METODE SIMULASI DAN RESAMPLING

Pika Silvianti

PENDAHULUAN

RENCANA PERKULIAHAN

REFERENSI:

1. Morgan, B. J. T. 1984. Elements of Simulation. Chapman and Hall.
2. Vasishth, S. & Broe, M. 2011. The Foundations of Statistics: A Simulation-based Approach. Springer Berlin Heidelberg
3. Voss, J. 2014. An Introduction to Statistical Computing: A Simulation-based Approach. Wiley series in computational statistics
4. Ross, S. R. 2012. Simulation 5th ed. Academic Press

1	Pengertian simulasi dan simulasi Statistika 1: Bab 1	5
2-3	Pembangkitan Bilangan acak tunggal 1: Bab 3 dan 4	15
4	Simulasi peubah acak 1: Bab 5	10
5-6	Simulasi Statistik 1: Bab 6 4: Bab 8	10
7	Pembangkitan Bilangan Acak Ganda 4: Bab 6.2	10

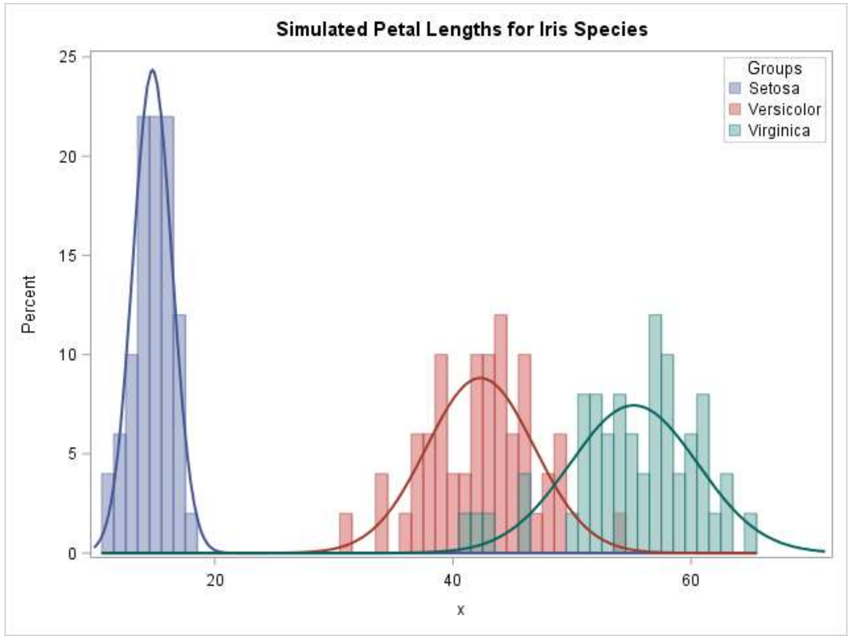
SIMULASI

Simulasi adalah suatu kegiatan yang dapat disesuaikan dengan objek atau rencana kegiatan yang ditentukan.

SIMULASI

SIMULASI STATISTIKA

Simulasi adalah cara untuk memodelkan peristiwa acak, sehingga hasil yang disimulasikan sangat mirip dengan hasil di dunia nyata.



TAHAPAN SIMULASI STATISTIKA

Langkah 1: Nyatakan masalahnya atau jelaskan eksperimennya.

Langkah 2: Nyatakan asumsinya

Langkah 3: Tetapkan angka untuk mewakili hasil.

Langkah 4: Simulasikan banyak pengulangan.

Langkah 5: Nyatakan Kesimpulan

https://accessdl.state.al.us/AventaCourses/APStatisticsA/Unit%205/Simulating%20Experiments.htm

CONTOH SIMULASI STATISTIKA

- Simulasi kepadatan penumpang krl jabodetabek
- Simulasi minat wisatawan yang menyaksikan lomba f1 boat race di danau toba
- Simulasi penambahan rute biskita dramaga – kota
- Simulasi jadwal penerbangan pada bandara baru
- Simulasi lama setiap warna pada lampu lalu lintas
- Simulasi pembangkitan data berdasarkan sebaran tertentu
- Simulasi pendugaan data hilang pada time series
- Simulasi pembuktian teori

ILUSTRASI SIMULASI 1

Melempar sebuah koin sebanyak 10 kali. Berapakah peluang munculnya 3 kepala atau ekor berturut-turut atau lebih?

TAHAPAN SIMULASI STATISTIKA: MELEMPAR KOIN 10X

Langkah 1: Nyatakan masalahnya atau jelaskan eksperimennya.

Melempar sebuah koin sebanyak 10 kali. Berapa peluang terjadinya paling sedikit 3 gambar berturut-turut atau 3 ekor berturut-turut?

Langkah 2: Nyatakan asumsinya

Ada dua:
- Kemungkinan terjadinya kepala atau ekor sama besarnya pada setiap pelemparan.
- Pelemparan tidak bergantung satu sama lain (yaitu, apa yang terjadi pada satu pelemparan tidak akan mempengaruhi pelemparan berikutnya).

https://accessdl.state.al.us/AventaCourses/APStatisticsA/Unit%205/Simulating%20Experiments.htm

Langkah 3: Tetapkan angka untuk mewakili hasil.

Dalam tabel bilangan acak, seperti Tabel Bilangan Acak, angka 0, 1, 2, 3, 4, 5, 6, 7, 8, dan 9 muncul dengan frekuensi relatif jangka panjang yang sama (1/10).

Kita juga tahu bahwa angka-angka yang berurutan dalam tabel adalah bebas.

Oleh karena itu, angka genap dan angka ganjil muncul dengan frekuensi relatif jangka panjang yang sama, 50%.

Berikut adalah salah satu penetapan angka untuk pelemparan koin:

- Satu digit menyimulasikan satu kali pelemparan koin.
- Angka ganjil melambangkan kepala; angka genap melambangkan ekor.

Digit yang berurutan dalam tabel mensimulasikan lemparan independen.

https://accessdl.state.al.us/AventaCourses/APStatisticsA/Unit%205/Simulating%20Experiments.htm

Langkah 4: Simulasikan banyak pengulangan.

Melihat 10 digit berurutan dalam Tabel Angka Acak mensimulasikan satu pengulangan.
Bacalah banyak kelompok yang terdiri dari 10 digit dari tabel untuk menyimulasikan banyak pengulangan.
Pastikan untuk mencatat apakah kejadian yang kita inginkan (lari 3 kepala atau 3 ekor) terjadi pada setiap pengulangan.
Berikut adalah tiga pengulangan pertama, dimulai dari baris 31 pada Tabel Angka Acak.

Gambar 3 atau lebih kepala atau ekor telah digarisbawahi.

Angka 4 1 6 9 2 4 0 5 8 1 | 9 3 0 5 0 4 8 7 3 4 | 3 4 6 5 2 4 1 5 7 7

Heads/TAils T H T H T T T H T H | H H T H T T T H H T | H T T H T H H H H

Jalankan(Run) 3 x

22repetisi tambahan dilakukan dengan total 25 repetisi;
23 di antaranya memang memiliki 3 kepala atau ekor atau lebih.

Langkah 5: Nyatakan Kesimpulan

- Kita menduga peluang *run* berdasarkan proporsi.
- Tentu saja, 25 pengulangan tidak cukup untuk yakin bahwa dugaan kita akurat.
- Sekarang setelah kita memahami cara melakukan simulasi, kita dapat memerintahkan komputer untuk melakukan ribuan pengulangan.
- Simulasi panjang (atau analisis matematis) menemukan bahwa probabilitas sebenarnya adalah sekitar 0,826.

SYNTAX R

Tentukan digit untuk Heads/Tails: 0 = Tails dan 1 = Heads

Syntax 1

```
> sample(0:1,15,rep=T)
 [1] 0 0 0 0 0 0 0 0 0 1 0 1 1 1 0
```

Syntax 2

```
> FlipCoin = function(n) sample(0:1,n,rep=T)
> e1=FlipCoin(30)
> e1
 [1] 1 1 0 0 1 1 0 1 1 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 1 1 1 1 1 1
```

```
sum(e1==0)
sum(e1==0)/30
sum(e1==1)
sum(e1==1)/30

hist(e1,breaks=c(-0.5,0.5,1.5), prob=T)
```

LATIHAN

- Lakukan percobaan melempar koin dengan R, isi table berikut
- Buat histogram pada pelemparan 100x, beri komentar anda

	Lempar koin 100 x	Lempar koin 500x
Frekuensi Relatif Heads		
Frekuensi Relatif Tails		

ILUSTRASI SIMULASI 2

Sepasang suami istri berencana untuk memiliki anak sampai mereka mempunyai anak perempuan atau sampai mereka mempunyai empat anak, mana saja yang lebih dulu. Seberapa besar kemungkinan mereka akan mempunyai anak perempuan di antara anak-anak mereka?

ILUSTRASI SMILULASI

- Kita akan menunjukkan cara menggunakan angka acak untuk menduga kemungkinan mereka akan mempunyai anak perempuan.
- Modelnya sama seperti melempar koin. Kita asumsikan bahwa setiap anak mempunyai peluang 0,5 untuk menjadi perempuan dan 0,5 untuk menjadi laki-laki, dan jenis kelamin anak-anak yang berurutan adalah independen.
- Menetapkan digit juga mudah. Satu digit mensimulasikan jenis kelamin satu anak:
  - 0, 1, 2, 3, 4 = Perempuan (G)
  - 5, 6, 7, 8, 9 = laki-laki (B)

- Untuk menyimulasikan satu kali pengulangan strategi melahirkan anak ini, bacalah angka-angka dari Tabel Angka Acak hingga pasangan tersebut mempunyai anak perempuan atau empat anak.
- Perhatikan bahwa jumlah digit yang diperlukan untuk mensimulasikan satu pengulangan bergantung pada seberapa cepat pasangan tersebut mendapatkan anak perempuan.
- Berikut simulasinya menggunakan baris 38 Tabel Angka Acak
- .Untuk menafsirkan angka-angka tersebut, G untuk anak perempuan dan B untuk anak laki-laki ditulis di bawahnya, spasi memisahkan pengulangan, dan di bawah setiap pengulangan “+” menunjukkan apakah seorang anak perempuan dilahirkan dan “–” menunjukkan ada yang tidak.

7854 54 92 0 1 0 53 2 91 4 1 82 1 0 971

BBBG BG BG G G G BG G BG G G BG G G BBG

+ + + + + + + + + + + + + + +

90 4 72 4 4 682 3 93 0 4 1 981 9557

BG G BG G G BBG G BG G G G BBG BBBB

+ + + + + + + + + + + + –

Dalam 28 kali pengulangan ini, seorang anak perempuan dilahirkan sebanyak 27 kali. Oleh karena itu, dugaan kita mengenai kemungkinan bahwa strategi ini akan menghasilkan anak perempuan adalah = 27/28 = 0,964

Beberapa hitungan matematika menunjukkan bahwa jika model peluang kita benar, kemungkinan sebenarnya untuk memiliki anak perempuan adalah 0,938.
Jawaban simulasi kita masuk akal.

Kecuali pasangan tersebut kurang beruntung, mereka akan berhasil mempunyai anak perempuan.

TABEL BILANGAN ACAK

Random Number Table

Row 20	39634 62349 74088 65564 16379 19713 39153 69459 17986 24537
	14595 35050 40469 27478 44526 67331 93365 54526 22356 93208
	30734 71571 83722 79712 25775 65178 07763 82928 31131 30196
	64628 89126 91254 24090 25752 03091 39411 73146 06089 15630
	42831 95113 43511 42082 15140 34733 68076 18292 69486 80468
Row 25	80583 70361 41047 26792 78466 03395 17635 09697 82447 31405
	00209 90404 99457 72570 42194 49043 24330 14939 09865 45906
	05409 20830 01911 60767 55248 79253 12317 84120 77772 50103
	95836 22530 91785 80210 34361 52228 33869 94332 83868 61672
	65358 70469 87149 89509 72176 18103 55169 79954 72002 20582
Row 30	72249 04037 36192 40221 14918 53437 60571 40995 55006 10694
	41692 40581 93050 48734 34652 41577 04631 49184 39295 81776
	61885 50796 96822 82002 07973 52925 75467 86013 98072 91942
	48917 48129 48624 48248 91465 54898 61220 18721 67387 66575
	88378 84299 12193 03785 49314 39761 99132 28775 45276 91816
Row 35	77800 25734 09801 92087 02955 12872 89848 48579 06028 13827
	24028 03405 01178 06316 81916 40170 53665 87202 88638 47121
	86558 84750 43994 01760 96205 27937 45416 71964 52261 30781
	78545 49201 05329 14182 10971 90472 44682 39304 19819 55799
	14969 64623 82780 35686 30941 14622 04126 25498 95452 63937
Row 40	58697 31973 06303 94202 62287 56164 79157 98375 24558 99241
	38449 46438 91579 01907 72146 05764 22400 94490 49833 09258
	62134 87244 73348 80114 78490 64735 31010 66975 28652 36166
	72749 13347 65030 26128 49067 27904 49953 74674 94617 13317

ROLLING DICE

Rolling dice

The probability of getting a number between 1 to 6 on a roll of a die is $1/6 = 0.1666667$. As above we can use **R** to simulate an experiment of rolling a die a number of times and compare our results with the theoretical probability. We can use the following command to tell **R** to roll a die 20 times:

```
> sample(1:6,20,rep=T)
[1] 3 3 4 1 1 2 2 5 1 2 4 4 3 2 1 5 2 6 5 2
```

As before we can write a function to roll a die n times:

```
> RollDie = function(n) sample(1:6,n,rep=T)
> d1=RollDie(50)
> d1
 [1] 3 4 5 5 6 5 1 6 3 3 1 3 5 4 4 3 2 1 5 2 1 1 2 2 3 1 6 2 6 1 5 1 4 1 4 4 4 6
[39] 2 1 5 5 2 6 1 3 6 3 1 6
```

Now we can use the `sum` command to compare the results from this experiment to the theoretical probabilities. For example in the above experiment the number of 3's and its relative frequency is:

```
> sum(d1==3)
[1] 8
> sum(d1==3)/50
[1] 0.16
```

The number 3 occurs 8 times and its relative frequency is 0.16 which is quite close to 1/6. **Note that you may get different answers**. We can plot a **relative** histogram using the command:

```
> hist(d1,breaks=c(0.5,1.5,2.5,3.5,4.5,5.5,6.5), prob=T)
```

LATIHAN 2

Questions

1. Use **R** to simulate an experiment of rolling a die 200 times. Print the relative histogram and write your name on it.
2. Find the relative frequency of the numbers 1 to 6 in your experiment and fill in the table on the next page.
3. Repeat 2 for rolling a die 1000 times (do not print histogram).

To Hand in

Fill in the next sheet with answers to above questions and hand it in along with one histograms each for "coin toss" and "rolling a die" with youor name on it.

Rolling Dice

	200 rolls	1000 rolls
Relative Frequency of 1		
Relative Frequency of 2		
Relative Frequency of 3		
Relative Frequency of 4		
Relative Frequency of 5		
Relative Frequency of 6		

https://www.math.csi.cuny.edu/abhijit/113/rlabs/lab2.pdf

TERIMA KASIH

Pembangkitan Bilangan Acak Tunggal

Pika Silvianti

Terima Kasih

Model Simulasi

Inti dari setiap model simulasi terdapat kemampuan untuk menciptakan angka-angka yang meniru angka-angka yang kita harapkan dalam kehidupan nyata.

Dalam pemodelan simulasi kita akan mengasumsikan bahwa proses tertentu akan menyebar menurut peubah acak tertentu.

Ilustrasi

- Asumsi: seorang karyawan di toko donat memerlukan waktu acak untuk melayani pelanggan yang didistribusikan menurut variabel acak Normal dengan rataan μ dan ragam σ^2
- Untuk kemudian melakukan simulasi, komputer perlu menghasilkan waktu penyajian acak.
- Hal ini sesuai dengan simulasi bilangan yang menyebar menurut sebaran tertentu.

Sebaran Peubah Acak

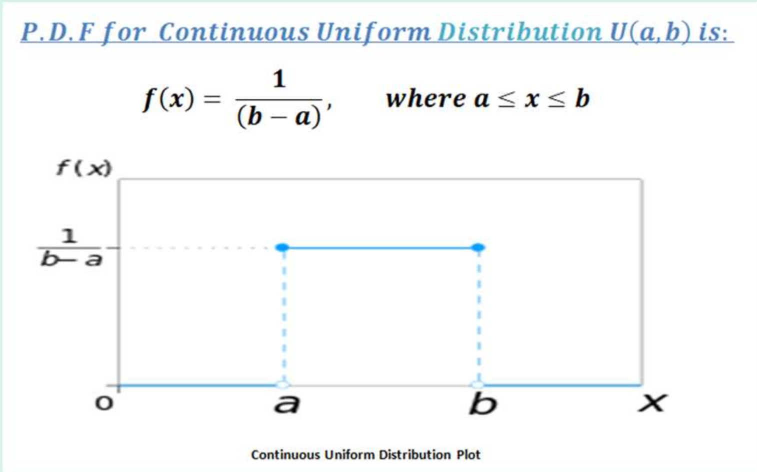
- Bernoulli
- Binomial
- Seragam Diskret
- Poisson
- Seragam
- Normal
- Eksponensial
- dan lain-lain...

Membangkitkan Bilangan Acak Seragam (uniform)

Pertemuan 2

Kenapa Sebaran Seragam?

1. Sebaran kontinu yang paling sederhana itumenggambarkan distribusi beberapa interval bilangan
2. Dari seragam, bisa ditransformasikan menjadi sebaran lain yang rumit



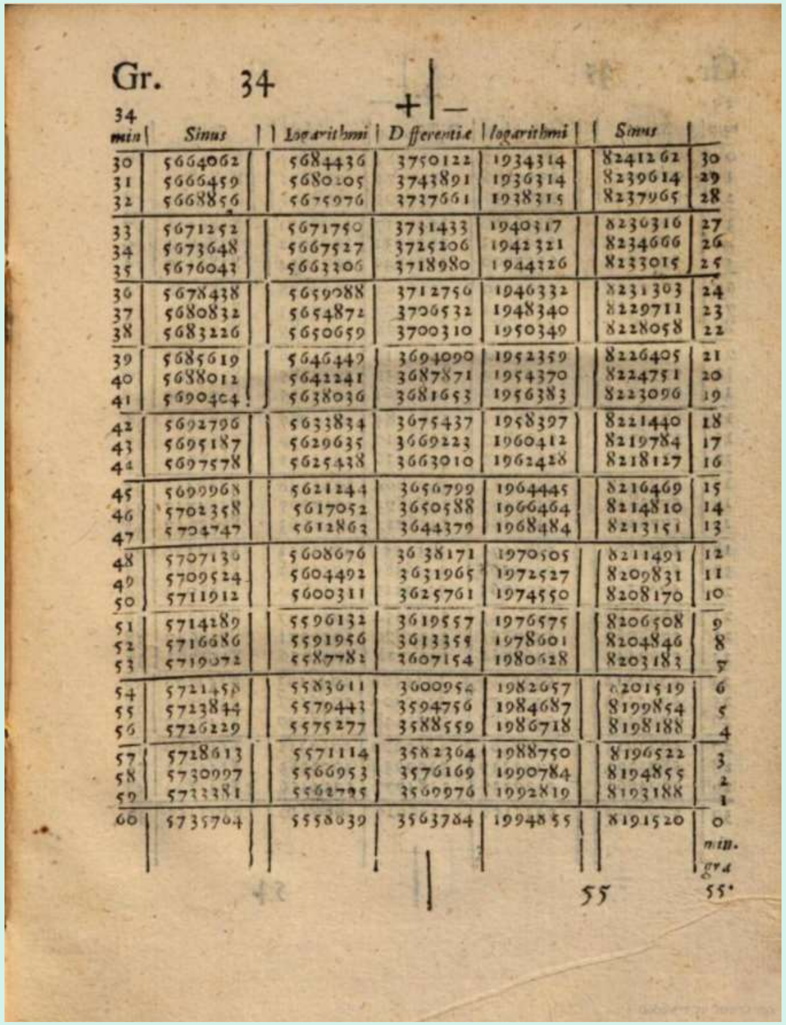
Proses apa pun di alam yang dianggap acak dapat digunakan untuk mencoba mensimulasikan bilangan acak seragam.

Kendall and Babbington-Smith (1939a)

Menggunakan rotating disk dengan 10 uniform segments, yang dihentikan secara acak

Tippett (1925)

menggunakan **digit yang dibaca dari tabel logaritma.**



Gr. 34

+ | –

34 min	Sinus	Logarithmi	Differentiae	logarithmi	Sinus	
30	5664062	5684436	3750122	1934314	8241262	30
31	5666459	5680205	3743891	1936314	8239614	29
32	5668856	5675976	3737661	1938315	8237965	28
33	5671252	5671750	3731433	1940317	8236316	27
34	5673648	5667527	3725206	1942321	8234666	26
35	5676043	5663306	3718980	1944326	8233015	25
36	5678438	5659088	3712756	1946332	8231303	24
37	5680832	5654872	3706532	1948340	8229711	23
38	5683226	5650659	3700310	1950349	8228058	22
39	5685619	5646449	3694090	1952359	8226405	21
40	5688012	5642241	3687871	1954370	8224751	20
41	5690404	5638036	3681653	1956383	8223096	19
42	5692796	5633834	3675437	1958397	8221440	18
43	5695187	5629635	3669223	1960412	8219784	17
44	5697578	5625438	3663010	1962428	8218127	16
45	5699968	5621244	3656799	1964445	8216469	15
46	5702358	5617052	3650588	1966464	8214810	14
47	5704747	5612863	3644379	1968484	8213151	13
48	5707136	5608676	3638171	1970505	8211491	12
49	5709524	5604492	3631965	1972527	8209831	11
50	5711912	5600311	3625761	1974550	8208170	10
51	5714289	5596132	3619557	1976575	8206508	9
52	5716686	5591956	3613355	1978601	8204846	8
53	5719072	5587782	3607154	1980628	8203183	7
54	5721458	5583611	3600954	1982657	8201519	6
55	5723844	5579443	3594756	1984687	8199854	5
56	5726229	5575277	3588559	1986718	8198188	4
57	5728613	5571114	3582364	1988750	8196522	3
58	5730997	5566953	3576169	1990784	8194855	2
59	5733381	5562795	3569976	1992819	8193188	1
60	5735764	5558639	3563784	1994855	8191520	0
						min. gra

55 55°

ERNIE

Merupakan komputer yang digunakan untuk memilih obligasi premium yang unggul di lotere warga negara Inggris, menggunakan 'suara' elektronik dari tabung neon.

Alat bantu Simulasi

- Dadu dan mesin tidak praktis kecuali untuk simulasi terkecil, yang mana sekarang dalam hal apapun mungkin dilakukan dengan bantuan tabel yang tersedial
- Simulasi skala besar biasanya dilakukan dengan menggunakan komputer, dan komputer di awal perkembangannya dilengkapi dengan generator nomor acak bawaan dari jenis fisik, menggunakan fitur elektronik acak, seperti di ERNIE
- kalkulator memiliki tombol RND untuk mensimulasikan U (0, 1) peubah acak

Masalah…

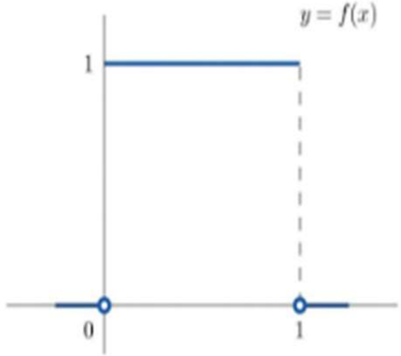
- perangkat tersebut menjadi tidak dapat diandalkan, karena perubahan pada perangkat seiring berjalannya waktu; jadi dadu, misalnya, bisa menjadi tidak rata, sehingga menimbulkan bias.
- Pemeriksaan terhadap angka-angka yang dihasilkan harus sering dilakukan

Pendekatan Modern

Pendekatan modern terhadap simulasi skala besar dilakukan untuk menghindari perlunya pemeriksaan yang sering dengan memproduksi barisan bilangan yang dapat ditunjukkan secara matematis mempunyai sifat tertentu sesuai fitur yang diinginkan.

Pendekatan ini juga mempunyai kelemahan

X ~ Uniform (0,1)



Jadi, bagaimana kita menghasilkan angka acak yang menyebar seragam?

Mulai dari sebaran seragam yang paling sederhana

X ~ Uniform (0,1)

So we have intervals of x_i, $0 \leq x_i \leq 1$

Congruential Generator :

$$X_{n+1} = aX_n + b \ (\mathrm{mod}\ m), n \geq 0$$

$$U_i = \frac{X_i}{m} \sim U(0,1)$$

Congruental Generator

Bilangan yang dihasilkan tidak benar-benar acak, Sehingga disebutr Pseudo-random Number

Maximum cycle

- b & m don't have same factors
- (a-1) (mod prime factor of m) = 0
- (a-1) (mod 4) = 0, if m (mod 4) = 0

So we got :

$\boxed{m = 2^k, k \geq 2}$ $\boxed{a = 4c + 1}$

$\boxed{b > 0, \text{odd numbers}}$

Independence of Observations

$$cov(X_i, X_j) \approx 0$$

$$\rho = \left[\frac{1}{a} - \frac{6b}{am}\left(1 - \frac{b}{m}\right)\right] \pm \frac{a}{m}$$

Congruential Generator (2)

Note:
Modulus adalah sisa pembagian dari suatu bilangan terhadap bilangan lainnya.
Contoh: 100 mod 9 = 1. Karena 100/9 = 11 dengan sisa sebesar 1.

Ilustrasi

$$X_{n+1} = aX_n + b \ (\mathrm{mod}\ m), n \geq 0$$
$$U_i = \frac{X_i}{m} \sim U(0,1)$$

$a = 1598$ $\quad X_0 = 78$

$b = 17$

$m = 1000$

i	$aX_{i-1} + b$	X_i	U_i
0		78	
1	124661	661	0.661
2	1056295	295	0.295
3	471427	427	0.427
4	682363	363	0.363
5	580091	91	0.091
6	145435	435	0.435
7	695147	147	0.147
8	234923	923	0.923
9	1474971	971	0.971
10	1551675	675	0.675
11	1078667	667	0.667
12	1065883	883	0.883
13	1411051	51	0.051

Aplikasi di R

```
x0<-78
n<-250
xi<-matrix(NA,n,3)
colnames(xi)<-c("aX(i-1)+b","Xi","Ui")
for (i in 1:n)
{
xi[i,1]<-(1598*x0+17)
xi[i,2]<-xi[i,1]%%1000
xi[i,3]<-xi[i,2]/1000
x0<-xi[i,2]
}
hist(xi[,3])
```

Other Generator Function

$$\boxed{U_{n+1} = (\pi + U_n)^5 (\mathrm{mod}\ 1), n \geq 0}$$

```
n<-1000
x1<-0.9
for (i in 2:n) x1[i]<-(pi+x1[i-1])^5%%1
hist(x1)
```

```
> cbind("i"=1:20, "Xi"=xi[1:20,2], "i"=101:120,
  "Xi"=xi[101:120,2], "i"=201:220,
  "Xi"=xi[201:220,2])
```

	i	Xi	i	Xi	i	Xi
[1,]	1	661	101	411	201	411
[2,]	2	295	102	795	202	795
[3,]	3	427	103	427	203	427
[4,]	4	363	104	363	204	363
[5,]	5	91	105	91	205	91
[6,]	6	435	106	435	206	435
[7,]	7	147	107	147	207	147
[8,]	8	923	108	923	208	923
[9,]	9	971	109	971	209	971
[10,]	10	675	110	675	210	675
[11,]	11	667	111	667	211	667
[12,]	12	883	112	883	212	883
[13,]	13	51	113	51	213	51
[14,]	14	515	114	515	214	515
[15,]	15	987	115	987	215	987
[16,]	16	243	116	243	216	243
[17,]	17	331	117	331	217	331
[18,]	18	955	118	955	218	955
[19,]	19	107	119	107	219	107
[20,]	20	3	120	3	220	3

Hasil pembangkitan data

- Pada hasil disamping terlihat series dari angka-angka yang berulang
- Hal ini menunjukkan kalua bilangan yang dibangkitkan memang tidak sepenuhnya random
- Hal ini misa diminimalisir dengan menggunakan aturan maximum cycle

$$X \sim \text{Uniform}\ (0,1) \xrightarrow{?} Y \sim \text{Uniform}\ (a,b)$$

If we have X~U 0,1 , how to generate Y~U(a, b) ?

X ~ Uniform (0,1)

$$f(x) = \begin{cases} 1\,, 0 \leq x \leq 0 \\ 0\ \ , x\ lainnya \end{cases}$$

Y ~ Uniform (a,b)

$$f(x) = \begin{cases} \dfrac{1}{b-a}\,, a \leq x \leq b \\ \quad 0 \quad\ , x\ lainnya \end{cases}$$

Y = (b-a)X + a

transformasi

Aplikasi di R

Y ~ Uniform (3,20)

```
Y<-(20-3)*x1+3
hist(Y)
```

Ilustrasi (soal 3.5)

100 kelereng bernomor membentuk populasi dalam percobaan pengambilan sampel. Siswa menduga berat rata-rata populasi(μ= **37,63 g**) dengan memilih 10 kelereng secara **acak menggunakan tabel bilangan acak (random)**, dan juga dengan memilih sampel 10 kelereng, **menggunakan penilaian subjektif mereka saja (judgement).** Hasil yang diperoleh dari 32 kelas diberikan di samping ini:

Judgement sample means	Random sample means
62.63	31.45
35.85	32.12
55.36	51.93
66.43	24.74
34.96	43.32
37.23	29.41
34.45	42.67
60.53	47.94
49.61	28.76
56.07	56.43
59.02	31.21
50.65	32.73
33.34	55.37
58.62	36.65
47.02	22.44
48.34	40.04
28.56	44.65
26.65	41.43
46.34	39.39
27.86	26.39
39.62	23.88
25.45	35.15
48.82	35.88
66.56	28.03
37.25	31.71
45.98	43.98
32.46	61.49
54.03	31.52
51.89	33.99
62.81	33.78
59.74	49.69
14.05	22.97

(lanjutan ilustrasi soal 3.5)

Hitung:

	Judgement	Random
$\bar{x}$		
s		

Diskusikan, dengan mengacu pada data ini, pentingnya pengambilan sampel secara acak.

IPB University
— Bogor Indonesia —
Department of Statistics

Pembangkitan Bilangan Acak SebaranDiskret

PEUBAH ACAK DISKRET

1

Let's start with the first topic of slides

Peubah Acak Bernoulli

Peubah Acak Bernoulli merupakan peubah acak yang digunakan untuk mewakili tindakan yang hanya memiliki dua buah kejadian yang diberi nilai 0 dan 1.

Nilai 1 => peluang sukses = p

Nilai 0 => peluang gagal= (1-p)

Fungsi Massa Peluang :

$$f(x) = p^x(1-p)^x$$

$$dengan\ x = 0{,}1$$

Nilai Harapan :

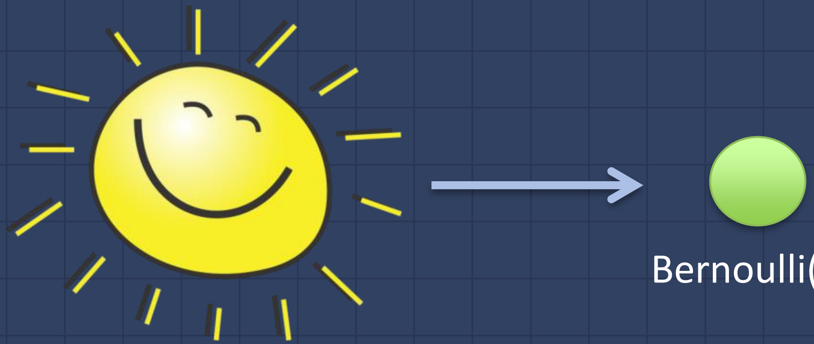
$$E(x) = p$$

Ragam :

$$V(x) = p(1-p)$$

Peubah Acak Bernoulli

Unifrom(0,1)

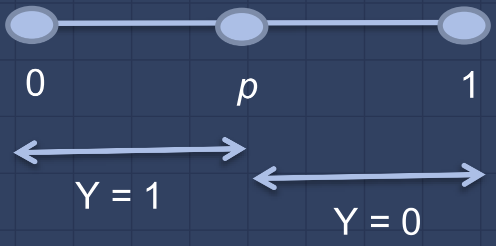


Bernoulli(p)

$X \sim Uniform(0,1)$

$0 \leq F_x(x) \leq 1$

$0 \leq X \leq 1$



$Y \sim Bernoulli(p)$

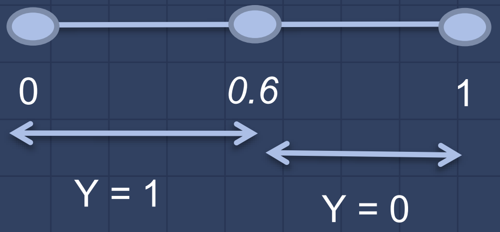
Ilustrasi

Bangkitkan peubah acak Bernoulli(0.6) dari sebaran Uniform(0,1) berdasarkan kasus diatas!

$$X \sim Uniform(0,1)$$

$$0 \leq F_x(x) \leq 1$$

$$0 \leq X \leq 1$$



$$Y \sim Bernoulli(p)$$

$$P(Y=y) = 0.6 \rightarrow Y = 1$$

$$P(Y=y) = 0.4 \rightarrow Y = 0$$

$$P(Y=1) = 0.6$$

$$P(Y=0) = 0.4$$

Ilustrasi

Bangkitkan peubah acak Bernoulli(0.8) dari sebaran Uniform(0,1) berdasarkan kasus diatas!

$$X \sim Uniform(0,1)$$

$$0 \leq F_x(x) \leq 1$$

$$0 \leq x \leq 1$$

0 0.8 1

Y = 1 Y = 0

$$Y \sim Bernoulli(p)$$

$$P(Y{=}y) = 0.8 \rightarrow Y = 1$$

$$P(Y{=}y) = 0.2 \rightarrow Y = 0$$

$$P(Y{=}1) = 0.8$$

$$P(Y{=}0) = 0.2$$

Peubah Acak Binomial

Peubah Acak Binomial merupakan peubah acak bernouli yang dilakukan berulang kali (representasi banyaknya "sukses" yang terjadi dari n buah tindakan Bernouli yang saling bebas).

Fungsi Massa Peluang :

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

dengan $x = 0, 1, \ldots, n$

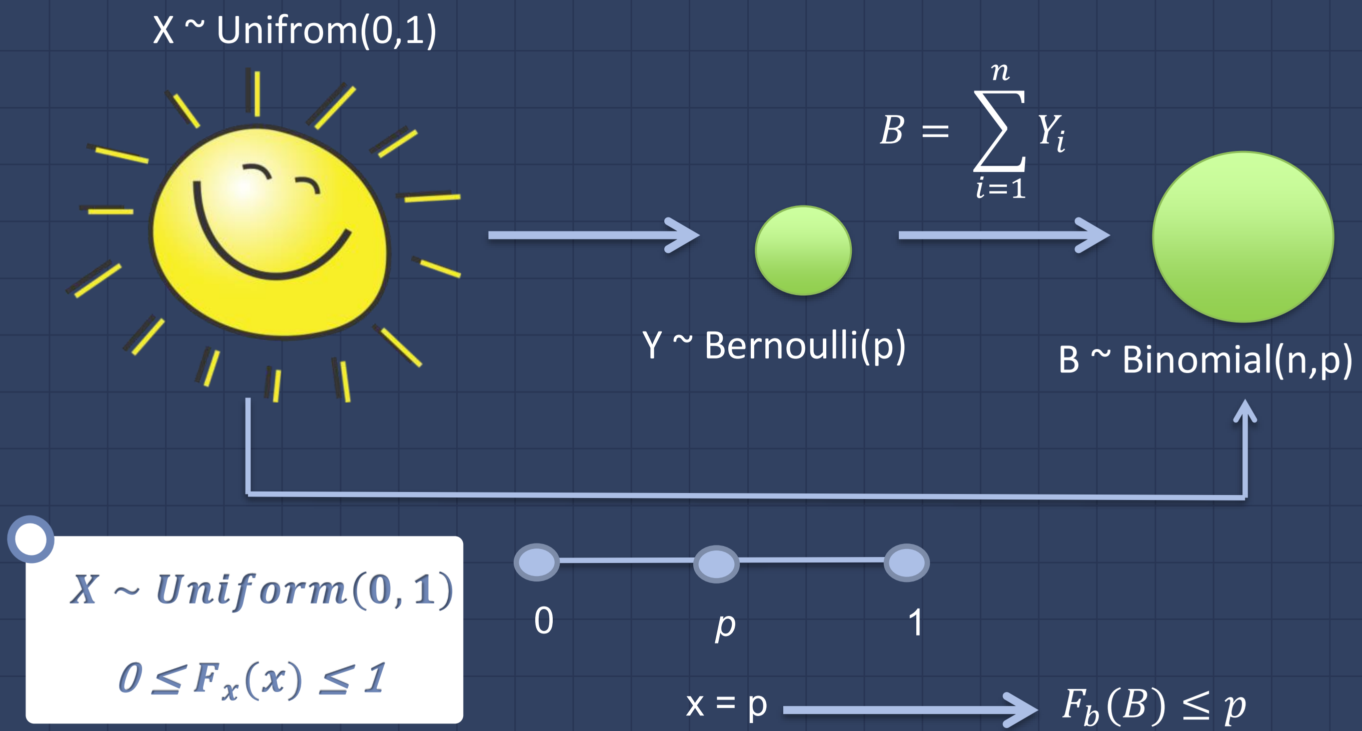
Nilai Harapan :

$$E(x) = np$$

Ragam :

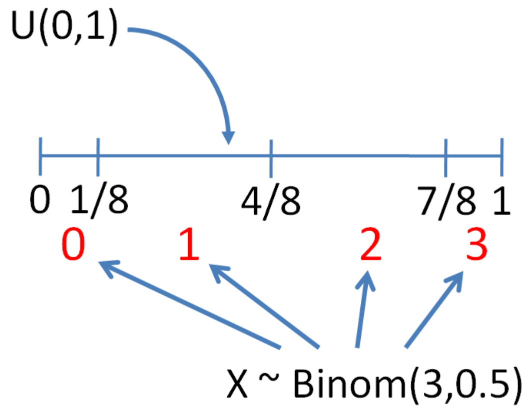
$$V(x) = np(1-p)$$

Peubah Acak Binomial



Ilustrasi

Bangkitkan bilangan acak Binomial(3,0.5) dari sebaran Uniform(0,1)



$$P(X=x)=\begin{cases}1/8;\, x=0,3\\ 3/8;\, x=1,2\\ 0;\, x \text{ lainnya}\end{cases}$$

Peubah Acak Poisson

Peubah Acak Poisson merupakan representasi dari peubah acak binomial yang peluangnya sangat kecil (jarang terjadi) dan banyaknya kejadian sangat besar (menuju tak hingga)

Fungsi Massa Peluang :

$$f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

dengan $x = 0,1,2,\ ...,$

Nilai Harapan :

$$E(x) = \lambda$$

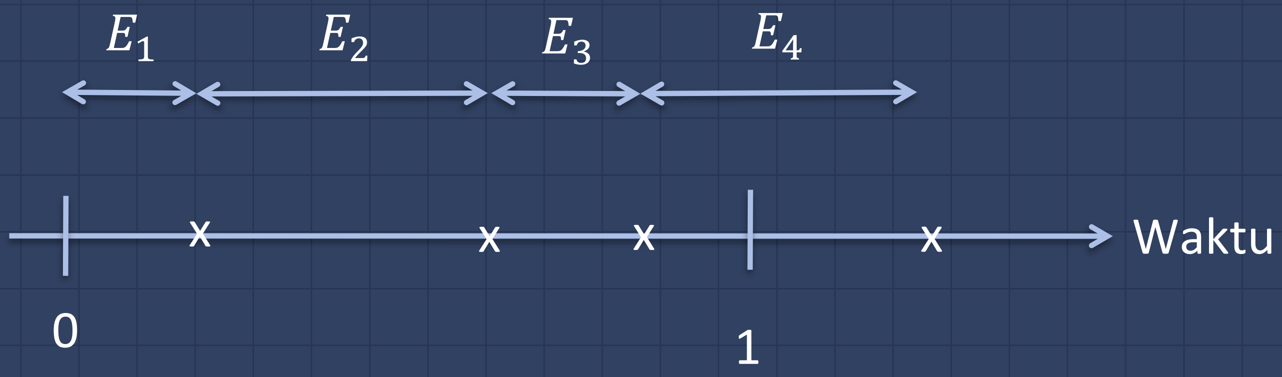
Ragam :

$$V(x) = \lambda$$

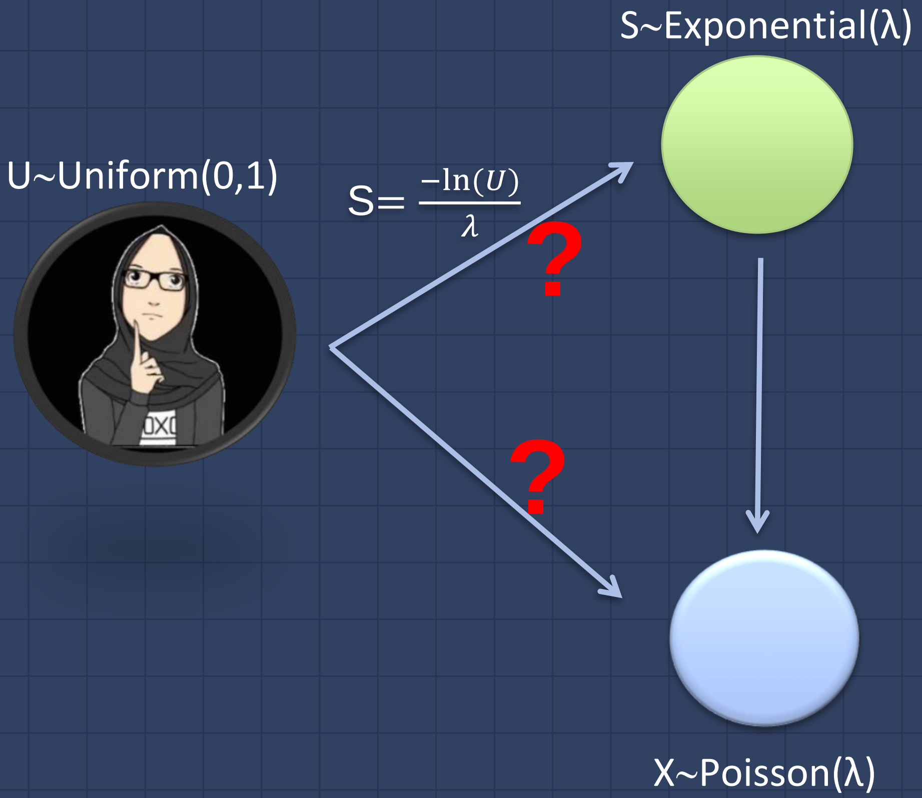
Peubah Acak Poisson

Proses Poisson dengan laju sebesar λ

- Waktu antar kejadian saling bebas menyebar Eksponensial (λ)
- Banyaknya kejadian pada selang waktu t menyebar Poisson (λt)



Peubah Acak Poisson

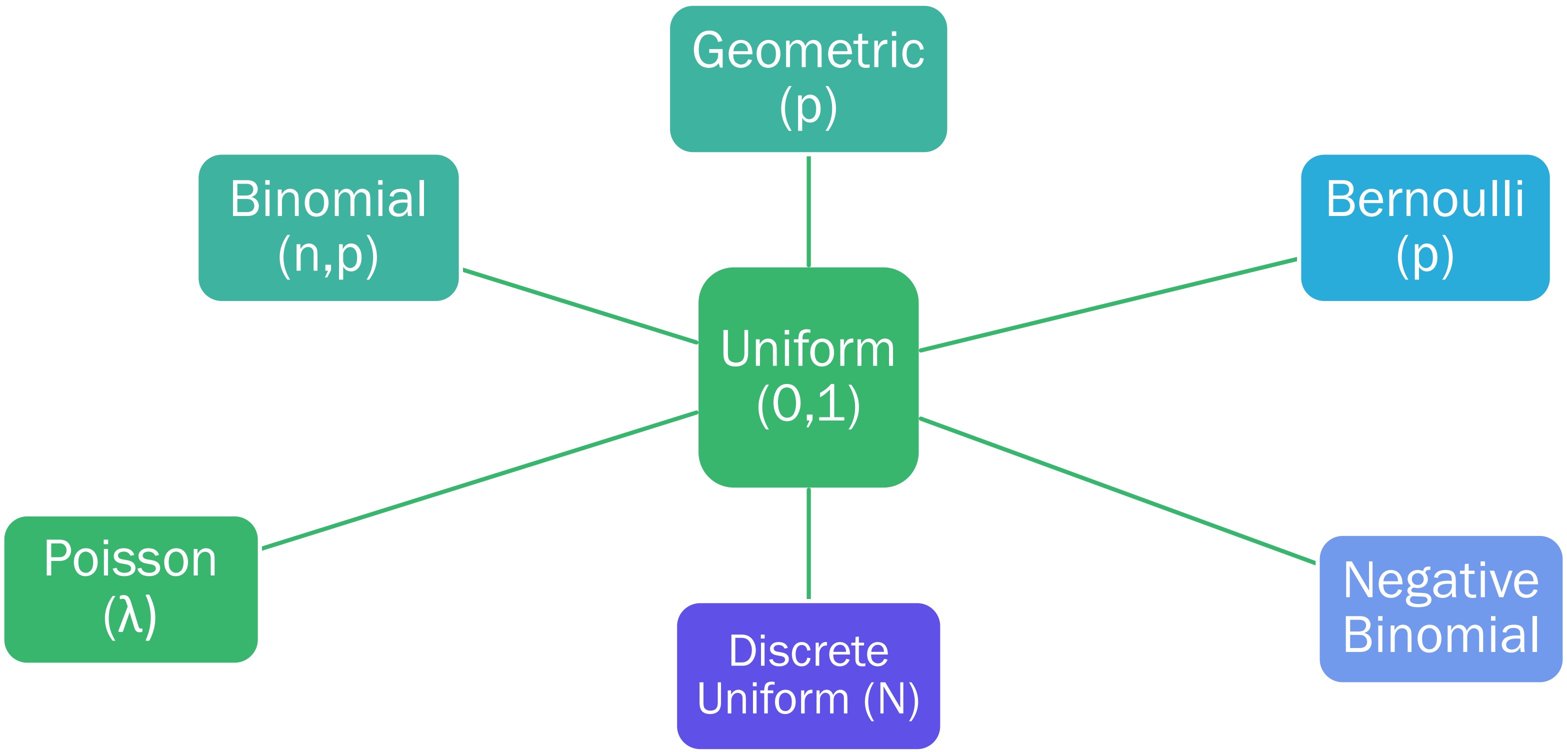


Terima Kasih

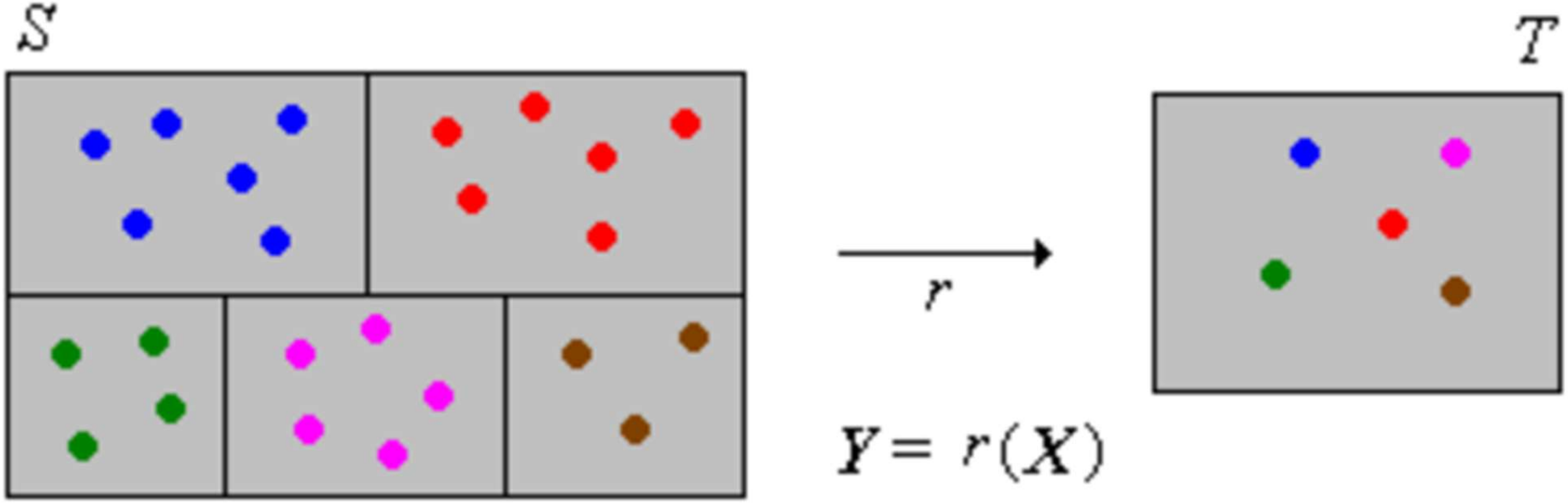
DISCRETE AND CONTINUE RANDOM NUMBERS

PIKA SILVIANTI

TRANSFORMASI PEUBAH ACAK DISKRET



TRANSFORMASI PEUBAH ACAK DISKRET



SEBARAN SERAGAM (0,1) ➔ SERAGAM (a,b)

$$X \sim Seragam\ (a, b)$$

Fungsi Kumulatifnya :

$$F(x) = \frac{(x-a)}{(b-a)}$$

$$\Downarrow$$

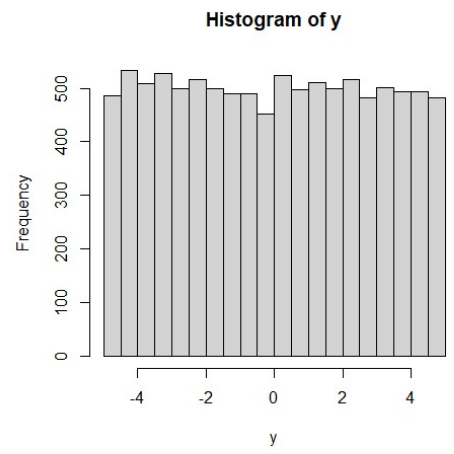
$$U = \frac{(x-a)}{(b-a)}$$

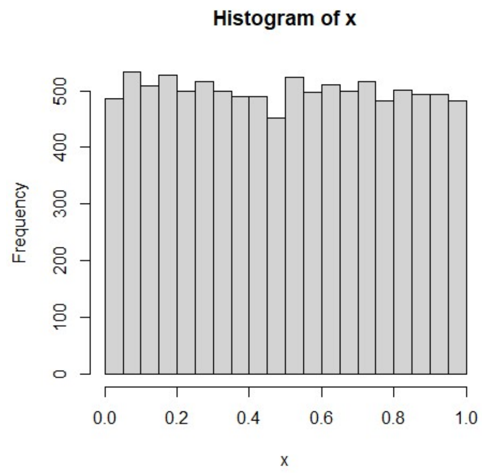
$$X = a + (b-a)U$$

SEBARAN SERAGAM (UNIFORM)

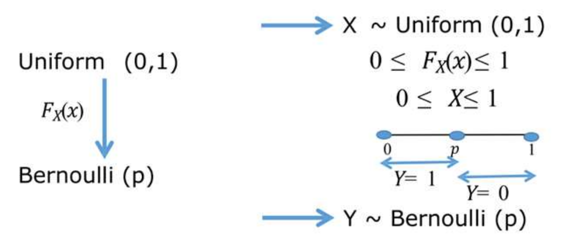
- Jika x merupakan peubah acak dengan sebaran uniform [0, 1], maka transformasi berikut untuk membangkitkan bilangan acak uniform (a,b) :
- y = (b − a)x + a
- Membangkitkan bilangan acak menyebar uniform [-5, 5] :

```
x <-runif (10000, 0, 1) ;
y <- 10*x - 5;
hist(y)
```



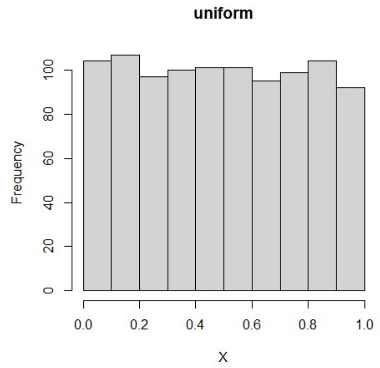


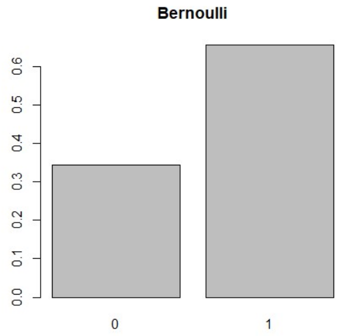
SEBARAN SERAGAM (0,1) ➔ SEBARAN BERNOULI (p)




```
i<-1000
p<-.65
X<-runif(i)
hist(X,main="uniform")
Y<-NULL
for (z in 1:i) ifelse (X[z]<=p,Y[z]<-1,Y[z]<-0)
(tabel<-table(Y)/length(Y))
barplot(tabel,main="Bernoulli")
```

Jika z<=p maka z=1, selainnya z=0





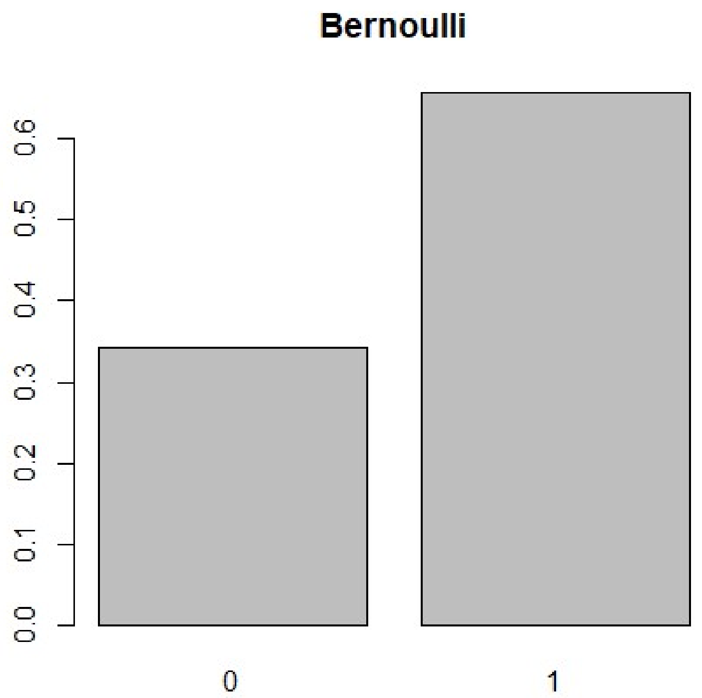

```
Y
    0     1
0.344 0.656
```

SEBARAN BERNOULI (p) LANJUTAN..

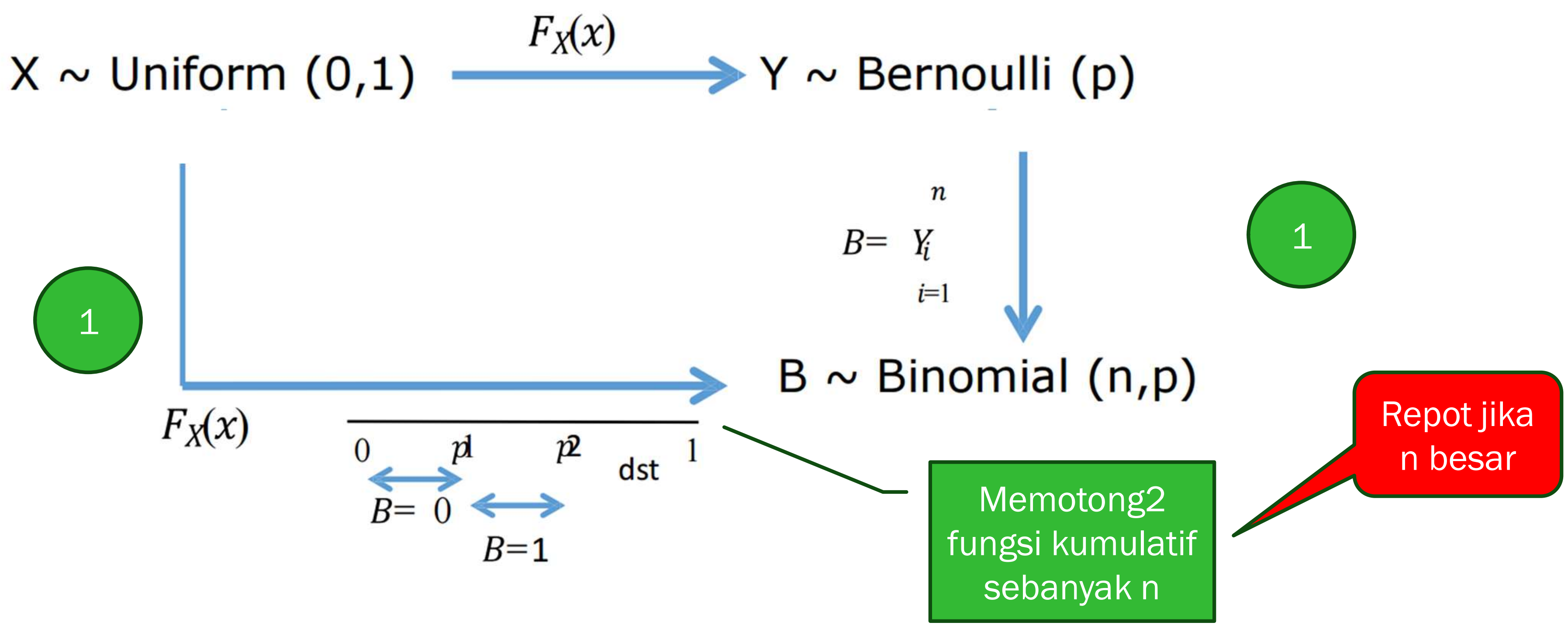
- Cara lain

```
i<-1000
p<-.65
X<-runif(i)
Y<-(X<=p)+0
(tabel<-table(Y)/length(Y))
barplot(tabel,main="Bernoulli")
```

Khusus di R



SEBARAN SERAGAM (0,1) ➔ BINOMIAL(n,p)



- https://math.stackexchange.com/questions/2717462/simulating-a-binomial-distribution-with-mathscru0-1

APLIKASI DI R

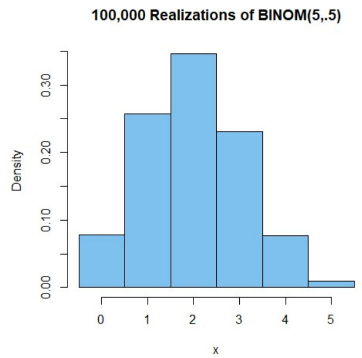
- runif (tanpa parameter tambahan) adalah sumber nilai pseudorandom dari seragam standar;
- dbinom ➔ binomial PDF (probability distribution function)
- pbinom ➔ binomial CDF (cumulative distribution function)
- qbinom ➔ fungsi kuantil (invese CDF).

Jika ingin membangkitkan m=100.000 pengamatan dari sebaran Binom(n=5,p=.4)

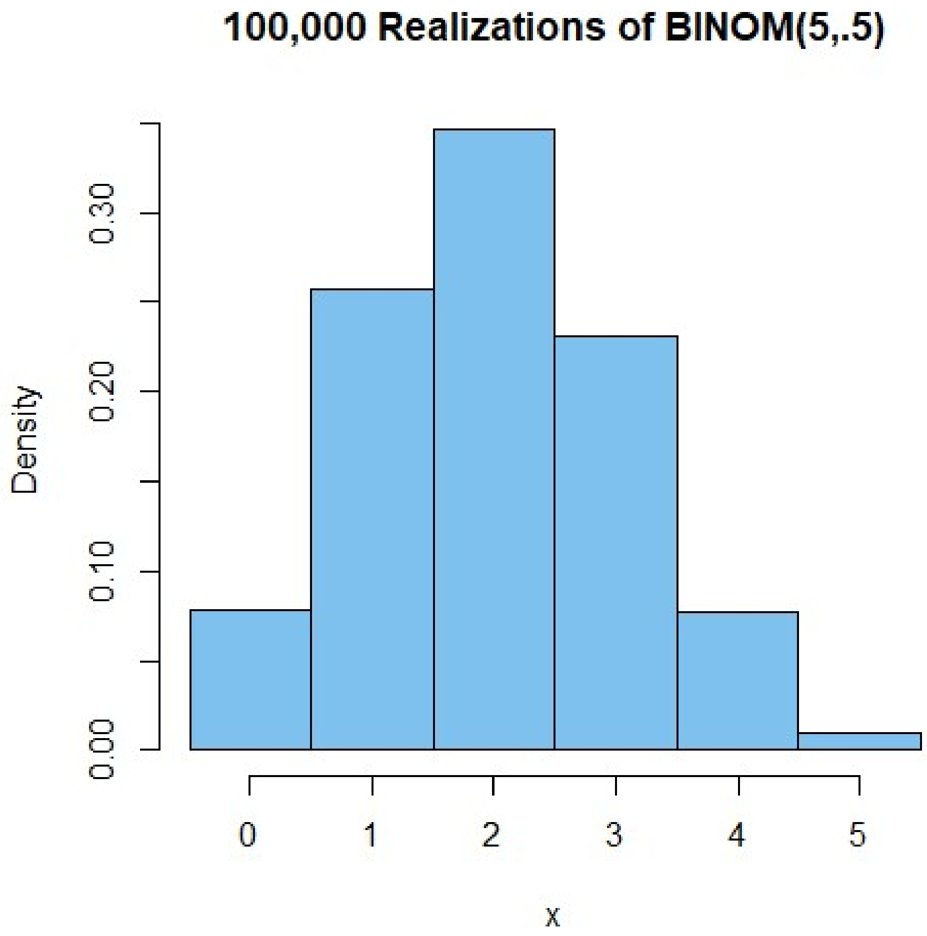
```
set.seed(4118)
m = 10^5;
u = runif(m)
x = qbinom(u, 5, .4) # inverse CDF transformation

table(x)/m
```

```
x
      0       1       2       3       4       5
0.07790 0.25775 0.34608 0.23105 0.07744 0.00978
```

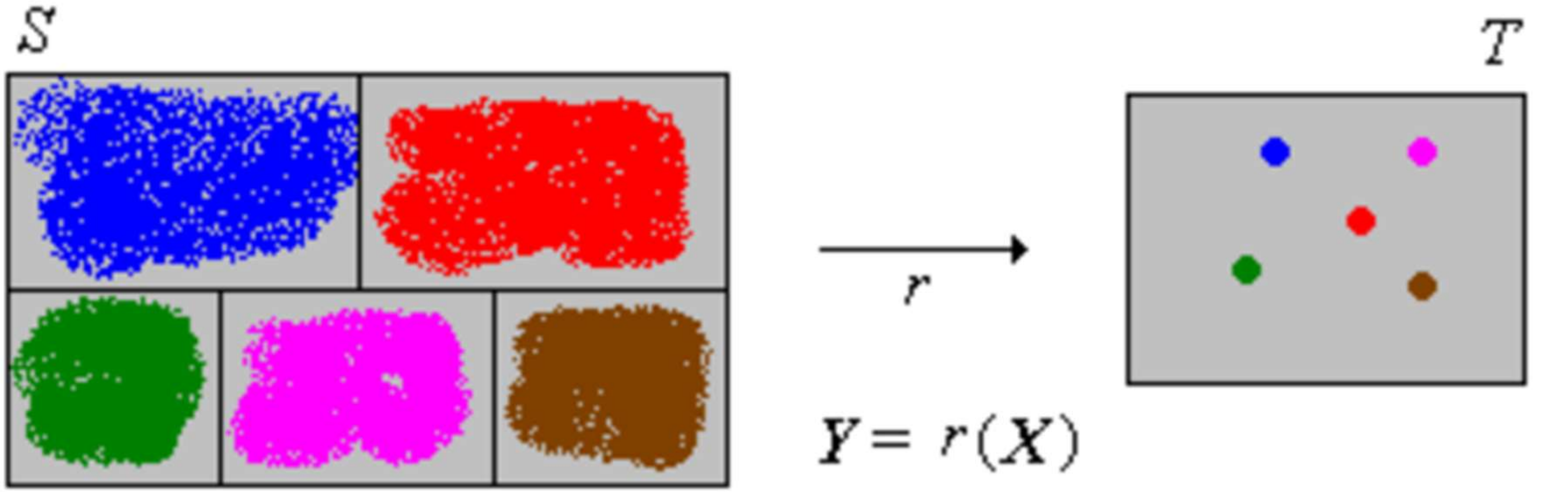
```
hist(x, prob=T, br=(0:6)-.5, col="skyblue2", main="100,000
Realizations of BINOM(5,.5)")
k = 0:5; pdf = dbinom(k, 5, .4)
points(k, pdf, col="red")
```



BINOM (5,0.5)



100,000 Realizations of BINOM(5,.5)
Density
0.00
0.10
0.20
0.30
0
1
2
3
4
5
x



SEBARAN PEUBAH ACAK KONTINU

INVERSE TRANSFORM METHOD

- Metode Transformasi Kebalikan
- Dikenal juga sebagai Look-Up Table Method
- Didasari pada kenyataan bahwa
  - jika U adalah bilangan acak Seragam(0, 1)
  - dan didefinisikan X = F-1(U), dengan F-1(U) adalah fungsi kebalikan dari F(X)
  - maka X akan memiliki sebaran yang diinginkan

INVERSE TRANSFORM METHOD (LANJUTAN)

Algoritma untuk mendapatkan bilangan acak X dengan sebaran tertentu

- Tentukan bentuk dari fungsi sebaran kumulatif X yang diinginkan, misal F(x)
- Cari fungsi kebalikan dari F(x), yaitu $F^{-1}(x)$
- Bangkitkan bilangan acak Seragam (0, 1), misal dilambangkan U
- Hitung $X = F^{-1}(U)$

INVERSE TRANSFORM METHOD SERAGAM (a, b)

- Ilustrasi untuk membangkitkan sebaran Seragam(a, b)

- X ~ Seragam(a, b)
  - F(x) = (x – a) / (b – a)
  - U = (x – a) / (b – a)
  - X = a + (b – a) U

Algoritma:

- Bangkitkan U, bilangan acak Seragam(0, 1)
- Hitung X = a + (b – a) * U
- Ulangi berkali-kali sesuai dengan banyaknya bilangan yangdiinginkan

LANJUT SLIDE PEUBAH ACAK KONTINU….

IPB University — Bogor Indonesia | Department of Statistics

Pembangkitan Bilangan Acak Sebaran Kontinyu

Dr. Ir. Erfiani, M.Si

PEUBAH ACAK KONTINU

2

Let's start with the first topic of slides

Peubah Acak Eksponensial

Peubah acak eksponensial adalah peubah acak yang digunakan untuk memodelkan waktu, misalnya waktu tunggu, hardware lifetime, dan waktu diantara panggilan telepon

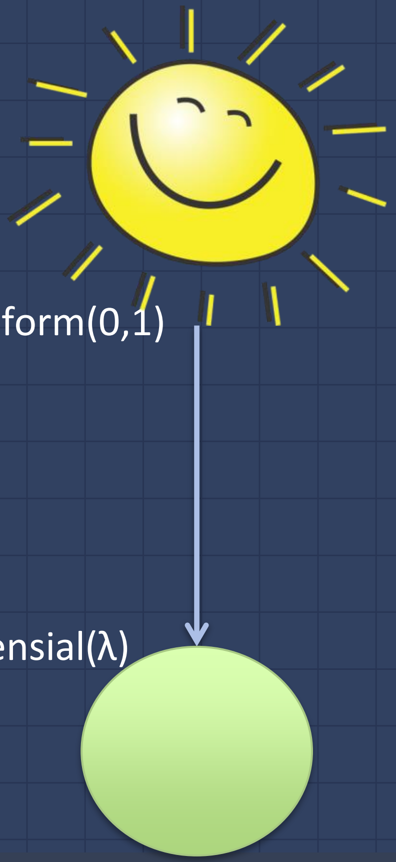
- Banyaknya kejadian yang terjadi pada selang waktu t ~ Poisson(λt)
- Waktu antar kejadian ~ Eksponensial (λ)

Fungsi Kepekatan Peluang	Nilai Harapan	Ragam
$\boldsymbol{f(x) = \lambda e^{-\lambda x}}$	$E(x) = \dfrac{1}{\lambda}$	$V(x) = \dfrac{1}{\lambda^2}$

Eksponensial(λ) adalah kasus khusus dari Gamma (1,λ)

Peubah Acak Eksponential

U ~ Uniform(0,1)



X ~ Eksponensial(λ)

$$0 \leq F_x(X) \leq 1$$
$$0 \leq U \leq 1$$
$$0 \leq X = F_x(X) \leq 1$$

$$
\begin{aligned}
U &= F_x(x) \\
U &= 1 - e^{-\lambda x} \\
e^{-\lambda x} &= 1 - U \\
\ln(e^{-\lambda x}) &= \ln(1-U) \\
-\lambda x &= \ln(1-U) \\
X &= \frac{-\ln(1-U)}{\lambda}
\end{aligned}
$$

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

U ~ Uniform(0,1)

1-U ~ Uniform(0,1)

Peubah Acak Gamma

Peubah acak gamma adalah peubah acak yang digunakan dalam teori antrian dan biasanya diaplikasikan dalam lamanya waktu untuk menyelesaikan pekerjaan.

$$X \sim Gamma\,(\alpha, \beta)$$

Fungsi Kepekatan Peluang

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

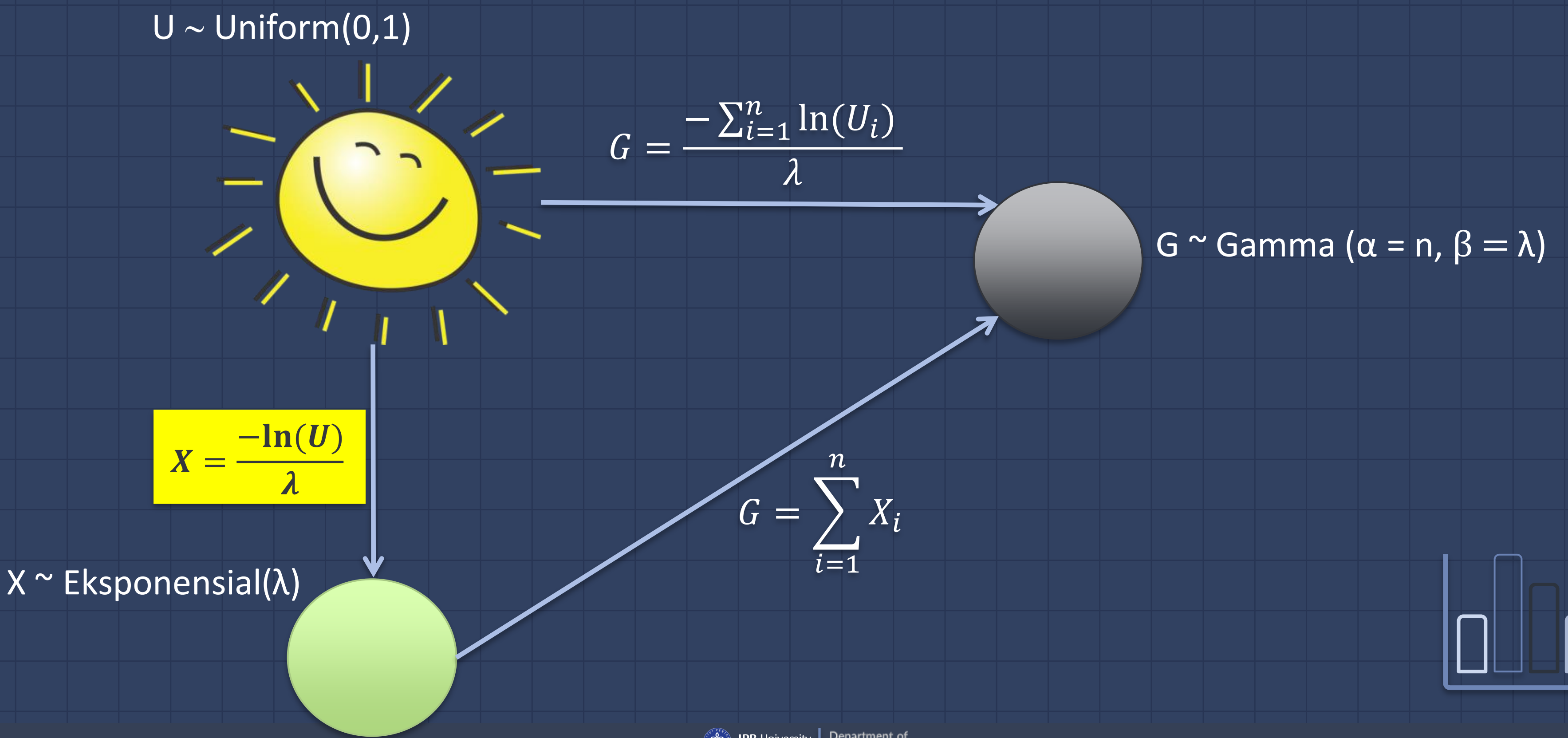
Nilai Harapan

$$E(x) = \alpha\beta$$

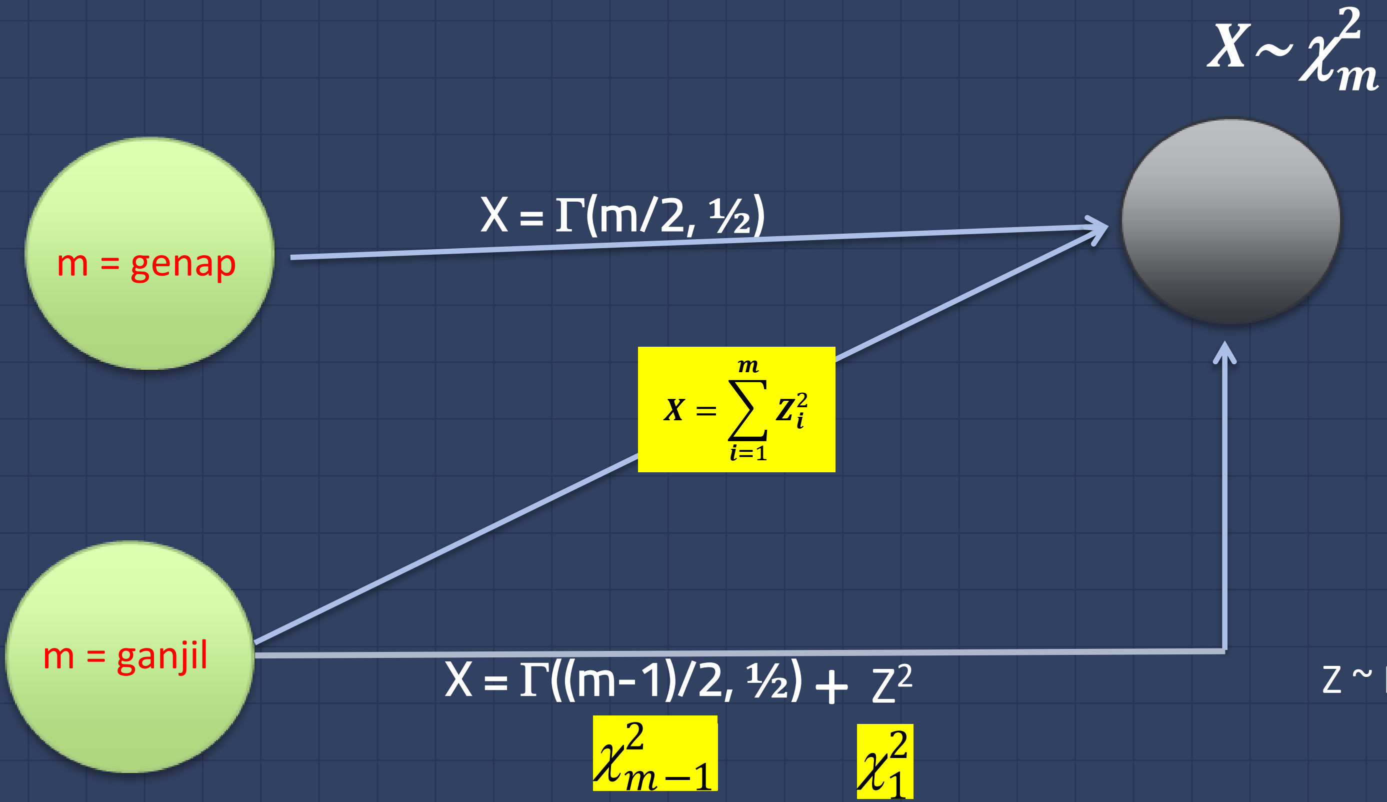
Ragam

$$V(x) = \alpha\beta^2$$

Peubah Acak Gamma



Peubah Acak Khi-kuadrat



Z ~ Normal (0,1)

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

1. Dalil Limit Pusat
2. Metode Box-Muller
3. Metode Polar Marsaglia

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Dalil Limit Pusat

$U_1, U_2, \ldots, U_n \sim \text{Uniform}(0,1)$ → $X = \sum_{i=1}^{n} U_i \sim \text{Normal}(\mu, \sigma^2)$

$E(U_i) = 1/2$

$Var(U_i) = 1/12$

$\xrightarrow{??}$ $X = \sum_{i=1}^{12} U_i - 6 \sim \text{Normal}(0, 1)$

$\downarrow ??$

$X = \sum_{i=1}^{??} U_i - ?? \sim \text{Normal}(\mu, \sigma^2)$

Dr. Ir. Erfiani, M.Si

See you next week

Next Week ....

Pembangkitan Peubah Acak

Normal

2 Metode Box-Muller ?

3 Metode Polar Marsaglia ?

t-student

Poisson

Simulasi Transformasi Peubah Acak

Dr. Ir. Erfiani, M.Si

Terima Kasih

Pembangkitan Bilangan Acak Normal

METODE SIMULASI DAN RESAMPLING

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

1. Dalil Limit Pusat

2. Metode Box-Muller — dari diagram kartesius ditransformasi ke diagram polar

3. Metode Polar Marsaglia

Dalil Limit Pusat

Pada limit pusat sebaran normal sering ditemui, dan menjadi dasar banyak teori statistik.

menggunakan dalil limit pusat untuk melakukan simulasi peubah acak normal.

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Dalil Limit Pusat

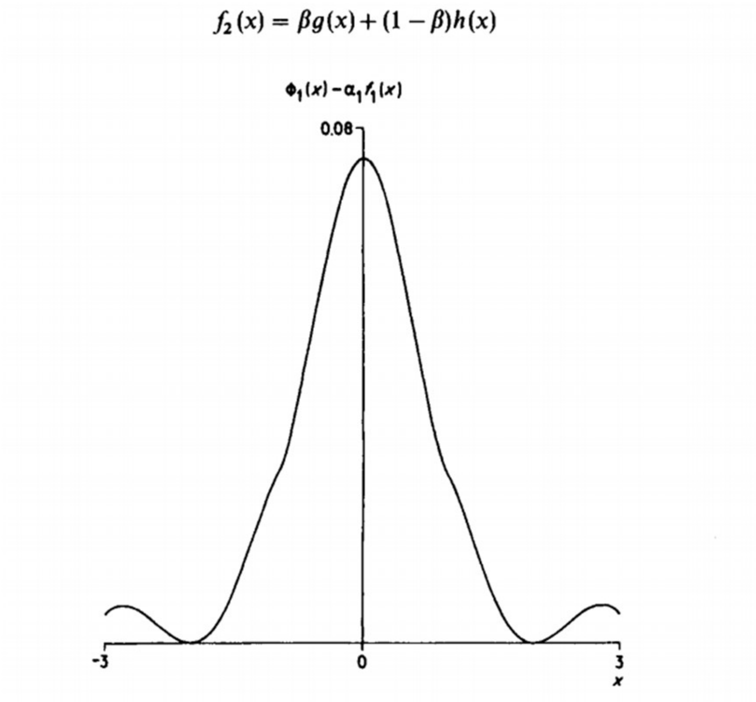
$n \rightarrow \infty$

Berapa besar n?

$U_1, U_2, \ldots, U_n \sim \text{Uniform}(0,1)$ → $X = \sum_{i=1}^{n} U_i \sim \text{Normal}(\mu, \sigma^2)$

tergntung pada kegunaannya, dan seberapa dekat aproksimasi yang diinginkan

Kasus n = 2 tidak cocok, karena
X mempunyai sebaran segitiga,



tetapi untuk n = 3, sebaran N
sudah 'berbentuk lonceng'
dengan baik

Sebaran Seragam (uniform)

Uniform(a, b)

pdf	$f(x\|a,b) = \frac{1}{b-a}, \quad a \le x \le b$
mean and variance	$\mathrm{E}X = \frac{b+a}{2}, \quad \mathrm{Var}\,X = \frac{(b-a)^2}{12}$
mgf	$M_X(t) = \frac{e^{bt}-e^{at}}{(b-a)t}$

$U_1, U_2, \ldots, U_n \sim$ Seragam(0,1)

$$E(U_i) = \frac{a+b}{2} = \frac{0+1}{2} = \frac{1}{2}$$

$$Var(U_i) = \frac{(b-a)^2}{12} = \frac{(1-0)^2}{12} = \frac{1}{12}$$

Sebaran Normal

Normal(μ, σ^2)

pdf	$f(x\|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$
mean and variance	$\mathrm{E}X = \mu, \quad \mathrm{Var}\,X = \sigma^2$
mgf	$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$

$E(X) = \mu$ dan $Var(X) = \sigma^2$

$$misalkan\ X = \sum_{i=1}^{n} U_i + k$$

$$Var(X) = Var\left(\sum_{i=1}^{n} U_i + k\right) = \sum_{i=1}^{n} Var(U_i) = \frac{n}{12} = \sigma^2 \qquad \boxed{n = 12\sigma^2}$$

$$E(X) = E\left(\sum_{i=1}^{n} U_i + k\right) = \sum_{i=1}^{n} E(U_i) + k = \frac{n}{2} + k = \mu \qquad \boxed{k = \mu - \frac{n}{2}}$$

Peubah Acak Normal

$U_1, U_2, \ldots, U_n \sim$ Uniform(0,1) $\rightarrow$ $X = \sum_{i=1}^{n} U_i \sim$ Normal (μ, σ^2)

Paling mudah n=12

$E(U_i) = 1/2$

$Var(U_i) = 1/12$

$\rightarrow$ $X = \sum_{i=1}^{n} U_i - 6$ ~ Normal (0, 1)

??

$X = \sum_{i=1}^{??} U_i - ??$ ~ Normal (μ, σ^2)

X~ Normal (0, 1)

$$X = \sum_{i=1}^{n} U_i - 6$$

Y~ Normal (5, 10)

$$n = 12\sigma^2 = 12x10 = 120$$

$$k = \mu - \frac{n}{2} = 10 - \frac{120}{2} = -55$$

$$Y = \sum_{i=1}^{120} U_i - 55$$

$$Y = \sqrt{10}N + 5$$

Buktikan!!

Peubah Acak Normal

Pembangkitan Peubah Acak Normal | Dalil Limit Pusat

Untuk sembarang peubah acak X_i yang saling bebas dan identik, $Y = \Sigma_{i=1}^{n} X_i$ akan menyebar normal untuk n cukup besar

$$\sum_{i=1}^{n} U_i \sim Normal(\mu, \sigma^2)$$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Metode Box-Muller

Metode ini memperoleh peubah acak normal melalui transformasi satu-ke-satu dari **dua peubah acak U(0, 1)**.

Untuk dua peubah acak U(0,1) yang saling bebas, U_1 dan U_2,

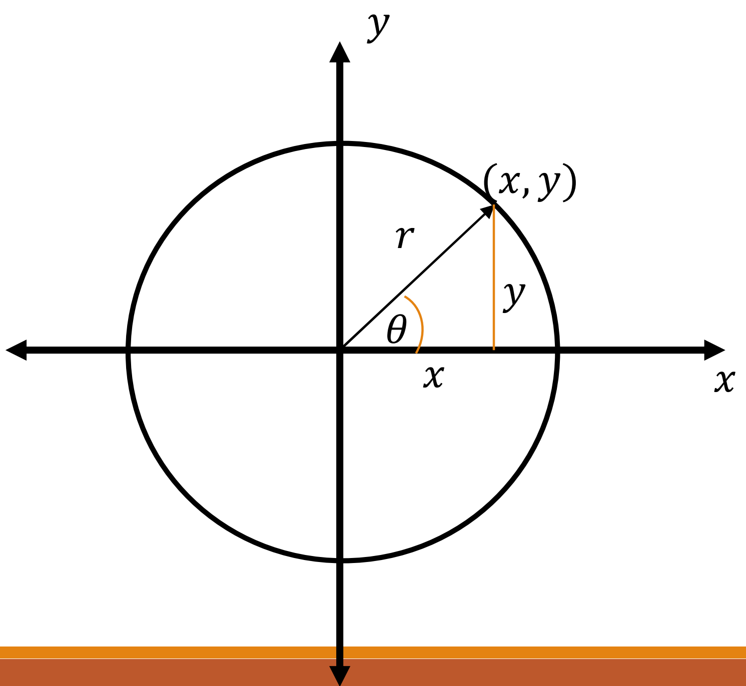
$$N_1 = (-2 \log_e U_1)^{1/2} \cos(2\pi U_2)$$
$$N_2 = (-2 \log_e U_1)^{1/2} \sin(2\pi U_2)$$

$N_1 \sim N(0,1)$ dan $N_2 \sim N(0,1)$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal — Metode Box-Muller

Jika kita mulai dengan peubah acak independen N(0, 1), X, dan Y , mendefinisikan sebuah titik (x, y) dalam dua dimensi dengan koordinat cartesius, dan kita ubah ke koordinat polar (R,Θ), maka:



- Fungsi sebelumnya bisa dituliskan menjadi:
  - x = r cos θ
  - y = r sin θ

$$r^2 = x^2 + y^2$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\sim\chi^2_{(2)} \qquad \sim\chi^2_{(1)} \qquad \sim\chi^2_{(1)}$$

$\downarrow$ ekuivalen

$$\sim exponential\left(\lambda = \tfrac{1}{2}\right) \xrightarrow[\text{negative log}]{\text{Bentuk}} -2 \log_e U_1$$

U_1~ uniform (0,1)

$r^2 = -2 \log_e U_1 \Rightarrow r = (-2 \log_e U_1)^{1/2}$

$\theta = 2\pi U_2$

U_2~ uniform (0,1)

$r = (-2 \log_e U_1)^{1/2}$

$\theta = 2\pi U_2$

Konversi ke koordinat cartesius

- $X = R \cos \Theta = (-2 \log_e U_1)^{1/2} \operatorname{Cos}(2\pi U_2)$
- $Y = R \sin \Theta = (-2 \log_e U_1)^{1/2} \operatorname{Sin}(2\pi U_2)$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Metode Box-Muller

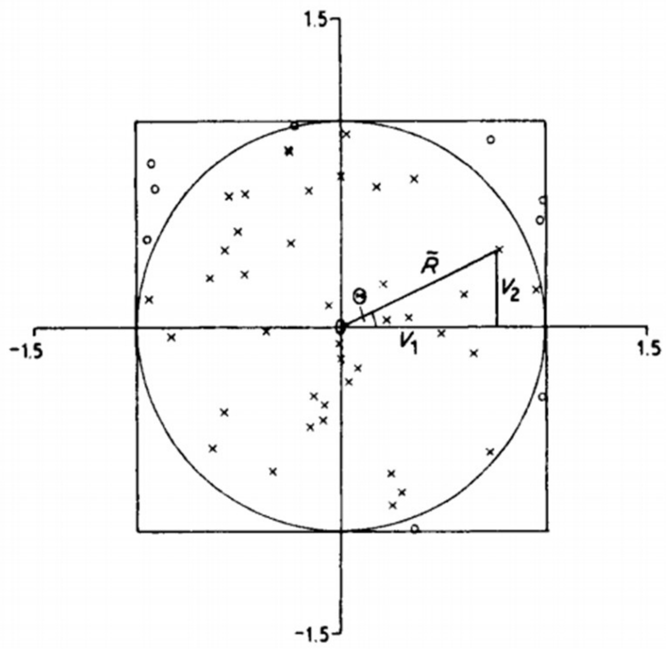
Jika kita mulai dengan peubah acak independen N(0, 1), X, dan Y , mendefinisikan sebuah titik (N1, N2) dalam dua dimensi dengan koordinat cartesius, dan kita ubah ke koordinat polar (R,Θ), maka:

- Fungsi sebelumnya bisa dituliskan menjadi:
  - $X = R \cos \Theta$
  - $Y = R \sin \Theta$
- dengan
  - $R = (-2 \log_e U_1)^{1/2}$ ~ Eksponensial (1/2)
  - $\Theta = 2\pi U_2$ ~ $U(0,2\pi)$

- $(X, Y) \leftrightarrow (R, \Theta)$
- Koordinat Cartesius $\leftrightarrow$ koordinat polar

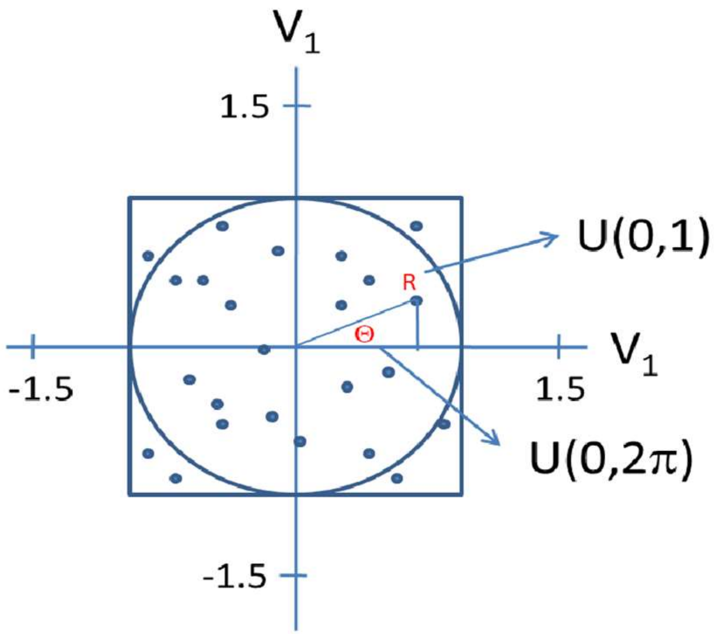
Metode Polar Marsaglia

- Cara menghindari penggunaan fungsi trigonometri adalah dengan membuat sinus dan cosinus sudut-sudut yang menyebar uniform secara langsung tanpa simulasi terlebih dahulu.
- Hal ini dapat dilakukan dengan metode *rejection* sebagai berikut:
  - Jika U~ U(0, 1),
  - maka 2U ~ U(0, 2),
  - dan V = 2U —1 ~ U(-1, 1)
- Jika kita memilih dua peubah acak independen U(-1, 1), V1 dan V2 , maka ini menentukan suatu titik secara acak dalam kuadrat Gambar di samping, dengan koordinat polar (R, Θ) diberikan oleh:
  - $R^2 = V_1^2 + V_2^2$
  - $\tan \Theta = V_2/V_1$

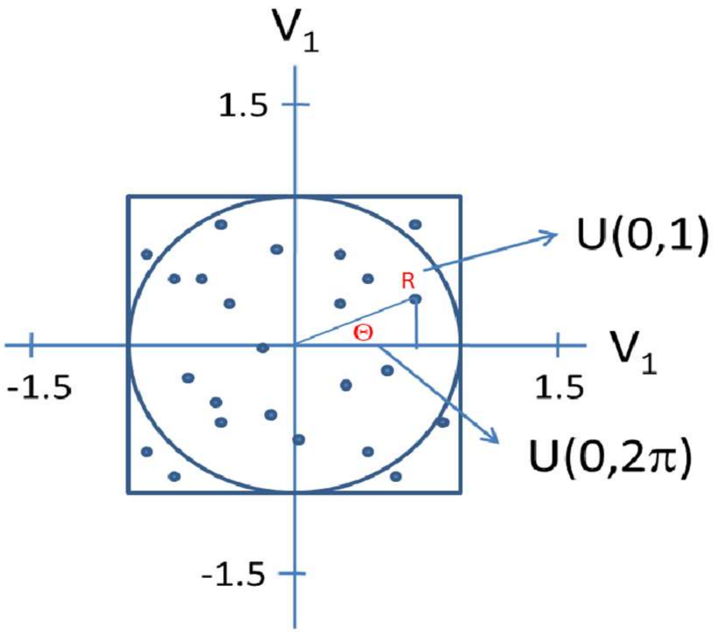


Peubah Acak Normal

Pembangkitan Peubah Acak Normal **Metode Polar-Marsaglia**



- Bila U ~ U(0,1)
  - 2U ~ U(0,2)
  - V = (2U – 1) ~ U(-1,1)
- V_1,V_2 ~ U(-1,1)
  - $R^2 = V_1^2 + V_2^2$
  - $\tan \Theta = V_2/V_1$



- Pemilihan titik-titik tersebut secara berulang-ulang menghasilkan penyebaran titik-titik secara acak di dalam persegi
- penolakan titik-titik di luar lingkaran yang ditunjukkan menyisakan penyebaran titik-titik acak yang seragam di dalam lingkaran.

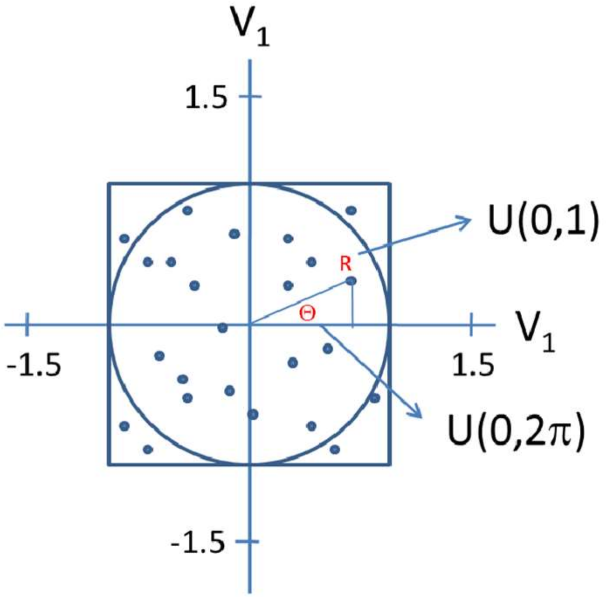
- koordinat kutub R dan Θ adalah peubah acak independen, dan selanjutnya Θ adalah peubah acak U(0, 2π).
- Selain itu R^2~U(0, 1) dan pasangan (R, Θ) adalah yang dibutuhkan oleh metode Box-Muller, dan di sini kita cukup menulis

$$\sin\Theta = \frac{V_2}{R} = V_2(V_1^2 + V_2^2)^{-1/2}$$

$$\cos\Theta = V_1(V_1^2 + V_2^2)^{-1/2}$$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal Metode Polar-Marsaglia



Dari Metode Box-Muller:

- $X = R \cos \Theta = (-2 \log_e U_1)^{1/2} \operatorname{Cos}(2\pi U_2)$
- $Y = R \sin \Theta = (-2 \log_e U_1)^{1/2} \operatorname{Sin}(2\pi U_2)$

R

$$\cos\Theta = \frac{V_1}{R} = \frac{V_1}{\sqrt{V_1^2 + V_2^2}}$$

$$\sin\Theta = \frac{V_2}{R} = \frac{V_2}{\sqrt{V_1^2 + V_2^2}}$$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal Metode Polar-Marsaglia

- $X = (-2 \log_e U_1)^{1/2} \operatorname{Cos}(2\pi U_2)$
- $Y = (-2 \log_e U_1)^{1/2} \operatorname{Sin}(2\pi U_2)$

- $X = (-2 \log_e R^2)^{1/2} V_1 (V_1^2 + V_2^2)^{-1/2}$
- $Y = (-2 \log_e R^2)^{1/2} V_2 (V_1^2 + V_2^2)^{-1/2}$

- $X = (-2 \log_e (V_1^2 + V_2^2))^{1/2} V_1 (V_1^2 + V_2^2)^{-1/2}$
- $Y = (-2 \log_e (V_1^2 + V_2^2))^{1/2} V_2 (V_1^2 + V_2^2)^{-1/2}$

- $X = (-2 \log_e W)^{1/2} V_1 (W)^{-1/2}$
- $Y = (-2 \log_e W)^{1/2} V_2 (W)^{-1/2}$
- $X = V_1 \{(-2 \log_e W)/W\}^{1/2}$
- $Y = V_2 \{(-2 \log_e W)/W\}^{1/2}$

- dengan $W = V_1^2 + V_2^2$

TERIMA KASIH

STK473 – Simulasi Statistika

Pembangkitan Bilangan Acak Normal

Dr. Ir. Erfiani, M.SI

Prodi Statistika dan Sains Data

Fakultas Matematika dan Ilmu Pengetahuan Alam

Institut Pertanian Bogor

IPB University
— Bogor Indonesia —

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

1. Dalil Limit Pusat
2. Metode Box-Muller
3. Metode Polar Marsaglia

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Dalil Limit Pusat

$U_1, U_2, \ldots, U_n \sim \text{Uniform}(0,1) \rightarrow X = \sum_{i=1}^{n} U_i \sim \text{Normal}(\mu, \sigma^2)$

$E(U_i) = 1/2$

$Var(U_i) = 1/12$

$\xrightarrow{??} X = \sum_{i=1}^{12} U_i - 6 \sim \text{Normal}(0, 1)$

$\downarrow ??$

$X = \sum_{i=1}^{??} U_i - ?? \sim \text{Normal}(\mu, \sigma^2)$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Dalil Limit Pusat

Untuk sembarang peubah acak X_i yang saling bebas dan identik, $Y = \mathbf{\Sigma}_{i=1}^{n} X_i$ akan menyebar normal untuk n cukup besar

$$X = \sum_{i=1}^{n} U_i \sim \text{Normal}(\mu, \sigma^2)$$

$U_i \sim \text{Uniform}(0,1)$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Metode Box-Muller

Untuk dua peubah acak U(0,1) yang saling bebas, U_1 dan U_2,

$N_1 = (-2 \log_e U_1)^{1/2} \cos(2\pi U_2)$

$N_2 = (-2 \log_e U_1)^{1/2} \sin(2\pi U_2)$

$N_1 \sim N(0,1)$ dan $N_2 \sim N(0,1)$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Metode Box-Muller

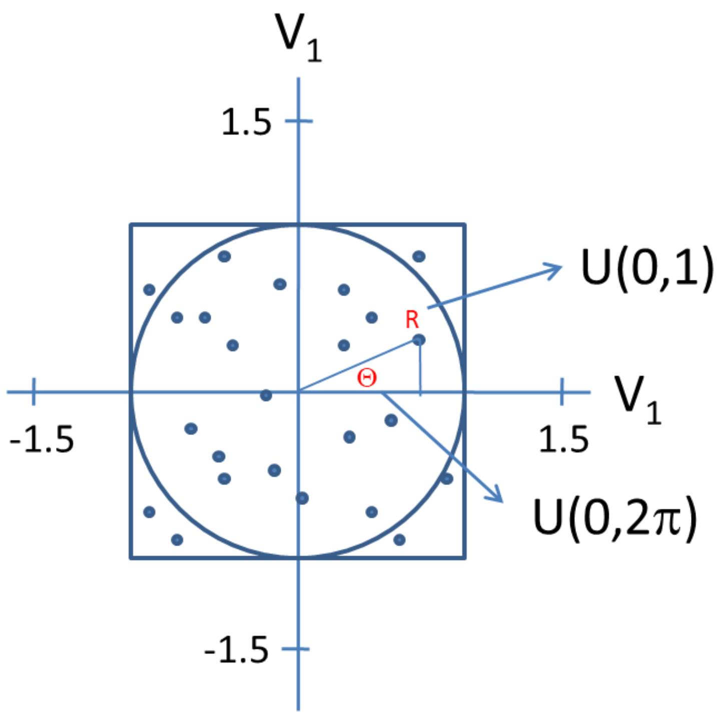
- Fungsi sebelumnya bisa dituliskan menjadi:
  - $N_1 = R \cos \Theta$
  - $N_2 = R \sin \Theta$
- dengan
  - $R = (-2 \log_e U_1)^{1/2}$ ~ Eksponensial (1/2)
  - $\Theta = 2\pi U_2$ ~ $U(0,2\pi)$

- $(N1, N2) \leftrightarrow (R, \Theta)$
- Koordinat Cartesius $\leftrightarrow$ koordinat polar

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Metode Polar-Marsaglia

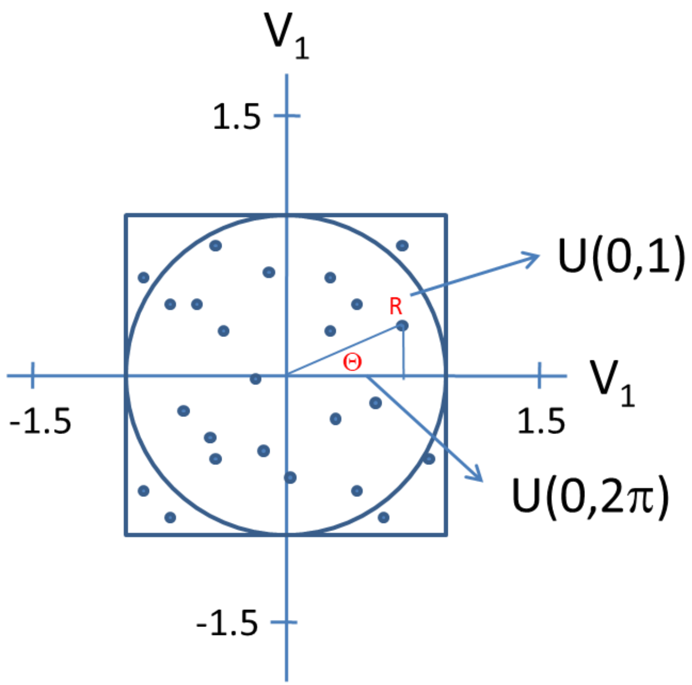


- Bila $U \sim U(0,1)$
  - $2U \sim U(0,2)$
  - $V = (2U - 1) \sim U(-1,1)$
- $V_1, V_2 \sim U(-1,1)$
  - $R^2 = V_1^2 + V_2^2$
  - $\tan \Theta = V_2/V_1$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal

Metode Polar-Marsaglia



- Dari metode Box-Muller
  - $N_1 = (-2 \log_e U_1)^{1/2} \cos(2\pi U_2)$
  - $N_2 = (-2 \log_e U_1)^{1/2} \sin(2\pi U_2)$

R

$\sin \Theta = V_2/R$

$\cos \Theta = V_1/R$

$V_1(V_1^2+V_2^2)^{-1/2}$

$V_2(V_1^2+V_2^2)^{-1/2}$

Peubah Acak Normal

Pembangkitan Peubah Acak Normal Metode Polar-Marsaglia

- $N_1 = (-2 \log_e U_1)^{1/2} \cos (2\pi U_2)$
- $N_2 = (-2 \log_e U_1)^{1/2} \sin (2\pi U_2)$

- $N_1 = (-2 \log_e R^2)^{1/2} V_1(V_1^2+V_2^2)^{-1/2}$
- $N_2 = (-2 \log_e R^2)^{1/2} V_2(V_1^2+V_2^2)^{-1/2}$

- $N_1 = (-2 \log_e (V_1^2+V_2^2))^{1/2} V_1(V_1^2+V_2^2)^{-1/2}$
- $N_2 = (-2 \log_e (V_1^2+V_2^2))^{1/2} V_2(V_1^2+V_2^2)^{-1/2}$

- $N_1 = V_1\{(-2 \log_e W)/W\}^{1/2}$
- $N_2 = V_2\{(-2 \log_e W)/W\}^{1/2}$

- dengan $W = V_1^2+V_2^2$

Dr. Ir. Erfiani, M.Si

See you next week

STK473 – Simulasi Statistika

Metode Pembangkitan Bilangan Acak

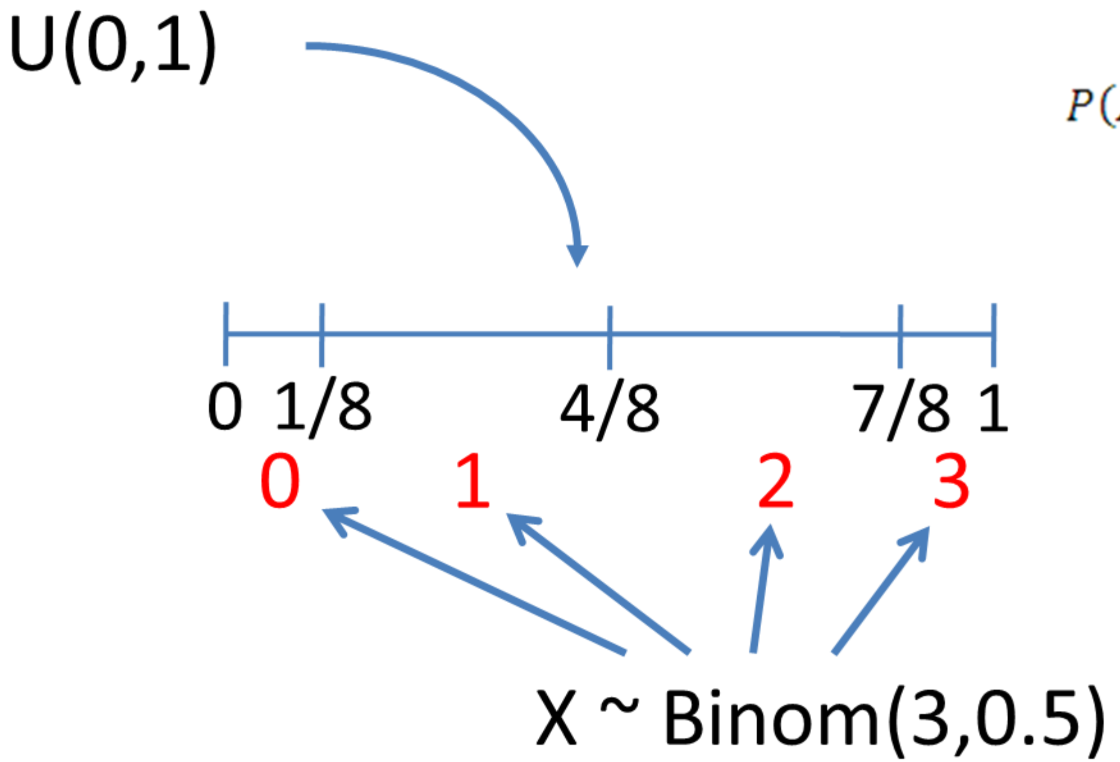
Dr. Ir. Erfiani, M.SI

Prodi Statistika dan Sains Data

Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Pertanian Bogor

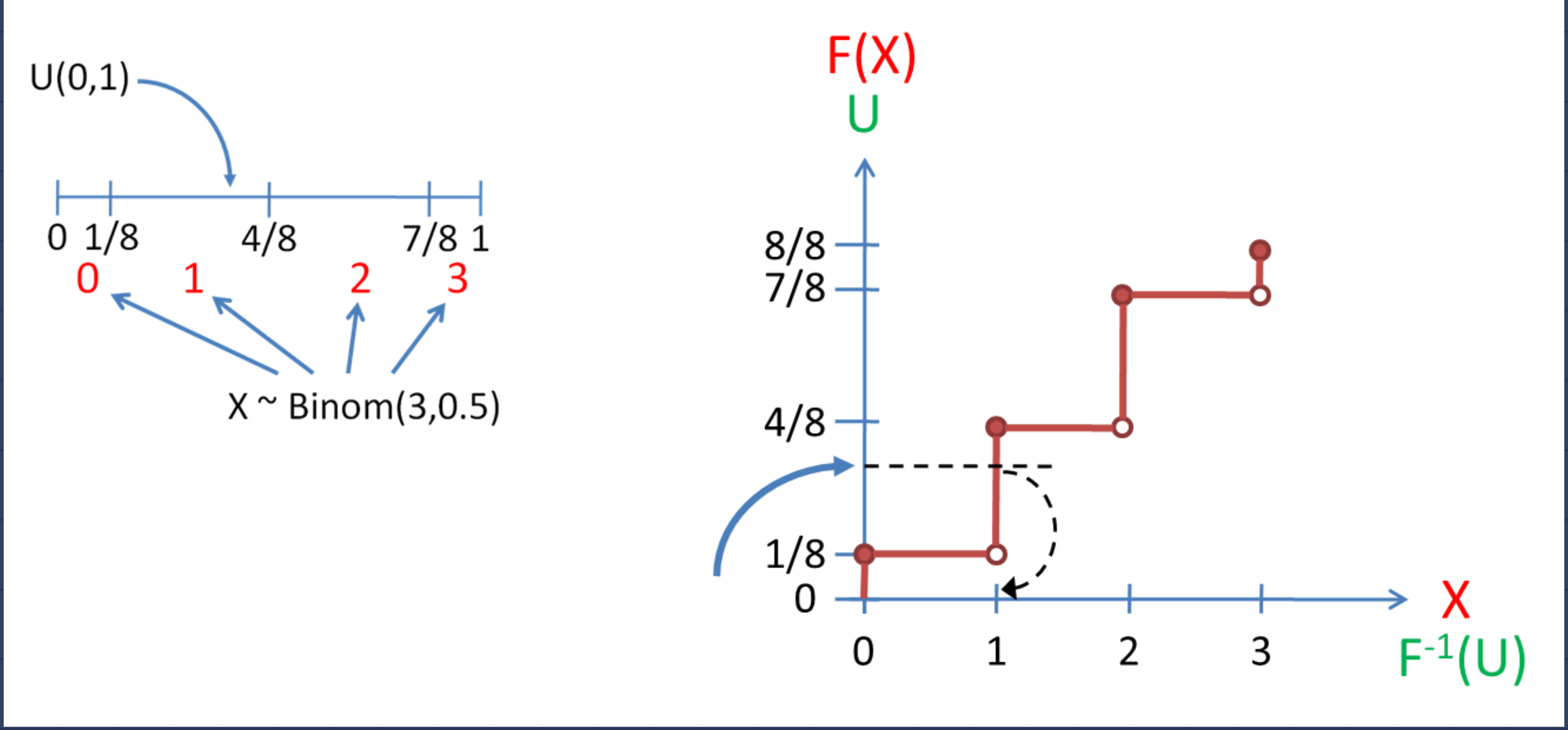
IPB University
— Bogor Indonesia —

Peubah Acak Binomial (n,p)



$$P(X = x) = \begin{cases} 1/8; x = 0,3 \\ 3/8; x = 1,2 \\ 0; x \text{ lainnya} \end{cases}$$

Peubah Acak Binomial (n,p)



Peubah Acak Eksponensial (λ)

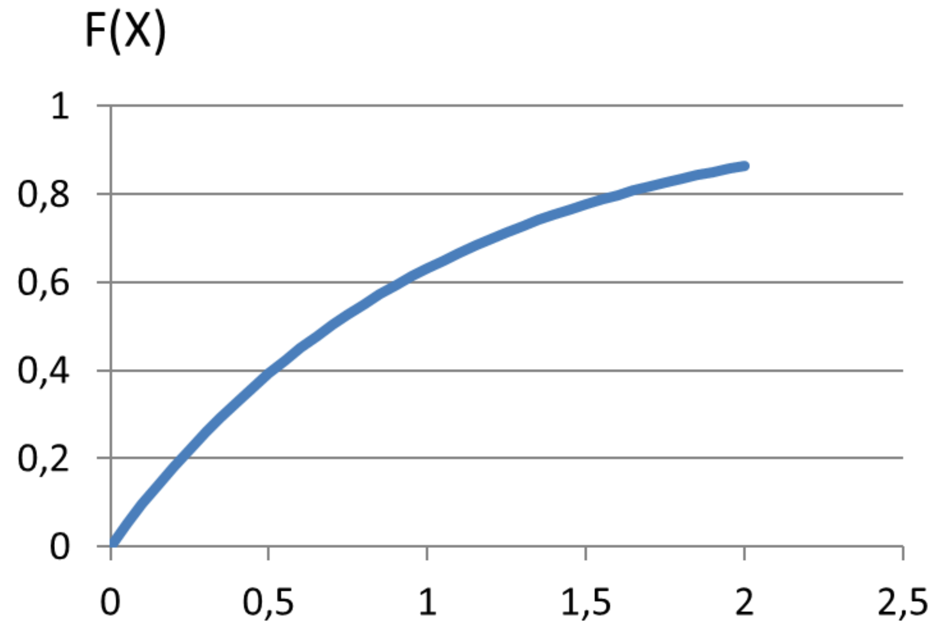
$f(x) = \lambda e^{-\lambda x}; x \geq 0$

$F(X) = 1 - e^{-\lambda x}$

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

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

$\rightarrow X = -1/\lambda \ln(1-U)$

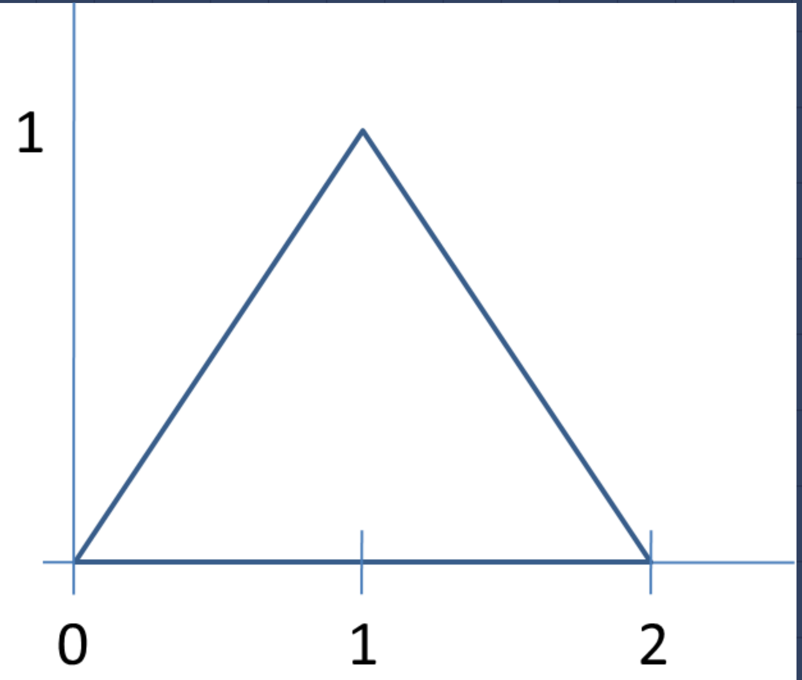


X ~ Exponensial(1)

Metode Tabel look-up atau Metode Kebalikan

Bagaimana dengan peubah acak berikut ini?

$$f(x) = \begin{cases} x;\ 0 \le x \le 1 \\ 2 - x;\ 1 \le x \le 2 \\ 0;\ x \text{ lainnya} \end{cases}$$

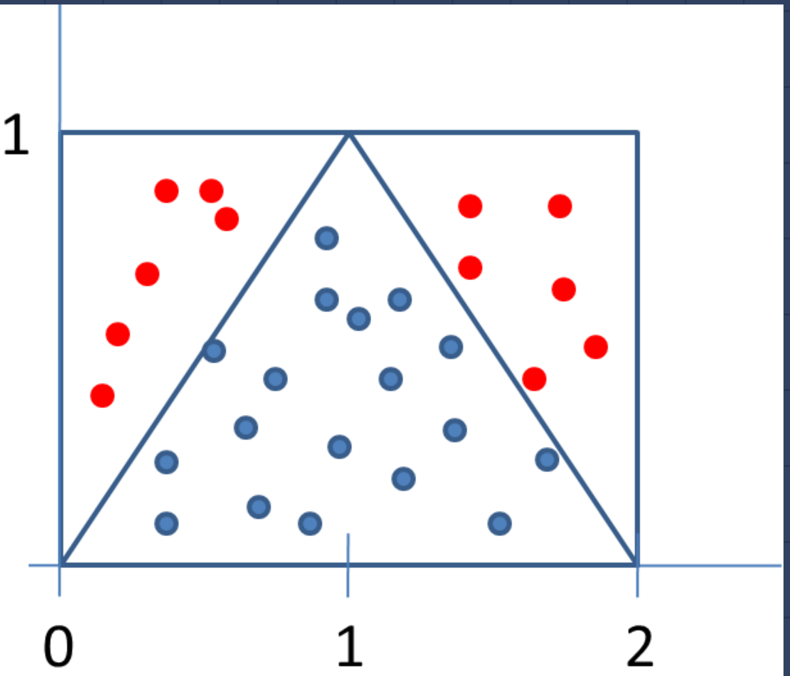


Metode Acceptance-Rejection

Kurva peluang target

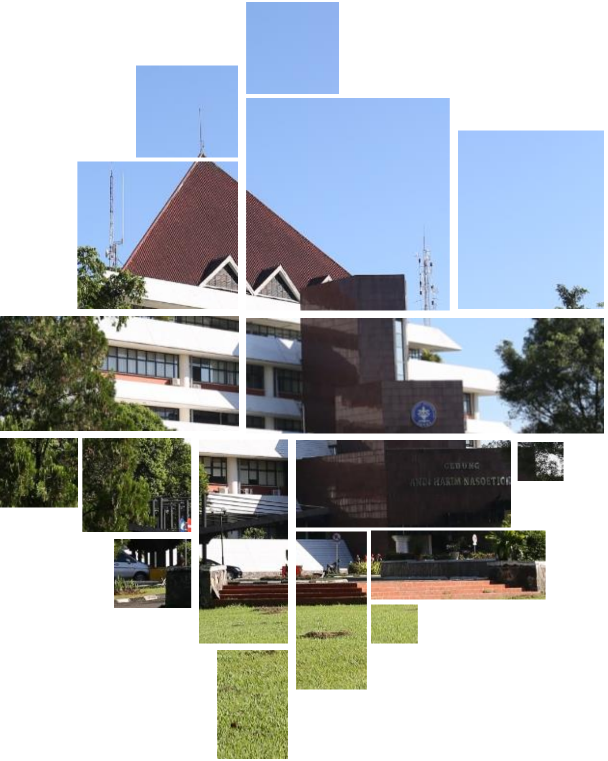
Kurva “bantuan”

- Satu titik dibangkitkan berdasarkan kurva “bantuan”
- Titik bangkitan:
  - ✓ di bawah kurva target → terima
  - ✓ di atas kurva target → tolak



Dr. Ir. Erfiani, M.Si

See you next week



STK473 – Simulasi Statistika

Simulasi Sifat Sebaran Percontohan Statistik

Dr. Ir. Erfiani, M.SI

Prodi Statistika dan Sains Data
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Pertanian Bogor

IPB University
Bogor Indonesia

Sebaran Percontohan

Misalkan sampel diambil dari populasi normal

- Sebaran dari rata-rata
  - Jika σ^2 diketahui maka rata-rata menyebar $N(\mu,\sigma^2/n)$
  - Jika σ^2 tidak diketahui maka rata-rata menyebar t-student dengan derajat bebas n-1
  - Berdasarkan dalil limit pusat, walau σ^2 tidak diketahui asalkan ukuran sampel besar (n>30) maka sebaran dari rata-rata dapat juga diaproksimasi dengan sebaran $N(\mu,s^2/n)$
- Sebaran dari $(n-1)s^2/\sigma^2$
  - $(n-1)s^2/\sigma^2$ menyebar khi-kuadrat dengan derajat bebas n-1
- Sebaran dari $(n_1-1)s_1^2/(n_2-1)s_2^2$
  - $(n_1-1)s_1^2/(n_2-1)s_2^2$ menyebar F dengan derajat bebas pembilang (n_1-1) dan derajat bebas penyebut (n_2-1).

Chi-square/chi-square = sebaran F

DALIL

Limit Pusat

Dalil Limit Pusat (*central limit theorem*)

Jika dari suatu populasi dengan nilai harapan μ dan ragam σ^2 ditarik contoh secara acak berukuran n yang besar maka rata-rata contoh akan:

1. memiliki sebaran yang mendekati normal jika ukuran contoh (n) semakin besar
2. nilai harapan rata-rata contoh adalah μ
3. ragam dari rata-rata contoh adalah σ^2/n

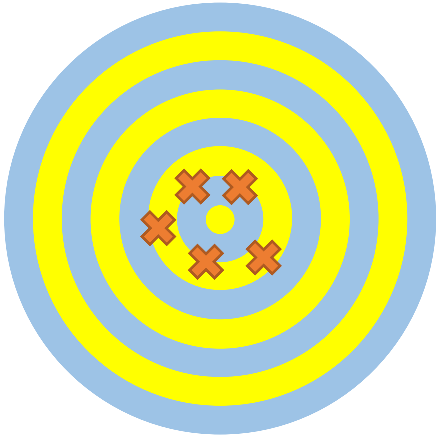
untuk n → ∞

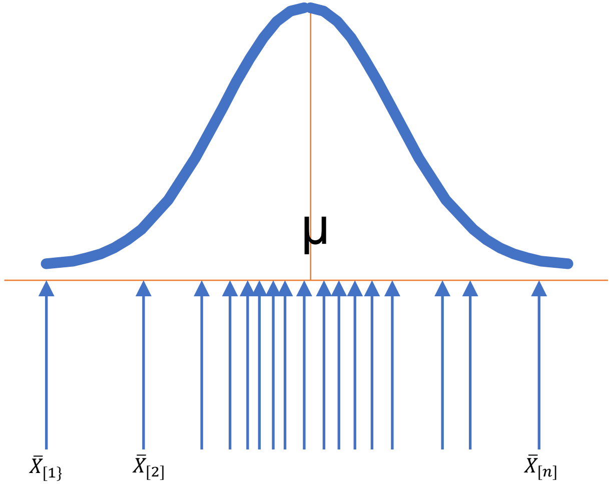
$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

Dalil limit pusat = X menyebar apapun, tapi jika n nya besar maka Xbar akan menyebar normal

nilai harapan rata-rata contoh adalah μ

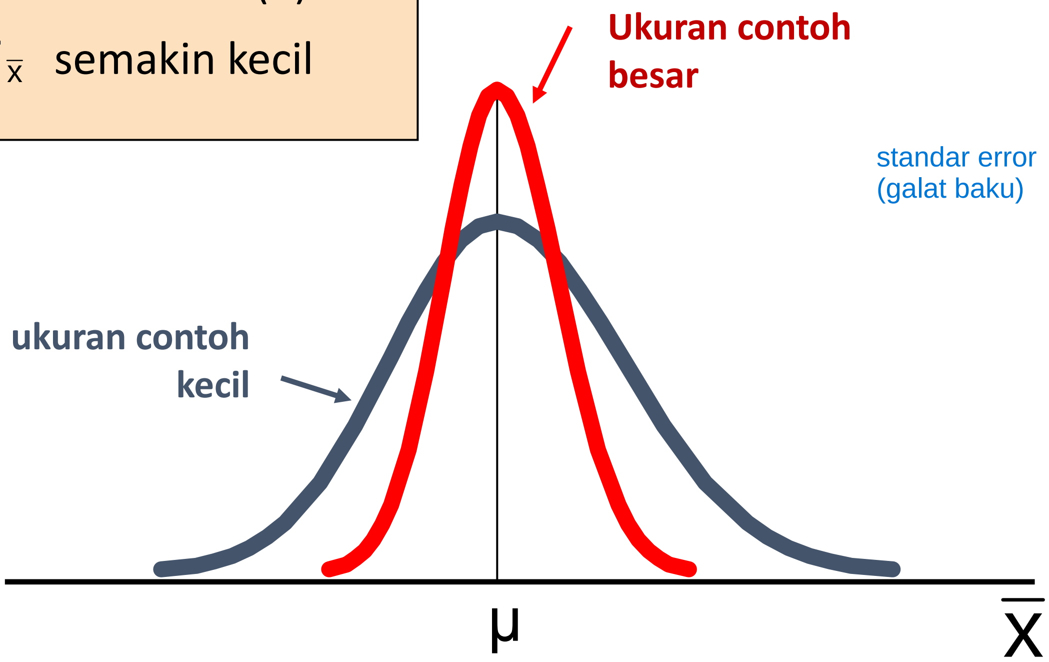
- nilai harapan rata-rata contoh sama dengan nilai harapan populasi
- rata-rata contoh adalah **penduga yang tak bias** bagi rata-rata populasi





Sifat keragaman rata-rata contoh

semakin besar
ukuran contoh (n)
$\sigma_{\bar{x}}$ semakin kecil



$\sigma_{\bar{x}}$ disebut standard error (galat baku)

standar error (galat baku)

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

standar deviasi (simpangan baku)

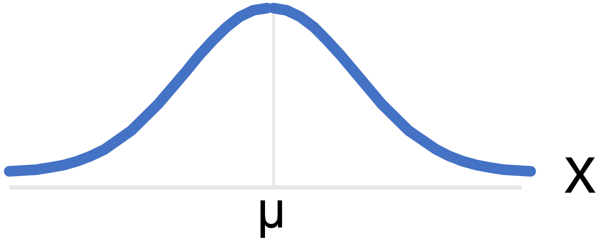
semakin besar n, simpangan rata-rata contoh terhadap μ cenderung lebih kecil

n > 30 besar hanya jika dalam dalil limit pusat
kalau untuk survei n 30 masih kecil, belum optimal

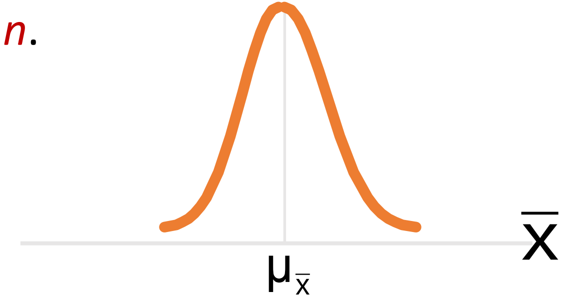
Yang normal itu rata-rata contohnya (xbar), bukan datanya
$\bar{x} \sim N(\mu, \sigma^2/n)$

Bentuk Sampling Distribution dari Rata-Rata Contoh

Jika contoh berasal dari populasi yang menyebar normal $X \sim N(\mu, \sigma^2)$

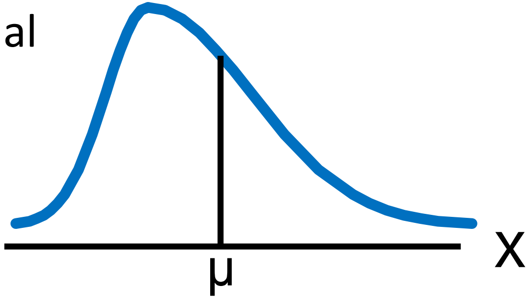


... maka rata-rata contoh akan menyebar normal, berapapun *n*.

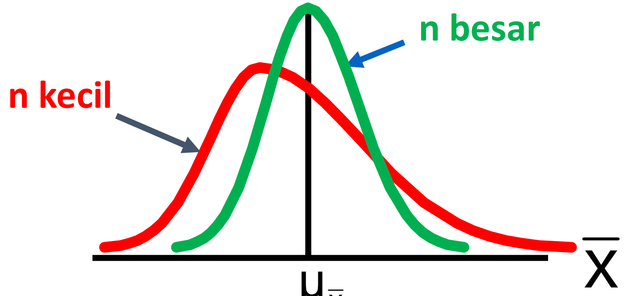


dalil limit pusat

Jika contoh berasal dari populasi yang menyebar tidak normal



... maka rata-rata contoh akan menyebar normal, asalkan *n* besar.

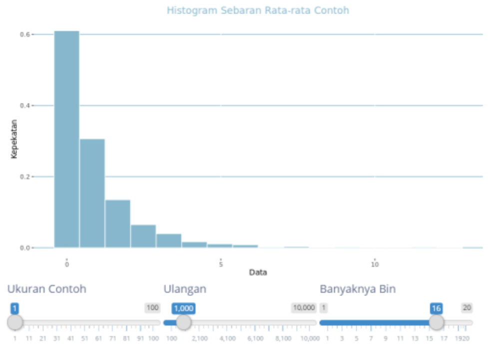


Seberapa besar n? Agar sebaran rata-rata contoh cukup dekat dengan sebaran normal…

- pada umumnya, untuk berbagai bentuk sebaran data populasi kita dapat mencapai itu ketika $n > 25$ (beberapa buku menyebut $n \geq 30$)

- pada sebaran data populasi yang sangat tidak simetris, diperlukan n yang lebih besar lagi

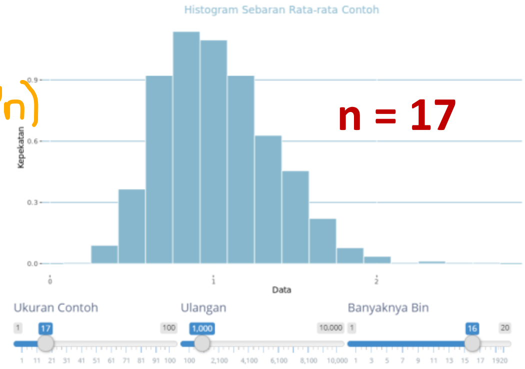
Seberapa besar n? Agar sebaran rata-rata contoh cukup dekat dengan sebaran normal…



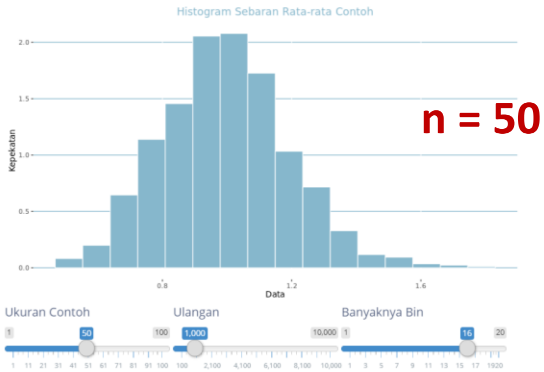
semakin besar nilai n
x nya tetep menyebar eksponensial
tapiii
xbar nya mulai menyebar normal

$\bar{X}_1$
$\bar{X}_2$
$\bar{X}_3$
$\vdots$
$\bar{X}_n$

$\bar{X} \sim N(\mu, \sigma^2/n)$



n = 17



n = 50

Teladan 1.

Tunjukan dengan menggunakan pendekatan simulasi, apakah dalil limit pusat berlaku pada sembarang sebaran populasi, sembarang ukuran populasi, dan sembarang ukuran contoh.

Distribusi t

Dalil Limit Pusat

Jika n besar, maka rata-rata contoh akan mengikuti sebaran normal dengan rata-rata μ dan ragam σ^2/n

diketahui

Syarat :

kondisi σ^2

Tidak diketahui

jika ragam dari populasi tidak diketahui, maka gunakan sebaran t

Sebaran t : σ^2 diduga dengan s^2.

$$\frac{\overline{X}-\mu}{s/\sqrt{n}} \sim \text{t-student db} = n-1.$$

sebaran t lebih bervariasi tergantung besarnya derajat bebas

untuk n yg besar maka akan mengikuti sebaran z atau normal

Teladan 2.

Tunjukan dengan menggunakan pendekatan simulasi

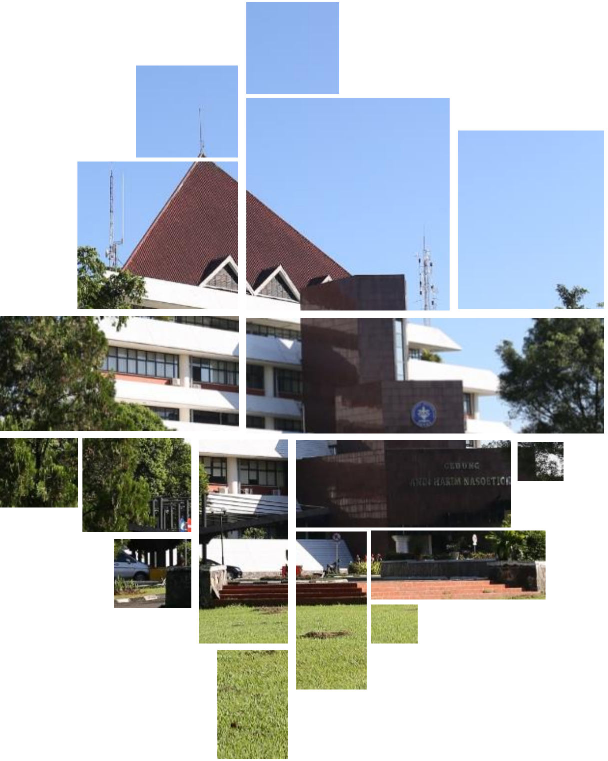
Jika sampel diambil dari populasi normal,
maka sebaran dari rata-rata jika σ^2 tidak diketahui adalah $N(\mu, S^2/n)$

IPB University

— Bogor Indonesia —

Inspiring Innovation with Integrity

in Agriculture, Ocean and Biosciences for a Sustainable World



STK473 – Simulasi Statistika

Berbagai Penerapan Simulasi Statistik

Dr. Ir. Erfiani, M.SI

Prodi Statistika dan Sains Data

Fakultas Matematika dan Ilmu Pengetahuan Alam

Institut Pertanian Bogor

IPB University

Bogor Indonesia

Outline

- Penerapan simulasi untuk melihat hubungan beberapa sebaran peubah acak
- Penerapan simulasi untuk pembuktian sifat statistic
- Penerapan Statistik dalam Pembuktian Dalil/Teorema
- Penerapan simulasi dalam Analisis Data

kalo ada unsur simulasi bukan pake data bangkitan tapi pake data rill dengan sedikit utak atik, maka nama judulnya membandingkan peforma/mencari yg terbaik

tapi kalau hanya mencari yg terbaik maka nama judulnya menerapkan model A dan B, di akhir kesimpulannya bukan model yg terbaik, tapi model yg paling sesuai dengan data saya

Penerapan simulasi untuk melihat hubungan beberapa sebaran peubah acak

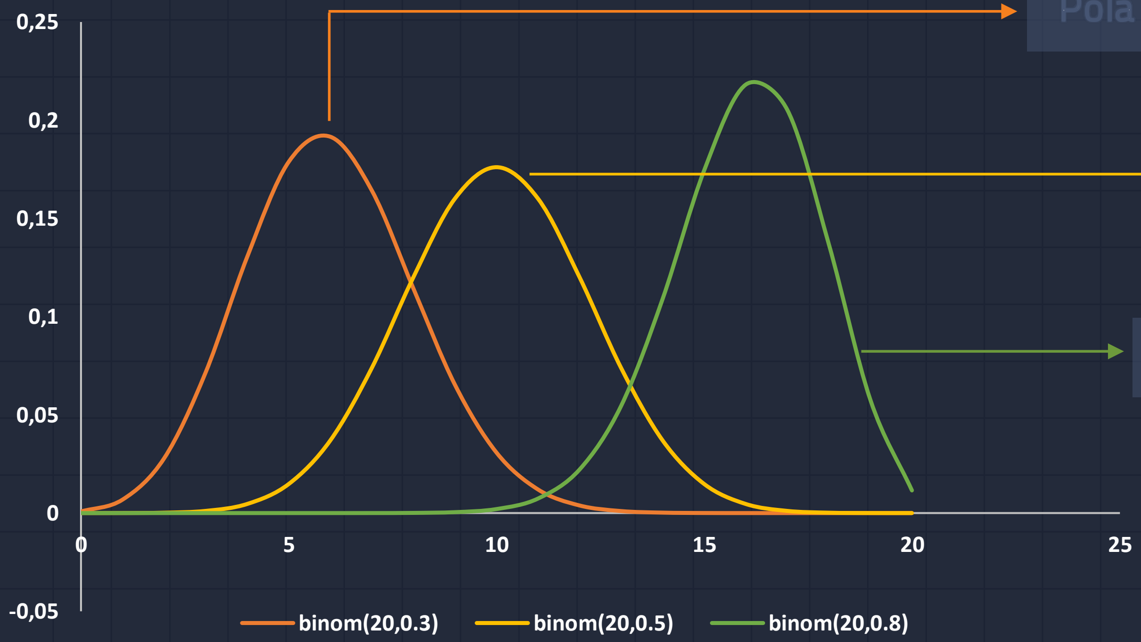
P≈0.5

n
besar??

Sebaran data

ILUSTRASI

X~Binomial(n,p)



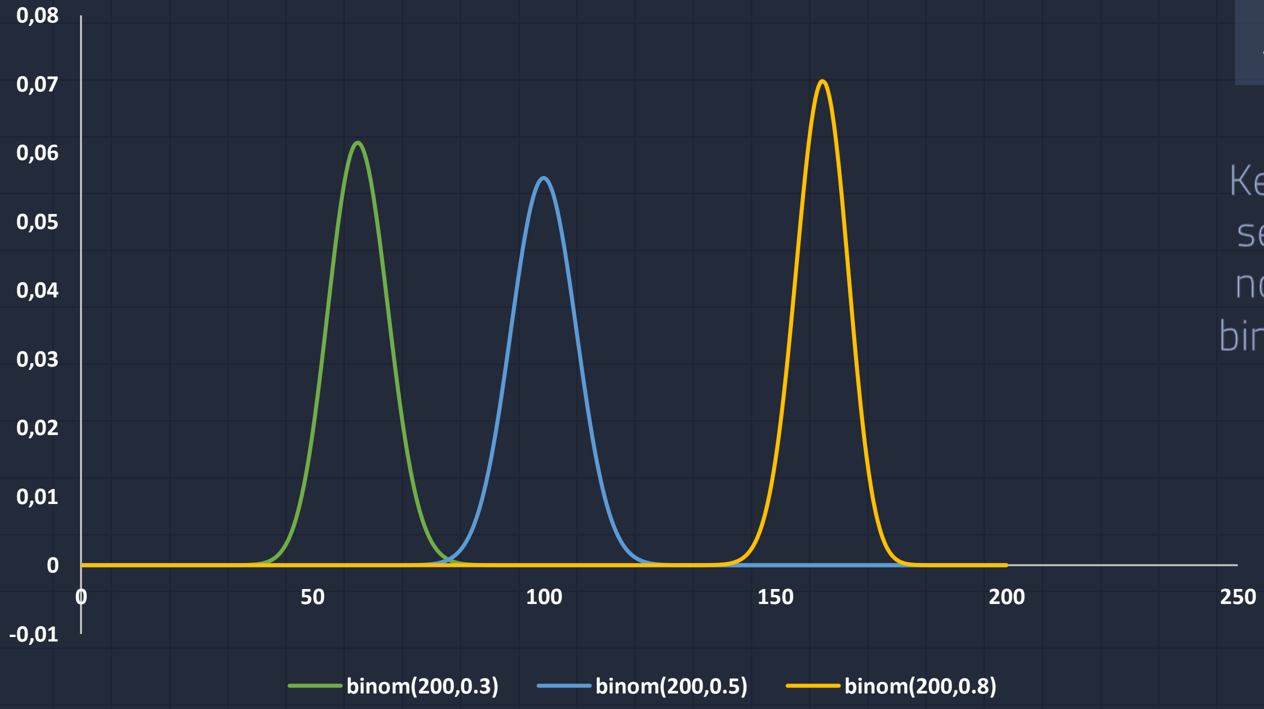
Pola yang terbentuk menjulur kanan

Pola yang terbentuk simetrik

Pola yang terbentuk menjulur kiri

ILUSTRASI

X~Binomial(n,p)



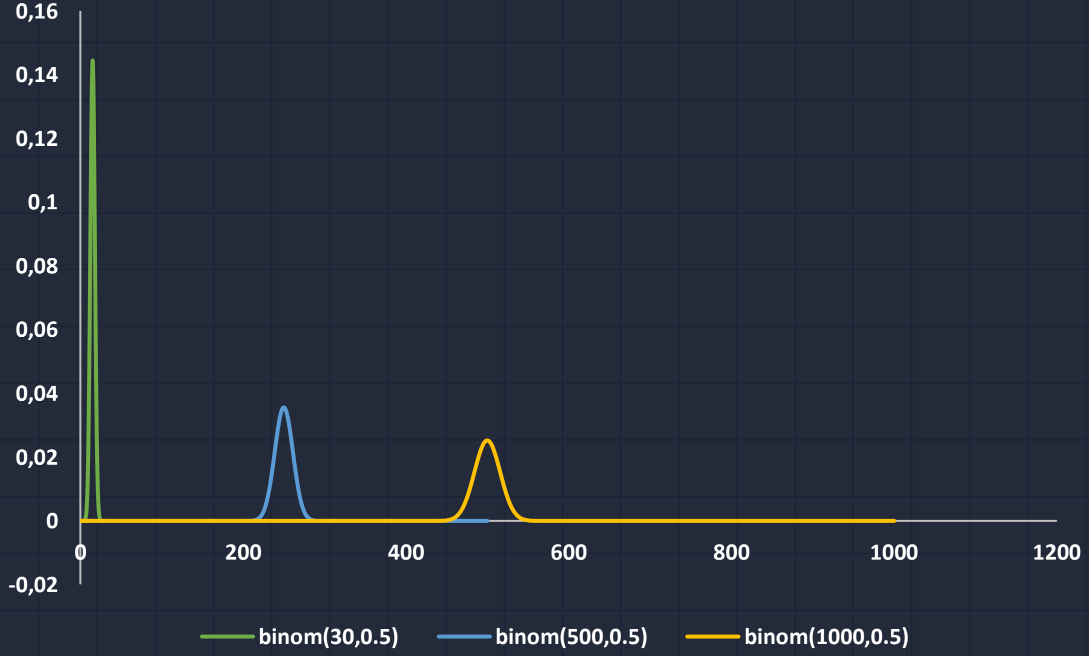
n besar tetap, p berubah

Ketika n semakin besar maka semakin mendekati sebaran normal, namun pada p = 0.5, binomial (n,p) = normal (μ, σ^2)

makin besar p, nilai tengahnya makin geser ke kanan. Ketiganya sama2 simetrik.
Yang oren dan hijau lebih lancip
Biru lebih mirip sebaran normal

ILUSTRASI

X~Binomial(n,p)



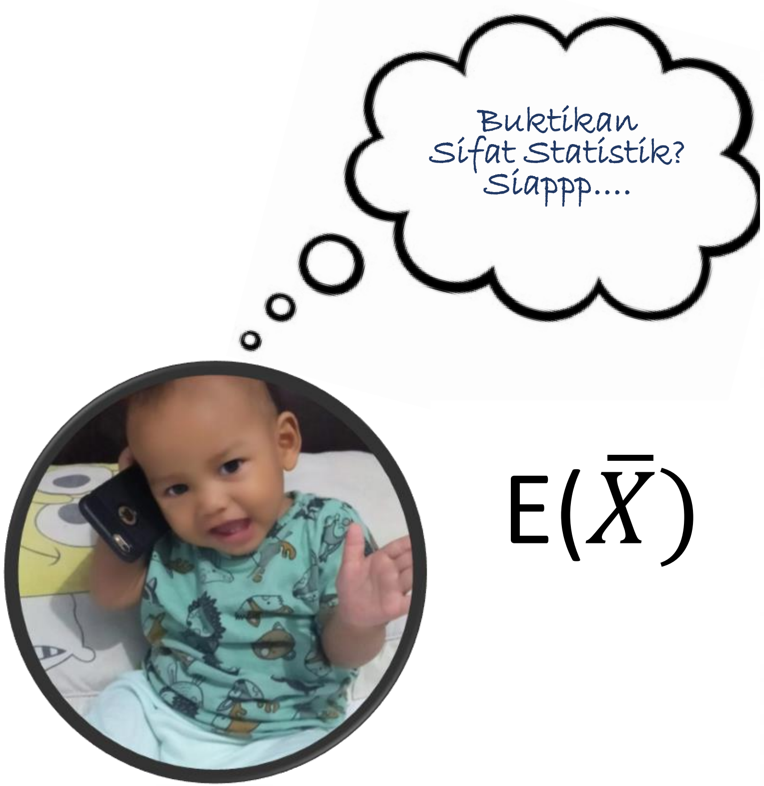
n berubah, p tetap

p = 0.5

Ketika n semakin besar dengan p = 0.5, grafik normal yang terbentuk semakin landai dengan nilai harapan yang bergerak ke kanan dan ragam yang semakin besar

Nilai harapannya bergeser
n terlalu sedikit makin lancip
n makin banyak makin mirip sebaran normal

Penerapan simulasi untuk pembuktian sifat statistic



$E(\bar{X})$

Var(x)

Peubah Acak Binomial Negatif

Peubah Acak Binomial Negatif merupakan representasi dari banyaknya tindakan peubah acak binomial yang saling bebas dan kejadian terakhir sukses

Fungsi Massa Peluang :

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

Nilai Harapan :

$$E(x) = \frac{r}{p}$$

Ragam :

$$V(x) = \frac{r(1-p)}{p^2}$$

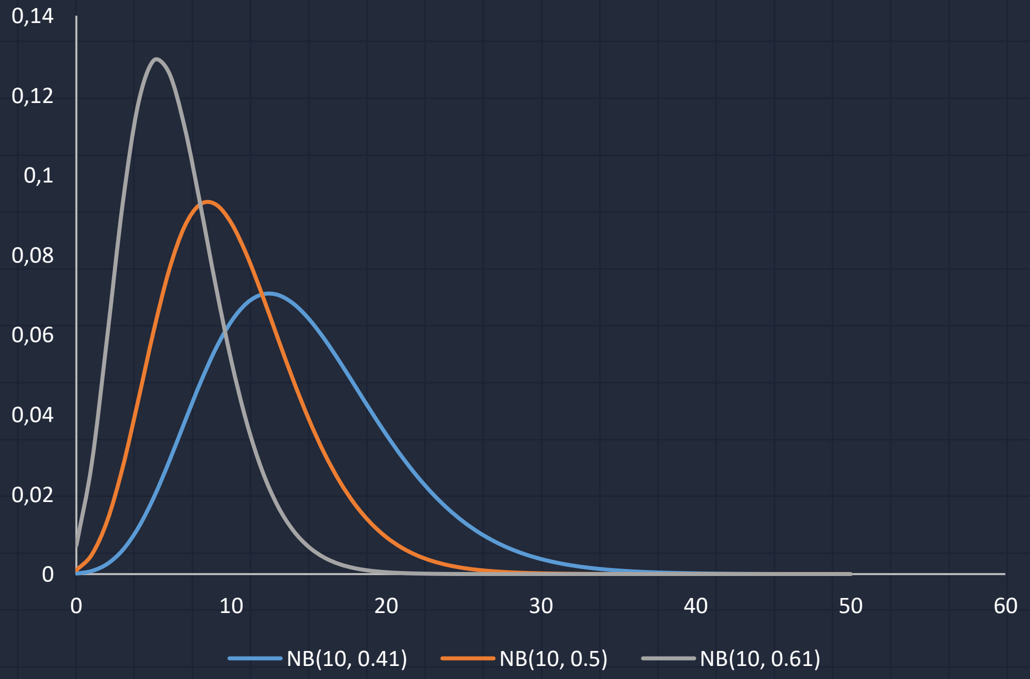
ILUSTRASI

X~Binomial Negatif(r,p)

peluangnya berubah, t tetap

r tetap, p berubah

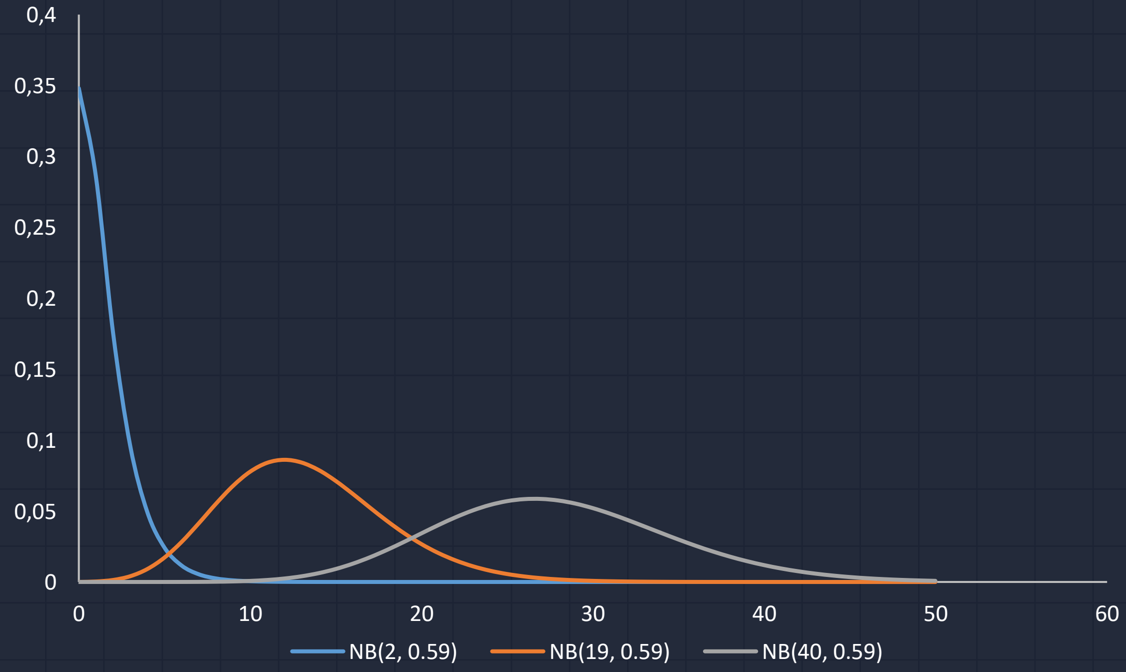
r = 10



Ketika p semakin besar maka kurva semakin curam, menunjukan ragam semakin kecil

ILUSTRASI

X~Binomial Negatif(r,p)



r berubah, p tetap

p = 0.59

Ketika r semakin besar maka kurva semakin landai, menunjukan ragam semakin besar

r semakin besar semakin landai

Penerapan simulasi untuk pembuktian sifat statistic

$\bar{X}$		μ
S^2	**Penduga Berbias**	σ^2
S		σ

Penerapan Statistik dalam Pembuktian Dalil/Teorema

Buktikan
Dalil Limit Pusat
Yuukkk.......

Aim
8 bln

Naureen
3 bln

Dalil Limit Pusat (*central limit theorem*)

Jika dari suatu populasi dengan nilai harapan μ dan ragam σ^2 ditarik contoh secara acak berukuran n yang besar maka rata-rata contoh akan:

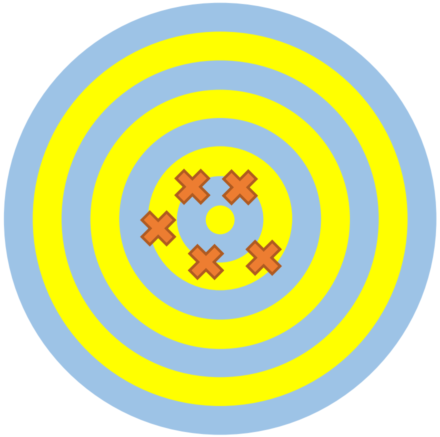
1. memiliki sebaran yang mendekati normal jika ukuran contoh (n) semakin besar
2. nilai harapan rata-rata contoh adalah μ
3. ragam dari rata-rata contoh adalah σ^2/n

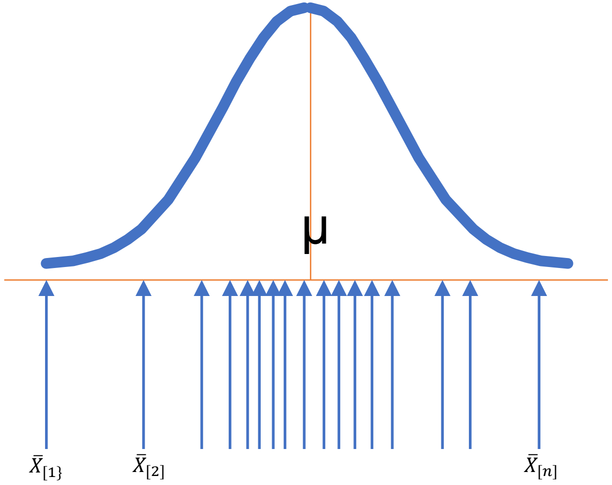
untuk n → ∞

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

nilai harapan rata-rata contoh adalah μ

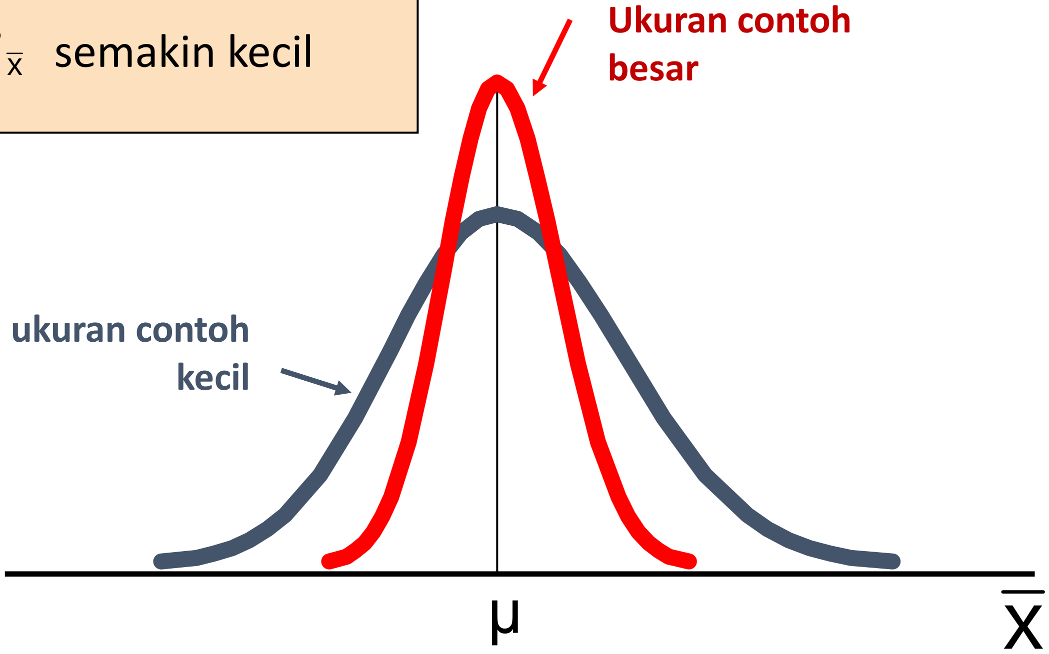
- nilai harapan rata-rata contoh sama dengan nilai harapan populasi
- rata-rata contoh adalah **penduga yang tak bias** bagi rata-rata populasi





Sifat keragaman rata-rata contoh

semakin besar
ukuran contoh (n)
$\sigma_{\bar{x}}$ semakin kecil



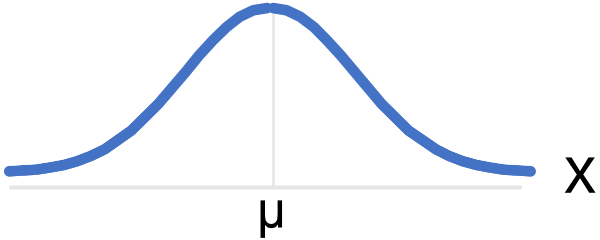
$\sigma_{\bar{x}}$ disebut standard error (galat baku)

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

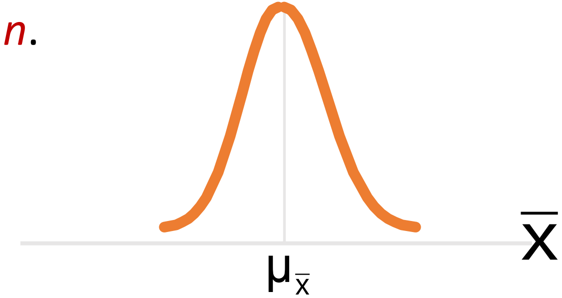
semakin besar n, simpangan rata-rata contoh terhadap μ cenderung lebih kecil

Bentuk Sampling Distribution dari Rata-Rata Contoh

Jika contoh berasal dari populasi yang menyebar normal

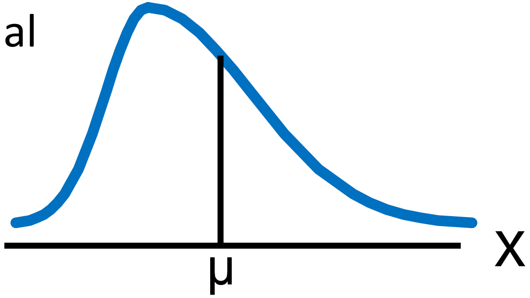


... maka rata-rata contoh akan menyebar normal, berapapun *n*.

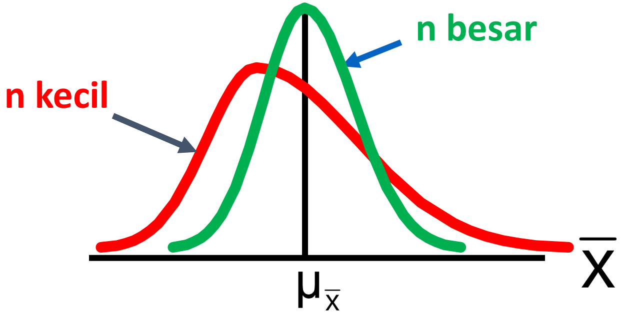


dalil limit pusat

Jika contoh berasal dari populasi yang menyebar tidak normal



... maka rata-rata contoh akan menyebar normal, asalkan *n* besar.



Penerapan Statistik dalam Pembuktian Dalil/Teorema

Langkah awal selalu dimulai dari penyusunan matriks desain simulasi

	Sebaran Simetrik					Sebaran Tidak SImetrik			
	Normal			Uniform	….	…	…	….	….
	σ kecil	σ sedang	σ besar						
n kecil	???								
n sedang									
n besar									

Penerapan simulasi dalam analisis data

- Menggunakan simulasi untuk pendekatan perhitungan integral
- Perhitungan luas daerah tidak beraturan
- Pendugaan Parameter
- Kasus Data Hilang
- Menggunakan simulasi untuk untuk menghitung p-value dan selang kepercayaan
- Menggunakan simulasi untuk menginvestigasi sifat dari prosedur statistik dan pendugaan
- Penerapan simulasi dalam Perancangan Percobaan

p-value daerah yg dibatasi oleh Zhitung

p-value digunakan untuk uji hipotesis di bagian keputusan. Tolak H0 jika p-value < taraf nyata 5%

Uji Hipotesis:
Hipotesis
Statistik uji
kaidah keputusan
Zhit > Zalfa/2
p-value < taraf nyata 5%

Penerapan simulasi dalam analisis data

Materi UAS

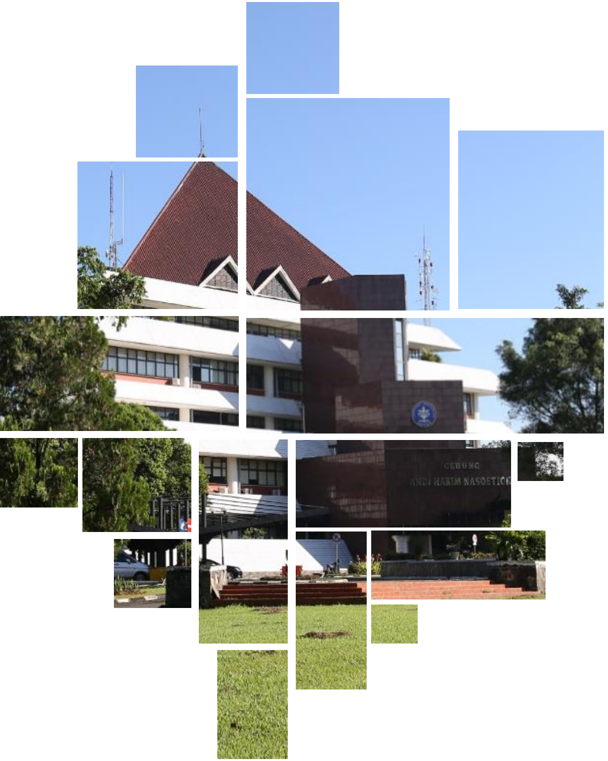
- Penerapan Simulasi dan Resampling Kuasa Uji dan Uji Hipotesis
- Penerapan Simulasi dalam Regresi Linear / Logistik
- Penerapan Simulasi dalam Analisis Data Time Series
- Menggunakan simulasi untuk mendemonstrasikan teorema
- Principal Component

INSTITUT PERTANIAN BOGOR

IPB University

— Bogor Indonesia —

Inspiring Innovation with Integrity
in Agriculture, Ocean and Biosciences for a Sustainable World



STK473 – Simulasi Statistika

Dr. Ir. Erfiani, M.SI

Teknik *Resampling*

Prodi Statistika dan Sains Data
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Pertanian Bogor

IPB University
Bogor Indonesia

Outline

- Simulasi untuk pendekatan perhitungan integral
- Teknik Resampling Bootstrap
- Teknik Resamplin Monte Carlo
- Teknik Resampling Jacknife

Simulasi untuk menghitung integral

Integral berbatas $M = \int_a^b h(x)dx$

dapat ditulis ulang $M = (b-a)\int_a^b h(x)\frac{1}{b-a}dx$

Jika X ~ Seragam(a, b) $\longrightarrow$ $f(x) = \frac{1}{b-a}$ untuk $a \le x \le b$

$$M = (b-a)\int_a^b h(x)f(x)dx = (b-a)E(h(X))$$

simulasi untuk menghitung integral

$$M = \int_a^b h(x)dx$$

luasan daerah yg dicari

$$M = (b-a)\int_a^b h(x)f(x)dx = (b-a)E(h(X))$$

X ~ Seragam(a, b)

$$E(Y) = E(g(X)) = \int_X g(x)f(x)dx$$

Untuk n sangat besar

$$\hat{E}(Y) = \frac{1}{n}\sum_{i=1}^{n} g(x_i)$$

Jadi, M dapat didekati dengan

$$\hat{M} = \frac{b-a}{n}\sum_{i=1}^{n} h(x_i)$$

untuk n yg sangat besar

dengan x1, ..., xn adalah bilangan acak Seragam(a,b)

simulasi untuk menghitung integral

Prosedur menghitung hasil pengintegralan

$$M = \int_a^b h(x)dx$$

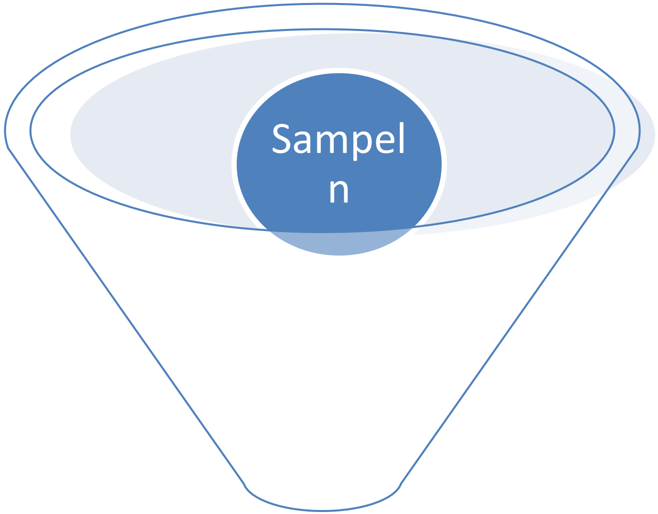
- Bangkitkan n buah peubah acak Seragam(a, b), beri nama x1, x2, …, xn
- Hitung nilai h(x1), h(x2), …, h(xn)
- Hitung rata-rata nilai h
- M didekati dengan nilai rata-rata h dikalikan (b-a)

TEKNIK *RESAMPLING BOOTSTRAP*

Ukuran n kecil mengakibatkan beberapa teori peluang dan statistika inferensia tidak terpenuhi.

Ukuran n kecil juga menjadi permasalahan dalam suatu pendugaan parameter yang berkaitan dengan tingkat keakuratan dan ketelitian

Resampling = solusi agar n membesar



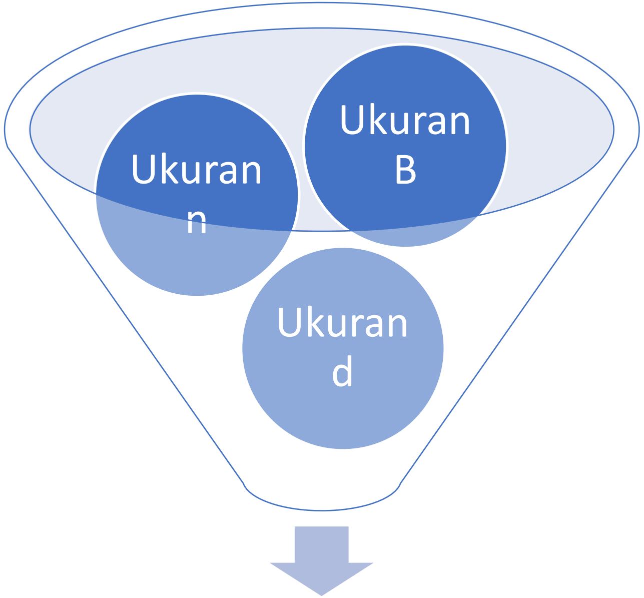
Pendugaan parameter?

Pengujian Hipotesis?

n kecil

Tidak ada informasi sebaran populasi

- Salah satu solusi untuk permasalahan tersebut adalah Metode Resampling Bootstrap
- Ide dasar dari bootstrap adalah membangun data semu dengan menggunakan informasi data asli.
- Metode ini menjadi sebuah metode yang efisien untuk mendapatkan hasil yang robust



RESAMPLING

Note :
n=ukuran data contoh
B=ulangan *Bootstrap*
d=ukuran contoh *Bootstrap*

Penentuan n, d, B

Desain simulasi

1. Data populasi
2. Data contoh (n) : kecil, sedang, besar
3. Ukuran contoh Bootstrap: $n<d$, $n=d$, $n>d$

   (d)=0.5n,0.75n,n,1.5n,2n

   lebih kecil sama dengan lebih besar
4. Pengulangan Bootstrap (B)=50,150,250,500

Tabulasi ukuran contoh Bootstrap pada berbagai ukuran data (n)

n	d				
	d<n		d=n	d>n	
	50%	75%	100%	150%	200%
10	5	8	10	15	20
15	8	11	15	23	30
20	10	15	20	30	40
30	15	23	30	45	60
50	25	38	50	75	100
100	50	75	100	150	200

B = repetisi pengambilan
d = banyak data yg diambil
n = banyak data awal

datanya diambil berulang kali, seperti mengambil bola dengan pengembalian. Jadi ada yg terambil lebih dari satu kali, ada juga yang sama sekali tidak terambil.

agar hasilnya tidak bias kita ulang berkali2 (sebanyak B). Setelah itu baru bisa dicari nilai harapan/ragamnya.

Simulasi Monte Carlo

Simulasi Monte Carlo

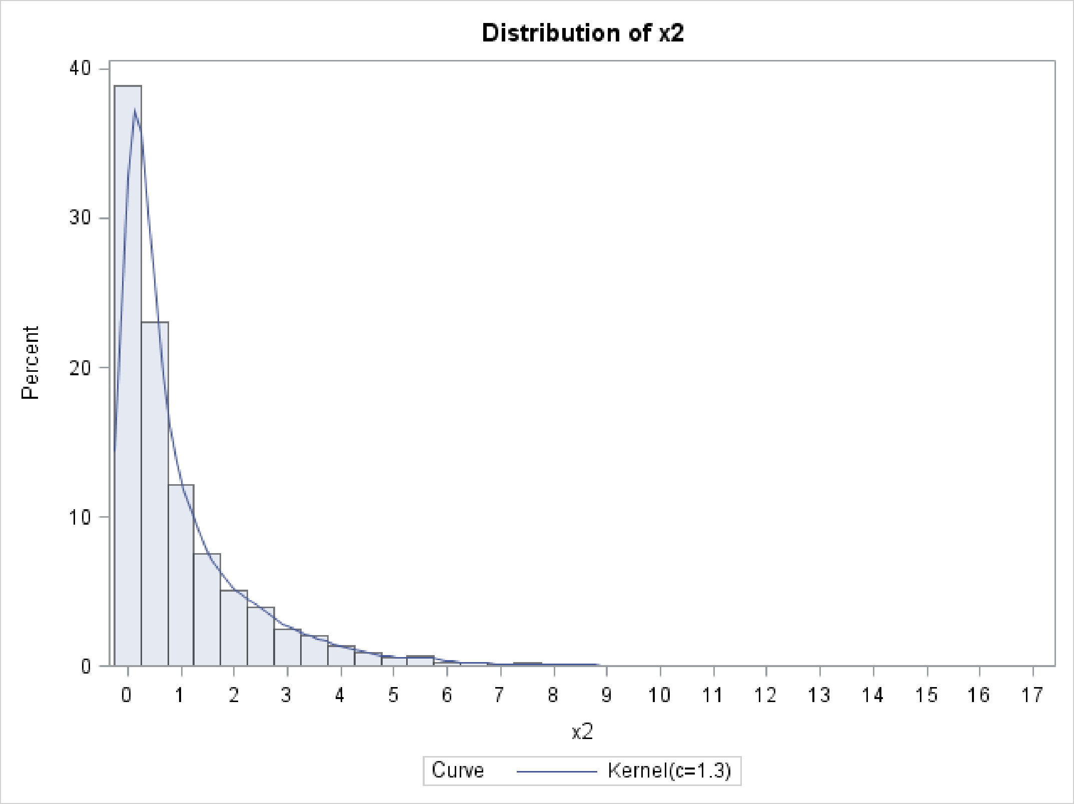
- Simulasi yang memanfaatkan informasi mengenai sebaran data yang **diketahui (dihipotesiskan, dianggap tahu)** dengan pasti. dianggap ikut menyebar apa dengan pasti
- Simulasi didasarkan pada pembangkitan bilangan acak dari sebaran hipotetik.
- Perilaku yang ingin dipelajari dapat diketahui jika proses diulang berkali-kali.
- Banyak digunakan untuk "mengetahui" sebaran dari statistik

Ilustrasi (1)

Bila X menyebar Normal(0, 1), bentuk sebarannya X^2 apa ya?

1. Bangkitkan X dengan sebaran Normal(0, 1)
2. Hitung dan simpan X^2
3. Ulangi 1 dan 2 sebanyak 10000 kali
4. Buat histogram dari X^2.

Ilustrasi (1)



Distribution of x2
Percent
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
x2
Curve
Kernel(c=1.3)

Ilustrasi (2)

Jika X menyebar Normal(0, 1)

Berapa $P(X^2 > 4)$

Menghitung menggunakan simulasi:

1. Bangkitkan X dengan sebaran Normal(0, 1)
2. Hitung dan simpan X^2
3. Ulangi 1 dan 2 sebanyak 10000 kali
4. Hitung berapa persentasi nilai X^2 yang lebih dari 4

Latihan

Andaikan contoh acak berukuran 15 diperoleh dari populasi yang menyebar Poisson.

3	8	4	2	5	5	6	7
4	1	3	3	4	4	3	

- Ujilah H_0: $\lambda = 5$
- Andaikan digunakan statistik uji
- Apa kesimpulan dari pengujian?

$$t = \left| \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \right|$$

Simulasi Jacknife

Refferensi:

Resampling Methods

A Practical Guide to Data Analysis

Third Edition

Phillip I. Good

Elements of Simulation

Byron J.T Morgan

Chapman and Hall

IPB University

— Bogor Indonesia —

Inspiring Innovation with Integrity

in Agriculture, Ocean and Biosciences for a Sustainable World

RUN TEST

Pika Silvianti

randomness

keacakan (randomness) data dari suatu sampel merupakan syarat yang harus dipenuhi dalam pengambilan sampel suatu populasi.

Oleh karena itu, diperlukan uji untuk mengetahui keacakan suatu data.

Uji Run digunakan untuk menguji apakah sederetan data yang terdiri dari dua kategori tersusun secara acak/random atau sistematik.

- Runtun (run) didefinisikan sebagai suatu urutan lambang-lambang yang sama, yang diikuti serta mengikuti lambang-lambang yang berbeda.

- Fungsi : menguji sederetan data yang terdiri atas dua kategori apakah tersusun secara random atau sistematik.

- Sebuah deretan simbol yang sama disebut satu runs.

Misal: + + - - - + - - - - + + - + Jadi, jumlah *run* = 7

$$\underbrace{+\,+}_{1}\;\underbrace{-\,-\,-}_{2}\;\underbrace{+}_{3}\;\underbrace{-\,-\,-\,-}_{4}\;\underbrace{+\,+}_{5}\;\underbrace{-}_{6}\;\underbrace{+}_{7}$$

For example, in simple data MMKMKKM there are 5 *runs* .

1st *run = MM*

2nd *Run = K*

3rd *Run = M*

4th *Run = KK*

Prosedur

1. Tentukan hipotesis nol (H0), hipotesis alternatif (H1) dan taraf signifikansinya (α)
2. Susunlah observasi-observasi n1 dan n2 menurut urutan terjadinya; kemudian hitunglah banyaknya run (r) berdasarkan pengamatan yang terjadi;
3. Hitunglah kemungkinan di bawah H0 yang dikaitkan dengan suatu nilai yang se-ekstrem r yang diobservasi. Jika probabilitas (p-value) itu sama atau kurang dari α, tolaklah H0. Nilai p-value bergantung pada ukuran n1 dan n2, yakni:
   a) Jika n1 dan n2 ≤20, pakailah Tabel run-F untuk α=0,05.
      - Tabel F1 menyajikan nilai-nilai r yang sedemikian kecilnya sehingga kemungkinannya di bawah H0.
      - Tabel F2 menyajikan nilai-nilai r yang sedemikian besarnya sehingga kemungkinannya di bawah H0
   b) Jika (n1 dan n2 > 20), hitunglah Z, lalu lihat tabel normal.
4. Keputusan: Jika p-value yang didapat sama atau kurang dari α, maka H0 ditolak.
5. Kesimpulan

Uji Run Sampel Kecil (n_1 dan n_2 ≤20)

- Data diubah dalam dua katagori
- Beri tanda kategori 1 dan kategori 2 dengan urutan tetap
- Hitung r (run) dengan urutan yang berbeda , n1 dan n2.
- Signifikansi gunakan tabel F1 dan F2 dan bandingkan tabel F1 dan F2

- Statistik Uji : hitung banyaknya runtun = r

- Tabel F_1 : nilai-nilai batas terkecil r untuk menolak H_0
- Tabel F_2 : nilai-nilai batas terbesar r untuk menolak H_0

- Jika r berada antara F_1 dan F_2 maka terima H_0 dan jika $r < F_1$ atau $r > F_2$ maka tolak H_0

Ilustrasi

Dalam suatu ritel ramayana, terdapat sekelompok karyawan yang sedang istirahat. Dari sekelompok karyawan itu ada 17 orang diambil secara random, selanjut diwawancarai, kapan akan mengambil cuti panjang. Dalam pertanyaan itu disediakan dua alternative jawaban yaitu akan mengambil cuti besar sebelum Idul fitri atau sesudah idul fitri. Wawancara dilakukan secara berurutan, yaitu mulai dari No.1 dan berakhir No.17

https://www.slideshare.net/NurSandy2/uji-run-keacakan

No	Jawaban	Tanda Run
1	1	+
2	1	+
3	1	+
4	0	-
5	1	+
6	0	-
7	0	-
8	1	+
9	0	-
10	1	+
11	0	-
12	0	-
13	1	+
14	0	-
15	0	-
16	0	-
17	1	+

- Keterangan :

1 : Mengambil cuti besar sebelum Idul fitri
0 : Mengambil cuti besar sesudah Idul Fitri

- Apakah data diatas tersusun acak?

Penyelesaian

- Hipotesis

Ho : data tersusun random

H1 : data tidak random

- Tingkat Signifikansi

$\alpha = 5\%$

➢ Dari data diperoleh :

R (run) = 11

n1 = 8

n2 = 9

No	Jawaban	Tanda Run
1	1	+
2	1	+
3	1	+
4	0	-
5	1	+
6	0	-
7	0	-
8	1	+
9	0	-
10	1	+
11	0	-
12	0	-
13	1	+
14	0	-
15	0	-
16	0	-
17	1	+

Nilai Tabel F1 dan F2

n1=8 n2=9

< F1

n1 \ n2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	4	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	8
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14

> F2

n1 \ n2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	16	16	17	18	19	19	20	20	21	21	21	22	22
13						15	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

- $F1 = 4$, $F2 = 15$

- Kriteria uji:

Jika r berada antara F_1 dan F_2 maka terima H_0 dan jika $r < F_1$ atau $r > F_2$ maka tolak H_0.

Karena $F_1 = 4$ dan $F_2 = 15$ maka r berada diantara F_1 dan F_2, sehingga H_0 diterima.

- Artinya data tersebut disusun secara random.

Uji Sampel Besar (n1 atau n2>20)

Untuk sampel besar, distribusi dari sampel R (run) mendekati distribusi normal Z dengan rata-rata μ_r dan varians σ_r^2.

$$Z = \frac{r - \mu_r}{\sigma_r}$$

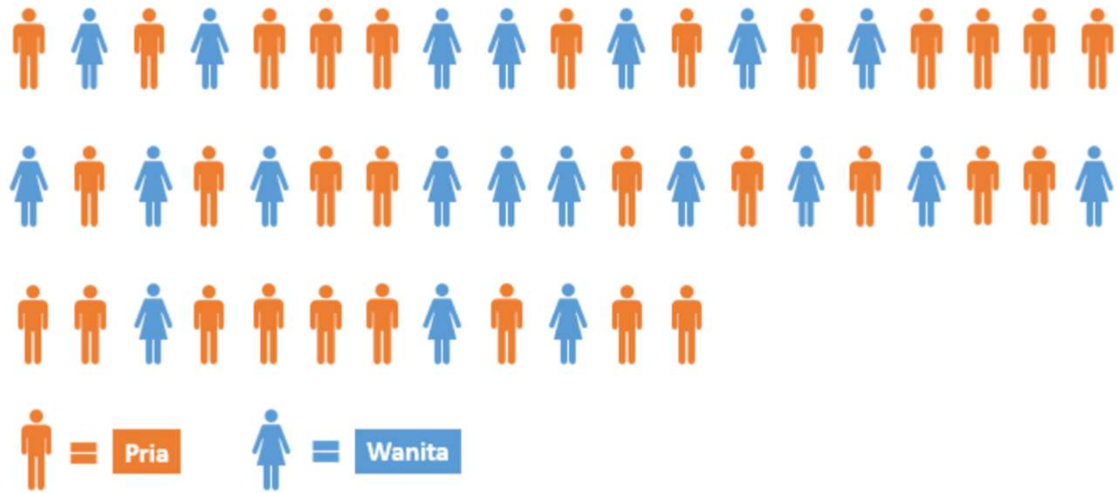
di mana:

$$\mu_r = \frac{2n_1n_2}{n_1 + n_2} + 1$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - N)}{N^2(N - 1)}}$$

Ilustrasi

- Seorang peneliti ingin mengetahui apakah pengaturan antrian pria dan wanita di depan box pembelian tiket teater mengikuti pengaturan random.
- Data diperoleh dengan melakukan tally jenis kelamin dari urutan 50 orang yang menuju box pembelian tiket.
- Urutan dari 30 pria (P) dan 20 wanita (W) adalah sebagai berikut:



Penyelesaian

Hipotesis:
H_0 : Urutan pria dan wanita dalam antrian adalah random
H_1 : Urutan pria dan wanita dalam antrian tidak random

Tingkat Signifikansi: α=0,05
Uji Statistik:
Karena bertujuan untuk menguji kerandoman data yang berasal dari sampel tunggal dan datanya berskala nominal, maka uji yang dipilih adalah uji run sampel tunggal.
Diketahui N = 50, P = 30, W = 20, dan R_{obs} = 35. Untuk sampel besar maka menggunakan pendekatan distribusi normal.

$$Z = \frac{r - \frac{2n_1n_2}{n_1+n_2} + 1}{\sqrt{\frac{2n_1n_2(2n_1n_2-N)}{N^2(N-1)}}}$$

$$= \frac{35 - \frac{2\cdot30\cdot20}{30+20} + 1}{\sqrt{\frac{2\cdot30\cdot20(2\cdot30\cdot20-50)}{50^2(50-1)}}}$$

$$= 2,98$$

Karena H1 tidak memprediksi arah dari kerandoman, maka digunakan uji dua arah.

Sehingga daerah penolakan mencakup nilai $Z_{(\alpha/2)}$ di luar ± 1,96; atau jika (p-value tabel normal x 2) ≤ α_{sig} maka Tolak H_0

Keputusan:

Karena Z_{obs} = 2,98 > $Z_{(\alpha/2)}$ = 1,96; maka tolak H_0. Atau karena (p-value tabel normal x 2) = 0,0014 x 2 ≤ α_{sig} = 0,05; maka tolak H_0.

Kesimpulan: Dengan tingkat signifikansi 5%, dapat disimpulkan bahwa urutan pria dan wanita dalam antrian adalah tidak acak.