



# DISTRIBUSI NORMAL

Oktavianus Ch Salim

Bagian Ilmu Kesehatan Masyarakat Fakultas Kedokteran Universitas Trisakti

## ABSTRACT

*The normal distribution is the most important of continuous distribution, used in statistical application with shown as a bell shape curve, because many type of data are approximately normally distribution, or can be transformed to another scala. This curve is symetric with the highest point at the center and the total area under the curve is unity, because it is symetric, the probability to the left of the mean is 1/2, as is the probability to the right of the mean. The mean is thus the same as the median and the mode. One particular member of the family of normal distribution is the standard normal curve, it has a mean of zero ( $\alpha=0$ ) and standard deviation of one ( $\beta=1$ ). (J Kedokter Trisakti 1999;18(2):107-111)*

*Key words: Statistics, normal distribution.*

## PENDAHULUAN

Data kuantitatif biasa dapat disajikan dalam bentuk suatu distribusi frekuensi atau histogram. Bila data yang didapat makin banyak dan intervalnya makin sempit, maka gambaran histogram akan terlihat sebagai suatu kurva yang halus. Selanjutnya kita dapat menggambarkan histogram tanpa terlihat gambaran batang-batangnya dan

tampak hanya sebagai sebuah kurva semata. Kurva yang didapat secara teoritis menggambarkan distribusi sebagian besar variabel yang bersifat kontinu. Salah satu bentuk kurva distribusi yang paling penting adalah Distribusi Normal<sup>1,6,4</sup>. (Lihat gambar 1).

Gambar 1



Salim

## DEFINISI DAN KONSEP DASAR <sup>2,4</sup> :

Distribusi Normal adalah suatu distribusi empirik atau teoritis, yang meskipun sudah banyak digunakan dalam bidang statistik tetapi masih merupakan suatu misteri pada banyak orang. Distribusi Normal disebut juga distribusi Gauss ( Carl Friedrich Gauss, seorang ahli matematik yang banyak memberikan andil pada pengembangannya pada permulaan abad 19). Kata 'normal' disini tidak diartikan sebagai kata-kata dalam bahasa inggris 'normal' yang berarti 'ordinary atau common' dan tidak juga seperti terminologi kedokteran sebagai 'tidak sakit',, namun merupakan suatu model matematik yang menggambarkan penyebaran probabilitas dari pengamatan yang tidak terbatas dan diukur terus menerus.

Ciri-ciri penting suatu distribusi normal :

1. Berbentuk lonceng berarti simetris di kanan dan kiri dari 'mean'
2. 'Mean' = 'median' = 'mode', nilai dari ketiga ukuran sentral ini terletak pada titik yang sama pada sumbu X dan hanya mempunyai satu 'mode' (unimodal).
3. Jumlah seluruh daerah diatas sumbu X dan dibawah kurva setara dengan satu atau seratus persen. Karena kurva Normal simetris,berbentuk lonceng dan unimodal maka daerah di di kanan dan di kiri garis tegak lurus diatas mean masing-masing besarnya 0,5 atau 50%.
4. Kurva ditetapkan oleh dua parameter yaitu 'mean' yang merupakan pusat atau konsentrasi distribusi dan standar deviasi yang menentukan penyebaran distribusi di sekitar 'mean'.
5. Ujung-ujung kurva meruncing dikanan dan kiri tetapi tidak pernah mennyentuh garis X (*asymptotic*), dan jarak keujung-ujungnya dari 'mean' menunjukkan tingkat frekuensi pengukuran.
6. Bila garis tegak lurus dibuat pada jarak satu standar deviasi di kanan dan di kiri 'mean' akan mencakup daerah seluas kira-kira 68% di dalamnya (antara garis tersebut, kurva dan sumbu bila dua standard deviasi 95%, bila tiga standar deviasi 99,7% dan area di luar tiga standar deviasi hampir selalu diabaikan.(Lihat gambar 2).

Gambar. 2



Distribusi normal ini menjadi sangat penting karena banyak variabel kontinu seperti kadar kolesterol, tekanan darah, berat badan dan tinggi badan mempunyai distribusi yang mendekati distribusi normal, apabila distribusi data tersebut tidak mengikuti distribusi normal maka perlu dilakukan transformasi menjadi skala lain misal dengan menggunakan logaritme atau akar sehingga distribusi data tersebut menjadi mendekati normal. Sebagai ilustrasi : Menurut laporan dari 'Joint National Committee on the Detection, Evaluation and Treatment of High Blood Pressure' tahun 1988, diagnosa hipertensi pada orang dewasa (18 tahun atau lebih) ditetapkan bila tekanan diastolik 90 mmHg atau lebih. Tekanan darah yang tergambar pada program skrining masyarakat luas biasanya terdistribusi secara simetris, kurva bentuk lonceng. Masalah yang sering dihadapi adalah menentukan titik potong dimana seseorang dinyatakan menderita atau tidak menderita hipertensi. Untuk ini biasanya digunakan titik potong dimana inferensi dari sampel pada populasi diperkirakan mencakup 95%, artinya diharapkan 95% dari populasi dalam keadaan normotensive.

Alasan lain yang mendasari pentingnya Distribusi Normal adalah 'Central Limit Theorem' yang menyebutkan bahwa , pengambilan sampel secara random dari sebuah distribusi ( tanpa menghiraukan jenis distribusinya ) maka distribusi dari sampel x akan mendekati distribusi normal, dan makin besar jumlah sampelnya distribusi akan makin mendekati distribusi normal. Menjadi pertanyaan bagi kita adalah, berapa jumlah sampel yang dibutuhkan sehingga kita dapat menggunakan theorem tersebut ? Ukuran dimana distribusi dari sampel mendekati distribusi normal tergantung pada bentuk distribusi dari populasi. Bila populasi terdistribusi secara normal, distribusi setiap ukuran sampel adalah tak selalu terlihat normal terutama bila jumlah sampel kecil. Sebaliknya untuk populasi yang terdistribusi sangat jauh dari normal, diperlukan jumlah sampel yang relatif lebih besar untuk

mendapatkan distribusi yang mendekati normal. Pada kenyataannya kita jarang mengetahui bentuk distribusi dari populasi, untuk ini dibutuhkan satu patokan yang praktis yang menyatakan bahwa sampel cukup besar untuk dapat menerapkan 'central limit theorem'. Di sini kita menjumpai apa yang disebut sebagai HUKUM 30 yang mengatakan bahwa : " secara umum, satu sampel yang terdiri dari 30 atau lebih elemen dapat dikatakan cukup besar agar 'central limit theorem' dapat diberlakukan". Di sini dikatakan 'secara umum' karena akan dibutuhkan jumlah sampel yang lebih besar untuk mendapatkan suatu distribusi yang mendekati normal bila distribusi pada populasi sangat jauh dari normal; sebaliknya jumlah sampel yang lebih sedikit akan dibutuhkan bila distribusi populasi mendekati distribusi normal. Konsekwensi dari ini sering timbul istilah sampel kecil bila jumlah sampel kurang dari 30 dan sampel besar bila jumlah sampel lebih besar dari 30. Penerapan hukum 30 ini juga harus hati-hati karena pada prinsipnya bila kita ingin informasi lebih banyak, kita ambil sampel yang lebih besar, sehingga ketidak pastian kita terukur dari standard deviasi) tentang parameter yang akan diestimasi akan berkurang.

**Distribusi Normal Standar <sup>3,5</sup> :**

Suatu distribusi yang paling penting, dengan mean = 0 dan standard deviasi = 1. Semua bentuk distribusi normal dapat ditransformasikan (transformasi Z) ke standard distribusi normal ini, yang disebut sebagai distribusi Z. Tranformasi ini menghasilkan suatu deviasi standar normal atau nilai Z, yang menunjukkan besarnya penyimpangan nilai observasi terhadap nilai mean (miu) dan dan dinyatakan dalam standar deviasi, yang kita hitung dengan cara :

$$Z = \frac{x - \mu}{SD}$$



Contoh penggunaan dalam praktek :

Skринing hipertensi : Misalnya kita mengukur tekanan darah pada 200.000 ribu partisipan pada suatu program skринing tekanan darah penduduk dalam skala besar. Hasilnya akan diperoleh suatu distribusi normal dengan 'mean 85 dan standard deviasi 13 mm Hg. Sehubungan dengan rekomendasi dari 'Joint National Committee on Detection, Evaluation and treatment of High Blood Pressure' tahun 1988 tentang hipertensi berdasarkan tekanan diastolik, dinyatakan bila tekanan diastolik 90 mmHg atau lebih harus dikonsultasi ke dokter. Beberapa pertanyaan bisa timbul :

1. Berapa proporsi individu dalam skринing yang mempunyai tekanan diastolik 90 mmHg atau lebih? Untuk ini kita dapat menghitung dengan menggunakan rumus Z .
2. Tekanan darah diastolik 90 mmHg terletak kira-kira 0,38 SD diatas 'mean' 85. Kemudian kita melihat dalam tabel Z, daerah yang sesuai dengan nilai Z = 0,38 yaitu 0,3520 = sekitar 35%. Dapat ditarik kesimpulan bahwa sekitar 35 % individu dalam sampel mempunyai tekanan darah diastolik 90 mmHg atau lebih atau bila secara random kita memilih partisipan akan didapati probabilitasnya 0,35 dengan tekanan darah 90 mmHg atau lebih. Ini berarti sekitar 0,35 X 200.000 orang harus berkonsultasi dengan dokter.
3. Berapa proporsi individu dengan tekanan darah diastolik 95 mmHg atau lebih ? Dengan cara yang sama didapatkan Nilai Z = 0,7692 = 0,77 yang setara dengan 0,2206 dalam tabel. Berarti 22% atau 44.120 partisipan mempunyai tekanan darah diastolik 95 mmHg atau lebih.
4. Berapa proporsi individu yang mempunyai tekanan diastolik antara 90 - 104 mmHg (hipertensi ringan) ? Kita harus menetapkan area diantara nilai Z = 0,38 (Tekanan diastolik 90

mmHg) dengan Z = 1,46 (tekanan diastolik 104 mmHg ) , didapatkan nilai 0,2799, yang berarti diperkirakan 28% individu atau 55.980 mempunyai hipertensi ringan.

5. Berapa proporsi individu yang mempunyai tekanan diastolik antara 105 dan 114 mmHg (hipertensi sedang) ? Ditetapkan area antara nilai Z = 1,54 (Tekanan diastolik 105 mmHg) dengan Z = 2,23 (tekanan diastolik 114 mmHg), didapatkan nilai 0,0489, yang berarti diperkirakan 4,89% individu atau 9800 mempunyai hipertensi sedang.
6. Berapa proporsi penderita yang mempunyai tekanan distolik 115 mmHg atau lebih ?. Dengan cara yang sama didapatkan nilai Z = 2,3077 atau 2,31 (tekanan diastolik 115 mmHg) yang setara dengan 0,0104 = 1.04% atau 2080 individu menderita hipertensi berat.
7. Dalam tekanan diastolik berapa, 95% individu (central limits 95%) ?. Karena kurva normal simetris, maka 95% sentral limit terbagi atas dua daerah 47.5% dikiri dan kanan mean yang sesuai dengan nilai Z = 1.96.

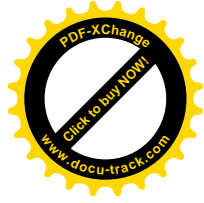
$\bar{x} = \alpha + 1.96 SD$

$$\begin{aligned} \bar{x} & \text{ atas/kanan} = \alpha + 1.96 SD \\ & = 85 + (1.96)(13) \\ & = 110.48 = 111 \text{ mmHg} \end{aligned}$$

$$\begin{aligned} \bar{x} & \text{ bawah/kiri} = \alpha - 1.96 SD \\ & = 85 - (1.96)(13) \\ & = 59.52 = 60 \text{ mmHg} \end{aligned}$$

95% peserta program skринing mempunyai tekanan darah diastolik antara 60 dan 111 mmHg.

Interval angka-angka diantara  $\bar{x} + 2 SD$  ini disebut sebagai 95% Confidence interval untuk  $\alpha$  yang mempunyai 2 aspek , yaitu suatu interval yang didapat dari perhitungan data dan tingkat kepercayaan yang memberikan kemungkinan bahwa cara ini menghasilkan yang dapat menggambarkan parameter.



Gambar.3

Dari contoh-contoh diatas terlihat bahwa kita dapat menggunakan kurva normal untuk menjawab berbagai pertanyaan tentang data kuantitatif. Banyak variabel dalam bidang kedokteran bila didapat dari kelompok individu yang besar, akan dapat digambarkan secara tepat dengan kurva normal. Implikasi penting dari hal-hal yang telah dibahas di atas antara lain dalam penghitungan 'confidence interval', karena pada dasarnya data yang menarik di bidang statistik (beda antara dua mean, proporsi)

akan mempunyai distribusi normal pada penarikan sampel berulang-ulang

### KESIMPULAN

Distribusi Normal yang merupakan suatu distribusi teoritis, yang tampaknya cukup sederhana, ternyata sangat berguna dalam penelitian-penelitian epidemiologis terutama dalam menentukan batas-batas nilai standar parameter.

### DAFTAR PUSTAKA

1. Altman, G.D., Bland, J.M.,1995. The Normal Distribution. *BMJ.*, 310:298 .
2. Bland, M. 1996. An introduction to medical statistics. 2nd ed. Oxford: Oxford University Press, 101 - 118.
3. Essex-Sorlie, D.1995. Medical Biostatistics & Epidemiology. Connecticut: A Simon & chuster Company, 71 - 85.
4. Moore, D.S., Mc Cabe, G.P.1986 Introduction to the practice of Statistics, New York: Freeman and Company, 80 - 5.
5. Rosner, B.1986. Fundamentals of Biostatistics, 2nd ed. Boston Duxbary Press, 102 - 110.
5. Simpson, J., Berry, G. 1998. Normal Distribution. Dalam: Kerr C et al ed's. Handbook of Public Health Methods. Sidney : The Mc Graw Hill Company. Inc, 212 - 216. 70.