PEMBANGKIT RANDOM VARIATE

Mata Kuliah Pemodelan & Simulasi

Pertemuan Ke- 8

Riani L.

Jurusan Teknik Informatika

Universitas Komputer Indonesia

Pendahuluan

- Random Variate merupakan nilai variabel random dari suatu distribusi tertentu.
- Pendekatan yang umum digunakan adalah :
  - Inverse Transform
  - Composition
  - Convulotion
  - Acceptance-Rejection

Metode Invers

- Membangkitkan sebuah variabel random X dari sebuah distribusi (umumnya distribusi kontinu).
- Jika yang dibangkitkan adalah suatu variabel random X yang kontinu, maka fungsi distribusinya : $0 < F(X) < 1$ yang artinya $X_1 < X_2$ dan $0 < F(X_1) \leq F(X_2) < 1$, sehingga dapat dimpulkan bahwa $F(X_1) < F(X_2)$
- Prosedur yang diperlukan untuk membangkitkan random variate dari fungsi distribusi kontinu :
  - Cari Cummulative Distribution Function (CDF) dari Probability Density Function (PDF) f(X) dengan cara mengintegralkan fungsi tersebut.
  - Generate random number U (0,1)
  - Hitung $X = F^{-1}(U)$, jika $U = F(X)$

- **Contoh :**
  Bangkitkan random variate distribusi kontinu dengan fungsi matematis sebagai berikut :

$$f(x) = \begin{cases} 2X\,, 0 < x < 1 \\ 0 \quad , x\ lainnya \end{cases}$$

Fungsi distribusi f(x) dijadikan CDF :

$$F(X) = \int_0^x 2X\,.dx = X^2$$

Sifat dari CDF adalah kontinu dan tidak menurun (increasing function).
Kemudian diambil 5 random number dengan salah satu metode (misal LCG), jika telah ditentukan :

$a = 19$
$Z_0 = 12357$
$C = 237$
$m = 128$

Maka diperoleh :

$Z_1 = (19 * 12357+237) \bmod 128 = 12 \rightarrow R_1 = {}^{12}/_{128} = 0{,}09375$

$Z_2 = (19 * 12+237) \bmod 128 = 81 \rightarrow R_2 = {}^{81}/_{128} = 0{,}6328125$

$Z_3 = (19 * 81+237) \bmod 128 = 112 \rightarrow R_3 = {}^{112}/_{128} = 0{,}8750$

$Z_4 = (19 * 112+237) \bmod 128 = 61 \rightarrow R_4 = {}^{61}/_{128} = 0{,}476525$

$Z_5 = (19 * 61+237) \bmod 128 = 116 \rightarrow R_5 = {}^{116}/_{128} = 0{,}90625$

Tranformasikan :

$$U = F(X) = X^2$$

$$X = \sqrt{U}$$

Jadi $F^{-1}(X) = \sqrt{U}$

Maka diperoleh :

$$X_1 = \sqrt{U_1} = \sqrt{0{,}09375} = 0{,}3062$$
$$X_2 = \sqrt{U_2} = \sqrt{0{,}6328125} = 0{,}7955$$
$$X_3 = \sqrt{U_3} = \sqrt{0{,}8750} = 0{,}9354$$
$$X_4 = \sqrt{U_4} = \sqrt{0{,}476525} = 0{,}6903$$
$$X_5 = \sqrt{U_5} = \sqrt{0{,}90625} = 0{,}9520$$

Jika dicari rata-ratanya maka :

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1+X_2+X_3+X_4+X_5}{5} = \frac{3{,}6794}{5} = 0{,}7359$$

Maka X optimalnya adalah 0,7359 untuk $0 < X < 1$

Random Variate Fungsi Kontinu & Diskrit

- Distribusi Kontinu Uniform
- Distribusi Eksponensial
- Distribusi Normal
  - Pendekatan Central Limit Theorem
  - Box Muller Method
- Distribusi Diskrit Uniform
- Distribusi Poisson

Distribusi Kontinu Uniform

Fungsi densitas uniform :

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, \alpha \le x \le \beta \\ \\ 0 \quad\quad , x\ lainnya \end{cases}$$

CDF-nya :

$$F(x) = \int_0^x \frac{1}{\beta - \alpha}.dx = \frac{x - \alpha}{\beta - \alpha}$$

Ditransformasikan :

$$U = F(x) = \frac{x - \alpha}{\beta - \alpha}$$

$$U(\beta\text{-}\alpha) = x - \alpha$$

$$x = U(\beta\text{-}\alpha) + \alpha \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Jadi $F^{-1}(x) = U(\beta\text{-}\alpha) + \alpha$

Distribusi Eksponensial

Fungsi densitas dengan mean β :

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta}, x \geq 0 \\ 0 \qquad , x \; lainnya \end{cases}$$

CDF-nya :

$$F(x) = \int_0^x \frac{1}{\beta} e^{-x/\beta} . dx = 1 - e^{-x/\beta}$$

Ditransformasikan :

$$U = F(x) = 1 - e^{-x/\beta}$$
$$e^{-x/\beta} = 1 - U$$
$$\ln (e^{-x/\beta}) = \ln (1 - U)$$
$$-\frac{x}{\beta} = \ln(1 - U)$$
$$x = -\beta \ln (1\text{-}U) = -\beta \ln U \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Jadi $F^{-1}(x) = -\beta \ln (1\text{-}U) = -\beta \ln U$

Jika diketahui bukan mean dari distribusi eksponensial, tapi tingkat pelayanan dan juga lamanya waktu pelayanan, maka :

t = waktu pelayanan

β= tingkat pelayanan (service rate) dalam unit waktu

Maka distribusi densitas eksponensialnya :

$$f(x) = \frac{1}{\beta} e^{-x/\beta} = \lambda e^{-\lambda x}$$

Menjadi,

$$f(t) = \lambda e^{-\lambda t} \quad , untuk\ t > 0$$

CDF-nya :

$$F(t) = \int_0^t \lambda e^{-\lambda t} . dt = 1 - \lambda e^{-\lambda t}$$

Ditransformasikan :

$F(t) = U = 1 - \lambda e^{-\lambda t}$

$\lambda e^{-\lambda t} = 1 - U$

$\ln (\lambda e^{-\lambda t}) = \ln (1 - U) \ , 0 \leq U \leq 1$

$-\lambda t = \ln U$

$$t = -\frac{1}{\lambda} \ln U = -\beta \ln U \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

Contoh :

Pada suatu sentral telepon ternyata distribusi penerimaan teleponnya adalah distribusi eksponensial dengan $\lambda = 0{,}1$ menit. Setiap kali mengecek dilakukan pengambilan sampel sebanyak 10 kali dengan menggunakan random number $Z_0 = 12357$, a = 173, dan m = 1237. Maka perhitungan waktu untuk pengecekan tersebut adalah :

- Distribusi desitas eksponensial :

$$f(t) = \lambda e^{-\lambda t} \quad , untuk\ t > 0$$

- CDF :

$$F(t) = \int_0^t \lambda e^{-\lambda t}\, . dt = 1 - \lambda e^{-\lambda t}$$

- Persamaan pembangkit random variate-nya :

$$t = -\frac{1}{\lambda} \ln U$$

- Random Number (dengan metode multiplivative karena tidak ada konstanta penggeser) :

  Z_1 = (173 x 12357) mod 1237 = 255 → U_1 = 255/1237 = 0,1819
  Z_2= (173 x 225) mod 1237 = 578 → U_2= 578/1237 = 0,4673
  Z_3= (173 x 578) mod 1237 = 1034 → U_3 = 1034/1237 = 0,8359
  Z_4= (173 x 1034) mod 1237 = 754 → U_4 = 754/1237 = 0,6095
  Z_5= (173 x 754) mod 1237 = 557 → U_5 = 557/1237 = 0,4503
  Z_6= (173 x 557) mod 1237 = 1112 → U_6 = 1112/1237 = 0,8989
  Z_7= (173 x 1112) mod 1237 = 641 → U_7 = 641/1237 = 0,5182
  Z_8= (173 x 1641) mod 1237 = 800 → U_8 = 800/1237 = 0,6467
  Z_9= (173 x 800) mod 1237 = 1093 → U_9 = 1093/1237 = 0,8836
  Z_{10}= (173 x 1093) mod 1237 = 1065→ U_{10} = 1065/1237 = 0,8609

- Random Variat untuk waktu pengecekan :
  t_1 = - (1/0,1) ln (0,1819) = 17,04 menit
  t_2 = - (1/0,1) ln (0,4673) = 7,61 menit
  t_3 = - (1/0,1) ln (0,8359) = 1,79 menit
  t_4 = - (1/0,1) ln (0,6095) = 4,95 menit
  t_5 = - (1/0,1) ln (0,4503) = 7,98 menit
  t_6 = - (1/0,1) ln (0,8989) = 1,07 menit
  t_7 = - (1/0,1) ln (0,5182) = 6,57 menit
  t_8 = - (1/0,1) ln (0,6467) = 4,56 menit
  t_9 = - (1/0,1) ln (0,8836) = 1,24 menit
  t_{10} = - (1/0,1) ln (0,8609) = 1,50 menit

- Total waktu pengecekan :

$$\sum_{i=1}^{10} t_i = t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7 + t_8 + t_9 + t_{10} = 54.31\ menit$$

Pendekatan Central Limit Theorem

Teori khusus dari suatu perkiraan yang mendekati distribusi normal dengan menggunakan rumus :

$$Zn = \frac{Xn - E(Xn)}{\sqrt{VAR(Xn)}} = \frac{Xn - \mu(X)}{\sqrt{\frac{\sigma^2(X)}{n}}}$$

dimana $E(Xn) = \mu(X)$ dan $VAR(Xn) = \sigma^2 (X) / n$
Jumlah random variate-nya : $Sn = X_1 + X_2 + \ldots + X_n$

Dengan random number U(0,1) : $T = U_1 + U_2 + \ldots + U_n$

Jika ekspektasi : $E(T) = n/2$

$$varian\ (T) = \sum_{i=1}^{n} VAR\ Ui = \frac{n}{12}$$

maka

$$Zn = \frac{x - \mu}{\sigma} = \frac{T - E(T)}{\sqrt{VAR(T)}} = \frac{T - {}^{n}/_{2}}{\sqrt{{}^{n}/_{12}}} = \frac{\left(\sum_{i=1}^{n} U_i - {}^{n}/_{2}\right)}{\sqrt{{}^{n}/_{12}}}$$

Untuk setiap distribusi normal dengan mean (μ) dan varian (VAR) = σ^2, akan diperoleh suatu random variate Y yang mengikuti jumlah n random number yang dihasilkan sebelumnya (hasil n yang baik yang memberikan cukup ketelitian n = 12) sehingga diperoleh :

$$\frac{Y-\mu}{\sigma} = \frac{\sum_{i=1}^{n} U_i - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$$

$$Y = \left(\mu + \sum_{i=1}^{n} U_i - \frac{n}{2}\right)\sigma$$

Jika distribusi standar normal N (0,1) maka :

$$Y = \sum_{i=1}^{n} U_i - \frac{n}{2} = \sum_{i=1}^{12} U_i - 6 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

Rumus tersebut digunakan untuk membangkitkan random variate Y dengan U_i = random number yang telah dibangkitkan sebelumnya.

Box Muller Method

Digunakan untuk distribusi normal dengan 2 variate yang tidak diketahui, dengan ketentuan :

$X_1 = N(0,1)$
$X_2 = N(0,1)$
$\mu = 0$
$\sigma = 1$

Rumus PDF-nya :

$$f(x_1) = \frac{1}{\sqrt{2\pi}} e^{-x_1/2}$$

$$f(x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_2/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Berarti :

$f(X_1,X_2) = f(X_1).\ f(X_2)$

$$= \frac{1}{2\pi} . e^{-\left[\frac{x_1^2}{2}+\frac{x_2^2}{2}\right]}$$

$$= \frac{1}{2\pi} e^{-\left[\frac{X_1^2+X_2^2}{2}\right]}$$

Jika diasumsikan : $Y = X_1^2 + X_2^2$
maka,

$$f(Y) = \frac{1}{2\pi} e^{-Y/2}$$

Sehingga

$$F(Y) = \int_0^x \frac{1}{2\pi} e^{-Y/2}.dy$$

$$= \frac{1}{2\pi}\left(-2e^{-Y/2}\right) \rightarrow Y = t$$

$$= \frac{-e^{-t/2}}{\pi} + \frac{1}{\pi}$$

Random variate-nya :

$$F(Y) = U = \frac{-e^{-t/2}}{\pi} + \frac{1}{\pi}$$

$$t = -2ln(1 - \pi U) \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

dari $Y = X_1^2 + X_2^2 = -2\ln(1-\pi U)$

Distribusi Diskrit Uniform

Jika f(x) = f(r) = p, dimana r = i, ….., j

dari fungsi massa diperoleh, $p = \frac{1}{j-i+1}$

Generate random number untuk mendapatkan X dengan cara :

$$X = i + \frac{U}{p} = i + \left[\frac{U}{1/j - i + 1}\right]$$

$$X = i + [U\,(j - i + 1)] \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

Contoh :

Sebuah perusahaan bakery membuat suatu kelompok jenis donat yang dijual di toko-toko dengan distribusi permintaan diskrit uniform dengan kebutuhan harian maksimum 100 unit dan minimal 40 unit. Tentukan random variate sebanyak lima kali dari distribusi tersebut jika diketahui a = 77, Z_0 = 12357, dan m = 127

Maka random variate distribusi diskrit uniform diperoleh dari persamaan (7), dimana :

$$f(x) = p = \frac{1}{j - i + 1}$$

dengan i = 40 unit dan j = 100 unit
sehingga diperoleh :

Z_1 = (77 x 12357) mod 127 = 5 → U_1 = 5/127 = 0,03937
Z_2= (77 x 5) mod 127 = 4 → U_1 = 4/127 = 0,0315
Z_3= (77 x 4) mod 127 = 54 → U_1 = 54/127 = 0,4252

Z_4= (77 x 54) mod 127 = 94 → U_1 = 94/127 = 0,74021
Z_5= (77 x 94) mod 127 = 126 → U_1 = 126/127 = 0,9921

random variate distribusi uniform dengan fungsi massa :

$$fx) = \begin{cases} \frac{1}{j - i + 1}, i \le x \le j + 1 \\ \\ 0 \quad , x\ lainnya \end{cases}$$

CDF-nya :

$$F(x) = \int_0^x \frac{1}{j-i+1}\,dx = \frac{x-i}{j-i+1}$$

dan

$$F(x) = U = \frac{x-i}{j-i+1}$$

$$x = I + U(j - i + 1)$$

sehingga diperoleh :

$X_1 = 40 + 0{,}03937\ (100\text{-}40+1) = 42{,}4016 \approx 42$ unit
$X_2 = 40 + 0{,}0315\ (100\text{-}40+1) = 40{,}5 \approx 41$ unit
$X_3 = 40 + 0{,}4252\ (100\text{-}40+1) = 65{,}937 \approx 66$ unit
$X_4 = 40 + 0{,}74021\ (100\text{-}40+1) = 85{,}152 \approx 85$ unit
$X_5 = 40 + 0{,}9921\ (100\text{-}40+1) = 100{,}518 \approx 100$ unit

Distribusi Poisson

Fungsi massa : $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, 2, \ldots\ldots$

Dari hubungan distribusi eksponensial (parameter β) apabila ada "waktu diantara dua kejadian adalah independen", maka untuk meng-generate random variate X sebagai suatu Poisson variate, dapat dirumuskan melalui pertabahan waktu t dengan batasan :

$$\sum_{i=1}^{n} t_i \leq t < \sum_{i=1}^{n+1} t_i$$

dimana t_i adalah random variate dari distribusi eksponensial ($t_i = -\frac{1}{\lambda} \ln U_i$) maka diperoleh :

$$\sum_{i=1}^{n} -\frac{1}{\lambda} \ln U_i \leq t < \sum_{i=1}^{n+1} -\frac{1}{\lambda} \ln U_i$$

$$-\frac{1}{\lambda} \sum_{i=1}^{n} \ln U_i \leq t < -\frac{1}{\lambda} \sum_{i=1}^{n+1} \ln U_i$$

$$\sum_{i=1}^{n} \ln U_i \geq -\lambda t > \sum_{i=1}^{n+1} \ln U_i$$

Jika :

$$\sum_{i=1}^{n} \ln U_i = \prod_{i=1}^{n+1} \ln U_i$$

maka diperoleh :

$$\prod_{i-1}^{n} \ln U_i \geq -\lambda t > \prod_{i-1}^{n+1} \ln U_i$$

diubah ke logaritma normal :

$$\prod_{i-1}^{n} U_i \geq e^{-\lambda t} > \prod_{i-1}^{n+1} U_i$$

Rumus di atas merupakan penentu simulasi untuk mendapatkan jumlah kedatangan dari distribusi Poisson dengan mean = λ per unit waktu.

Contoh :

Suatu percobaan penelitian untuk meng-generate distribusi Poisson melalui simulasi dengan parameter mean = 2 kejadian per jam selama periode waktu 1,4 jam. Berapakah jumlah kedatangan yang diharapkan di proses simulasi ini, apabila menggunakan data random number sebagai berikut :

a = 7
m = 128
Z_0 = 12357

Maka dari distribusi Poisson diketahui :

λ = 2 kejadian per jam

t = 1,4 jam

Berarti ,

λt = 2 x 1,4 = 2,8 kejadian

$e^{-\lambda t} = (2{,}7182)^{-2{,}8} = 0{,}0608$

Random Number yang dibangkitkan dengan metode Multiplicative karena tidak ada konstanta penggeser :

$Z_1 = (7 \times 12357) \bmod 128 = 99 \rightarrow U_1 = 99/128 = 0{,}7734$

$Z_2 = (7 \times 99) \bmod 128 = 53 \rightarrow U_2 = 53/128 = 0{,}4141$

$Z_3 = (7 \times 53) \bmod 128 = 115 \rightarrow U_3 = 115/128 = 0{,}8984$

$Z_4 = (7 \times 115) \bmod 128 = 37 \rightarrow U_4 = 37/128 = 0{,}2891$

$Z_5 = (7 \times 37) \bmod 128 = 3 \rightarrow U_5 = 3/128 = 0{,}0234 \Leftarrow$ nilai U pertama yang lebih kecil dari 0,0608

$Z_6 = (7 \times 3) \bmod 128 = 21 \rightarrow U_6 = 21/128 = 0{,}1641$

$Z_7 = (7 \times 21) \bmod 128 = 19 \rightarrow U_7 = 19/128 = 0{,}1484$

$$\prod_{i-1}^{n} U_i \geq 0{,}0608 > \prod_{i-1}^{n+1} U_i$$

Dari random number yang diambil, kemudian dimanipulasi :

$$\prod_{i-1}^{4} U_i = U_1 + U_2 + U_3 + U_4$$

$$= 0{,}7734 + 0{,}4141 + 0{,}8984 + 0{,}2891$$
$$= 0{,}0832$$

$$\prod_{i-1}^{4+1} U_i = U_1 + U_2 + U_3 + U_4 + U_5$$

$$= 0{,}7734 + 0{,}4141 + 0{,}8984 + 0{,}2891 + 0{,}0234$$
$$= 0{,}5019$$

Maka dari kedua perhitungan di atas, diperoleh n = 4. Dengan demikian n hasil perhitungan ini merupakan jumlah kedatangan atau pemunculan yang optimal.

Ada metode lain yang dapat meng-generate distribusi Poisson melalui simulasi tertentu. Untuk metode yang kedua, dari teori probabilitas dapat ditemukan sejumlah kejadian yang muncul dan dinyatakan dalam distribusi Poisson. Sedangkan waktu antara kedua kejadian tersebut dinyatakan dalam parameter distribusi eksponensial.

Hasil akhir distribusi Poisson dinyatakan dalam jumlah kejadian n yang muncul dalam waktu atau periode t. Dengan demikian untuk mengambil sampel simulasi dari distribusi poisson dilakukan dengan :

rata-rata = $\mu p = \lambda t$

dimana λ = parameter eksponensial dan t = jumlah waktu

Dari distribusi eksponensial diketahui random variabelnya (persamaan 3) :

$$t_i = -\frac{1}{\lambda} \ln U_i \quad \rightarrow untuk\ \mu e = \frac{1}{\lambda}$$

Dengan mengambil nilai t_i berulang kali sampai jumlah random variate eksponensial melampaui nilai waktu t (jumlah waktu) untuk pertama kalinya.

Contoh :

Suatu percobaan penelitian untuk meng-generate distribusi Poisson melalui simulasi dengan parameter mean = 2 kejadian per jam selama periode waktu 1,4 jam. Berapakah jumlah kedatangan yang diharapkan di proses simulasi ini, apabila menggunakan data random number sebagai berikut :

a = 7
m = 128
Z_0 = 12357

Berarti,

$$\mu e = \frac{1}{\lambda} = \frac{1}{3} jam$$

maka,

$$t_i = -\frac{1}{\lambda} ln U_i$$

Perhitungan selanjutnya disajikan dalam Tabel Poisson Number :

i	U_i	t_i	$\sum_{i=1}^{n} t_i$	
1	0,7734	0,1285	0,1285	
2	0,4141	0,4408	0,5693	
3	0,8984	0,0536	0,6229	
4	0,2891	0,6205	1,2434	⇐ Hasil maksimum
5	0,0234	1,8775	3,1209	⇐ Hasil akhir melampaui nilai 6 (t = 1,4 jam)
6	0,1641	0,9036	4,2450	

Hasil akhir perhitungan random variate merupakan penjumlahan dari t_i (yaitu $\sum_{i=1}^{n} t_i$) untuk dapat mencapai/melampaui t = 1,4 jam.
Berarti,

$$\sum_{i=1}^{n} t_i \geq t = 1{,}4\ jam$$

Untuk yang pertama pada n = 5. Maka berdasarkan distribusi poisson, nilai poisson yang diperoleh lebih besar sama dengan n = 6 – 1 = 5, dan hal ini tercatat sama dengan hasil pada metode yang pertama.

TUGAS

1. Bangkitkan bilangan acak sebanyak 10 kali.
   - Dibangkitkan dengan metode Linier Congruential Generator
   - Jika diketahui a = 5, c = 2, m = 128, dan nilai awal (Z_0) = 12003 digit terakhir NIM masing-masing
2. Bila diketahui fungsi densitas distribusi kontinu adalah :

$$f(X) = \begin{cases} 5X, & 0 < X < 1 \\ 0, & x < 0 \end{cases}$$

a. Tentukan fungsi distribusi kumulatifnya !

b. Bangkitkan random variate distribusi tersebut, bila digunakan metode LCG untuk membangkitkan 10 bilangan random. Jika ditentukan bahwa konstanta pengali = 197, konstanta penggeser = 2375, modulus = 1387, dan nilai awal 12003

c. Berapakah nilai optimal dari X untuk $0 < X < 1$?

3. Distribusi penerimaan telepon pada suatu sentral telepon adalah distribusi eksponensial. Dimana penerimaan teleponnya terjadi setiap 0,1 menit. Setiap kali pengecek dilakukan pengambilan sampel sebanyak 10 kali dengan menggunakan random number yang dibangkitkan dengan metode LCG dengan ketentuan konstanta pengali = 127, konstanta penggeser = 2375, modulus = 1237, dan nilai awal didefinisikan 12003

   a. Hitung waktu untuk pengecekan tersebut ?

   b. Bila pengecekan dilakukan kembali secara berturut-turut, berapa lamakah waktu yang diperlukan ?