

$$\left. \begin{cases} u \sim \mathcal{U}(S_g) \\ u' \sim \mathcal{U}(0,m) \ \Rightarrow \ [u \mid u' < g(u)] \sim g \end{cases} \right\} \qquad (AR1)$$

Θα χρειαστούμε το παρακάτω αποτέλεσμα:

$$P\{h(X,Y) \in A\} = \mathbb{E}\left[\mathcal{I}\left(h(X,Y) \in A\right)\right] = \mathbb{E}\left[\mathbb{E}\left[\mathcal{I}\left(h(X,Y) \in A\right) \mid X\right]\right]$$
$$= \int_{\mathbb{R}} \mathbb{E}\left[\mathcal{I}\left(h(X,Y) \in A\right) \mid X = x\right] f_X(x)dx = \int_{\mathbb{R}} P\{h(x,Y) \in A \mid X = x\} f_X(x)dx \ ,$$

και εάν X και Y είναι ανεξάρτητες τ.μ.

$$P\{h(X,Y) \in A\} = \int_{\mathbb{R}} P\{h(x,Y) \in A\} f_X(x)d \ .\!.$$

όπου $P\{h(x,Y) \in A\} = \int_{h(x,y) \in A} f_Y(y)dy$.

Ας συμβολίζουμε το ενδεχόμενο της αποδοχής (acceptance region) με

$$\mathbb{A} = \{\omega \in \Omega : u'(\omega) < g(u(\omega))\} = \{u' < g(u)\},$$

η πιθανότητα αποδοχής όταν $S_g = (a,b)$ θα είναι:

$$P(\mathbb{A}) = P\{u' < g(u)\} = \int_{S_g} P\{u' < g(u) | u = x\} \mathcal{U}(x \mid S_g)d \ \ x$$

$$= \int_{x=a}^{b} P\{u' < g(u) \mid u = x\} \mathcal{U}(x \mid a,b)dx = \frac{1}{b-a} \int_{x=a}^{b} P\{u' < g(x) \mid u = x\} dx \ ,$$

και επειδή οι u και u' είναι ανεξάρτητες τ.μ. παίρνουμε

$$P(\mathbb{A}) = \frac{1}{b-a} \int_{x=a}^{b} P\{u' < g(x)\} dx \ ,$$

αλλά $u' \sim \mathcal{U}(0,m)$ και $u' < g(x) \le m$ εφόσον $\sup_{x \in S} g(x) \le m$, που δίνει:

$$P\left\{u' < g(x)\right\} = \int_{u'=0}^{g(x)} \mathcal{U}(u' \mid 0, m)\, du' = \frac{g(x)}{m},$$

και έτσι $P(\mathbb{A}) = \dfrac{1}{m(b-a)} \displaystyle\int_{S_g} g(x)\, dx = \dfrac{1}{m(b-a)}$.

Για να δείξουμε ότι $u|\mathbb{A} \sim g$, αρκεί να δείξουμε ότι $P\left\{u \le x|\mathbb{A}\right\} = G(x)$, όπου G η αθροιστική συνάρτηση κατανομής της g

$$P\left\{u \le x|\mathbb{A}\right\} = \frac{P\left(\left\{u \le x\right\} \cap \mathbb{A}\right)}{P(\mathbb{A})} = \frac{P\left\{u \le x, u' < g(u)\right\}}{P(\mathbb{A})}$$

$$= \frac{1}{P(\mathbb{A})} \int_{\mathbb{R}} P\left\{u \le x, u' < g(u) | u = y\right\} \mathcal{U}(y \mid a, b)\, dy$$

$$= \frac{1}{P(\mathbb{A})} \int_{\mathbb{R}} P\left\{y \le x, u' < g(y) | u = y\right\} \mathcal{U}(y \mid a, b)\, dy$$

$$= \frac{1}{P(\mathbb{A})} \int_{\mathbb{R}} P\left\{y \le x, u' < g(y)\right\} \mathcal{U}(y \mid a, b)\, dy$$

$$= \frac{1}{P(\mathbb{A})} \int_{\mathbb{R}} P\left\{y \le x \mid u' < g(y)\right\} P\left\{u' < g(y)\right\} \mathcal{U}(y \mid a, b)\, d\ \ y$$

$$= \frac{1}{(b-a) P(\mathbb{A})} \int_{y=a}^{b} \mathcal{I}(y \le x) P\left\{u' < g(y)\right\} dy$$

$$= \frac{1}{(b-a) P(\mathbb{A})} \int_{y=a}^{x} \left\{\frac{g(y)}{m}\right\} dy$$

$$= \int_{y=a}^{x} g(y)\, dy = G(x)$$

<u>Εναλλακτικά με τη χρήση διδιάστατης ομοιόμορφης</u>

Εάν γνωρίζουμε ότι $g(x) \le m < \infty$ για κάθε $x \in S_g$, $S_g = (a, b)$ και ότι η $G(x)$ είναι η αθροιστική συνάρτηση κατανομής που αντιστοιχεί στην $g(x)$, θέτουμε σαν από κοινού πυκνότητα των τ.μ. X και Y την ομοιόμορφη στο $(a, b) \times (0, m)$

$$f_{X,Y}(x,y)=\begin{cases}\dfrac{1}{m(b-a)} & a<x<b,\ 0<y<m\\ 0 & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}.$$

Εμφανώς $f_{X,Y}(x,y)=\mathcal{U}(x\,|\,S_g)\mathcal{U}(y\,|\,0,m)=f_X(x)\,f_Y(y)$,

δηλαδή X και Y είναι ανεξάρτητες με $f_X(x)=\mathcal{U}(x\,|\,S_g)$ και $f_Y(y)=\mathcal{U}(y\,|\,0,m)$,

Τότε η πιθανότητα αποδοχής θα είναι:

$$P(\mathbb{A})=P\{Y<g(X)\}=\iint\limits_{y<g(x)} f_{X,Y}(x,y)\,dx\,dy$$

$$=\int_{x=a}^{b}\int_{y=0}^{g(x)}\frac{1}{m(b-a)}dy\,dx=\frac{1}{m(b-a)}\int_{S_g}g(x)\,dx=\frac{1}{m(b-a)},$$

ενώ ισχύει ότι $\left[X\,|\,Y<g(X)\right]\sim g$ εφόσον

$$P\{X\le x\,|\,Y<g(X)\}=\frac{P\{X\le x,Y<g(X)\}}{P\{Y<g(X)\}}$$

$$=m(b-a)\iint\limits_{\substack{u\le x\\ v<g(u)}} f_{X,Y}(u,v)\,du\,dv$$

$$=m(b-a)\int_{u=a}^{x}\int_{v=0}^{g(u)}\frac{1}{m(b-a)}dv\,du$$

$$=\int_{u=a}^{x}g(u)\,du=G(x).$$

<u>Παράδειγμα</u>

Να δοθεί το σχήμα δειγματοληψίας accept-reject για την πυκνότητα $g(x,y)$, εάν γνωρίζουμε ότι έχει στήριγμα $S_g=(a,b)\times(c,d)$ με πεπερασμένο μέτρο Lebesgue $\lambda(S_g)=(b-a)(c-d)$ και ότι $g(x,y)\le m<\infty,\ \forall(x,y)\in S_g$.

Θέτουμε

$$f_{X,Y,Z}(x,y,z)=\begin{cases}\dfrac{1}{m\lambda(S_g)} & (x,y)\in S_g,\ z\in(0,m)\\ 0 & \alpha\lambda\lambda o\acute{\upsilon}\end{cases}.$$

Τότε η πιθανότητα αποδοχής είναι:

$$\begin{aligned}
&P(\mathbb{A})=P\left\{Z<g(X,Y)\right\}\\
&=\iiint\limits_{z<g(x,y)} f_{X,Y,Z}(x,y,z)\,dx\,dy\,dz\\
&=\frac{1}{m\lambda(S_g)}\iint\limits_{(x,y)\in S_g}\left\{\int\limits_{z=0}^{g(x,y)} dz\right\}dx\,dy\\
&=\frac{1}{m\lambda(S_g)}\iint\limits_{S_g} g(x,y)\,dx\,dy\\
&=\frac{1}{m\lambda(S_g)}.
\end{aligned}$$

Ισχύει ότι $\left[X,Y\,|\,Z<g(X,Y)\right]\sim g$, πράγματι

$$P\left\{X\le x,Y\le y\,|\,Z<g(X,Y)\right\}=\frac{P\left\{X\le x,\,Y\le y,\ Z<g(X,Y)\right\}}{P\left\{Z<g(X,Y)\right\}}$$

$$=\frac{\iiint\limits_{u\le x,\,v\le y,\,t<g(u,v)} f_{X,Y,Z}(u,v,t)\,dudvdt}{\iiint\limits_{t<g(u,v)} f_{X,Y,Z}(u,v,t)\,dudvdt}=m\lambda(S_g)\iint\limits_{\substack{u\le x\\ v\le y}}\left\{\int\limits_{t=0}^{g(u,v)}\frac{dt}{m\lambda(S_g)}\right\}dudv$$

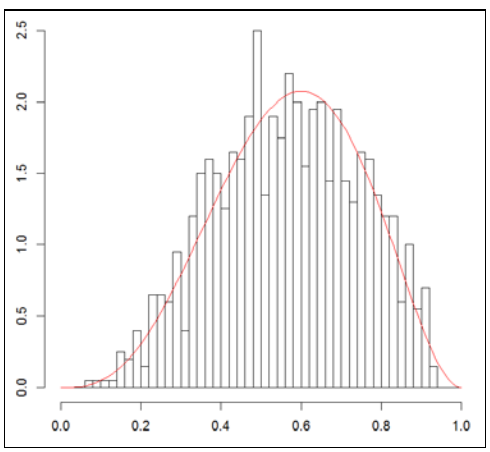
$$=\iint\limits_{\substack{u\le x\\ v\le y}} g(u,v)\,dudv=G(x,y)\ .$$

Δειγματοληψία από την $f_X(x)=\mathcal{B}e(x\,|\,4,3)=60x^3(1-x)^2$

Δειγματοληψία από την $\mathcal{B}e(a,b)$ για $a>1,b>1$.

```
sampledensity1 <- function(SS=100, a=4, b=3) {

   sample <- NULL

   m <- dbeta(x=(a-1)/(a+b-2), shape1=a, shape2=b)
```

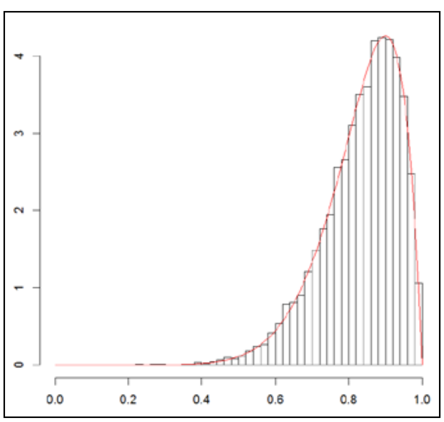
```
   for(i in 1:SS) {
        repeat{
            u <- runif(1); uprime <- m*runif(1)
            if(uprime < dbeta(x=u, shape1=a, shape2=b)) break
        }
        sample <- c(sample, u)
   }
   return(sample)
}
createsample1 <- function(mysize=100){
    x<-sampledensity1(SS=mysize)
    mymbreaks <- seq(from=0, to=1, by=1/50)
    hist(x, freq=FALSE, breaks=mymbreaks)
    curve(dbeta(x, shape1=3, shape2=4), add=TRUE, xlim=c(0, 1), col="red")
    return(x)
}
set.seed(1)
createsample1(1000)
```



#Δειγματοληψία από την $Be(x \mid p,q)$ για $p>1,q>1$.

```
g <- function(x, p, q) return(dbeta(x, shape1=p, shape2=q))
```

```
sampledensity <- function(SS, a, b){
   sample <- NULL
   m <- g(x=(a-1)/(a+b-2), p=a, q=b)
   for(i in 1:SS) {
        repeat{
            u <- runif(1); uprime <- m*runif(1)
            if(uprime < g(x=u, p=a, q=b)) break
        }
        sample <- c(sample, u)
   }
   return(sample)
}
createsample <- function(mysize, myshape1=3, myshape2=4){
    u <- sampledensity(SS=mysize, a=myshape1, b=myshape2)
    mymbreaks <- seq(from=0, to=1, by=1/50)
    hist(u, freq=FALSE, breaks=mymbreaks)
    curve(g(x, p=myshape1, q=myshape2), add=TRUE, xlim=c(0, 1), col="red")
    return(x)
}
set.seed(1)
createsample(mysize=10000, myshape1=10, myshape2=2)
```

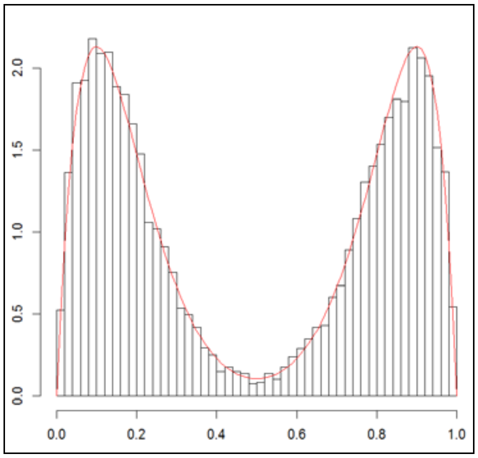


#Δειγματοληψία από την μίξη $\pi Be\left(x \mid p_1, q_1\right) + \left(1-\pi\right) Be\left(x \mid p_1, q_1\right)$

```
gmix <- function(x, pr, p, q)
```

```
       return(pr* dbeta(x, shape1=p[1], shape2=q[1])
        +(1-pr)* dbeta(x, shape1=p[2], shape2=q[2]))

g <- function(x, r, s) return(dbeta(x, shape1=r, shape2=s))

sampledensity <- function(SS, a, b) {
   sample <- NULL
   m <- g(x=(a-1)/(a+b-2), r=a, s=b)
   for(i in 1:SS) {
        repeat{
            u <- runif(1); uprime <- m*runif(1)
            if(uprime < g(x=u, r=a, s=b)) break
        }
        sample <- c(sample, u)
   }
   return(sample)
}

createsample <- function(mysize, p, q, pr){
    set.seed(1); u <- NULL
    for(i in 1:mysize){
       if(runif(1)<=pr)
            x <- sampledensity(SS=1, a=p[1], b=q[1])
         else
            x <- sampledensity(SS=1, a=p[2], b=q[2])
        u <- c(u, x)
    }
    mymbreaks <- seq(from=0, to=1, by=1/50)
    hist(u, freq=FALSE, breaks=mymbreaks)
    curve(gmix(x, pr=pr, p=p, q=q), add=TRUE, xlim=c(0, 1), col="red")
    return(u)
```

```
}
createsample(mysize=10000,  p=c(2,10), q=c(10,2), pr=1/2)
```



<u>Άσκηση</u>

Να γίνει δειγματοληψία από την πυκνότητα:

$$g(x) = \frac{1}{3} Be(x \mid 2,10) + \frac{1}{3} Be(x \mid 10,10) + \frac{1}{3} Be(x \mid 10,2).$$

<u>Άσκηση</u>

Έστω ότι στο διωνυμικό μοντέλο έχουμε $n = 75$ Bernoulli παρατηρήσεις $\{x_i : 1 \le i \le n\}$ με $x_i \in \{0,1\}$ και αντίστοιχη διωνυμική παρατήρηση $\{x = 34\}$. Εάν a-priori γνωρίζουμε ότι $\pi(\vartheta) = Be(\vartheta \mid p_0, q_0)$, $p_0 = 4.4,\ q_0 = 6.6$

1. Να βρεθεί η εκτίμηση του ϑ κάτω από τετραγωνική συνάρτηση απώλειας, καθώς και το posterior variance προσομοιώνοντας δείγμα από την posterior $\pi(\vartheta \mid x = 34)$ (από την θεωρία γνωρίζουμε ότι και αυτή είναι beta).

2. Να προσομοιωθεί δείγμα από την prior predictive (από την θεωρία γνωρίζουμε ότι εάν $\vartheta \sim \pi(\cdot) = Be(\cdot \mid p_0, q_0)$ και $x \mid \vartheta \sim Bin(\cdot \mid 75, \vartheta)$ τότε
$$x \sim \int_{\vartheta=0}^{1} Bin(\cdot \mid 75, \vartheta) Be(\vartheta \mid p_0, q_0) d\vartheta).$$

3. Εάν $y \mid \vartheta \sim Bin(10,\vartheta)$, να προσομοιωθεί δείγμα από την posterior predictive (από την θεωρία γνωρίζουμε ότι εάν $\vartheta \sim \pi(\cdot \mid x=34)$ και $y \mid \vartheta \sim Bin(\cdot \mid 75,\vartheta)$ τότε $x \sim \int_{\vartheta=0}^{1} Bin(\cdot \mid 75,\vartheta) Be(\vartheta \mid p_0,q_0) d\vartheta$).

Παράδειγμα

Δίνεται ότι X είναι ο αριθμός των δοκιμών σε σχήμα δειγματοληψίας accept – reject έως ότου επιτευχθεί αποδοχή και ότι η πιθανότητα αποδοχής είναι $P(\mathbb{A})$. Να βρεθεί η μέση τιμή της τ.μ. X. Πότε το σχήμα αποδοχής-απόρριψης γίνεται βέλτιστο;

Το ενδεχόμενο

$\{X=k\} = \{$ έχουμε αποδοχή μετά από k δοκιμές $\}$

$= \{$ γίνονται $k-1$ απορρίψεις και τελικά αποδοχή στην k δοκιμή $\}$,

έχει πιθανότητα

$P\{X=k\} = P(\mathbb{A})\left(1-P(\mathbb{A})\right)^{k-1} \mathcal{I}(k \ge 1)$, από όπου και $X \sim Geo\left(P(\mathbb{A})\right)$

και $\mathbb{E}(X) = \dfrac{1}{P(\mathbb{A})} = m(b-a)$.

Το σχήμα αποδοχής-απόρριψης γίνεται βέλτιστο όταν $m = \sup\limits_{x\in S} g(x)$.

Παρατήρηση:

$$\Pi_X(t) = \mathbb{E}\left(t^X\right) = \sum_{k=1}^{\infty} t^k P(\mathbb{A})\left(1-P(\mathbb{A})\right)^{k-1} = t\,P(\mathbb{A}) \sum_{k=1}^{\infty} \left\{t\left(1-P(\mathbb{A})\right)\right\}^{k-1} = \frac{t\,P(\mathbb{A})}{1-t\left(1-P(\mathbb{A})\right)}$$

από όπου και $\mathbb{E}(X) = \Pi'_X(1) = \dfrac{1}{P(\mathbb{A})}$.

Δειγματοληψία από την πυκνότητα $g(x)$ με accept – reject, όταν το στήριγμα S_g της πυκνότητας $g(x)$ δεν έχει γενικά πεπερασμένο μήκος.

Έστω ότι γνωρίζουμε πώς να κάνουμε δειγματοληψία από την πυκνότητα f αλλά είναι δύσκολο να κάνουμε δειγματοληψία από την πυκνότητα g.

Υποθέτουμε ότι:

1. Οι δύο πυκνότητες f και g, έχουν το ίδιο στήριγμα $S = S_f = S_g$ που μπορεί όμως να μην είναι πεπερασμένο.
2. Μπορούμε να βρούμε θετική σταθερά $a \geq 1$ τέτοια ώστε $g(x) < a\, f(x)$, για $\forall\, x \in S_g$, είτε ότι (ισχυρότερα) ότι a ικανοποιεί $a = \sup\limits_{x \in S} \dfrac{g(x)}{f(x)}$.

Τότε εάν πάρουμε δείγμα y από την f, δηλαδή $y \sim f$, και στην συνέχεια δοθέντος του y, δείγμα u' από την ομοιόμορφη με στήριγμα το $(0, a\, f(y))$ δηλαδή $u'|y \sim \mathcal{U}(0, a\, f(y))$ και ισχύει ότι $u' < g(y)$, αποδεχόμαστε το y σαν δείγμα από την g.

Ουσιαστικά αυτό που κάνουμε είναι δειγματοληψία από την ομοιόμορφη που ορίζεται στο χωρίο $\Gamma_{a f}$ κάτω από το γράφημα της $a\, f(x)$ που όμως σε αυτήν την περίπτωση μπορεί να έχει στήριγμα με άπειρο μήκος. Εάν $(y, u') \sim \mathcal{U}(\Gamma_g)$ θα έχουμε και $y \sim g$.

Το accept – reject σχήμα δειγματοληψίας είναι:

$$\left.\begin{array}{l} y \sim f \\ u'|y \sim \mathcal{U}(0, a\, f(y)) \end{array}\right\} \Rightarrow (y \,|\, u' < g(y)) \sim g$$

Ορίζουμε το acceptance region $\mathbb{A} = \{u' < g(y)\}$. Τότε πιθανότητα αποδοχής θα είναι

$$P(\mathbb{A}) = P\{u' < g(y)\} = \int\limits_{x \in S} P\{u' < g(y) | y = x\}\, f(x)\, d\; x$$

$$= \int\limits_{x \in S} P\{u' < g(x) | y = x\}\, f(x)\, dx$$

$$= \int\limits_{x \in S} \left\{ \int\limits_{t'=0}^{g(x)} \frac{dt'}{a\, f(x)} \right\} f(x)\, dx = \frac{1}{a} \int\limits_{x \in S} g(x)\, dx = \frac{1}{a}$$

εφόσον $0 < u' < g(x) < a\, f(x)$.

Θα δείξουμε τώρα ότι $y|\mathbb{A} \sim g$. Πράγματι

$$
\begin{aligned}
P\{y \le t|\mathbb{A}\} &= \frac{P\{y \le t, u' < g(y)\}}{P(\mathbb{A})} \\
&= a\int_S P\{y \le t, u' < g(y)|y = x\}\, f(x)\,dx \\
&= a\int_S P\{x \le t, u' < g(x)|y = x\}\, f(x)\,dx \\
&= a\int_S P\{x \le t\}P\{u' < g(x)|y = x\}\, f(x)\,dx \\
&= a\int_S \mathcal{I}(x \le t)P\{u' < g(x)|y = x\}\, f(x)\,dx
\end{aligned}
$$

Εάν $S = (\underline{S}, \overline{S})$ είναι υποσύνολο του $\mathbb{R}$ (πιθανώς απείρου μήκους), και επειδή $P\{u' < g(x)|y = x\} = \int_{t'=0}^{g(x)} \frac{dt'}{a\, f(x)}$, το προηγούμενο ολοκλήρωμα γίνεται

$$
P\{y \le t|\mathbb{A}\} = a\int_{\underline{S}}^{t} \left(\frac{g(x)}{a\, f(x)}\right) f(x)\,dx = \int_{\underline{S}}^{t} g(x)\,dx = G(t).
$$

Εναλλακτικά με τη χρήση δισδιάστατης ομοιόμορφης

Εάν γνωρίζουμε ότι υπάρχει θετική σταθερά $a > 1$ τέτοια ώστε $g(x) \le a\, f(x),\ \forall x \in S$, και ότι η G είναι η αθροιστική συνάρτηση κατανομής της g, θέτουμε σαν από κοινού πυκνότητα των τ.μ. X και Y την ομοιόμορφη στο χωρίο κάτω από την $y = a\, f(x)$, δηλαδή

$$
f_{X,Y}(x, y) = \begin{cases} C & x \in S,\ y < a\, f(x) \\ 0, & elsewhere \end{cases}
$$

$$
1 = \iint_{x\in S,\ y<a\, f(x)} C\,dy\,dx = C\int_{x\in S}\int_{y=0}^{a\, f(x)} dy\,dx = C\,a\int_{x\in S} f(x)\,dx = C\,a,
$$

από όπου και

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{a} & x \in S,\ y < a\ f(x) \\ 0 & \alpha\lambda\lambda o\acute{\upsilon} \end{cases} = \frac{1}{a}\mathcal{I}\left(x \in S,\ y < a\ f(x)\right),$$

με περιθώριες πυκνότητες:

$$\boxed{f_X(x) = \int_{y<a\ f(x)} f_{X,Y}(x,y)\,dy = \frac{1}{a}\int_{y=0}^{a\ f(x)} dy = f(x),\ \ x \in S}$$

$$\boxed{f_{Y|X}(y\,|\,x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\frac{1}{a}\mathcal{I}\left(y < a\ f(x)\right)}{f(x)} = \mathcal{U}\left(y|0, a\ f(x)\right),\ \ x \in S}$$

Η πιθανότητα αποδοχής $P(\mathbb{A})$ είναι:

$$P(\mathbb{A}) = P\{Y < g(X)\} = \iint_{y<g(x)} \frac{1}{a}\mathcal{I}\left(y < a\ f(x), x \in S\right) dy\,dx$$

$= \frac{1}{a}\int_{x\in S}\int_{y=0}^{g(x)} dy\,dx$, εφόσον $g(x) \le a\ f(x)$, για $\forall\ x \in S$

$$= \frac{1}{a}\int_S g(x)\,dx = \frac{1}{a}$$

Ενώ $G(x) = P\{X \le x\,|\,Y < g(X)\}$ εφόσον

$$P\{X \le x\,|\,Y < g(X)\} = \frac{P\{X \le x, Y < g(X)\}}{P\{Y < g(X)\}} = a\iint_{\substack{u\le x \\ v<g(u)}} f_{X,Y}(u,v)\,du dv$$

$$= a\int_{u=-\infty}^{x}\int_{v=0}^{g(u)} \frac{1}{a}\mathcal{I}\left(v < a\ f(u), u \in S\right) dv\,du = \int_{u=-\infty}^{x} g(u)\,du = G(x)$$

επειδή $g(u) \le a\ f(u)$, για $\forall\ u \in S$.

Χρησιμοποιώντας την δισδιάστατη ομοιόμορφη έχουμε ότι:

Όταν $g(x) < a\ f(x),\ \forall\ x \in S$

$$\left.\begin{cases} x \sim f_X = f \\ y|x \sim f_{Y|X}(\cdot|x) = \mathcal{U}(\cdot|0, a\, f(x)) \end{cases}\right\} \Rightarrow \left(x \mid y < g(x)\right) \sim g(\cdot)\,.$$

Παρατήρηση 1:

αντικαθιστώντας $x \leftarrow y$ και $y \leftarrow u$ παίρνουμε το σχήμα δειγματοληψίας

$$y \sim f(\cdot)$$

$$u|y \sim \mathcal{U}(\cdot|0, a\, f(y)) \;\Rightarrow\; \left(y \mid u < g(y)\right) \sim g(\cdot)\,.$$

Παρατήρηση 2:

Παρατηρήστε ότι για $g(x) \le m,\; x \in S$ με $\lambda(S) < \infty$, έχουμε

$$g(x) \le \underbrace{m\,\lambda(S)}_{a} \cdot \underbrace{\mathcal{U}(x \mid S)}_{f(x)}$$

$$x \sim f_X(\cdot) = f(\cdot) = \mathcal{U}(\cdot|S)$$

$$y|x \sim f_{Y|X}(\cdot|x) = \mathcal{U}(\cdot|0, a\, f(x)) = \mathcal{U}(\cdot|0, m)$$

$$\Leftrightarrow y|x \overset{d}{=} y \sim \mathcal{U}(\cdot|0, m) \Rightarrow \left(x \mid y < g(x)\right) \sim g(\cdot)$$

αντικαθιστώντας $x \leftarrow u$ και $y \leftarrow u'$ παίρνουμε το σχήμα δειγματοληψίας $(AR1)$:

Όταν $g(x) \le m,\; x \in S$

$$u \sim \mathcal{U}(\cdot|S)$$

$$u' \sim \mathcal{U}(\cdot|0, m) \;\Rightarrow\; u \mid u' < g(u) \sim g\,.$$

Παρατήρηση 3:

$$f_Y(y) = \int_{x\in\mathbb{R}} \frac{1}{a}\mathcal{I}\left(y < a\ f(x),\ x\in S\right)dx = \int_{x\in S} \frac{1}{a}\mathcal{I}\left(y < a\ f(x)\right)dx$$
$$= \frac{1}{a}\int_{x\in S\cap\{x:\ y<a\ f(x)\}} dx = \frac{1}{a}\lambda\left(S\cap\left\{x:\ y<a\ f(x)\right\}\right).$$

Για παράδειγμα, εάν η f είναι μονοκόρυφη θα έχουμε ότι

$$\left\{x\in\mathbb{R}:\ f(x) > \frac{y}{a}\right\} = \left(x_1(y), x_2(y)\right),$$

όπου $x_1(y)$ και $x_2(y)$ είναι οι ρίζες της εξίσωσης $f(x) = \frac{y}{a}$, ενώ για το y θα πρέπει να ισχύει ότι $0 < y \le a\ f(x)$, έτσι έχουμε

$$f_Y(y) = \frac{1}{a}\int_{x\in S} \mathcal{I}\left(x_1(y) < x < x_2(y)\right)dx = \frac{1}{a}\lambda\left(S\cap\left(x_1(y), x_2(y)\right)\right),\ y < a\ f(x)$$

και επειδή $\left(x_1(y), x_2(y)\right)$ υποσύνολο του S, $f_Y(y) = \frac{1}{a}\left(x_2(y) - x_1(y)\right)$.

Υπό-συνθήκη πυκνότητα $f_{X|Y}(x\,|\,y) = \dfrac{\mathcal{I}\left(f(x) > \frac{y}{a}\right)}{\int_{x\in S}\mathcal{I}\left(f(x) > \frac{y}{a}\right)dx}$.

#==

Παράδειγμα

Εάν $g(x) = \mathcal{N}(x\,|\,0,1)$ και $f(x) = \frac{1}{2}e^{-|x|}$ για κάθε $x\in\mathbb{R}$, τότε $g(x) \le \sqrt{\frac{2e}{\pi}}f(x)$.

Αναπαριστούμε τις πυκνότητες g και f σαν διακριτές μίξεις με στηρίγματα $\mathbb{R}^{\le 0}$ και $\mathbb{R}^{>0}$

$$g(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\mathcal{I}(-\infty < x < \infty) = \left(\frac{1}{2}\right)\sqrt{\frac{2}{\pi}}e^{-x^2/2}\mathcal{I}(-\infty < x \le 0) + \left(\frac{1}{2}\right)\sqrt{\frac{2}{\pi}}e^{-x^2/2}\mathcal{I}(0 < x < \infty),$$
$$f(x) = \frac{1}{2}e^{-|x|}\mathcal{I}(-\infty < x < \infty) = \left(\frac{1}{2}\right)e^{x}\mathcal{I}(-\infty < x \le 0) + \left(\frac{1}{2}\right)e^{-x}\mathcal{I}(0 < x < \infty).$$

1. Για $-\infty < x \le 0$ έχουμε:

$$\frac{g(x)}{f(x)} = \sqrt{\frac{2}{\pi}}\exp\left(-\frac{1}{2}x^2 + x\right) = \sqrt{\frac{2e}{\pi}}\exp\left(-\frac{1}{2}(x-1)^2\right) \le \sqrt{\frac{2e}{\pi}}.$$

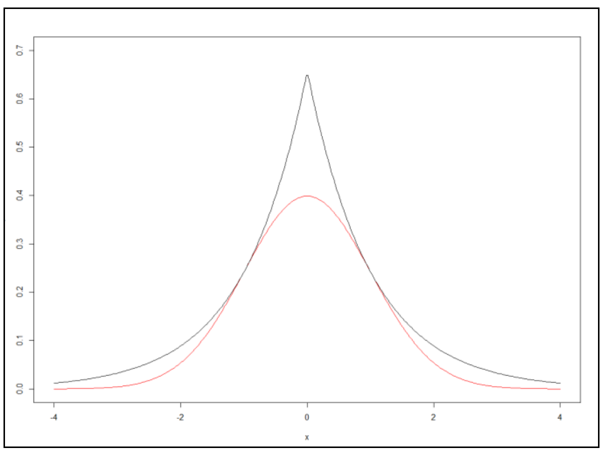
2. Για $0 < x < \infty$ έχουμε:

$$\frac{g(x)}{f(x)} = \sqrt{\frac{2}{\pi}}\exp\left(-\frac{1}{2}x^2 - x\right) = \sqrt{\frac{2e}{\pi}}\exp\left(-\frac{1}{2}(x^2+2x+2)\right) \le \sqrt{\frac{2e}{\pi}}$$

επειδή $x^2 + 2x + 2 > 0$ για κάθε πραγματικό x.

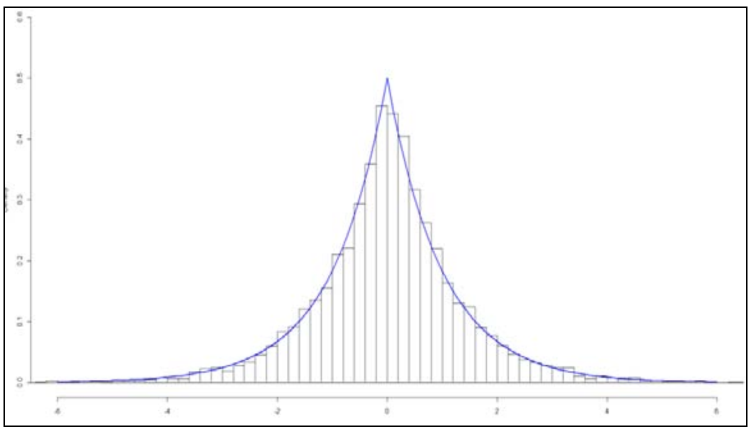
Έτσι για κάθε $x \in \mathbb{R}$, έχουμε $g(x) \le \sqrt{\frac{2e}{\pi}} f(x)$

```
f <- function(x) return(0.5*exp(-abs(x))) # GDE(x| p=0.5, λ1=1, λ2=1)=(1/2)exp(-|x|)
g <- function(x) return(dnorm(x,mean=0,sd=1))

curve(g(x), from=-4, to=4, n=300, xlab="x", ylab="y", lwd=1, col="red", xlim=c(-4,4),
      ylim=c(0,0.7))

curve((2*exp(1)/pi)^0.5*f(x), from=-4, to=4, n=300, xlab="x", ylab="y", lwd=1,
      col="black", xlim=c(-4,4), add=T)
```



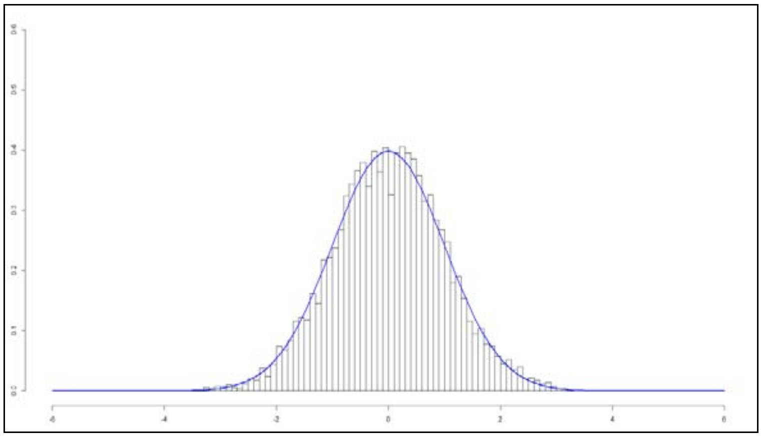

```
#==================================================================================

# sample from the generalized Double exponential distribution GDE(p, λ1, λ2).
# GDE(x| p=0.5, λ1=λ, λ2=λ)=(λ/2)exp(-λ|x|)
#----------------------------------------------------------------------------
SampleDGExp<-function(SS, p=0.5,  lambda1=1, lambda2=1){
  v<-c(1:SS)
  for(i in 1:SS){
    u1<-runif(1);u2<-runif(1)
    if(u1<p){
      v[i] <- (1/lambda1)*log(u2)
    }
    else{
      v[i] <-(-1/lambda2)*log(u2)
```

```
    }
  }
  return(v)
}

## The density of the generalized Double exponential distribution
#-------------------------------------------------------------------
GDEXP<-function(x, p=0.5,  lambda1=1, lambda2=1){
  v <- c()
  for(i in 1:length(x)) {
    if(x[i] < 0)
      branch <- p*lambda1*exp(lambda1*x[i])
    else
      branch <- (1-p)*lambda2*exp(-lambda2*x[i])
    v <- append(v, branch)
  }
  return(v)
}

set.seed(11)
v <- SampleDGExp(SS=1000, p=0.5,  lambda1=1, lambda2=1)
hist(v, breaks=100, freq=FALSE, ylim=c(0, 0.6), xlim=c(-6, 6),
     main="SS=10000, De(0.5, 1, 1)", col="red")
curve(GDEXP(x,p=0.5,lambda1=1,lambda2=1), from=-6 ,to=6, col="blue", lwd=2, add=TRUE)
```



```
#================================================================================
# Sample from the unit normal distribution N(0,1) using GDE(0.5,1,1)
#--------------------------------------------------------------------------------


# Define g as the unit normal distribution N(0,1).
#-------------------------------------------------------
g <- function(x) return(dnorm(x,mean=0,sd=1))

# Sample from the unit normal distribution N(0,1) using
# the Double Exponential distribution.
#-------------------------------------------------------
sampleUnitNormal <- function(SS){
  v<-c(1:SS)
  for(i in 1:SS){
    repeat{
    y <- SampleDGExp(SS=1, p=0.5,  lambda1=1, lambda2=1)
    #cat("y=",y,"\n")
    uprime <- (2*exp(1)/pi)^0.5*GDEXP(y,p=0.5,lambda1=1,lambda2=1)*runif(1)
    if (uprime < g(y)) break
    }
    v[i] <- y
  }
  return(v)
}
```

```
set.seed(11)
v<-sampleUnitNormal(5000)
hist(v, breaks=50, freq=FALSE, ylim=c(0, 0.6), xlim=c(-6, 6))
curve(g(x), from=-6 ,to=6, col="blue", lwd=2, add=TRUE)
```



#==

Το σχήμα δειγματοληψίας accept-reject για συνάρτηση μάζας πιθανότητας

Εάν X και Y διακριτές τ. μ. με κοινό στήριγμα το $S_X = S_Y = \mathbb{N}_0$, και συναρτήσεις μάζας πιθανότητας $p_X(\cdot) = P\{X = \cdot\}$ και $p_Y(\cdot) = P\{Y = \cdot\}$, αντιστοίχως, τότε ισχύει

$$\left.\begin{array}{l} y \sim p_Y(\cdot) \\ x \mid y \sim \mathcal{U}\left(0,\, a\, p_Y(y)\right) \end{array}\right\} \Rightarrow \left[\, y \mid x < p_X(y) \right] \sim p_X(\cdot)$$

Το ενδεχόμενο $\mathbb{A} = \{X < p_X(Y)\}$ αποδοχής έχει πιθανότητα

$$P(\mathbb{A}) = P\{X < p_X(Y)\} = \mathbb{E}\{P\{X < p_X(Y) \mid Y\}\} = \sum_{y \in \mathbb{N}_0} P\{X < p_X(Y) \mid Y = y\}\, p_Y(y)$$

$$= \sum_{y \in \mathbb{N}_0} P\{X < p_X(y) \mid Y = y\}\, p_Y(y) = \sum_{y \in \mathbb{N}_0} \left(\int_{u=0}^{p_X(y)} \frac{du}{a\, p_Y(y)} \right) p_Y(y) = \frac{1}{a} \sum_{y \in \mathbb{N}_0} p_X(y) = \frac{1}{a}$$

ενώ

$$P\left\{Y \le t \mid X < p_X(Y)\right\} = a\,P\left\{Y \le t, X < p_X(Y)\right\} = a\,\mathbb{E}\left\{P\left\{Y \le t, X < p_X(Y) \mid Y\right\}\right\}$$
$$= a\sum_{y\in\mathbb{N}_0} P\left\{Y \le t, X < p_X(Y) \mid Y = y\right\} p_Y(y) = a\sum_{y\in\mathbb{N}_0} P\left\{y \le t, X < p_X(y) \mid Y = y\right\} p_Y(y)$$
$$= a\sum_{y\le t} P\left\{X < p_X(y) \mid Y = y\right\} p_Y(y) = a\sum_{y\le t}\left(\int_{u=0}^{p_X(y)} \frac{du}{a\,p_Y(y)}\right) p_Y(y) = \sum_{y\le t} p_X(y) = P\left\{X \le t\right\}.$$

Δηλαδή έχουμε ότι

$P\left\{Y \le t \mid X < p_X(Y)\right\} = P\left\{X \le t\right\}$ οπότε και $\left[Y \mid X < p_X(Y)\right] \overset{\mathcal{D}}{=} X$.

Accept-Reject με άγνωστη σταθερά κανονικοποίησης

Δειγματοληψία από την πυκνότητα g με accept – reject, όταν το στήριγμα S της πυκνότητας g δεν έχει γενικά πεπερασμένο μήκος, και η σταθερά κανονικοποίησης της g είναι άγνωστη.

Έστω ότι

$$g(x) \propto \tilde{g}(x) \le a\, f(x),\ \forall\, x \in S \text{, και } g(x) = C_g\, \tilde{g}(x),$$

για κάποια γενικά άγνωστη θετική σταθερά C_g. Όπου $\frac{1}{C_g} = \int_S \tilde{g}(x)dx$, και $a C_g \ge 1$.

Το ολοκλήρωμα $\int_S \tilde{g}(x)dx$ μπορεί να είναι δύσκολο να υπολογιστεί (ειδικότερα όταν η διάσταση του x είναι μεγάλη). Τότε

$$\begin{cases} x \sim f_X(\cdot) = f(\cdot) \\ y|x \sim f_{Y|X}(\cdot|x) = \mathcal{U}\left(\cdot|0, a\, f(x)\right) \end{cases} \Rightarrow\ \left(x \mid y < \tilde{g}(x)\right) \sim g(\cdot) \qquad (AR3)$$

Απόδειξη

Θέτουμε $f_{X,Y}(x,y) \propto \mathcal{I}\left(y < a\, f(x),\ x \in S\right)$ (όπως και στην προηγούμενη περίπτωση όπου $g(x) \le a\, f(x)$) από όπου και

$$f_{X,Y}(x,y) = \frac{1}{a}\mathcal{I}\left(0 < y < a\, f(x), x \in S\right) = \begin{cases} \dfrac{1}{a} & x \in S,\ 0 < y < a\, f(x) \\ 0 & \alpha\lambda\lambda o\acute{\upsilon} \end{cases},$$

με

$$f_X(x) = \int\limits_{y\in\mathbb{R}} \frac{1}{a}\mathcal{I}\left(0<y<a\ f(x),\, x\in S\right)dy = \frac{1}{a}\int\limits_{y=0}^{a\,f(x)} dy = f(x),\ \ x\in S\ ,$$

$$f_{Y|X}(y\,|\,x) = \frac{\frac{1}{a}\mathcal{I}\left(0<y<a\ f(x),\, x\in S\right)}{f(x)}$$

$$= \left\{\frac{1}{a\ f(x)}\mathcal{I}\left(0<y<a\ f(x)\right)\right\}\mathcal{I}(x\in S) = \mathcal{U}\left(y|0,a\ f(x)\right),\ \forall\ x\in S\ .$$

Η πιθανότητα αποδοχής είναι:

$$P\{Y<\tilde{g}(X)\} = \iint\limits_{y<\tilde{g}(x)} \frac{1}{a}\mathcal{I}\left(0<y<a\ f(x),\, x\in S\right)dy\,dx$$

$= \frac{1}{a}\int\limits_{x\in S}\int\limits_{y=0}^{\tilde{g}(x)} dy\,dx = \frac{1}{a}\int\limits_{x\in S}\tilde{g}(x)\,dx = \frac{1}{a\,C_g}$, εφόσον $\tilde{g}(x)\le a\ f(x)$, για $\forall\ x\in S$.

Έτσι τελικά παίρνουμε

$$P\{Y<\tilde{g}(X)\} = \frac{1}{a\,C_g}\,.$$

Δηλαδή σε αυτή την περίπτωση *δεν γνωρίζουμε ακριβώς την πιθανότητα αποδοχής*. Γνωρίζουμε όμως ότι $P\{Y<\tilde{g}(X)\}\propto\frac{1}{a}$.

$$P\{X\le x\,|\,Y<\tilde{g}(X)\} = \frac{P\{X\le x, Y<\tilde{g}(X)\}}{P\{Y<\tilde{g}(X)\}}$$

$$= a\,C_g\int\limits_{u=-\infty}^{x}\int\limits_{v=0}^{\tilde{g}(u)}\frac{1}{a}\mathcal{I}\left(0<v<a\ f(u),\, u\in S\right)dv\,du$$

$= \int\limits_{u=-\infty}^{x} C_g\tilde{g}(u)\,du = \int\limits_{u=-\infty}^{x} g(u)\,du = G(x)$, εφόσον $\tilde{g}(u)\le a\ f(u)$, για $\forall\ u\in S$,

που δίνει το σχήμα δειγματοληψίας $(AR3)$.

Δειγματοληψία με majorization

Θέλουμε να κάνουμε δειγματοληψία από την πυκνότητα $g(x)$ $g(x)$ και για κάθε $x \in S$ ισχύει

$g(x) < a\, f(x)$ με $f(x)$ πυκνότητα (που δίνει $a > 1$). Ισοδύναμα $0 < \dfrac{g(x)}{a\, f(x)} < 1$. Εάν θέσουμε $h(x) = \dfrac{g(x)}{a\, f(x)}$, έχουμε εκφράσει την πυκνότητα σαν:

$$g(x) = a\, h(x)\, f(x) \text{ για } a > 1 \text{ και κάποια συνάρτηση } 0 < h(x) < 1,\ \forall\, x \in S.$$

Πρόταση

Εάν για την πυκνότητα $g(x)$ ισχύει $g(x) < a\, f(x)$ με $f(x)$ πυκνότητα, για $\forall\, x \in S$, όπου S το κοινό στήριγμα των πυκνοτήτων g και f. Τότε ισοδύναμα έχουμε ότι $g(x) = ah(x)\, f(x)$ με $0 < h(x) < 1$ για κάθε $x \in S$ και $h(x) = \dfrac{g(x)}{a\, f(x)}$, που οδηγεί στο σχήμα δειγματοληψίας

$$\left.\begin{cases} x \sim f \\ y \sim \mathcal{U}(0,1) \\ \mathbb{A} = \{y < h(x)\} \end{cases} \Rightarrow \quad (x \,|\, \mathbb{A}) \sim g \right\}. \qquad (AR4)$$

Πράγματι, εάν ορίσουμε την από κοινού $f_{X,Y}$

$$f_{X,Y}(x,y) = f(x)\, \mathcal{U}(y \,|\, 0,1) = \begin{Bmatrix} f(x), & 0 < y < 1,\ x \in S \\ 0, & elsewhere \end{Bmatrix},$$

οι τ.μ. X και Y είναι ανεξάρτητες και

$$P(\mathbb{A}) = P\{Y < h(X)\} = \int_{x \in S} \int_{y=0}^{h(x)} f(x)\, dy\, dx \text{ επειδή } 0 < h(x) < 1$$

$$= \int_{x \in S} f(x) h(x)\, dx = a^{-1} \int_{x \in S} g(x)\, dx = a^{-1}.$$

Θέτοντας σαν αθροιστική συνάρτηση κατανομής της g την G, δεν είναι δύσκολο να δείξουμε ότι

$$G(x) = P\left\{X \le x \mid Y < h(X)\right\}.$$

Πράγματι

$$P\left\{X \le x \mid Y < h(X)\right\} = a\,P\left\{X \le x, Y < h(X)\right\} = a \int_{s=-\infty}^{x} \int_{t=0}^{h(s)} f(s)\,dt\,ds$$

$$= a \int_{s=-\infty}^{x} \frac{1}{a} g(s)\,ds = \int_{s=-\infty}^{x} g(s)\,ds = G(x).$$

Δηλαδή έχουμε $\left[X \mid Y < h(X)\right] \sim g$ με πιθανότητα αποδοχής $P\left\{Y < h(X)\right\} = a^{-1}$.

<u>Πρόταση (majorization με άγνωστη σταθερά κανονικοποίησης)</u>

Εάν για την πυκνότητα $g(x)$ ισχύει $g(x) \propto \tilde{g}(x) < a\, f(x)$ με $f(x)$ πυκνότητα, για $\forall\, x \in S$, όπου S το κοινό στήριγμα των πυκνοτήτων g και f. Τότε ισοδύναμα έχουμε ότι $\tilde{g}(x) = a\tilde{h}(x)\, f(x)$ με $0 < \tilde{h}(x) < 1$ για κάθε $x \in S$ και $\tilde{h}(x) = \dfrac{\tilde{g}(x)}{a\, f(x)}$, που οδηγεί στο σχήμα δειγματοληψίας

$$\left\{\begin{array}{l} x \sim f \\ y \sim \mathcal{U}(0,1) \\ \mathbb{A} = \left\{y < \tilde{h}(x)\right\} \end{array}\right. \quad \Rightarrow \quad (x \mid \mathbb{A}) \sim g. \qquad (AR5)$$

Πράγματι εάν ορίσουμε την από κοινού $f_{X,Y}$

$$f_{X,Y}(x,y) = f(x)\,\mathcal{U}(y \mid 0,1) = \left\{\begin{array}{ll} f(x), & 0 < y < 1,\ x \in S \\ 0, & elsewhere \end{array}\right\},$$

οι τ.μ. x και y είναι ανεξάρτητες και

$$P(\mathbb{A}) = P\left\{Y < \tilde{h}(X)\right\} = \int_{x \in S} \int_{y=0}^{\tilde{h}(x)} f(x)\,dy\,dx \text{ επειδή } 0 < \tilde{h}(x) < 1$$

$$= \int_{x\in S} f(x)\tilde{h}(x)dx = a^{-1}\int_{x\in S}\tilde{g}(x)dx = \left(C_g a\right)^{-1}.$$

Θέτοντας σαν αθροιστική συνάρτηση κατανομής της g την G, δεν είναι δύσκολο να δείξουμε ότι

$$G(x) = P\left\{X \le x \mid Y < h(X)\right\}.$$

Πράγματι

$$P\left\{X \le x \mid Y < \tilde{h}(X)\right\} = a\,P\left\{X \le x, Y < \tilde{h}(X)\right\} = a\int_{s=-\infty}^{x}\int_{t=0}^{\tilde{h}(s)} f(s)\,dt\,ds$$

$$= C_g a\int_{s=-\infty}^{x}\frac{1}{a}\tilde{g}(s)ds = \int_{s=-\infty}^{x} C_g\tilde{g}(s)ds = \int_{s=-\infty}^{x} g(s)ds = G(x).$$

Δηλαδή έχουμε $\left[X \mid Y < \tilde{h}(X)\right] \sim g$ με πιθανότητα αποδοχής $P\left\{Y < \tilde{h}(X)\right\} = \left(C_g a\right)^{-1}$.

Adaptive Rejection Sampling

(με majorization και άγνωστη σταθερά κανονικοποίησης)

Έχουμε διαπιστώσει ότι εάν θέλουμε να κάνουμε δειγματοληψία από την $g(x)$ και ισχύει:

$g(x) \propto \tilde{g}(x) < a\, f(x)$ με $f(x)$ πυκνότητα, για $\forall\, x \in S$,

όπου S το κοινό στήριγμα των πυκνοτήτων g και f, τότε θέτοντας

$\tilde{g}(x) = a\tilde{h}(x)\, f(x)$ με $0 < \tilde{h}(x) < 1,\ \forall\, x \in S$ και $\tilde{h}(x) = \dfrac{\tilde{g}(x)}{a\, f(x)}$,

παίρνουμε το γνωστό σχήμα δειγματοληψίας $(AR5)$

$\left.\begin{cases} x \sim f \\ y \sim \mathcal{U}(0,1) \ \Rightarrow\ \left(x \mid y < \tilde{h}(x)\right) \sim g \end{cases}\right\}$, με πιθανότητα αποδοχής $P\left\{y < \tilde{h}(x)\right\} = \left(C_g\, a\right)^{-1}$.

Όταν η $\tilde{g}(x)$ είναι **logconcave** δηλαδή έχουμε ότι η $\log\tilde{g}(x)$ **είναι μονοκόρυφη**, ή ισοδύναμα ότι

$r(x) = \log \tilde{g}(x)$ και $r''(x) < 0,\ \forall\ x \in S$,

μπορούμε να προσδιορίσουμε κατά τόπους γραμμική (piecewise linear) συνάρτηση $u_k(x)$ (στο log – space) με την ιδιότητα:

$$r(x) = \log\left(\tilde{g}(x)\right) \le u_k(x),\ \forall\ x \in S \Leftrightarrow \tilde{g}(x) = \exp\left(r(x)\right) \le \exp\left(u_k(x)\right),\ \forall\ x \in S\ ,$$

ή ισοδύναμα ότι η majorizing συνάρτηση είναι

$$\tilde{g}(x) = \exp\left(r(x)\right) \le \underbrace{\left(\frac{1}{C_{u_k}}\right)}_{a}\underbrace{\left(C_k \exp\left(u_k(x)\right)\right)}_{f(x)},\ \forall\ x \in S \text{ με } C_{u_k}^{-1} = \int_{t \in S} \exp\left(u_k(t)\right) dt\ ,$$

τότε

$$\tilde{h}(x) = \frac{\tilde{g}(x)}{a\, f(x)} = \frac{\exp\left(r(x)\right)}{\exp\left(u_k(x)\right)} = \exp\left(r(x) - u_k(x)\right).$$

Από τα προηγούμενα και το σχήμα δειγματοληψίας $(AR5)$ παίρνουμε

$$\left\{\begin{array}{l} x \sim C_{u_k} \exp\left(u_k(x)\right) \\ y \sim \mathcal{U}(0,1)\ \Rightarrow\ \left(x \mid y < \exp\left(r(x) - u_k(x)\right)\right) \sim g \end{array}\right\},$$

με πιθανότητα αποδοχής ανάλογη του C_{u_k} εφόσον

$$P\left\{y < \tilde{h}(x)\right\} = \left(C_g a\right)^{-1} = \frac{C_{u_k}}{C_g} \propto C_{u_k}\ .$$

Adaptive Rejection Sampling (with majorization, minorization and unknown normalizing constant)

Εάν η $\tilde{g}(x) = \exp\left(r(x)\right)$ είναι πολύπλοκη (για παράδειγμα περιέχει την μεταβλητή της σε gamma συνάρτηση) για να επιταχύνουμε την διαδικασία μπορούμε να ορίσουμε και minorizing συνάρτηση $l_k(x)$ στο log – space, δηλαδή

$$l_k(x) \le r(x) = \log\left(\tilde{g}(x)\right) \le u_k(x),\ \forall\ x \in S$$

$$\Leftrightarrow \exp\left(l_k(x)\right) \le \tilde{g}(x) \le \exp\left(u_k(x)\right),\ \forall\ x \in S\ .$$

Τότε θα έχουμε

$$\exp\left(l_k(x)\right) \le \tilde{g}(x)$$

$$\Leftrightarrow \frac{\exp\left(l_k(x)\right)}{a\,f(x)} = \frac{\exp\left(l_k(x)\right)}{\exp\left(u_k(x)\right)} = \exp\left(l_k(x) - u_k(x)\right) \le \frac{\tilde{g}(x)}{a\,f(x)} = \tilde{h}(x) = \exp\left(r(x) - u_k(x)\right)$$

$$\Leftrightarrow \left\{ y < \frac{\exp\left(l_k(x)\right)}{a\,f(x)} \right\} \subset \left\{ y < \frac{\exp\left(u_k(x)\right)}{a\,f(x)} \right\}$$

Δηλαδή το ενδεχόμενο αποδοχής έχει σαν υποσύνολο του το ενδεχόμενο $\left\{ y < \frac{\exp\left(l_k(x)\right)}{a\,f(x)} \right\}$ που σημαίνει ότι εάν $x \sim C_{u_k} \exp\left(u_k(x)\right)$ και $y \sim \mathcal{U}(0,1)$, και ισχύει ότι $y < \frac{\exp\left(l_k(x)\right)}{a\,f(x)}$ τότε $y \sim g$. Εάν όμως έχουμε ότι $y \ge \frac{\exp\left(l_k(x)\right)}{a\,f(x)}$ θα πρέπει να ελέγξουμε εάν ισχύει $y < \tilde{h}(x)$. Στην καταφατική περίπτωση θα έχουμε $y \sim g$, ενώ εάν $y \ge \tilde{h}(x)$ το y απορρίπτεται.

Το σχήμα δειγματοληψίας με squeeze γίνεται:

$$\left\{ \begin{array}{l} x \sim C_{u_k} \exp\left(u_k(x)\right) \\ y \sim \mathcal{U}(0,1) \;\Rightarrow\; \mathit{If}\;\; \left(y < \exp\left(l_k(x) - u_k(x)\right)\right) \mathit{then}\;\; x \sim g \\ \qquad\qquad\qquad\quad \mathit{else\ if}\;\; \left(y < \exp\left(r(x) - u_k(x)\right)\right) \mathit{then}\;\; x \sim g \\ \qquad\qquad\qquad\quad \mathit{else}\;\; \mathit{reject}\;\; x. \end{array} \right\}$$

Παρατηρήστε ότι η πιθανότητα αποδοχής από τον προκαταρκτικό έλεγχο, είναι

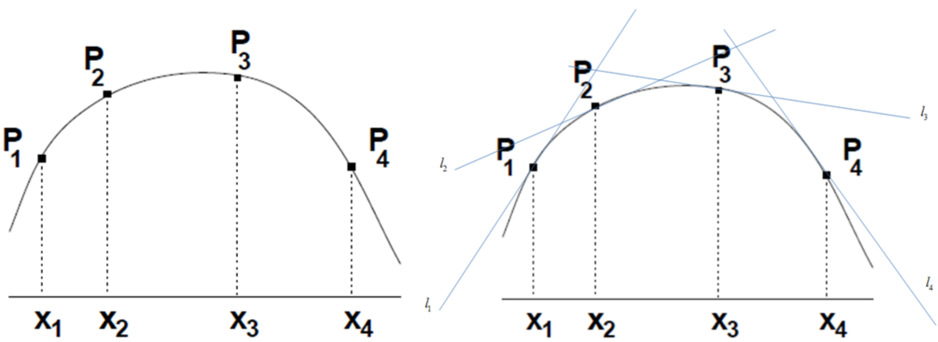
$$P\left\{ y < \exp\left(l_k(x) - u_k(x)\right) \right\} = P\left\{ y < \frac{\exp\left(l_k(x)\right)}{\exp\left(u_k(x)\right)} \right\} = P\left\{ y < \frac{\exp\left(l_k(x)\right)}{a\,f(x)} \right\}$$

$$= \int_{x\in S} \int_{y=0}^{\frac{\exp\left(l_k(x)\right)}{a\,f(x)}} f(x)\,dy\,dx = \frac{1}{a} \int_{x\in S} \exp\left(l_k(x)\right)dx = \frac{C_{u_k}}{C_{l_k}}.$$

Κατασκευή της enveloping συνάρτησης $u_k(x)$

Έστω $S_k = \{x_0, x_1, \dots, x_k, x_{k+1}\} \subset S$ με $x_i < x_{i+1}$ για $0 \le i \le k$ και $x_0 = \inf S$, και $x_{k+1} = \sup S$

$$\mathbb{P}_k = \left\{\left(x_1, \log\left(\tilde{g}\left(x_1\right)\right)\right), \left(x_2, \log\left(\tilde{g}\left(x_2\right)\right)\right), \dots, \left(x_k, \log\left(\tilde{g}\left(x_k\right)\right)\right)\right\} = \{P_1, P_2, \dots, P_k\}$$



Τότε

$$\begin{aligned} u_k(x) &= T(x;x_1)1\left(x_{0,1} < x \le x_{1,2}\right) + T(x;x_2)1\left(x_{1,2} < x \le x_{2,3}\right) + \cdots + T(x;x_i)1\left(x_{i-1,\,i} < x \le x_{i,i+1}\right) \\ &\quad + \cdots + T(x;x_{k-1})1\left(x_{k-2,\,k-1} < x \le x_{k-1,\,k}\right) + T(x;x_k)1\left(x_{k-1,\,k} < x \le x_{k,\,k+1}\right) \\ &= \sum_{i=1}^{k} T(x;x_i)1\left(x_{i-1,\,i} < x \le x_{i,i+1}\right) \end{aligned}$$

όπου $T(x;x_i) = r_i + (x - x_i)r_i'$ με $r_i = r(x_i)$ και $r' = r'(x_i)$.

Εάν η άγνωστη πυκνότητα έχει στήριγμα το , τότε $x_{0,1} = x_1 - \frac{r_1}{r_i'}$ είναι η $x-$συντεταγμένη της τομής της (l_1) με τον $x-$άξονα όταν $x_1 - \frac{r_1}{r_i'} > 0$ και $x_{0,1} = 0$ όταν $x_1 - \frac{r_1}{r_i'} \le 0$.

$$x_{0,1} = \begin{Bmatrix} x_1 - \frac{r_1}{r_i'}, & x_1 - \frac{r_1}{r_i'} > 0 \\ 0, & x_1 - \frac{r_1}{r_i'} < 0 \end{Bmatrix}$$

Παράδειγμα

Να γίνει δειγματοληψία με majorization από την πυκνότητα $g(x) = Be(x|4,3)$.

Εδώ έχουμε $g(x) = 60x^3(1-x)^2 \cdot 1(0 < x < 1)$, $S = (0,1)$ και
$g'(x) = 60x^2(5x^2 - 8x + 3) = 60x^2(x-0.6)(x-1)$. Επειδή $g'(0.6) = 0$ με $g''(0.6) < 0$ έχουμε
$\max_{x \in S} g(x) = g(0.6)$.

Έτσι η g έχει την majorization αναπαράσταση:

$$g(x) = g(0.6)\left\{\frac{Be(x|4,3)}{g(0.6)}\right\}\mathcal{I}(0 < x < 1) \text{ με } 0 < \frac{Be(x|4,3)}{g(0.6)} < 1 \text{ για κάθε } x \in (0,1),$$

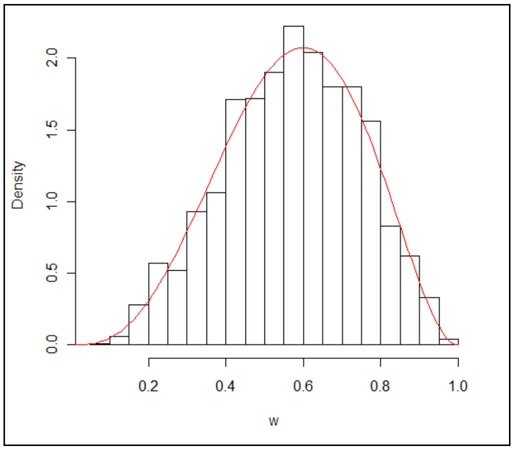
με $a = g(0.6) > 1$, $h(x) = \dfrac{Be(x|4,3)}{g(0.6)}$ και $f(x) = \mathcal{U}(x|0,1)$.

Το σχήμα δειγματοληψίας λοιπόν είναι: $\left\{\begin{array}{ll} x \sim f = \mathcal{U}(0,1) & \\ y \sim \mathcal{U}(0,1) & \Rightarrow \quad x|\mathbb{A} \sim g \\ \mathbb{A} = \{y < h(x)\} & \end{array}\right\}$.

#R – Script για δειγματοληψία με majorization από την $beta(4,3)$.

```
mydensity <- function(x) return(60*x^3*(1-x)^2)


sampledensity <- function(SS=100) {
  sample <- NULL
  for(i in 1:SS) {
      repeat {
          y<-runif(1); u<-runif(1)
          if(u <= mydensity(y)/mydensity(0.6)) { break }
      }
      sample <- c(sample, y)
  }
  return(sample)
```

```
}
# Experiment
w <- sampledensity(SS=2000)
hist(w, freq=FALSE, breaks=30)
curve(mydensity(x), add=TRUE, xlim=c(0, 1), col="red")
```



Άσκηση

Να γίνει δειγματοληψία με majorization από την $g(x) = Be(x \mid p,q)$ όταν $p > 1$ και $q > 1$ (για $p > 1$ και $q > 1$ το mode της beta ορίζεται).

Δεν είναι δύσκολο να δείξουμε ότι το mode της g υπάρχει και είναι $x_{Mode} = \dfrac{p-1}{p+q-2}$, όταν $p > 1$ και $q > 1$. Τότε η f έχει την majorization αναπαράσταση

$$g(x) = g(x_{Mode})\left\{\frac{Be(x \mid p,q)}{g(x_{Mode})}\right\}1(0 < x < 1) \text{ με } 0 < \frac{Be(x \mid p,q)}{g(x_{Mode})} < 1 \text{ για κάθε } x \in (0,1), \text{ ή}$$

ισοδύναμα

$$g(x) = ah(x)f(x) \text{ με } a = g(x_{Mode}) > 1,\ h(x) = \frac{Be(x \mid p,q)}{g(x_{Mode})} \text{ και } f(x) = \mathcal{U}(x \mid 0,1).$$

Η δειγματοληψία γίνεται ακριβώς όπως και στο προηγούμενο παράδειγμα. Όταν $0 < p < 1$ και $0 < q < 1$, για να κάνουμε δειγματοληψία από την beta κατανομή, μπορούμε να χρησιμοποιήσουμε (όπως θα δούμε στα επόμενα) την gamma κατανομή. Στο επόμενο παράδειγμα δείχνουμε πώς να κάνουμε δειγματοληψία από την gamma με majorization

Δειγματοληψία από την gamma κατανομή

Να γίνει AR δειγματοληψία από την οικογένεια πυκνοτήτων. $f = Ga(shape = a, rate = \lambda)$ όπου $scale = rate^{-1}$.

1. Πρώτα θα κάνουμε δειγματοληψία από την **fractional gamma** κατανομή $X \sim Ga(shape = \alpha, rate = 1)$ με $0 < \alpha = \{a\} = a - \lfloor a \rfloor < 1$. Όπου $\lfloor a \rfloor$ η συνάρτηση floor στο a, δηλαδή $\alpha = \{a\}$ είναι το δεκαδικό κομμάτι του a.

Τότε $g(x) \propto x^{\alpha-1}e^{-x}$, και παρατηρούμε ότι επειδή:

- για $0 < x < 1$ έχουμε: $x^{\alpha-1}e^{-x} < x^{\alpha-1}\sup\limits_{0<x<1} e^{-x} = x^{\alpha-1}$,
- και για $x \geq 1$ και επειδή $0 < \alpha < 1$ έχουμε: $x \geq 1 \Leftrightarrow x^{\alpha-1} \leq 1 \Leftrightarrow x^{\alpha-1}e^{-x} \leq e^{-x}$,

η πυκνότητα $g(x) = Ga(\alpha,1)$ φράσσεται με τον εξής τρόπο:

$$g(x) = \frac{1}{\Gamma(\alpha)}x^{\alpha-1}e^{-x} < \frac{1}{\Gamma(\alpha)}\begin{Bmatrix} x^{\alpha-1}, 0 < x < 1 \\ e^{-x}, \quad x \geq 1 \end{Bmatrix},$$

$$= \frac{1}{\Gamma(\alpha)}\left[x^{\alpha-1}\,1(0 < x < 1) + e^{-x}\,1(x \geq 1)\right].$$

Θέλουμε να βρούμε πυκνότητα f, τέτοια ώστε

$$f(x) \propto \frac{1}{\Gamma(\alpha)}\left[x^{\alpha-1}1(0 < x < 1) + e^{-x}1(x \geq 1)\right].$$

Εάν λοιπόν $K > 0$ είναι η σταθερά κανονικοποίησης θα έχουμε:

$$f(x) = K\frac{1}{\Gamma(\alpha)}\left[x^{\alpha-1}1(0 < x < 1) + e^{-x}1(x \geq 1)\right],$$

ή ισοδύναμα

$$K^{-1} = \frac{1}{\Gamma(\alpha)}\int_{\mathbb{R}^+}\left\{x^{\alpha-1}1(0 < x < 1) + e^{-x}1(x \geq 1)\right\}dx,$$

από όπου και $K = \dfrac{\alpha\, e\,\Gamma(\alpha)}{\alpha + e}$, που δίνει

$$f(x) = \frac{\alpha\, e}{\alpha + e}\left[x^{\alpha-1}1(0 < x < 1) + e^{-x}1(x \geq 1)\right].$$

Για το αντίστοιχο majorization σχήμα δειγματοληψίας έχουμε $f(x) < c\,h(x)$, ή ότι:

$$f(x) = c\,g(x)h(x) \text{ για κάθε } x \in S = (0,\infty) \text{ όπου } c = K^{-1}$$

$$h(x) = \frac{x^{\alpha-1}e^{-x}}{x^{\alpha-1}1(0<x<1)+e^{-x}1(x\geq 1)} = e^{-x}1(0<x<1)+x^{\alpha-1}1(x\geq 1) \text{ με } 0<g(x)<1 \text{ και}$$

$$f(x) = \frac{\alpha e}{\alpha+e}\left[x^{\alpha-1}1(0<x<1)+e^{-x}1(x\geq 1)\right],$$

από όπου

$$\begin{Bmatrix} x \sim f \\ y \sim \mathcal{U}(0,1) \end{Bmatrix} \Rightarrow \left[x \,|\, y < h(x)\right] \sim g\,.$$

Για να κάνουμε δειγματοληψία από την πυκνότητα f, την γράφουμε σαν μίξη των πυκνοτήτων:

$$f_1(x) = \alpha x^{\alpha-1}1(0<x<1) = Be(x\,|\,\alpha,1),$$

$$f_2(x) = e^{1-x}1(x\geq 1) \propto Exp(x\,|1)1(x\geq 1),$$

δηλαδή $f(x) = \frac{e}{\alpha+e}f_1(x)+\frac{\alpha}{\alpha+e}f_2(x)$, με αθροιστική συνάρτηση κατανομής

$$F(x) = \frac{e}{\alpha+e}\int_{-\infty}^{x} f_1(u)du + \frac{\alpha}{\alpha+e}\int_{-\infty}^{x} f_2(u)du = \frac{e}{\alpha+e}\cdot F_1(x)+\frac{\alpha}{\alpha+e}\cdot F_2(x)$$

$$= \begin{Bmatrix} 0 & x \leq 0 \\ \frac{e}{\alpha+e}x^{\alpha} & 0<x<1 \\ \frac{e}{\alpha+e}+\frac{\alpha}{\alpha+e}\left(1-e^{1-x}\right) & x \geq 1 \end{Bmatrix},$$

όπου

$$F_1(x) = \begin{Bmatrix} 0 & x\leq 0 \\ x^{\alpha} & 0<x<1 \\ 1 & x\geq 1 \end{Bmatrix} \text{ και } F_2(x) = \begin{Bmatrix} 0 & x\leq 1 \\ 1-e^{1-x} & x\geq 1 \end{Bmatrix},$$

ενώ έχουμε ότι

$F_1^{-1}(y) = y^{1/\alpha}$ και $F_2^{-1}(y) = 1 - \log(1-y)$.

Τελικά λοιπόν παίρνουμε το σχήμα:

$$\left\{ \begin{array}{l} y \sim f(\cdot) = \dfrac{e}{\alpha+e} f_1(\cdot) + \dfrac{\alpha}{\alpha+e} f_2(\cdot) \\ u \sim \mathcal{U}(\cdot \mid 0,1) \end{array} \right\}$$

$$\Rightarrow \left[x \mid y < e^{-x} 1(0 < x < 1) + x^{\alpha-1} 1(x \geq 1) \right] \sim Ga(\cdot \mid \alpha, 1)$$

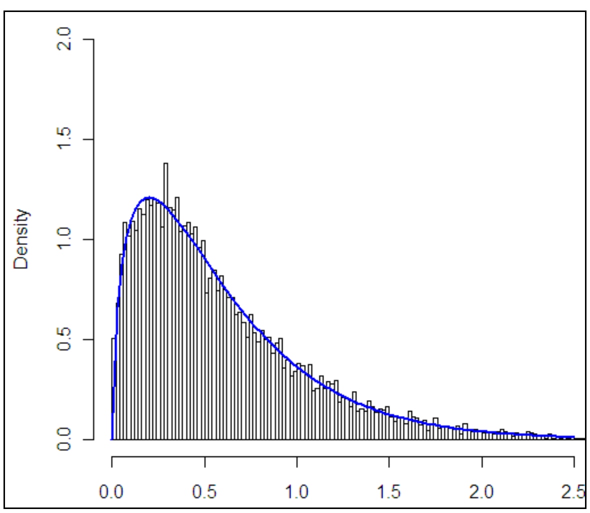
```
# sample (explicit looping) from gamma(shape=a, rate=1) when 0<a<1.
gammafractional <- function(SS=1000, a=0.5) {
   sample <- NULL
   for(i in 1:SS){
      while(TRUE){
         u <- runif(3); e <- exp(1)
         if(u[1] < e/(a+e)) x=u[2]^(1/a) else x=1-log(u[2])
         if(((x < 1) & u[3] < exp(-x)) | ((x >= 1) & u[3] < x^(a-1))) break
      }
      sample <- c(sample, x)
    }
    return(v)
}
```

2. Για την γενική περίπτωση, δηλαδή για να κάνουμε δειγματοληψία από την πυκνότητα $g = Ga(a, \lambda)$ για γενικά $a > 0$ και $\lambda > 0$, παρατηρούμε ότι εάν $n = \lfloor a \rfloor$ και $\alpha = a - n$ τότε

$$\left\{ \begin{array}{l} X \sim Ga(\alpha, 1) \Rightarrow Y = \lambda^{-1} X \sim Ga(\alpha, \lambda) \\ Z_i \overset{iid}{\sim} Ga(1, \lambda) = Exp(\lambda) \Rightarrow Z = \sum_{i=1}^{n} Z_i \sim Ga(n, \lambda) \end{array} \right\} \Rightarrow Y + Z \sim Ga(\alpha + n, \lambda).$$

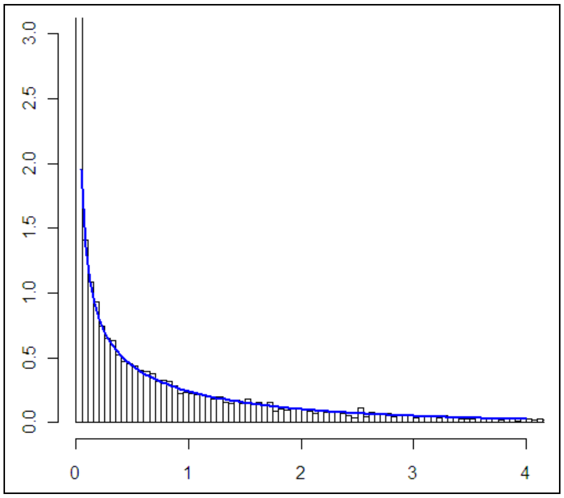
Sample a SS vector of $Exp(\lambda)$ random deviates.

```
myexpsample <- function(SS=100, lambda=2) return((-1/lambda)*log(runif(SS)))

# Sample from gamma(shape=a, rate=b) for a general a>0, b>0.
mygammasample <- function(SS=100, a=1.5, lambda=2) {
    n <- floor(a); v <- gammafractional(SS=SS, a=a-n)/lambda
    if(n==0)
         return(v)
      else {
         w<-rep(0, SS)
         for(i in 1:n) w <- w + myexpsample (SS=SS, lambda=lambda)
         return(v + w)
         }
}
# Experiment I
w <- mygammasample(SS=10000, a=1.5, lambda=2.5)
hist(w, breaks=200, freq=FALSE, ylim=c(0, 2), xlim=c(0, 2.5))
curve(dgamma(x, shape=1.5, scale=1/2.5), from=0, to=2.5, col="blue",
                lwd=2.2, add=TRUE)
```




```
# Experiment II
```

```
w <- mygammasample(SS=10000, a=0.5, lambda=0.5)
hist(w, breaks=200, freq=FALSE, ylim=c(0, 3), xlim=c(0, 4))
curve(dgamma(x, shape=0.5, scale=1/0.5), from=0, to=4, col="blue",
      lwd=2.2, add=TRUE)
```



Δειγματοληψία από την beta κατανομή $g = beta\left(shape1 = a_1, shape2 = a_2\right)$ **για γενικά** $a_1 > 0$ **και** $a_2 > 0$ **με την χρήση της gamma κατανομής**.

Θα δείξουμε ότι εάν $x_i \overset{ind}{\sim} Ga\left(a_i, b\right)$ για $i = 1,2$ τότε

$$\left\{\begin{array}{l} u = x_1 + x_2 \sim Ga\left(a_1 + a_2, b\right) \\ v = \dfrac{x_1}{x_1 + x_2} \sim Be\left(a_1, a_2\right) \end{array}\right\}$$ και τα u και v είναι ανεξάρτητα.

Πράγματι

$$T : \left\{\begin{array}{l} u = x_1 + x_2 \\ v = \dfrac{x_1}{x_1 + x_2} \end{array}\right\} \Leftrightarrow T^{-1} : \left\{\begin{array}{l} x_1 = uv \\ x_2 = u\left(1 - v\right) \end{array}\right\}$$ ενώ $$\left\{\begin{array}{l} 0 < u = x_1 + x_2 < \infty \\ 0 < v = \dfrac{x_1}{x_1 + x_2} < 1 \end{array}\right\}$$

$$Jac\left(T^{-1}\right) = \begin{vmatrix} v & u \\ 1 - v & -u \end{vmatrix} = u \Rightarrow \pi\left(x_1, x_2\right) \propto \left(x_1^{a_1 - 1} e^{-bx_1}\right)\left(x_2^{a_2 - 1} e^{-bx_2}\right) = x_1^{a_1 - 1} x_2^{a_2 - 1} e^{-b\left(x_1 + x_2\right)}$$

από όπου και

$$f(u,v) \propto (uv)^{a_1-1}\left(u(1-v)\right)^{a_2-1} e^{-bu} \,|\,Jac\left(T^{-1}\right)| = \left(u^{a_1+a_2-1}e^{-bu}\right)\left(v^{a_1-1}(1-v)^{a_2-1}\right)$$

$$\propto Ga\left(u|a_1+a_2,\, b\right) Be\left(v|a_1,\, a_2\right).$$

Δηλαδή για να πάρουμε ένα δείγμα από την $v \sim Be(a_1, a_2)$ για οποιοδήποτε $a_1 > 0$ και $a_2 > 0$ αρκεί να προσομοιώσουμε ανεξάρτητα $x_1 \sim Ga(a_1, b)$ και $x_2 \sim Ga(a_2, b)$, (όπου συνήθως θέτουμε το rate $b=1$). Τότε $v = \dfrac{x_1}{x_1+x_2} \sim Be(a_1, a_2)$.

Δειγματοληψία από $N\left(x\,|\,\mu,\sigma^2\right)$ με majorization

Αρχικά θέτουμε $g = N(0,1)$ και αναπαριστούμε την f σαν διακριτή μίξη

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\mathcal{I}(-\infty < x < \infty) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\mathcal{I}(x \le 0) + \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\mathcal{I}(x > 0)$$

$$= \frac{1}{2}\underbrace{\sqrt{\frac{2}{\pi}}e^{-x^2/2}\mathcal{I}(x \le 0)}_{\mathcal{N}^-(x|0,1)} + \frac{1}{2}\underbrace{\sqrt{\frac{2}{\pi}}e^{-x^2/2}\mathcal{I}(x > 0)}_{\mathcal{N}^+(x|0,1)},$$

όπου $\mathcal{N}^+(0,1)$ η τυπική half normal πυκνότητα και $\mathcal{N}^-(0,1)$ η τυπική αρνητική half normal κατανομή. Παρατηρούμε ότι

$$\begin{aligned} g(x) &= \frac{1}{2}\sqrt{\frac{2}{\pi}}e^{-x^2/2}\mathcal{I}(x \le 0) + \frac{1}{2}\sqrt{\frac{2}{\pi}}e^{-x^2/2}\mathcal{I}(x > 0) \\ &\le \frac{1}{2}\sqrt{\frac{2e}{\pi}}e^{x}\mathcal{I}(x \le 0) + \frac{1}{2}\sqrt{\frac{2e}{\pi}}e^{-x}\mathcal{I}(x > 0) \end{aligned}$$

εφόσον

$$\sqrt{\frac{2}{\pi}}e^{-x^2/2} \le \sqrt{\frac{2e}{\pi}}e^{x} \Leftrightarrow e^{-x^2/2} \le e^{x+1/2} \Leftrightarrow (x+1)^2 \ge 0, \text{ όταν } x \le 0,$$

$$\sqrt{\frac{2}{\pi}}e^{-x^2/2} \le \sqrt{\frac{2e}{\pi}}e^{-x} \Leftrightarrow e^{-x^2/2} \le e^{-x+1/2} \Leftrightarrow (x-1)^2 \ge 0, \text{ όταν } x > 0.$$

Ισοδύναμα έχουμε ότι

$$g(x) = N(x\,|\,0,1) \le \sqrt{\frac{2e}{\pi}} \cdot \frac{1}{2} e^{-|x|} = \sqrt{\frac{2e}{\pi}}\, h(x)$$, όπου $\sqrt{\frac{2e}{\pi}} \approx 1.3155$

με $g(x) = \frac{1}{2} e^{-|x|} = GDe\left(x\,|\,\frac{1}{2},1,1 \right)$, όπου

$GDe(x\,|\,p,\lambda_1,\lambda_2) = p\lambda_1 e^{\lambda_1 x} \mathcal{I}(x \le 0) + (1-p)\lambda_2 e^{-\lambda_2 x} \mathcal{I}(x > 0)$.

Έχουμε ότι

$$\begin{aligned}
h(x) = \frac{g(x)}{c\,f(x)} &= \frac{\frac{1}{2}\left\{ \sqrt{\frac{2}{\pi}} e^{-x^2/2} \mathcal{I}(x \le 0) \right\} + \frac{1}{2}\left\{ \sqrt{\frac{2}{\pi}} e^{-x^2/2} \mathcal{I}(x > 0) \right\}}{\sqrt{\frac{2e}{\pi}} \left\{ \frac{1}{2} e^{x} \mathcal{I}(x \le 0) + \frac{1}{2} e^{-x} \mathcal{I}(x > 0) \right\}} \\
&= \frac{e^{-x^2/2} \mathcal{I}(x \le 0) + e^{-x^2/2} \mathcal{I}(x \ge 0)}{e^{x+1/2} \mathcal{I}(x \le 0) + e^{-x+1/2} \mathcal{I}(x \ge 0)} \\
&= e^{-x^2/2 - x - 1/2} \mathcal{I}(x \le 0) + e^{-x^2/2 + x - 1/2} \mathcal{I}(x \ge 0) \\
&= e^{-\frac{1}{2}(x+1)^2} \mathcal{I}(x \le 0) + e^{-\frac{1}{2}(x-1)^2} \mathcal{I}(x \ge 0) = e^{-\frac{1}{2}(|x|-1)^2}
\end{aligned}$$

με πιθανότητα αποδοχής $P(\mathbb{A}) = \frac{1}{a} = \sqrt{\frac{\pi}{2e}} \approx 0.7602$. Ουσιαστικά λοιπόν έχουμε την majorization αναπαράσταση για την τυπική κανονική:

$$N(x\,|\,0,1) = \underbrace{\sqrt{2e/\pi}}_{a} \cdot \underbrace{e^{-\frac{1}{2}(|x|-1)^2}}_{h(x)} \cdot \underbrace{e^{-|x|}/2}_{f(x)}.$$

Το σχήμα δειγματοληψίας λοιπόν είναι:

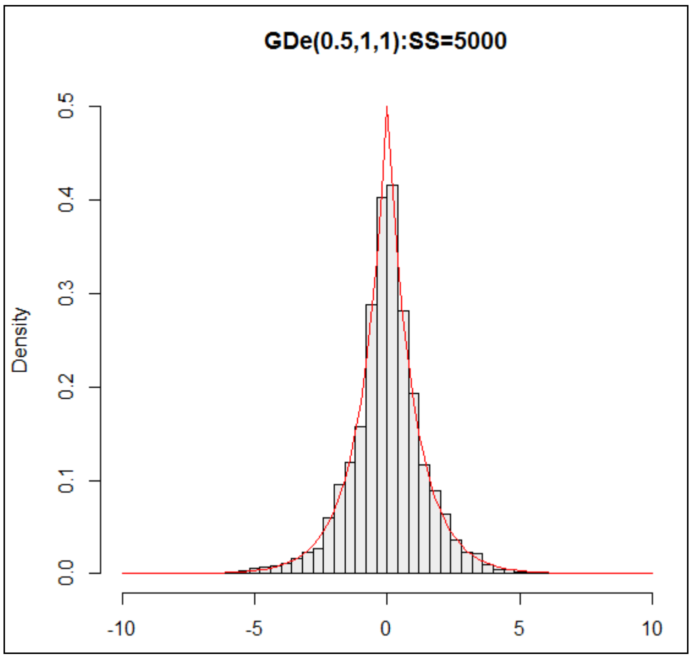
$$\left\{ \begin{array}{ll} x \sim GDe\left(\frac{1}{2},1,1\right) & \\ y \sim \mathcal{U}(0,1) & \Rightarrow\ x\,|\,\mathbb{A} \sim \mathcal{N}(0,1) \\ \mathbb{A} = \left\{ y < \exp\left(-\frac{1}{2}\left(|x|-1\right)^2 \right) \right\} & \end{array} \right\}.$$

Εάν έχουμε δείγμα y από την $N(0,1)$, το μετατρέπουμε σε δείγμα από την $N(\mu,\sigma^2)$ μέσω του μετασχηματισμού $z = \mu + \sigma y \sim N(\mu,\sigma^2)$ και συνολικά έχουμε:

$$\left\{\begin{array}{c} x \sim De\left(\frac{1}{2},1,1\right) \\ y \sim \mathcal{U}(0,1) \\ \mathbb{A} = \left\{ y < \exp\left(-\frac{1}{2}\left(|x|-1\right)^2\right)\right\} \end{array} \Rightarrow \mu + \sigma x \mid \mathbb{A} \sim N\left(\mu, \sigma^2\right)\right\}.$$

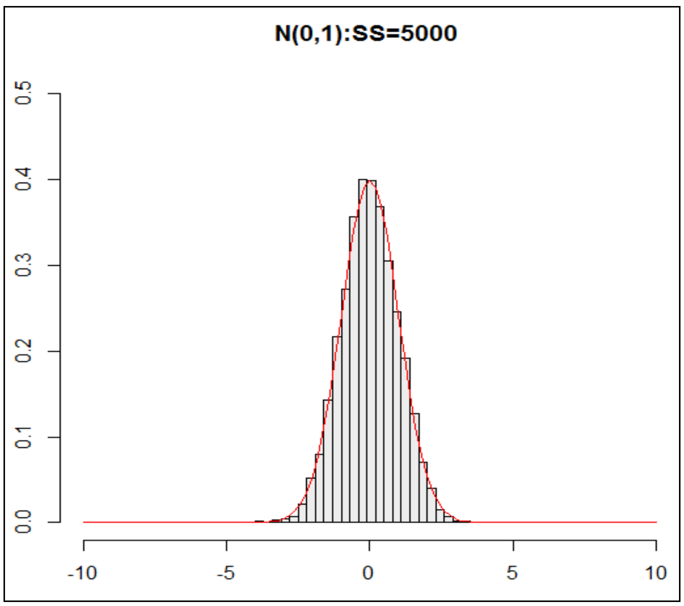
```
# sample from the generalized Double exponential distribution De(0.5, 1, 1)
SampleDe <- function(SS=10000, p=0.5,  lambda1=1, lambda2=1) {
   v <- NULL
   for(i in 1:SS) {
        u1<-runif(1); u2<-runif(1)
        if(u1<p)
            v <- c(v, (1/lambda1)*log(u2))
        else
            v <- c(v, (-1/lambda2)*log(u2))
   }
   return(v)
}
# Define the generalized Double exponential density De(p,l1,l2).
DEXP <- function(x, p=0.5, lambda1=1, lambda2=1){
   v <- c()
   for(i in 1:length(x)) {
       if(x[i]<0)
           branch <- p*lambda1*exp(lambda1*x[i])
         else
           branch <- (1-p)*lambda2*exp(-lambda2*x[i])
       v <- c(v, branch)
   }
   return(v)
}
# Experiment
```

```
v <- SampleDe(SS=5000)
a <- 0.4; h <- 0.5
L <- 10; mybreaks <- seq(from = -L, to = L, by = a)

hist(v, breaks=mybreaks, freq=FALSE, ylim=c(0, h),
        main="GDe(0.5,1,1):SS=5000", col="gray93", xlab="")
curve(DEXP(x), xlim=c(-L,L), col=2, lwd=1.5, add=TRUE)
```




```
# Sample from the standard normal N(0,1) by majorization.
 SampleStdNormal <- function(SS=100){
  sample <- c()
  for(i in 1:SS){
     repeat {
        y <- SampleDe(SS=1, p=0.5,  lambda1=1, lambda2=1)
        u<-runif(1)
        if(u <= exp(-0.5*(abs(y)-1)^2)) break
     }
     sample <- c(sample, y)
```

```
  }
  return(sample)
}
# Experiment
 v <- SampleStdNormal(SS=5000)
 a <- 0.3; h <- 0.5
 L <- 10; mybreaks <- seq(from = -L, to = L, by = a)
 hist(v, breaks=mybreaks, freq=FALSE, ylim=c(0, h),
      main="N(0,1):SS=5000", col="gray93", xlab="")
 curve(dnorm(x), xlim=c(-L,L), col=2, lwd=1.5, add=TRUE)
```



Δειγματοληψία από $\mathcal{N}(0,1)$ με την πολική μέθοδο (μετασχηματισμός των Box – Muller)

$$\left\{\begin{array}{c} \\ \vartheta \sim \mathcal{U}(0,2\pi) \\ r^2 \sim Exp\left(\frac{1}{2}\right) \Leftrightarrow r \sim \mathcal{W}\left(r \mid \frac{\sqrt{2}}{2}, 2\right) \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} x_1, x_2 \overset{iid}{\sim} \mathcal{N}(0,1) \\ x_1 = r\cos\vartheta \\ x_2 = r\sin\vartheta \end{array}\right.$$

Εάν υποθέσουμε ότι $x, y \overset{iid}{\sim} N(0,1)$ έχουμε ότι

$f_{X,Y}(x,y) = \mathcal{N}(x \mid 0,1)\mathcal{N}(y \mid 0,1) = \frac{1}{2\pi}\exp\left\{-\frac{1}{2}\left(x^2+y^2\right)\right\}$ και θέτοντας

$$T^{-1}: \left\{\begin{array}{c} x = r\cos\vartheta \\ y = r\sin\vartheta \\ 0 < r < \infty,\ 0 < \vartheta < 2\pi \end{array}\right. ,$$

παίρνουμε

$$f_{R,\Theta}(r,\vartheta) = f_{X,Y}(x,y)\left|Jac\left(T^{-1}\right)\right|$$

$$= f_{X,Y}(r\cos\vartheta, r\sin\vartheta) \left| \begin{vmatrix} \cos\vartheta & -r\sin\vartheta \\ \sin\vartheta & r\cos\vartheta \end{vmatrix} \right|$$

$$= r\exp\left(-r^2/2\right)\mathcal{I}(r>0)\cdot\frac{1}{2\pi}\mathcal{I}(0<\vartheta<2\pi) = f_R(r)\, f_\Theta(\vartheta),$$

και επειδή $\mathcal{W}(r \mid \lambda,\rho) = \lambda\rho(\lambda r)^{\rho-1} e^{-(\lambda r)^\rho}\mathcal{I}(r>0)$ έχουμε

$$\left\{\begin{array}{l} f_R(r) = r\exp\left(-r^2/2\right)\mathcal{I}(r>0) = \mathcal{W}\left(r \mid \frac{\sqrt{2}}{2}, 2\right) \\ f_\Theta(\vartheta) = \frac{1}{2\pi}\mathcal{I}(0<\vartheta<2\pi) = \mathcal{U}(\vartheta \mid 0, 2\pi). \end{array}\right. ,$$

Έχουμε ότι $r \sim f_R(\cdot) \Rightarrow r^2 \sim Exp\left(\frac{1}{2}\right)$.

Θέτοντας $y = T(r) = r^2$ παίρνουμε $f_Y(y) = f_R\left(\sqrt{y}\right)\left|\frac{d}{dy}\sqrt{y}\right| = \frac{1}{2}e^{-y/2} = Exp\left(y|\frac{1}{2}\right)$.

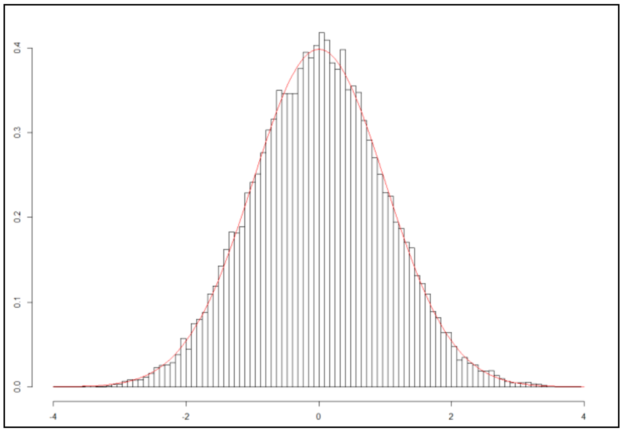
Το σχήμα δειγματοληψίας λοιπόν γίνεται:

$$\begin{cases} u_1 \sim \mathcal{U}(0,1) \Rightarrow r^2 \sim -2\log u_1 \overset{\mathcal{D}}{=} Exp\left(\frac{1}{2}\right) \Leftrightarrow r \sim \sqrt{-2\log u_1}, \\ u_2 \sim \mathcal{U}(0,1) \Rightarrow \vartheta \sim 2\pi u_2 \overset{\mathcal{D}}{=} \mathcal{U}(0,2\pi), \end{cases}$$

και το ανεξάρτητο ζεύγος $u_1 \sim \mathcal{U}(0,1)$ και $u_2 \sim \mathcal{U}(0,1)$, δίνει το ανεξάρτητο ζεύγος $x_1 \sim N(0,1)$ και $x_2 \sim N(0,1)$ σύμφωνα με τον μετασχηματισμό:

$$\begin{cases} x_1 = \sqrt{-2\log u_1}\cos(2\pi u_2) \sim \mathcal{N}(0,1) \\ x_2 = \sqrt{-2\log u_1}\sin(2\pi u_2) \sim \mathcal{N}(0,1). \end{cases}$$

```
# Define g as the unit normal distribution N(0,1).
#----------------------------------------------------------
g <- function(x) return(dnorm(x,mean=0,sd=1))

#The Polar method (Box-Muller) Ver.1.1
#---------------------------------------------------
samplenormal<-function(mu=0, sigma=1, SS=1000){
  sample<-c(1:SS)
  for(i in 1:(SS/2)){
    theta <- runif(1)
    u <- runif(1)
    z <- -2*log(u)
    x1 <- z^(0.5)*cos(2*pi*theta)
    x2 <- z^(0.5)*sin(2*pi*theta)
    sample[2*i-1] <- mu+sigma*x1
    sample[2*i] <- mu+sigma*x2
  }
  return(sample)
}

set.seed(101)
inf <- -4; sup <- 4
w<-samplenormal(sigma=1, SS=20000)
w1 <- w[(w>inf) & (w<sup)]
w2 <- w[(w<inf) | (w>sup)] # equivalent to: w2 <- c(w[w<=inf],w[w>=sup])
mybreaks <- seq(from=inf, to=sup, by=(sup-inf)/100)
hist(w1, freq=FALSE, breaks=mybreaks)
curve(dnorm(x,mean=0,sd=1), add=T, col="red")

> w2
[1] -4.027260  4.182819
```



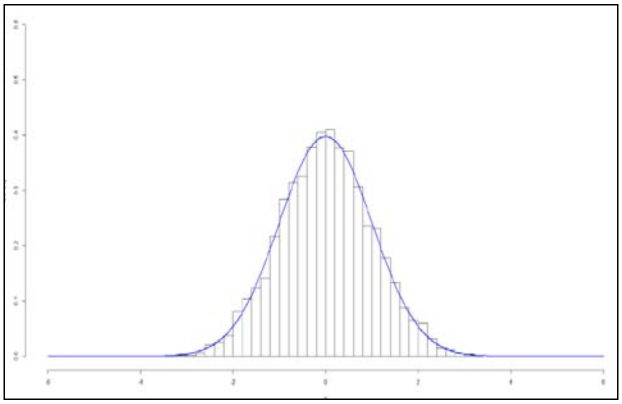
Προσεγγιστική δειγματοληψία από την τυπική κανονική κάνοντας χρήση του Κεντρικού Οριακού Θεωρήματος

Γνωρίζουμε ότι $Z_n = \dfrac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} Z \sim \mathcal{N}(0,1)$

όπου $S_n = \sum_{i=1}^{n} X_i$ και $X_i \overset{i.i.d.}{\sim} f_X(\cdot)$ για $i \geq 1$ με $\mathbb{E}(X_i) = \mathbb{E}(X) = \mu$ και $Var(X_i) = Var(X) = \sigma^2$.

Θέτοντας $X_i \overset{i.i.d.}{\sim} \mathcal{U}(0,1)$ έχουμε $Z_{12} = \dfrac{S_{12} - 12\frac{1}{2}}{\frac{1}{\sqrt{12}}\sqrt{12}} = S_{12} - 6 \overset{\mathcal{D}}{\approx} \mathcal{N}(0,1)$

```
# Define g as the unit normal distribution N(0,1).
#-----------------------------------------------------------
g <- function(x) return(dnorm(x,mean=0,sd=1))

sampleUnitNormal2 <- function(SS){
  Z<-c(1:SS)
  for(i in 1:SS) Z[i]=sum(runif(12))-6
  return(Z)
}


set.seed(11)
v<-sampleUnitNormal2(5000)
hist(v, breaks=50, freq=FALSE, ylim=c(0, 0.6), xlim=c(-6, 6))
curve(g(x), from=-6 ,to=6, col="blue", lwd=2, add=TRUE)
```



Δειγματοληψία σπουδαιότητας (importance sampling)

Αυτή η μέθοδος υπολογισμού ολοκληρωμάτων και πιθανοτήτων γενικεύει την μέθοδο Monte – Carlo κατά την έννοια του ότι μπορούμε να χρησιμοποιήσουμε στη θέση της ομοιόμορφης κατανομής, οιαδήποτε κατανομή που ικανοποιεί συγκεκριμένες συνθήκες

Για παράδειγμα εάν θέλουμε να υπολογίσουμε το ολοκλήρωμα $I=\int_A g(x)dx$ και f είναι η προτεινόμενη πυκνότητα (proposal density) με στήριγμα το A, θα έχουμε:

$$I=\int_A g(x)dx=\int_A \frac{g(x)}{f(x)}f(x)dx=\mathbb{E}\left\{\frac{g(X)}{f(X)}\right\},\quad X\sim f$$

Δηλαδή εάν γνωρίζουμε τρόπο δειγματοληψίας για την f και μπορούμε να παράγουμε δείγμα $x_i \overset{iid}{\sim} f,\ 1\le i\le n$ τότε μια εκτίμηση $\hat{I}_{f,n}$ για το I θα είναι, σύμφωνα με τον ισχυρό νόμο των μεγάλων αριθμών

$$\hat{I}_{f,n}=\frac{1}{n}\sum_{i=1}^{n}\frac{g(x_i)}{f(x_i)},\quad x_i \overset{iid}{\sim} f,\ 1\le i\le n.$$

Πιο γενικά, θα πρέπει, το σύνολο θετικότητας της g να είναι υποσύνολο του στηρίγματος της πυκνότητας f

$$I=\int_A g(x)dx=\int_A w(x)f(x)dx=\int_A w(x)P_X(dx),\text{ όπου } w(x)=\frac{g(x)}{f(x)}.$$

<u>Παράδειγμα</u>

Εάν $Z \sim N(0,1)$ θέλουμε να υπολογίσουμε την πιθανότητα $I = P\{Z > 3\} = \int_{z=3}^{\infty} N(z \mid 0,1)\, dz$.

Μια απλοϊκή μέθοδος θα ήταν η εξής

$$I = \int_{\mathbb{R}} 1(z > 3) N(z \mid 0,1)\, dz = \int_{\mathbb{R}} 1(z > 3) P_Z(dz) = \mathbb{E}\left[1(Z > 3)\right]$$

έτσι θα είχαμε την εκτίμηση

$$\hat{I}_n = \frac{1}{n}\sum_{i=1}^{n} 1(z_i > 3), \quad z_i \overset{iid}{\sim} N(0,1) \text{ με } P\left\{\lim_{n\to\infty} \hat{I}_n = I\right\} = 1.$$

Μια λύση χρησιμοποιώντας importance sampling και proposal density $X \sim N(4,1)$ θα ήταν:

$$I = \int_{\mathbb{R}} 1(z > 3) \frac{N(z \mid 0,1)}{N(z \mid 4,1)} N(z \mid 4,1)\, dz = \int_{\mathbb{R}} \left\{1(z > 3) w(z)\right\} P_X(dz)$$

όπου $w(z) = \dfrac{N(z \mid 0,1)}{N(z \mid 4,1)} = \exp(8 - 4z)$.

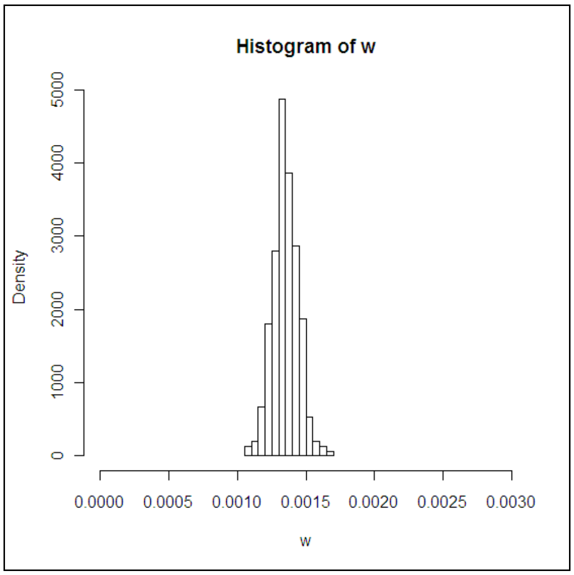
Έτσι

$$I = \mathbb{E}\left\{1(X > 3) w(X)\right\} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \left\{1(x_i > 3) w(x_i)\right\}, \quad x_i \overset{iid}{\sim} N(4,1)$$

<u>Τα R – Scripts είναι:</u>

```
ImportanceSampling <- function(mu=4, sigma=1, SS=1000){
    x <- rnorm(SS, mean=mu, sd=sigma)
    x <- x[x>3]
    w <- dnorm(x, mean=0, sd=1)/dnorm(x, mean=mu, sd=sigma)
    return(sum(w)/SS)
}
EstimatorDistribution <- function(SSEstimator=300, mu=4, sigma=1,
                                  SSImportance=1000){
```

```
    sample <- c()
    for(i in 1:SSEstimator)
        sample <- c(sample, ImportanceSampling(mu=mu, sigma=sigma,
        SS=SSImportance))
    return(sample)
}
w<-EstimatorDistribution()
hist(w, freq=FALSE, breaks=10, xlim=c(0,0.003))
```




```
mean(w);var(w)
 [1] 0.001348459
[1] 8.91818e-09
```

Στο προηγούμενο R – script έχουμε χρησιμοποιήσει την ικανότητα της R για πράξεις με διανύσματα. Εναλλακτικά η συνάρτηση ImportanceSampling() θα μπορούσε να δοθεί με την πιο παραδοσιακή μορφή

```
ImportanceSampling<-function(SS=1000){
    sum<-0
    for(i in 1:SS){
```

```
        x<-rnorm(1, mean=4, sd=1)
        if(x>3) sum<sum+exp(8-4*x)
    }
    return(sum/SS)
}
```

Παράδειγμα

Δίνεται ότι $Y \sim f$ με $f(y) \propto N\left(y \mid 0, \lambda^{-1}\right) 1(3<y<4)$. Να βρεθεί εκτίμηση της $\mathbb{E}[Y]$.

Πρώτα εκτιμούμε την σταθερά c κανονικοποίησης της f

$$f(y) = c \cdot N\left(y \mid 0, \lambda^{-1}\right) \cdot 1(3<y<4) \Rightarrow c^{-1} = \int_{\mathbb{R}} 1(3<y<4) N\left(y \mid 0, \lambda^{-1}\right) d\ y$$

Χρησιμοποιώντας σαν proposal density $Y \sim N\left(3.5, \lambda^{-1}\right)$, έχουμε

$$c^{-1} = \int_{\mathbb{R}} 1(3<y<4) w(y) N\left(y \mid 3.5, \lambda^{-1}\right) dy$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1(3<y_i<4) w(y_i), \quad y_i \overset{iid}{\sim} N\left(3.5, \lambda^{-1}\right) \text{ και } w(y_i) = \frac{N\left(y_i \mid 0, \lambda^{-1}\right)}{N\left(y_i \mid 3.5, \lambda^{-1}\right)}.$$

Ενώ για την $\mathbb{E}[Y]$ έχουμε:

$$\mathbb{E}[Y] = \int_{\mathbb{R}} c 1(3<y<4) y N\left(y \mid 0, \lambda^{-1}\right) dy = c \int_{\mathbb{R}} 1(3<y<4) y w(y) N\left(y \mid 3.5, \lambda^{-1}\right) dy$$

$$= c \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1(3<y_i<4) y_i w(y_i) = \lim_{n\to\infty} \frac{\sum_{i=1}^{n} 1(3<y_i<4) y_i w(y_i)}{\sum_{i=1}^{n} 1(3<y_i<4) w(y_i)}, \quad y_i \overset{iid}{\sim} N\left(3.5, \lambda^{-1}\right).$$

Παράδειγμα

Εάν $X \sim N(0,1)$ και $Y = T(X) = \sigma|X| + \mu$, δείξτε ότι

$$f_Y(y) = \begin{Bmatrix} \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}, & y > \mu \\ 0, & otherwise \end{Bmatrix}.$$

Επίσης δείξτε ότι σε αυτή περίπτωση έχουμε $f_Y(y) \propto N(y \mid \mu, \sigma^2) 1(y > \mu)$. Δηλαδή η τ.μ. Y είναι η περικομμένη (truncated) κανονική κατανομή στο διάστημα (μ, ∞).

Επειδή, $y = T(x) = \sigma|x| + \mu$, έχουμε $\begin{Bmatrix} x > 0 \Rightarrow x_+ = T_+^{-1}(y) = \frac{y-\mu}{\sigma} \\ x < 0 \Rightarrow x_- = T_-^{-1}(y) = -\frac{y-\mu}{\sigma} \end{Bmatrix}$, και $y = \sigma|x| + \mu > \mu$,

έτσι παίρνουμε:

$$f_Y(y) = f_X\left(T_+^{-1}(y)\right)\left|\frac{d}{dy}T_+^{-1}(y)\right| + f_X\left(T_-^{-1}(y)\right)\left|\frac{d}{dy}T_-^{-1}(y)\right|$$

$$= \frac{1}{\sigma}N\left(\frac{y-\mu}{\sigma} \mid 0,1\right) + \frac{1}{\sigma}N\left(-\frac{y-\mu}{\sigma} \mid 0,1\right)$$

$$= \frac{2}{\sigma}N\left(\frac{y-\mu}{\sigma} \mid 0,1\right) = \begin{Bmatrix} \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}, & y > \mu \\ 0, & otherwise \end{Bmatrix}.$$

Έστω ότι $f_Y(y) \propto N(y \mid \mu, \sigma^2) 1(y > \mu)$, τότε υπάρχει $C > 0$, τέτοιο ώστε $f_Y(y) = C \cdot N(y \mid \mu, \sigma^2) 1(y > \mu)$. Ολοκληρώνοντας στο $\mathbb{R}$ έχουμε:

$$1 = \int_{\mathbb{R}} f_Y(y)\,dy = C \cdot \int_{\mathbb{R}} N(y \mid \mu, \sigma^2) 1(y > \mu)\,dy$$

$$= C \int_{y=\mu}^{\infty} N(y \mid \mu, \sigma^2)\,dy = \frac{C}{2} \Rightarrow C = 2$$, από όπου και

$$f_Y(y) = 2 \cdot N\left(y \mid \mu, \sigma^2\right) 1(y > \mu) = \left\{ \begin{array}{cc} \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\}, & y > \mu \\ 0, & otherwise \end{array} \right\}.$$