Universitas Gadjah Mada
Fakultas Teknik
Departemen Teknik Sipil dan Lingkungan

DISTRIBUSI PROBABILITAS VARIABEL RANDOM

Statistika dan Probabilitas

Distribusi Probabilitas Variabel Random

- Distribusi probabilitas variabel random diskrit
  - Distribusi Hipergeometrik
  - Proses Bernoulli
    - Distribusi Binomial
    - Distribusi Geometrik
    - Distribusi Binomial Negatif
  - Proses Poisson
    - Distribusi Poisson
    - Distribusi Exponensial
    - Distribusi Gamma
  - Distribusi Multinomial
- Distribusi probabilitas variabel random kontinu
  - Distribusi Normal
  - Distribusi t
  - Distribusi Chi-kuadrat (χ^2)
  - Distribusi F

Distribusi Probabilitas Variabel Random Diskrit

Distribusi Hipergeometrik

Distribusi Hipergeometrik

- Situasi
  - Mengambil sampel (random) berukuran n tanpa pengembalian dari suatu populasi berukuran N
  - Elemen-elemen di dalam populasi tersebut terbagi kedalam dua kelompok, masing-masing berukuran k dan $(N - k)$
- Contoh
  - Suatu populasi berupa
    - hari hujan dan hari tak hujan
    - stasiun dengan data baik dan stasiun dengan data jelek
    - sukses dan gagal

Distribusi Hipergeometrik

- Persamaan/rumus
  - Jumlah cara/hasil dari memilih n elemen dari N objek adalah sebuah kombinasi

$$\binom{N}{n} = \frac{N!}{(N-n)!n!}$$

  - Jumlah cara/hasil dari memilih/memperoleh x sukses dan $(n - x)$ gagal dari suatu populasi yang terdiri dari k sukses dan $(N - k)$ gagal adalah:

$$\binom{k}{x}\binom{N-k}{n-x} = \frac{k!}{(k-x)!x!}\,\frac{(N-k)!}{(N-k-n+x)!(n-x)!}$$

Distribusi Hipergeometrik

- Jadi probabilitas mendapatkan $X = x$ sukses dalam sampel berukuran n yang diambil dari suatu populasi berukuran N yang memiliki k elemen sukses adalah

$$f_X(x;N,n,k) = \binom{k}{x}\binom{N-k}{n-x} \Big/ \binom{N}{n}$$

- Distribusi kumulatif dari probabilitas mendapatkan x sukses atau kurang adalah:

$$F_X(x;N,n,k) = \sum_{i=0}^{x} \binom{k}{i}\binom{N-k}{n-i} \Big/ \binom{N}{n}$$

Distribusi Hipergeometrik

- Nilai rerata (mean) suatu distribusi hipergeometrik adalah

$$E(X) = \frac{nk}{N}$$

- Varian

$$\mathrm{Var}(X) = \frac{nk(N-n)(N-n)}{N^2(N-1)}$$

- Catatan:

$$x \le k;\ \ x \le n;\ \ k \le N;\ \ n \le N;\ \ n - x \le N - k$$

Distribusi Hipergeometrik

- Contoh
  - Suatu DAS memiliki 12 stasiun penakar curah hujan dan diketahui bahwa 2 diantaranya dalam keadaan rusak.
  - Manajemen telah memutuskan untuk mengurangi jumlah stasiun menjadi 6 saja.
  - Apabila 6 stasiun dipilih secara acak dari 12 stasiun tersebut, berapakah peluang terpilihnya stasiun rusak sejumlah 2, 1, atau tidak ada sama sekali?

Hypergeometric Distributions

- Penyelesaian
  - populasi, $N = 12$
  - jumlah stasiun rusak, $k = 2$
  - ukuran sampel, $n = 6$
  - peluang (probabilitas) mendapatkan stasiun rusak sejumlah $x = 2, 1, 0$ dalam sampel berukuran $n = 6$ yang diambil dari populasi berukuran $N = 12$ yang memiliki stasiun rusak sejumlah $k = 2$ adalah:

$$f_X(x;N,n,k) = \binom{k}{x}\binom{N-k}{n-x} \Big/ \binom{N}{n}$$

Distribusi Hipergeometrik

$$f_X(x;N,n,k)=\binom{k}{x}\binom{N-k}{n-x}\Big/\binom{N}{n}$$

$$x=2:\quad f_X(2;12,6,2)=\binom{2}{2}\binom{12-2}{6-2}\Big/\binom{12}{6}=0.2273$$

$$x=1:\quad f_X(1;12,6,2)=\binom{2}{1}\binom{12-2}{6-1}\Big/\binom{12}{6}=0.5454$$

$$x=0:\quad f_X(0;12,6,2)=\binom{2}{0}\binom{12-2}{6-0}\Big/\binom{12}{6}=0.2273$$

Distribusi Hipergeometrik

- Ekspektasi jumlah stasiun rusak yang ada di dalam sampel adalah:

$$\mathrm{E}(X) = \frac{nk}{N} = \frac{6 \times 2}{12} = 1$$

- atau

$$M_1 = \sum_{i=0}^{2} x_i \, f_X(x_i) = 0 \times 0.2273 + 1 \times 0.5454 + 2 \times 0.2273 = 1$$

Distribusi Hipergeometrik

$$f_X(x;N,n,k)=\binom{k}{x}\binom{N-k}{n-x}\Big/\binom{N}{n}$$ =HYPGEOM.DIST(x,n,k,N,FALSE)

$$F_X(x;N,n,k)=\sum_{i=0}^{x}\binom{k}{i}\binom{N-k}{n-i}\Big/\binom{N}{n}$$ =HYPGEOM.DIST(x,n,k,N,TRUE)

f_X(2;12,6,2)=HYPGEOM.DIST(2,6,2,12,FALSE) = 0.2273

f_X(1;12,6,2)=HYPGEOM.DIST(1,6,2,12,FALSE) = 0.5454

f_X(0;12,6,2)=HYPGEOM.DIST(0,6,2,12,FALSE) = 0.2273

Distribusi Probabilitas Variabel Random Diskrit

Distribusi Binomial

Contoh Ilustrasi

- Investigasi thd suatu populasi
  - karakteristik populasi → variabel
  - nilai variabel
    - nilai ujian: 0 s.d. 100
    - status perkawinan: tidak kawin, kawin, cerai, duda/janda
    - usia: 0 s.d. ...
    - cuaca: cerah, berawan, hujan

Contoh Ilustrasi

- Contoh lain
  - Jawaban pertanyaan:
    - ya / tidak
    - benar / salah
    - menang / kalah
    - lulus / tak-lulus
    - sukses / gagal

SUKSES
vs
GAGAL

Distribusi Binomial

- Jika
  - variabel hanya memiliki 2 kemungkinan hasil
  - probabilitas (peluang) kedua hasil tersebut tidak berubah (tetap) apapun hasil eksperimen sebelumnya

→ Distribusi Binomial

- Probabilitas hasil suatu distribusi binomial
  - prob(sukses) $= p$
  - prob(gagal) $= q = 1 - p$

Distribusi Binomial

- Ilustrasi
  - Peluang sukses (S) dalam suatu eksperimen adalah $p \rightarrow \text{prob}(S) = p$
  - Peluang gagal (G) adalah $q = 1 - p \rightarrow \text{prob}(G) = q$
  - 1x eksperimen:
    - peluang sukses p
    - peluang gagal q
  - 2x eksperimen:
    - peluang sukses kmd sukses (S,S): pp
    - peluang sukses kmd gagal (S,G): pq
    - peluang gagal kmd sukses (G,S): qp
    - peluang gagal kmd gagal (G,G): qq

Sukses-Gagal dalam 2x Eksperimen

Jumlah sukses	Cara sukses	Jumlah cara sukses	Probabilitas sukses	
2	*SS*	1	*pp*	1 p^2q^0
1	*SG* atau *GS*	2	*pq*+*qp*	2 p^1q^1
0	*GG*	1	*qq*	1 p^0q^2

Sukses-Gagal dalam 3x Eksperimen

Jumlah sukses	Cara sukses	Jumlah cara sukses	Probabilitas sukses	
3	SSS	1	ppp	$1\ p^3q^0$
2	SSG, SGS, GSS	3	$ppq+pqp+qpp$	$3\ p^2q^1$
1	SGG, GSG, GGS	3	$pqq+qpq+qqp$	$3\ p^1q^2$
0	GGG	1	qqq	$1\ p^0q^3$

Sukses-Gagal dalam 3x atau 5x Eksperimen

- 3x eksperimen:
  - peluang sukses pada eksperimen ke-3: qqp
  - peluang sukses di salah satu eksperimen: $pqq + qpq + qqp$
- 5x eksperimen:
  - peluang sukses 2x: $ppqqq + pqpqq + \ldots + qqqpp$

Distribusi Binomial

- Jika
  - peluang sukses p dan peluang gagal $q = 1 - p$
  - probabilitas sukses p tidak berubah apapun hasil eksperimen yang lain
- Maka
  - peluang mendapatkan x kali sukses dalam n kali eksperimen adalah:

$$f_X(x;n,p) = \underbrace{\binom{n}{x}}_{\text{koefisien binomial}} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

koefisien binomial → =COMBIN(n,x)

Distribusi Binomial

- Distribusi binomial dan distribusi binomial kumulatif

$$f_X(x;n,p)=\binom{n}{x}p^x(1-p)^{n-x},\quad x=0,1,2,\ldots,n$$ =BINOM.DIST(*x*,*n*,*p*,FALSE)

$$F_X(x;n,p)=\sum_{i=0}^{x}\binom{n}{i}p^i(1-p)^{n-i}$$ =BINOM.DIST(*x*,*n*,*p*,TRUE)

Distribusi Binomial

- Nilai rerata dan varian

$$E(X) = np$$

$$VAR(X) = npq$$

- Koefisien skewness

$$c_s = \frac{q-p}{\sqrt{npq}}$$

$p = q$ → simetris

$q > p$ → negative skew

$q < p$ → positive skew

Distribusi Binomial

- Contoh #1
  - Setiap tahun dalam 5 tahun dilakukan pemilihan acak untuk menetapkan alokasi dana kepada 1 dari 4 kegiatan (A,B,C,D).
  - Setiap kali dilakukan pemilihan, masing-masing kegiatan memiliki peluang yang sama untuk terpilih (mendapatkan dana).
  - Berapa persen peluang kegiatan A mendapatkan dana 3x?
  - Berapa persen peluang kegiatan A mendapatkan dana 5x, 4x, 3x, 2x, 1x, 0x?

Distribusi Binomial

- Setiap kali pemilihan
  - prob(As) = probabilitas kegiatan A terpilih
    prob(As) = ¼ = 0.25 = p
  - prob(Ag) = probabilitas kegiatan A tak terpilih
    prob(Ag) = 1 – p = 0.75 = q
- Dalam 5x pemilihan, maka probabilitas kegiatan A mendapatkan dana 3x adalah:

$$f_X(x;n,p) = f_X(3;5,0.25) = \binom{5}{3} 0.25^3 (1-0.25)^{5-3} = 0.0879$$

=BINOM.DIST(3,5,0.25,FALSE)

Distribusi Binomial

Jumlah sukses	Jumlah cara sukses	Probabilitas
0	1	0.237
1	5	0.396
2	10	0.264
3	10	0.088
4	5	0.015
5	1	0.001
	Jumlah =	**1.000**

Distribusi Probabilitas Variabel Random Diskrit

Distribusi Poisson

Distribusi Poisson

- Situasi
  - Proses Bernoulli dalam suatu selang waktu → *p* adalah probabilitas terjadinya suatu event dalam selang waktu tersebut.
  - Jika selang waktu *t* sangat pendek, sedemikian hingga probabilitas *p* menjadi kecil dan jumlah pengamatan (eksperimen) *n* bertambah, sedemikian hingga *np* konstan, maka
    - ekspektasi jumlah kejadian dalam selang waktu total → tetap

Distribusi Poisson

- Sifat
  - Proses Poisson adalah suatu proses diskrit pada skala waktu kontinu.
    - Oleh karena itu, distribusi probabilitas jumlah event dalam suatu waktu *T* adalah sebuah distribusi diskrit,
    - akan tetapi distribusi probabilitas waktu antar events serta waktu sampai ke event ke-*n* adalah distribusi kontinu.

Distribusi Poisson

- Probabilitas distribusi Poisson

$$f_X(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots \quad \text{dan} \quad \lambda = np > 0$$

- Distribusi Poisson kumulatif

$$F_X(x;\lambda) = \sum_{i=0}^{x} \frac{\lambda^i e^{-\lambda}}{i!}$$

Distribusi Poisson

- Nilai rerata dan varian

$$E(X) = \lambda \qquad VAR(X) = \lambda$$

- Skewness coefficient

$$c_s = \lambda^{-1/2}$$

Distribusi Poisson

- Contoh
  - Suatu printer membuat kesalahan secara random pada kertas cetak rerata 2 kesalahan per halaman.
  - Hitunglah probabilitas terjadi satu salah cetak dalam satu halaman.
- Penyelesaian

$\lambda = 2, x = 1$

$$f_X(x;\lambda) = f_X(1;2) = \frac{2^1\, e^{-2}}{1!} = \frac{2}{e^2} = 0.2707$$

Distribusi Probabilitas Variabel Random Kontinu

Distribusi Normal

Flash-back Distribusi Binomial

- Ilustrasi contoh pemilihan kegiatan
  - Setiap tahun dalam 5 tahun dilakukan pemilihan secara acak untuk menetapkan alokasi dana kepada 1 dari 4 kegiatan (A,B,C,D).
  - Setiap kali dilakukan pemilihan, masing-masing kegiatan memiliki peluang yang sama untuk terpilih (mendapatkan dana).
  - Berapakah probabilitas kegiatan A mendapatkan dana 5×, 4×, 3×, 2×, 1×, 0×?

Flash-back Distribusi Binomial

Distribusi Binomial

memilih 1 di antara 4 kegiatan untuk diberi dana

Histogram distribusi probabilitas kegiatan A mendapatkan dana

Flash-back Distribusi Binomial

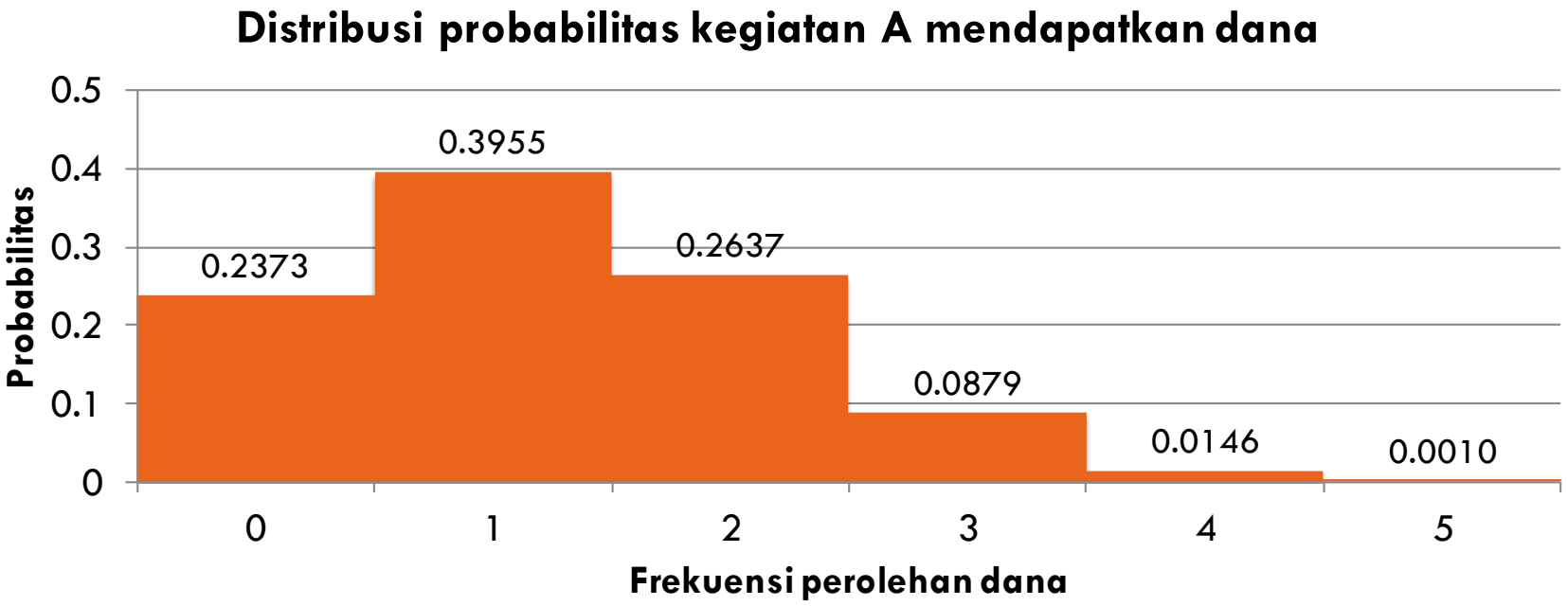
- Setiap kali pemilihan
  - prob(As) = probabilitas kegiatan A terpilih
    prob(As) = ¼ = 0.25 = p
  - prob(Ag) = probabilitas kegiatan A tak terpilih
    prob(Ag) = 1 – p = 0.75 = q
- Dalam 5x pemilihan, maka probabilitas kegiatan A mendapatkan dana 3x adalah:

$$f_X(x;n,p) = f_X(3;5,0.25) = \binom{5}{3} 0.25^3 (1-0.25)^{5-3} = 0.0879$$ =BINOM.DIST(3,5,0.25,FALSE)

Flash-back Distribusi Binomial

Jumlah sukses	Jumlah cara sukses	Probabilitas
0	1	0.2373
1	5	0.3955
2	10	0.2637
3	10	0.0879
4	5	0.0146
5	1	0.0010
	Jumlah =	**1.0000**

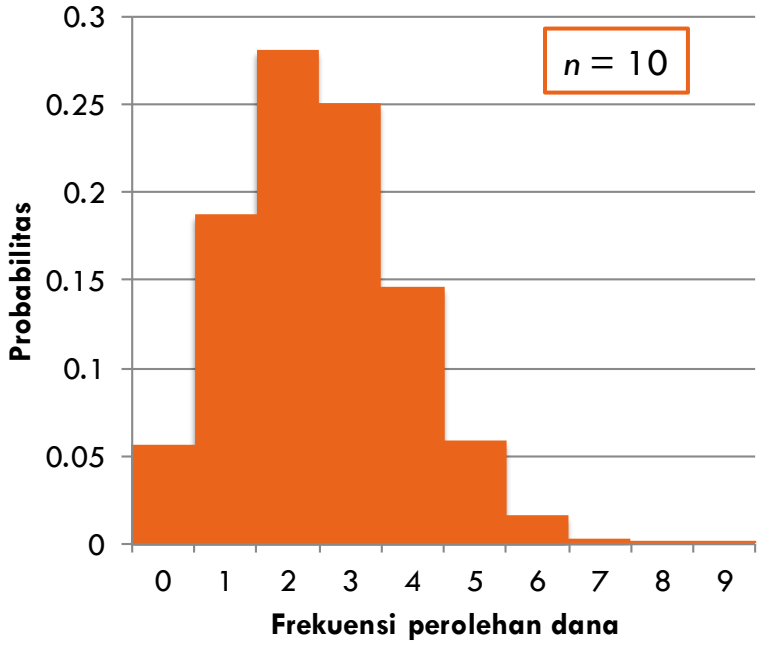
Flash-back Distribusi Binomial

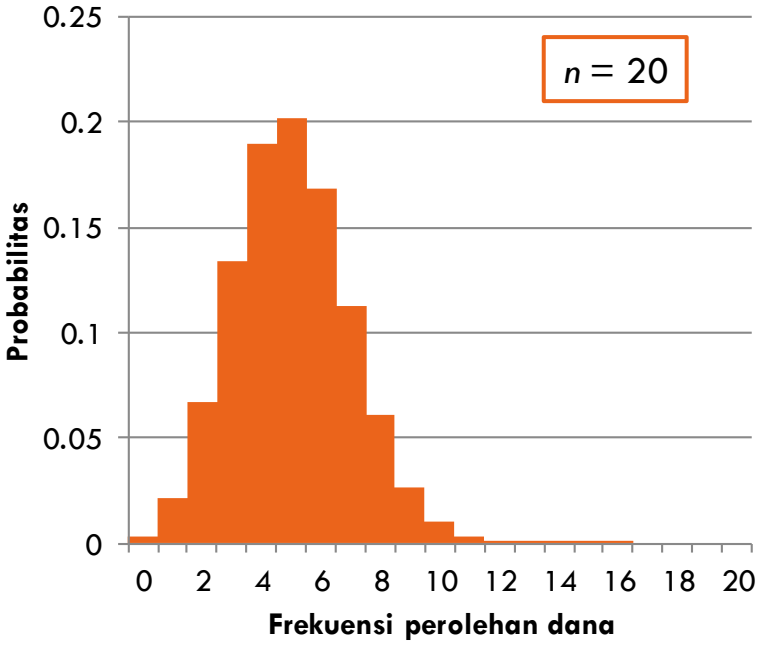


Flash-back Distribusi Binomial

- Apabila pemilihan dilakukan untuk waktu yang lebih panjang
  - 10 tahun
  - 20 tahun
  - *n* tahun
    - diperoleh $n + 1$ kemungkinan hasil
    - Kegiatan A dapat memperoleh dana sejumlah n kali, $n - 1$ kali, ... 0 kali

Flash-back Distribusi Binomial

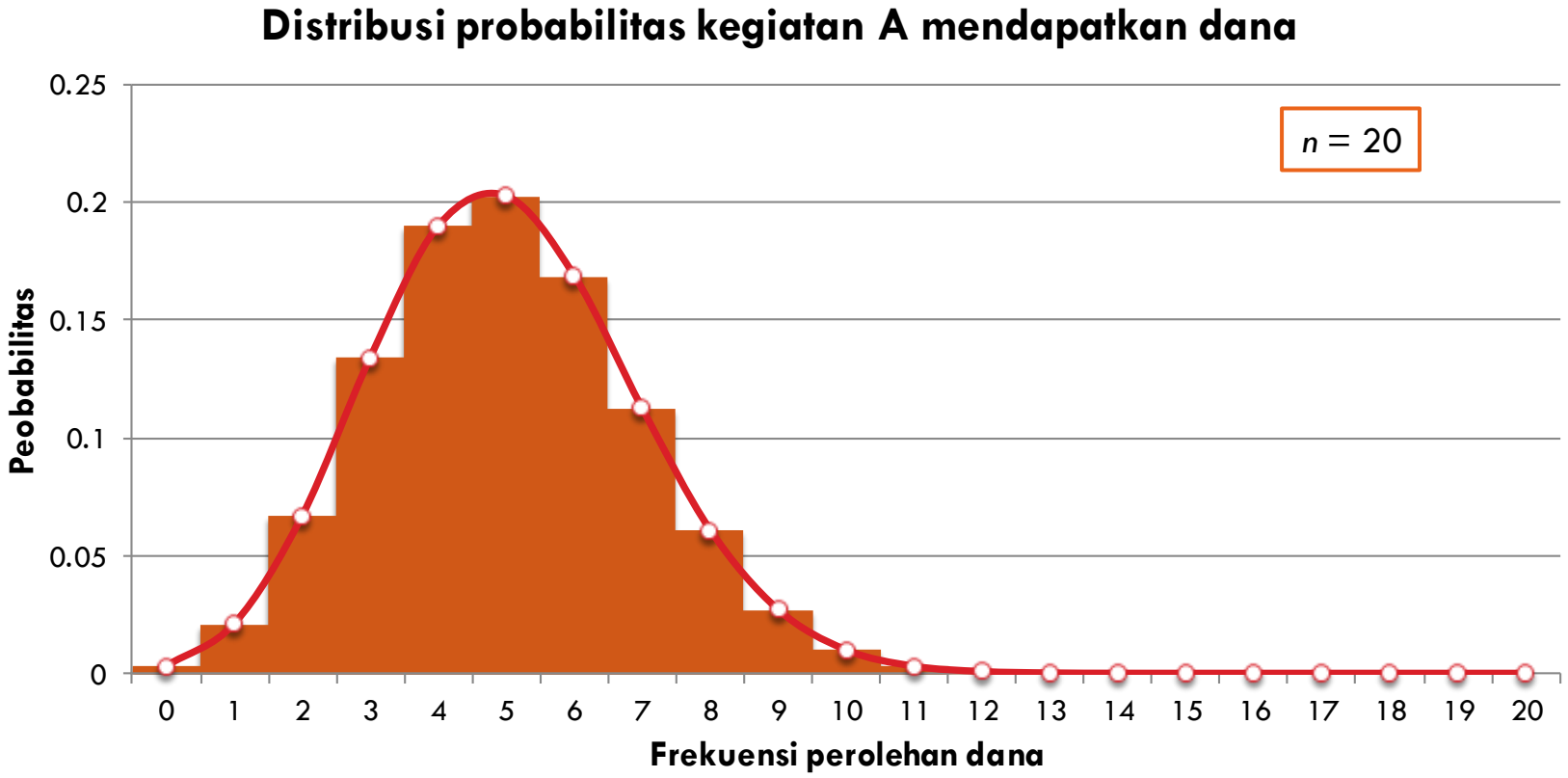


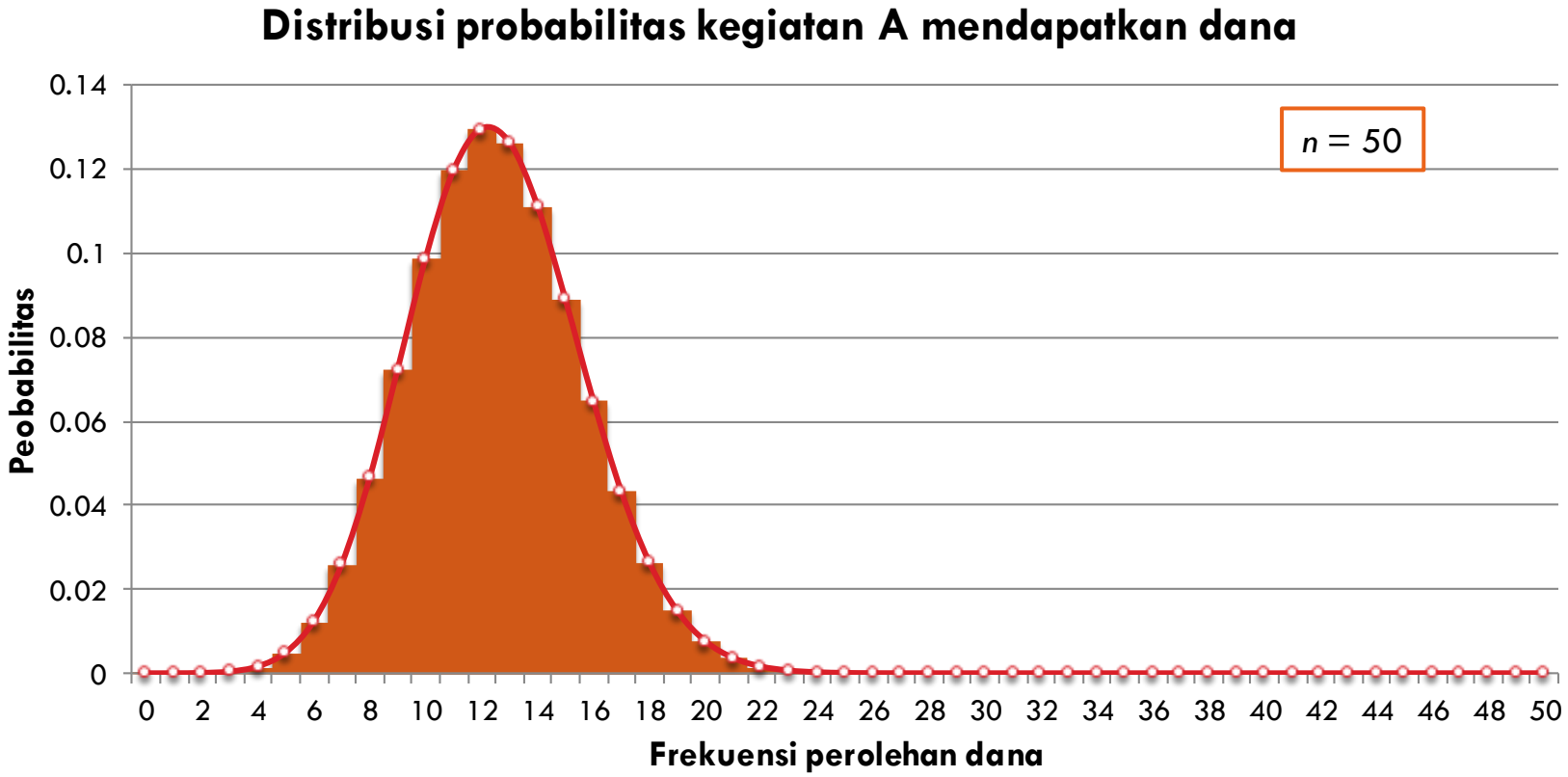


Distribusi Binomial vs Kurva Normal

- Apabila pemilihan (eksperimen) dilakukan sejumlah *n* kali dan *n* »
  - histogram distribusi probabilitas sukses (Kegiatan A memperoleh dana) memiliki selang (interval) kecil
  - garis yang melewati puncak-puncak histogram ⟶ kurva mulus berbentuk seperti lonceng

Kurva Normal


Distribusi probabilitas kegiatan A mendapatkan dana
n = 20
Peobabilitas
0.25
0.2
0.15
0.1
0.05
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Frekuensi perolehan dana


Distribusi probabilitas kegiatan A mendapatkan dana
n = 50
Peobabilitas
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50
Frekuensi perolehan dana

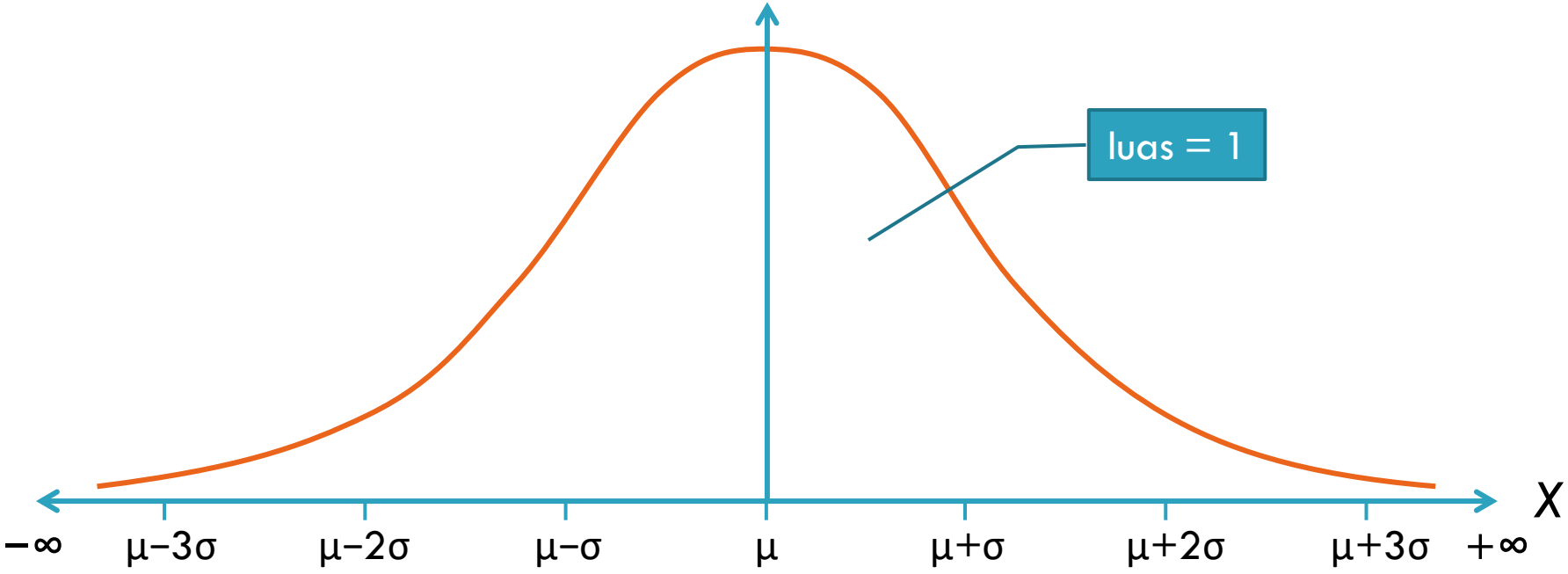
Kurva Normal dan Distribusi Normal

- Kurva Normal
  - berbentuk seperti lonceng dengan karakteristika tertentu
  - tidak setiap kurva berbentuk seperti lonceng adalah kurva normal
- Kurva Normal menggambarkan suatu distribusi yang disebut Distribusi Normal
- Permasalahan distribusi binomial dapat diselesaikan dengan pendekatan distribusi normal
- Distribusi normal lebih mudah dilakukan daripada distribusi binomial karena karakteristika distribusi normal telah diketahui (didefinisikan)
  - tabel distribusi normal
  - perintah/fungsi dalam MSExcel

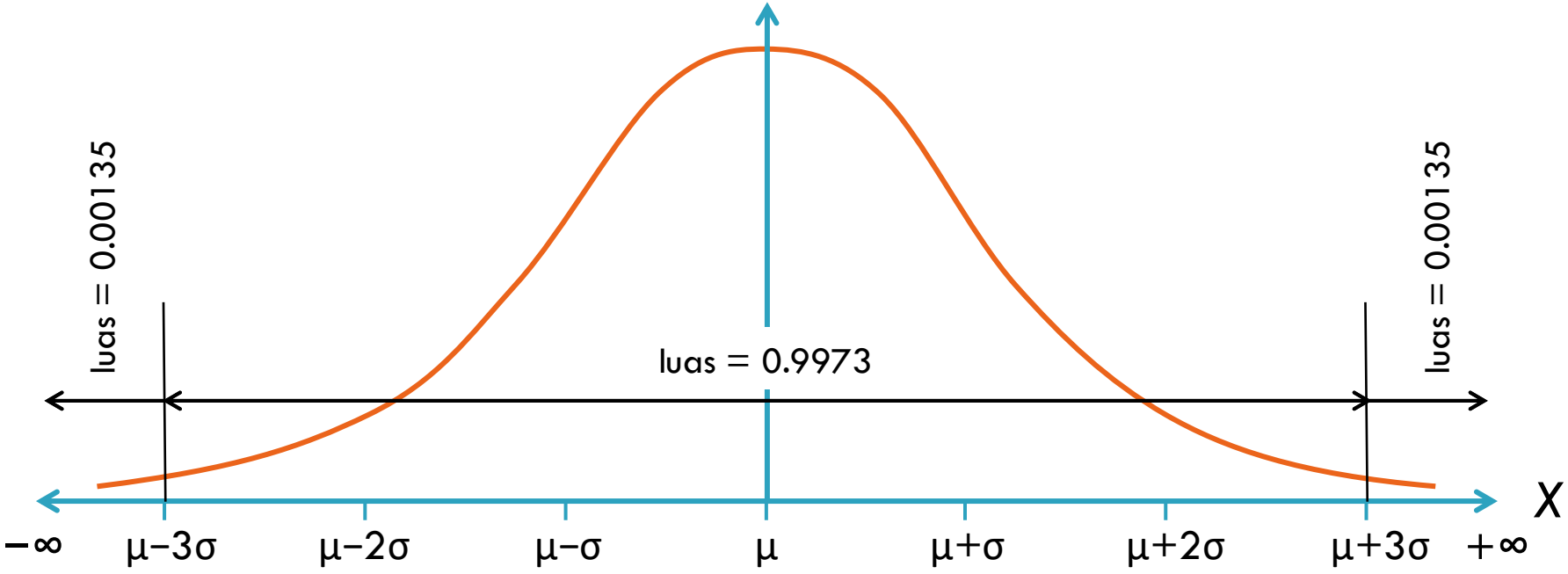
Distribusi Normal

- Karakteristika distribusi normal
  - simetris terhadap nilai rerata (mean)
  - score mengumpul di sekitar nilai rerata
  - kisaran score tak terbatas, namun sangat sedikit yang berada di luar kisaran tiga kali simpangan baku dari nilai rerata

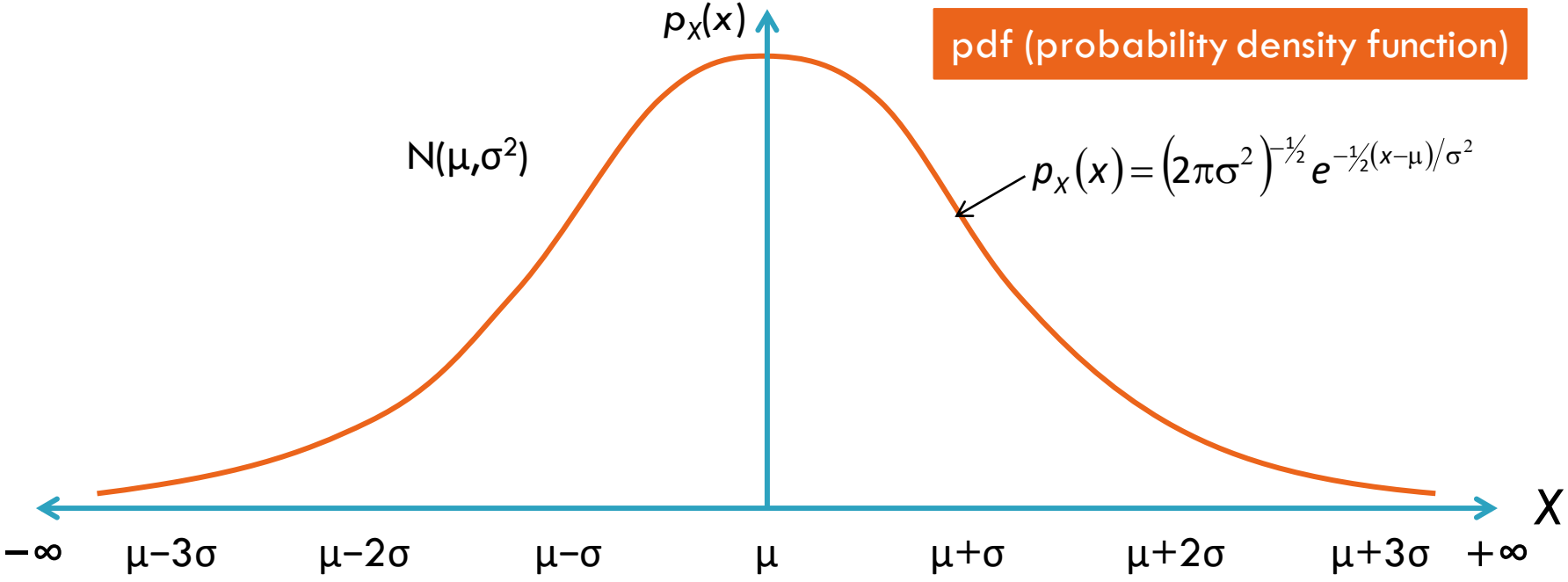
Distribusi Normal



Distribusi Normal



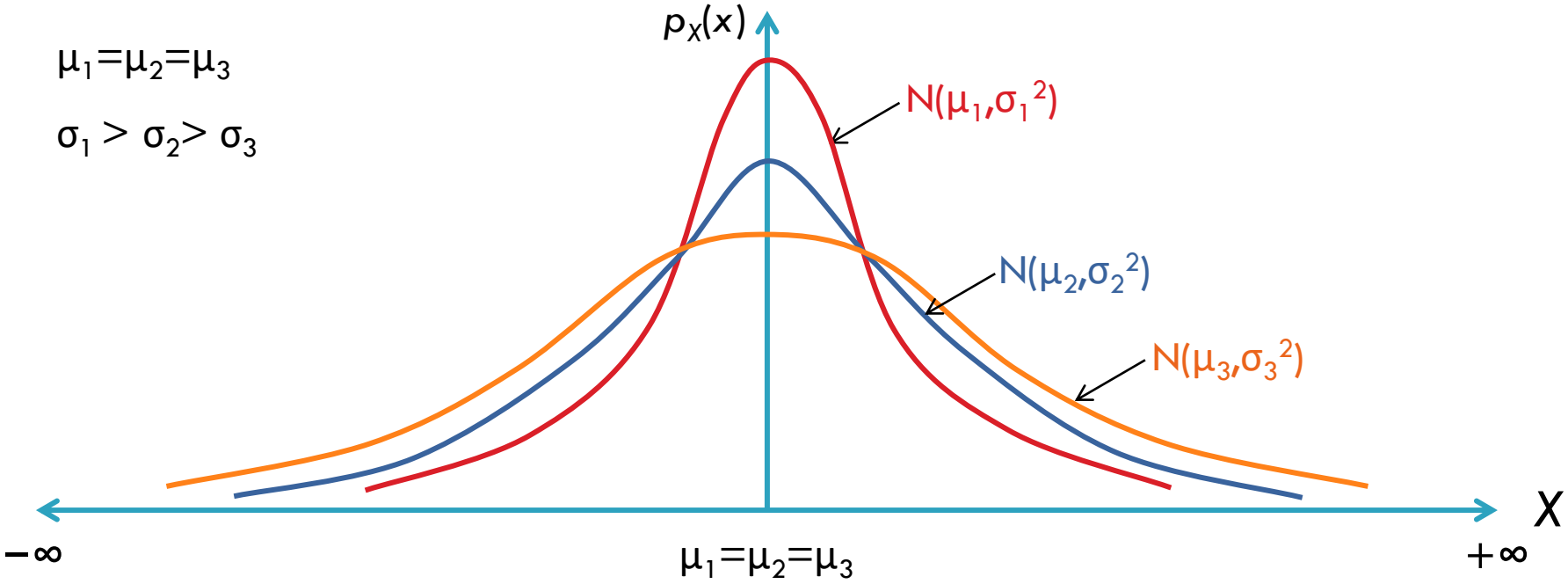
pdf Distribusi Normal



pdf Distribusi Normal

$\mu_1 = \mu_2 = \mu_3$

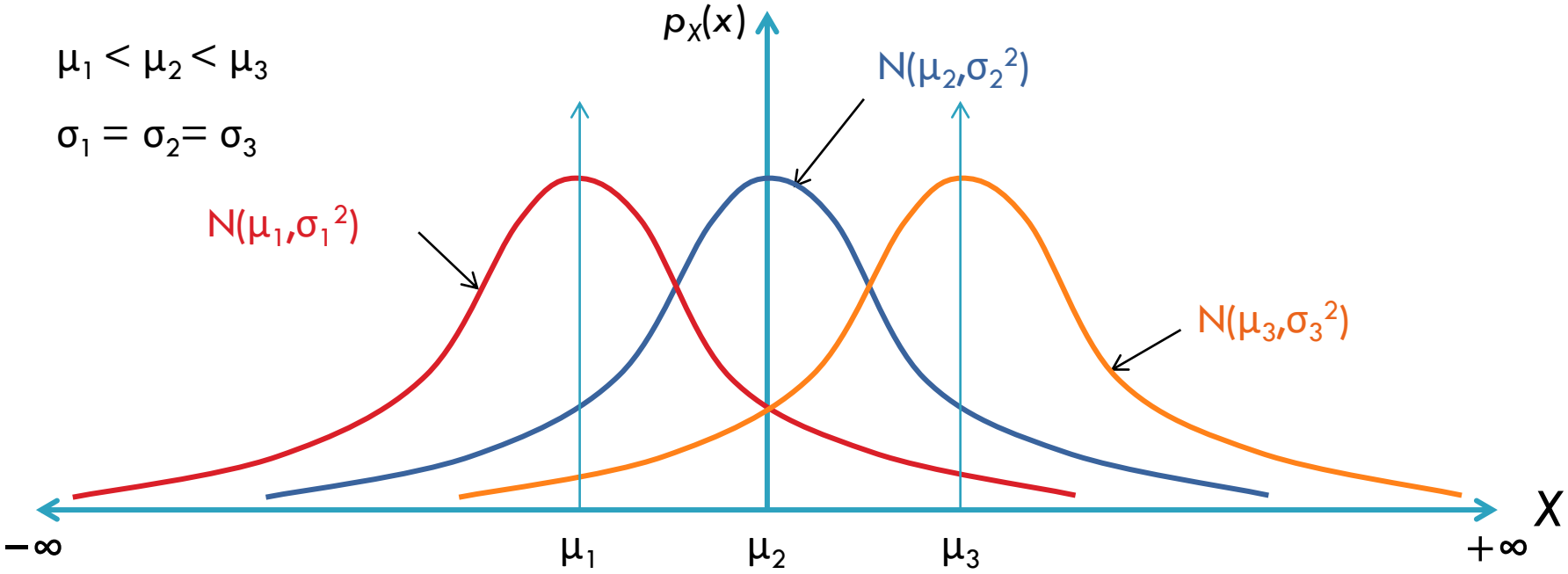
$\sigma_1 > \sigma_2 > \sigma_3$



pdf Distribusi Normal

$\mu_1 < \mu_2 < \mu_3$

$\sigma_1 = \sigma_2 = \sigma_3$

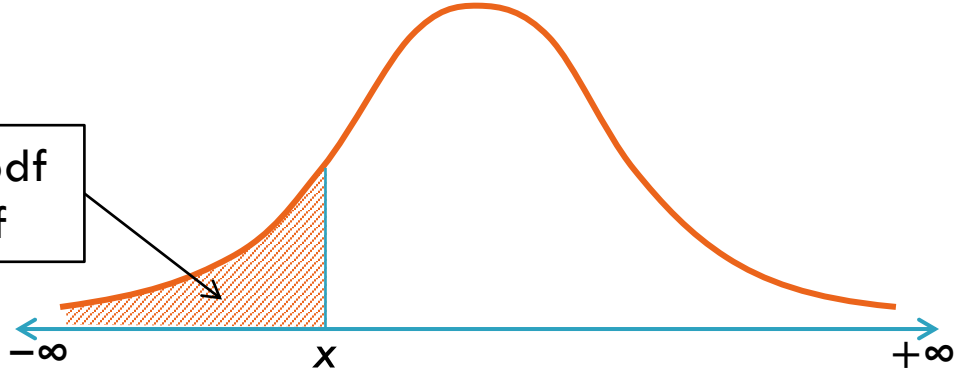


Distribusi Normal

- Jika X berdistribusi normal, N(⌈ , ⌈^{2}), maka prob($X \leq x$) dapat dicari dengan:

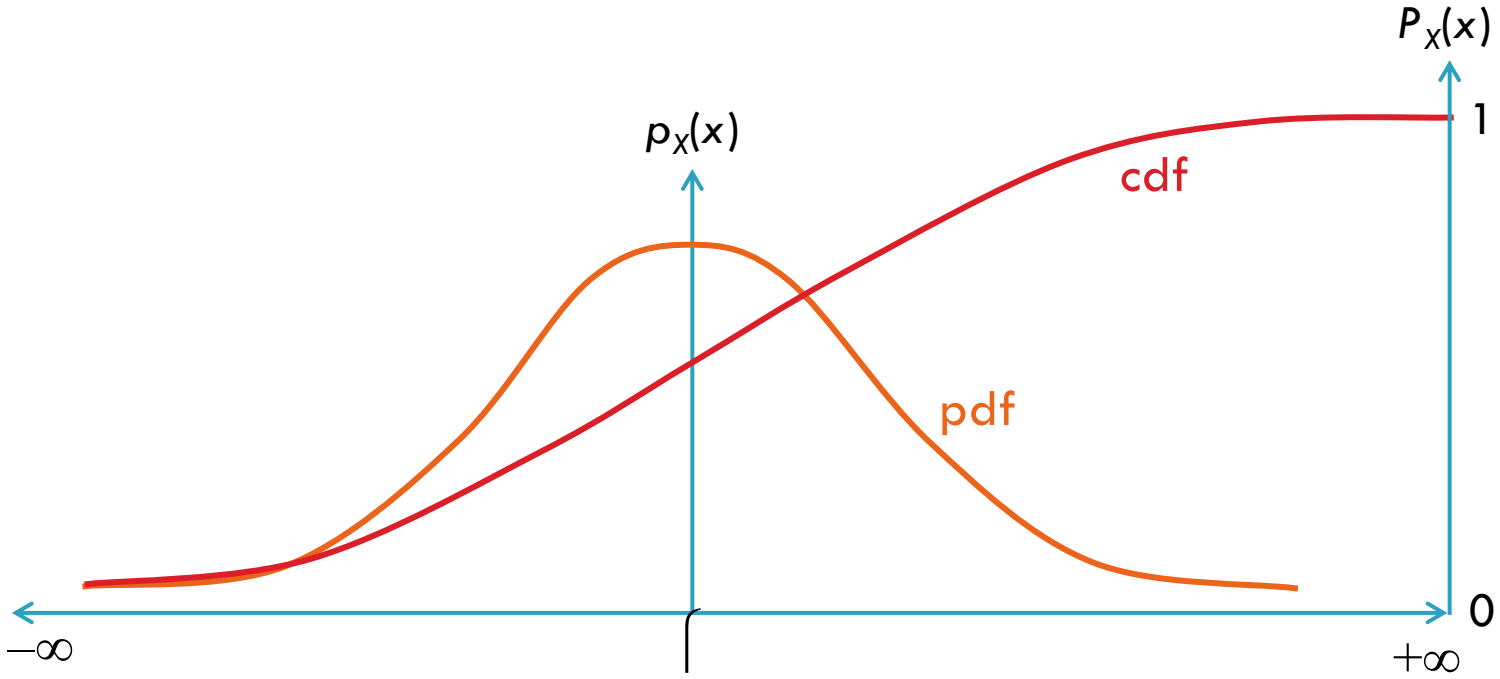
$$\text{prob}(X \leq x) = P_X(x) = \int_{-\infty}^{x} p_X(t)\,dt = \int_{-\infty}^{x} \left(2\pi\sigma^2\right)^{-1/2} e^{-\frac{1}{2}(t-\mu)/\sigma^2}\,dt$$

luas di bawah kurva pdf (dari $-\infty$ s.d. x) → cdf



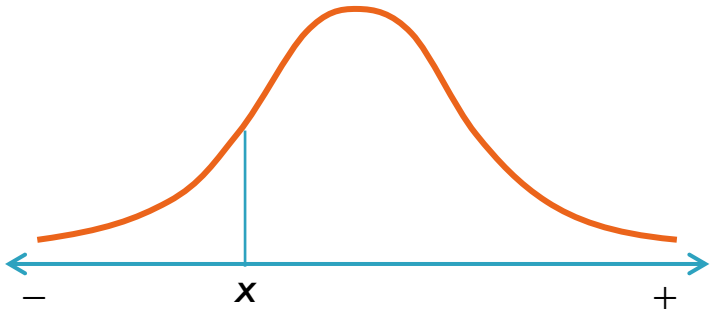
cdf (cumulative distribution function)

pdf - cdf



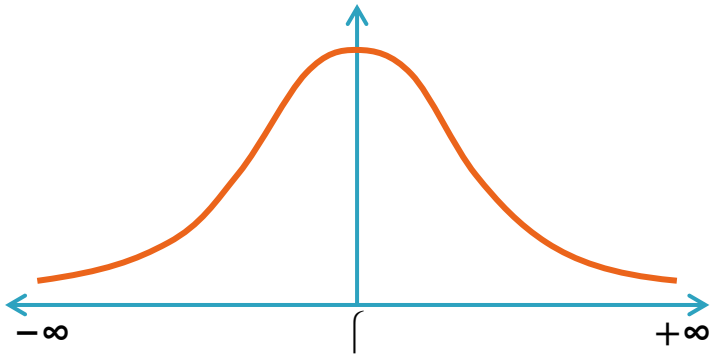
Distribusi Normal

- Luas di bawah kurva pdf
  - menunjukkan probabilitas suatu event
  - menunjukkan percentile rank
  - prob($X \leq x$) = prob($- \leq X \leq x$)
    = luas di bawah kurva antara $-$ s.d. x
  - prob($- \leq X \leq +$) = 1
    = luas di bawah kurva antara $-$ s.d. $+$
  - prob($X \geq x$) = prob($x \leq X \leq +$)
    = luas di bawah kurva antara x s.d. $+$
    = 1 – prob($X \leq x$)



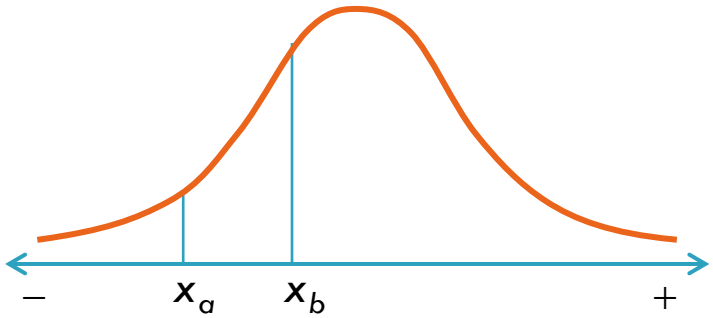
Distribusi Normal

- Probabilitas
  - prob($X \leq \mu$) = prob($X \geq \mu$) = 0.50
  - prob($\mu - x \leq X \leq \mu$) = prob($\mu \leq X \leq \mu + x$)



Distribusi Normal

- Probabilitas
  - $\text{prob}(X = x)$ = luas di bawah kurva antara x s.d. x
    $= 0$
  - $\text{prob}(X \leq x) = \text{prob}(X < x)$
  - $\text{prob}(X \geq x) = \text{prob}(X > x)$
  - $\text{prob}(x_a \leq X \leq x_b) = \text{prob}(x_a < X < x_b)$



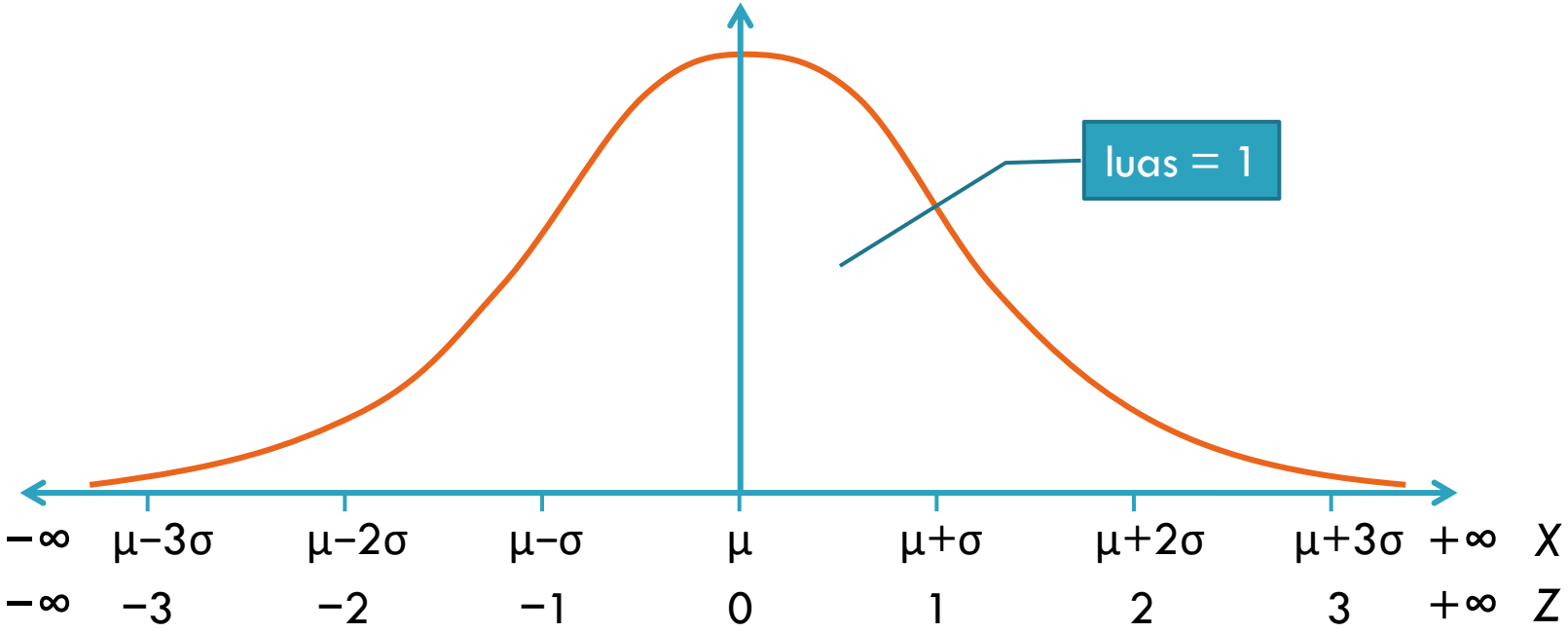
Distribusi Normal Standar

- Distribusi normal biasa disajikan dalam bentuk distribusi normal standar
  - dinyatakan dalam variabel Z

$$Z_x = \frac{X - \mu}{\sigma}$$

  - Z_x berdistribusi normal dengan $\mu = 0$ dan $\sigma = 1$, $N(0,1)$
    → disebut dengan nama **distribusi normal standar**

Distribusi Normal Standar



Tabel Distribusi Normal Standar

- Tabel z vs ordinat kurva normal standar
  - z vs ordinat pdf (probability density function) → file1
- Tabel z vs luas di bawah kurva
  - z vs cdf (cumulative distribution function)
  - luas kurva dari 0 s.d. z_x → file2
  - luas kurva dari $-\infty$ s.d. z_x → file3

Tabel Distribusi Normal Standar

- Contoh
  - Suatu variabel random X berdistribusi normal, serta memiliki nilai rerata 12, dan simpangan baku 3
  - prob($X < 15$)= prob($Z < 1$)= … (lihat tabel)
  - prob($X < 9$)= prob($Z < -1$)= … (lihat tabel)
  - prob($9 < X < 15$)= prob($-1 < Z < 1$)= … (lihat tabel)

Perintah (Fungsi) MS Excel

- Distribusi Normal
  - NORM.DIST(x,mean,standard_dev,cumulative)
    - x = nilai yang diinginkan untuk dicari distribusi normalnya
    - mean = nilai rerata (aritmetik)
    - standard_dev = nilai simpangan baku
    - cumulative = TRUE atau FALSE; TRUE jika ingin menghitung cdf, FALSE jika ingin menghitung pdf
  - NORM.INV(probability,mean,standard_dev)
    - probability = probabilitas suatu distribusi normal
    - mean = nilai rerata (aritmetik)
    - standar_dev = nilai simpangan baku

Perintah (Fungsi) MS Excel

- Distribusi Normal Standar
  - NORM.S.DIST(z,TRUE)
    - menghitung nilai cdf distribusi normal standar
  - NORM.S.DIST(z,FALSE)
    - menghitung nilai pdf distribusi normal standar
  - NORM.S.INV(probability)
    - kebalikan dari NORM.S.DIST(z)
    - mencari nilai z apabila probabilitasnya diketahui
- Ingat
  - Distribusi Normal Standar
    - mean = 0
    - simpangan baku = 1

Perintah (Fungsi) MS Excel

- Contoh 1
  - NORM.DIST(15,12,3,TRUE)
    - rerata = 12
    - simpangan baku = 3
    - prob(X < 15) = NORM.DIST(15,12,3,TRUE) = 0.8413
  - NORM.INV(0.8,12,3)
    - prob(X < x) = 0.8
    - x = NORM.INV(0.8,12,3) = 14.5249

Perintah (Fungsi) MS Excel

- Contoh 2
  - NORM.S.DIST(3,TRUE)
    - rerata = 0
    - simpangan baku = 1
    - prob($Z < 3$) = NORM.S.DIST(3,TRUE) = 0.9987
    - prob($0 < Z < 3$) = NORM.S.DIST(3,TRUE) – 0.5 = 0.4987
    - prob($1 < Z < 3$) = NORM.S.DIST(3,TRUE) – NORM.S.DIST(1,TRUE)
    - prob($Z > 1.5$) = 1 – NORM.S.DIST(1.5,TRUE)
  - NORM.S.INV(0.65)
    - prob($Z < z$) = 0.65
    - z = NORM.S.INV(0.65) = 0.3843

Terima kasih