
8

DISTRIBUSI PELUANG KONTINU

Dalam bab ini, akan dibahas tentang distribusi peluang kontinu. Ada beberapa distribusi peluang kontinu yang dikenal, yaitu : distribusi normal, distribusi gamma, distribusi eksponensial, distribusi Chi-kuadrat, distribusi Beta dan lain-lain. Tapi, pada kesempatan ini, hanya akan dibahas distribusi yang paling banyak digunakan, yaitu distribusi normal.

8.1. DISTRIBUSI NORMAL

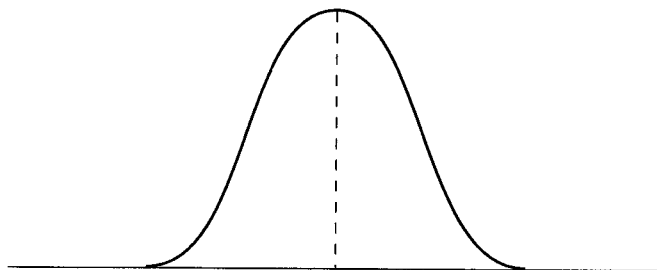
Distribusi normal sering disebut juga dengan distribusi Gauss , inilah distribusi peluang kontinu yang terpenting dan paling banyak digunakan. Grafiknya disebut kurva normal, berbentuk seperti lonceng, Pada gambar 8.1 kurva normal menggambarkan kumpulan data yang muncul dalam berbagai penelitian.

Pada tahun 1733, De Moivre menemukan persamaan matematika untuk kurva normal yang menjadi dasar dalam banyak teori statistika induktif.

Suatu perubah acak X dengan rata-rata μ dan varians σ^2 mempunyai fungsi densitas

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \dots\dots\dots 8(1)$$

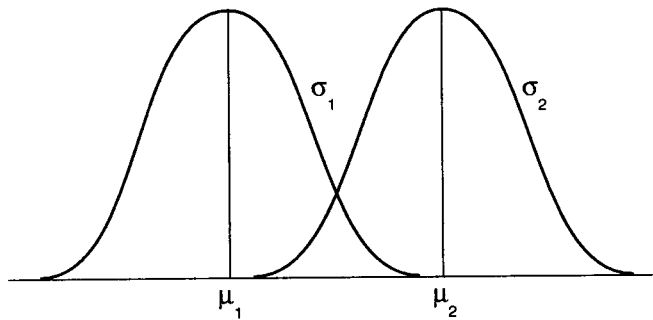
Sehingga, dengan μ dan σ^2 yang diketahui, maka seluruh kurva normal dapat diketahui.



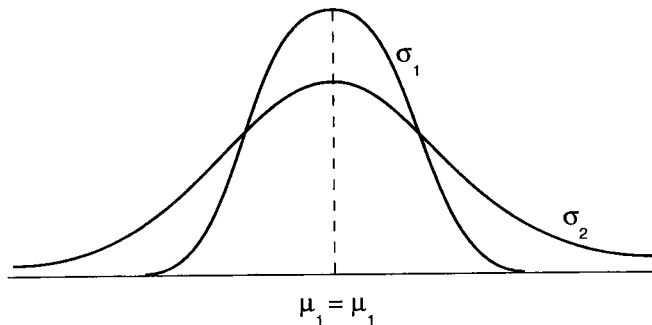
Gambar 8.1. Kurva Normal

Ada beberapa bentuk kurva normal, dibandingkan dalam 2 kurva yang berbeda:

1. Dua kurva berbeda dalam rata-rata dan simpangan baru
2. Dua kurva dengan simpangan baku berbeda tapi rata-rata sama
3. Dua kurva normal baik rata-rata maupun simpangan baku berbeda



Gambar 8.2.
Kurva Normal dengan $\mu_1 \leq \mu_2$ dan $\sigma_1 \leq \sigma_2$



Gambar 8.3.
Kurva Normal dengan $\mu_1 \leq \mu_2$ dan $\sigma_1 \leq \sigma_2$

Sifat-sifat penting distribusi normal:

1. Grafiknya selalu ada diatas gambar datar x
2. Bentuknya simetris terhadap $x = \mu$
3. Mempunyai satu modus, jadi kurva unimodal, tercapai pada $x = \mu$ sebesar $\frac{0,3989}{\sigma}$
4. Grafiknya mendekati (berasimtotkan) sumbu datar x , mulai dari $x = \mu + 3\sigma$ ke kanan dan $x = \mu - 3\sigma$ kekiri
5. Luas daerah grafik selalu sama dengan satu unit persegi.

Untuk tiap pasang μ dan σ , sifat-sifat diatas selalu dipenuhi, hanya bentuk kurvanya saja yang berlainan. Jika σ makin besar kurvanya makin rendah dan sebaliknya.

Berkaitan dengan sifat yang berlaku untuk sebuah fungsi densitas, dalam distribusi normal berlaku pula :

$$1. \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

$$2. P(a < X < b) = \int_{-\infty}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} dx$$

8.2. LUAS DIBAWAH KURVA NORMAL

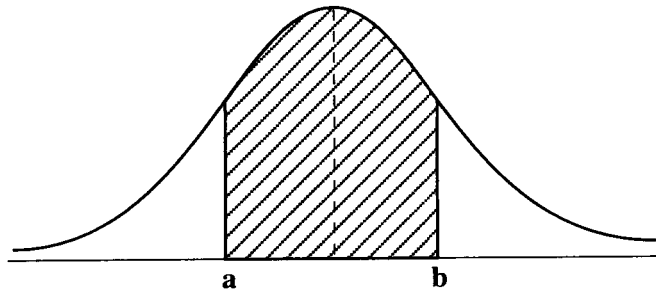
Sebuah kurva normal, sangat penting artinya dalam menghitung peluang, sebab luas daerah yang ada dalam kurva tersebut menunjukkan besarnya peluang.

Misalnya, suatu peubah acak X , mempunyai harga masing-masing $X = a$, dan $X = b$. Ingin dicari $P(a < X < b)$.

Maka

$$P(a < X < b) = \int_{-\infty}^{\infty} \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

dinyatakan oleh luas daerah yang diarsir :



Gambar 8.4.

P(a < X < b) = luas daerah yang diarsir

Untuk kepentingan praktis, rumus diatas sudah disusun dalam sebuah daftar, sehingga memudahkan para praktisi. Daftar yang dimaksud adalah daftar distribusi normal standar (baku). Distribusi normal standar adalah distribusi normal dengan rata-rata μ dan simpangan baku σ .

Ini diperoleh dari transformasi

$$Z = \frac{\bar{x} - \mu}{\sigma} \dots\dots\dots 8(4)$$

Sehingga fungsi densitasnya berbentuk :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \dots\dots\dots 8(5)$$

untuk z dalam daerah $-\infty < z < \infty$

Setelah kita memperoleh distribusi normal baku maka kita mencari luas daerah dibawah kurva normal baku tersebut. Caranya adalah sebagai berikut:

1. Hitung z hingga 2 desimal.
2. Gambarkan kurvanya
3. Letakkan harga z pada gambar datar, lalu tarik garis vertikal hingga memotong kurva
4. Luas daerah yang tertera dalam daftar, adalah luas daerah antara garis ini dengan garis tegak dititik nol.
5. Dalam daftar distribusi normal baku, cari harga z pada kolom paling kiri hanya 1 desimal, dan desimal keduanya dicari pada baris paling atas.
6. Dari z dikolom kiri, maju kekanan dan dari z da baris atas turun ke bawah, maka didapat bilangan yang merupakan luas daerah yang dicari. Bilangan yang didapat, ditulis dalam bentuk 0,xxxx (4 desimal).

Karena luas seluruh kurva adalah 1, dan kurva simetris di $m = 0$, maka luas dari garis tegak pada titik nol ke kiri ataupun ke kanan adalah 0,5.

Untruk mencari kembali z , jika luasnya diketahui, maka dilakukan langkah sebaliknya. Misalnya, jika luas = 0,4931, maka dalam badan daftar dicari 4931, lalu menuju ke pinggir sampai pada kolom z , didapat 2,4 dan menuju ke atas sampai batas z , dan didapat 6. Jadi harga $z = 2,46$.

Contoh 8.1.

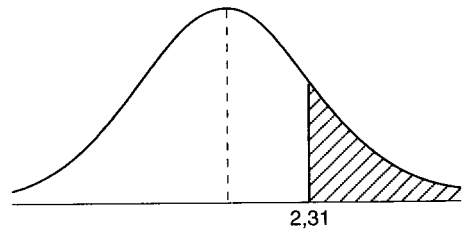
Berat bayi yang baru lahir rata-rata 3750 gram dengan simpangan baku 325 gram. Jika berat bayi berdistribusi normal, maka tentukan :

- a. Berapa % bayi yang beratnya lebih dari 4500 gram.
- b. Berapa banyak bayi yang beratnya antara 3500 gram sampai 4500 gram, jika semua ada 10000 bayi.
- c. Berapa bayi yang beratnya lebih kecil atau sama dengan 4.000 gram jika semuanya ada 10.000 bayi ?
- d. Berapa bayi yang beratnya 4.250 gram jika semuanya ada 5.000 bayi ?

Jawab :

a. dengan transformasi z :

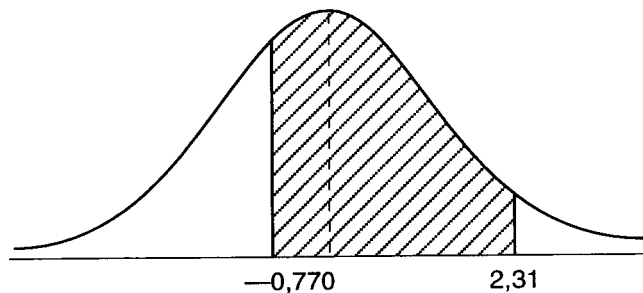
$$z = \frac{4500 - 3750}{325} = 2,31$$



Berat yang lebih dari 4500 gram, pada grafiknya ada di sebelah kanan $z = 2,31$

Luas daerah ini = $0,5 - 0,4896 = 0,0104$

Jadi ada 1,04 % dari bayi yang lahir beratnya lebih dari 4500 gram



b. dengan $x_1 = 3500$ dan $x_2 = 4500$

didapat $z_1 = -0,77$ dan $z_2 = 2,31$

luas daerah yang diarsir $L(z_1) = 0,2794$ dan $L(z_2) = 0,4896$

Jadi, luas daerah keseluruhan adalah $L(z_1) + L(z_2) = 0,7690$

Jadi banyaknya bayi yang beratnya antara 3500 gram dan 4500 gram diperkirakan ada $0,7690 (10000) = 7690$

c. Beratnya lebih kecil atau sama dengan 4.000 gram, setelah penyesuaian maka beratnya harus lebih kecil dari 4.000,5 gram.

$$Z = \frac{4000,5 - 3750}{325} = 0,77$$

Peluang berat bayi lebih kecil atau sama dengan 4.000 gram adalah $0,5 + 0,2794 = 0,7794$

Banyaknya bayi = $(0,7794) (10.000) = 7794$.

- d. Berat 4250 gr, setelah penyesuaian berarti antara 4.249,5 gram dan 4250,5 gram. Jadi untuk $X = 4,249,5$ dan $X = 4.250,5$ didapat :

$$z = \frac{4.249,5 - 3750}{325} = 1,53 \quad \text{luas daerah} = 0,4382$$

$$z = \frac{4250,5 - 3750}{325} = 1,54 \quad \text{luas daerah} = 0,4370$$

Peluang adalah $0,4382 - 0,4370 = 0,0012$

Banyaknya bayi = $(0,0012) (5.000) = 6$.

8.3. PENDEKATAN NORMAL TERHADAP BINOMIAL

Antara distribusi Binomial dan distribusi normal terdapat hubungan tertentu. Jika untuk fenomena yang berdistribusi binomial berlaku kondisi berikut :

- ukuran N cukup besar
- $P(A)$; peluang terjadinya peristiwa A tidak terlalu dekat ke nol, maka distribusi binomial dapat didekati oleh distribusi normal, dengan rata-rata $m = Np$ dan simpangan baku $s = \sqrt{Np(1-p)}$ dengan X adalah perubah acak diskrit yang menyatakan terjadinya peristiwa A . Karena disini telah terjadi perubahan dari perubah acak diskret ke kontinu, maka nilai-nilai X perlu mendapat penyesuaian. Yang dipakai adalah dengan jalan menambah atau mengurangi dengan 0,5.

Pendekatan normal terhadap binomial sangat memudahkan dalam perhitungan.

Contoh 8.2.

10 % dari penduduk tergolong kategori A. Sebuah sampel acak terdiri dari 400 orang telah diambil. Tentukan peluang akan terdapat :

- a. Paling banyak 30 orang tergolong kategori A
- b. Antara 30 dan 50 orang tergolong kategori A
- c. 55 orang atau lebih termasuk kategori A.

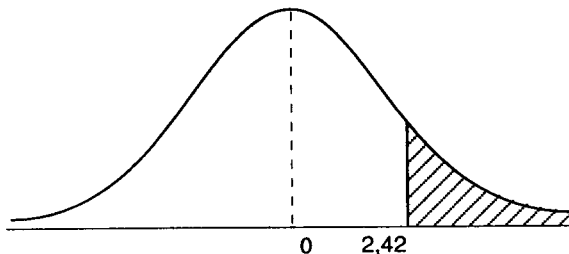
Jawab :

- c). 55 orang atau lebih untuk distribusi binom membenrkan $X > 54,5$, untuk distribusi normal

$$\text{Maka } z = \frac{54,5 - 40}{6} = 2,42$$

Sehingga kita perlu luas daerah seperti dalam gambar yang diarsir

Peluangnya adlah $0,5 - 0,4922$



Jawab :

misalkan X = banyak penduduk termasuk kategori A
dari data diatas diperoleh bahwa

$$\mu = 0,1 (400) = 40$$

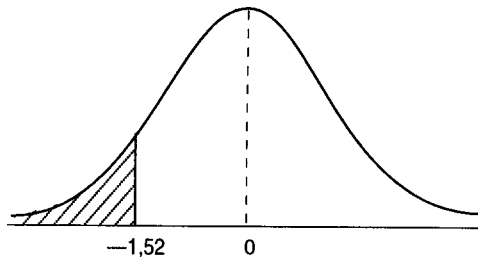
$$\sigma = 6$$

a. $z_1 = \frac{-0,5 - 40}{6} = -6,57$

$$z_1 = \frac{30,5 - 40}{6} = -6,57$$

Luas daerah yang diarsir adalah $0,5 - 0,4429 = 0,0571$.

Jadi peluang terdapat paling banyak 30 orang yang termasuk kategori A adalah 0,0571



- b. untuk distribusi normal, disini berlaku $30,5 < X < 49,5$. Angka bakunya masing-masing $z_1 = -1,58$ dan $z_2 = 1,58$

Dari daftar distribusi normal baku terdapat peluang yang dicari = $2(0,4429)$ = 0,8858

8.4. UJI DISTRIBUSI NORMAL

Untuk keperluan analisis selanjutnya, dalam statistika, ternyata model distribusi harus diketahui bentuknya terlebih dahulu. Dalam teori menaksir dan uji hipotesis misalnya, perhitungan dilakukan berdasarkan asumsi bahwa populasi berdistribusi normal. Jika asumsi ini tidak dipenuhi, artinya ternyata populasi berdistribusi normal, maka kesimpulan berdasarkan teori itu tidak berlaku. Karenanya, sebelum teori lebih lanjut digunakan dan kesimpulan diambil, terlebih dahulu diselidiki apakah asumsi normal itu dipenuhi atau tidak.

Ada beberapa cara pengujian distribusi normal, antara lain secara parametrik, uji Liliefors (non parametrik) dan kertas peluang.

Uji Secara Parametrik

Untuk keperluan pengujian, kita harus menghitung frekwensi teoritik E_i dan frekwensi hasil pengamatan O_i . Selanjutnya statistik χ^2 dihitung berdasarkan rumus :

$$\chi^2 = \frac{\sum (O_i - E_i)^2}{E_i}$$

dengan hipotesis :

H_0 : populasi berasal dari distribusi normal

H_1 : populasi bukan berasal dari distribusi normal

Kriteria uji :

Tolak H_0 jika $\chi^2 \text{ hit} \geq \chi^2_{(1-\alpha)(k-3)}$

dk = k - 3

dengan taraf kepercayaan α

Contoh 8.3.

Pengukuran terhadap tinggi mahasiswa tingkat pertama dilakukan dan diambil sampelnya secara acak sebanyak 100 orang. hasilnya dicatat dalam dist frekuensi sebagai berikut :

tinggi (Cm)	f
140-144	7
145-149	10
150-154	16
155-159	23
160-164	21
165-164	21
170-174	6
Jumlah	10

Setelah dihitung $X = 157,8$ dan $S = 8,09$.

Selanjutnya perlu ditentukan batas-batas kelas interval untuk menghitung luas dibawah kurva normal bagi tiap interval. Kelas interval kesatu dibatasi oleh 139,5 dan 144,5. Jika ditransformasikan dalam angka baku Z, interval kesatu dibatasi oleh -2,26 dan -1,64. Luas dibawah kurva normal untuk interval ke satu adalah $L = 0,4881 - 0,4495 = 0,0386$. Sehingga frekuensi harapan (E_i) untuk kelas interval kesatu adalah $0,0386 \times 100 = 3,9$. Demikian untuk selanjutnya tinggi diperoleh hasil sebagai berikut :

Batas kelas (X)	z	luas kelas int	frekuensi harapan (E_i)	frekuensi Pengamatan (si)
139,5	-2,26	0,0386	3,9	7
144,5	-1,64	0,1010	10,1	10
149,5	-1,03	0,1894	18,9	16
154,5	-0,41	0,2423	24,2	23
159,5	0,21	0,2135	21,4	21
164,5	0,83	0,1298	13,0	17
169,5	1,45	0,0538	5,4	6
174,5	2,06			

$$x^2_{hitung} = \frac{(7-3,9)^2}{3,9} + \frac{(10-10,1)^2}{10,1} + \dots + \frac{(17-13,0)^2}{13} + \frac{(6-5,4)^2}{5,4}$$

$$x^2_{hitung} = 4,27$$

Dari daftar distribusi frekuensi, diperoleh banyak kelas $k = 7$, sehingga derajat kebebasan ($dk = 4$). Dari tabel z^2 diperoleh

$$X_{20,95}(4) = 9,49$$

Sehingga jelas bahwa $x^2_{hitung} < z^2_{tabel}$. Kesimpulannya μ_0 diterima. Jadi Populasi berasal dari distribusi normal.

Uji Lilliefors

Uji Lilliefors adalah uji normalitas secara non parametrik

Misalkan kita mempunyai sampel acak dengan hasil pengamatan X_1, X_2, \dots, x_n .

Berdasarkan sampel ini akan diuji hipotesis :

H_0 = sampel berasal dari populasi berdistribusi normal

H_1 = sampel bukan berasal dari populasi berdistribusi normal.

Prosedur pengujiannya adalah sebagai berikut :

- a). Pengamatan x_1, x_2, \dots, x_n dijadikan bilangan baku z_1, z_2, \dots, z_n dengan rumus

$$Z_i = \frac{x_i - \bar{X}}{S}$$

- b). Untuk tiap bilangan baku ini, dihitung frekuensi kumulatif $F(z_i) = P(Z \leq z_i)$

- c). Selanjutnya hitung proporsi z_1, z_2, \dots, z_n yang lebih kecil atau sama dengan z_i . Jika proporsi ini dinyatakan oleh $S(z_i)$, maka

$$S(z_i) = \frac{\text{banyaknya } z_1, z_2, \dots, z_n \text{ yang } \leq z_i}{n}$$

- d). Hitung selisih $F(z_i) - S(z_i)$ kemudian tentukan harga mutlaknya.
- e). Ambil harga mutlak terbesar diantara harga-harga mutlak lainnya. Sebutlah harga terbesar ini sebagai L_0 .

Untuk menerima atau menolak hipotesis, bandingkan L_0 dengan nilai kritis L dalam daftar uji Lilliefors untuk taraf nyata L yang dipilih

Kriteria uji :

Tolak H_0 jika $L_0 > L$ daftar. untuk hal lain, H_0 diterima.

Contoh 8.4.

Misalkan kita mempunyai data penghasilan (dalam ribuan) 10 keluarga di kec. Lingkar Selatan Bandung sebagai berikut :

52 48 62 49 69 85 54 87 61 35

Apakah data diatas bersumber dari populasi yang berdistribusi normal ?

Jawab :

$$X = 60,02$$

$$S = 16,43$$

Berdasarkan prosedur pengujian uji Lilliefors, diperoleh :

Xi	Z	F(Zi)	S(Zi)	F(Zi) - S(Zi)
35	-1,522	0,064	0,1	0,036
48	-0,743	0,2296	0,2	0,0296
49	-0,682	0,248	0,3	0,052
52	-0,409	0,312	0,4	0,088
54	-0,377	0,356	0,5	0,144
61	0,049	0,520	0,6	0,08
62	0,11	0,544	0,7	0,156
69	0,54	0,705	0,8	0,095
85	1,52	0,935	0,9	0,035
87	1,63	0,949	1,0	0,051

$$Z_1 = \frac{35 - 60,02}{16,43} = -1,52$$

$$\text{Luas } (Z_1) = 0,5 - 0,4357$$

$$\text{Peluang} = 0,0643$$

dan seterusnya untuk Z2, ..., Z10.

Dari Daftar untuk L = 5% , n = 10 diperoleh L kritis = 0,258.

Diketahui bahwa $L_0 < L$ kritis.

Kesimpulan, H_0 diterima. Jadi Sampel berasal dari distribusi Normal.

8.5. SOAL LATIHAN

1. Misalkan tinggi mahasiswa berdistribusi normal dengan rata-rata 167,5 cm dan simpangan baku 4,6 cm. Semuanya ada 200.000 mahasiswa. Tentukan ada berapa mahasiswa yang tingginya :
 - a. lebih dari 175 cm
 - b. lebih dari 160 cm
 - c. antara 158 cm dan 170 cm

2. Kekeliruan yang terjadi sebagai hasil pengukuran panjang semacam alat ternyata berdistribusi normal dengan nilai ekspektasi nol dan simpangan baku 1,5 cm. Bepara peluang hasil pengukuran akan lebih dari 3 cm ?

3. Peubah acak X berdistribusi normal baku dengan rata-rata 25 dan simpangan baku 3,6. Tiap pengamatan dikalikan 4 dan kemudian ditambah 15. Tentukan rata-rata dan simpangan baku yang baru !

4. Daya pakai dua merek semacam alat masing-masing berdistribusi normal. Yang pertama dengan rata-rata 80 jam dan simpangan baku 12 jam, sedangkan yang kedua dengan rata-rata 90 jam dan simpangan baku 6 jam. Jika ingin menggunakan alat untuk periode 100 jam, anda pilih mana ? Mengapa ?

5. 5% dari penduduk suatu daerah menderita penyakit A. Sampel acak diambil sebanyak 1.000 orang. Tentukanlah :
 - a. Ada berapa diharapkan yang menderita penyakit A ?
 - b. Berapa variansnya
 - c. Tentukan berapa peluangnya akan terdapat lebih dari 65 orang yang menderita penyakit A.