Analisa Data Secara Statistik

Fungsi Distribusi, Distribusi Normal, Distribusi Binomial, Distribusi Poisson

Rita Prasetyowati
Fisika FMIPA UNY
2012

Makna fungsi distribusi

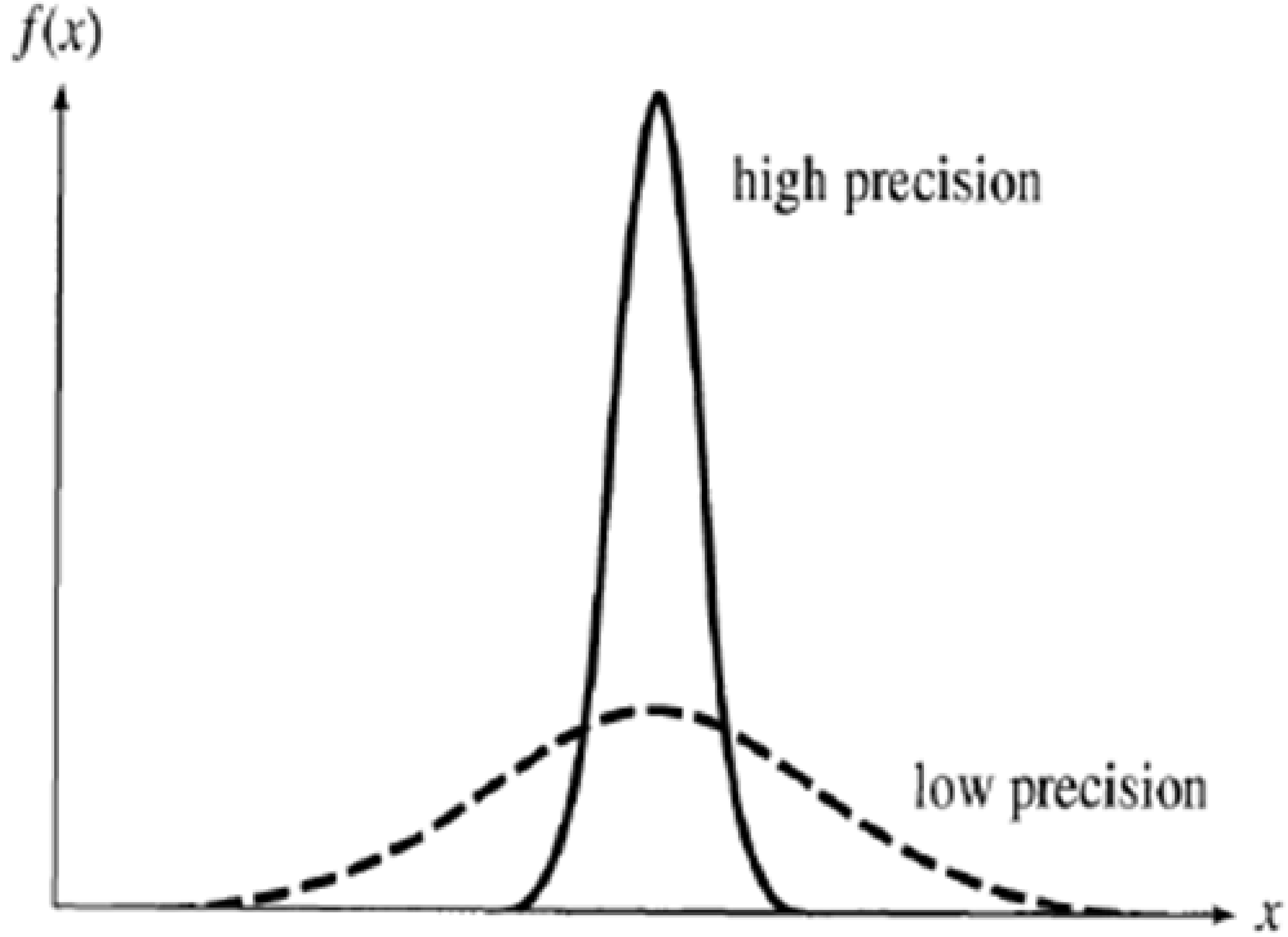
Fungsi distribusi *f(x)* memiliki makna bahwa suatu hasil ukur bernilai *x* mempunyai kebolehjadian untuk dapat ditemukan:

- Terletak antara *x* dan $x+\Delta x \rightarrow f(x) \equiv F_i$
- Terletak antara A dan B $\rightarrow \int_A^B f(x)dx$
- Terletak antara $-\infty$ dan $+\infty$, yang berarti $\int_{-\infty}^{+\infty} f(x)dx \equiv 1 \equiv \sum F_i$

  $\rightarrow$ ternormalisasi

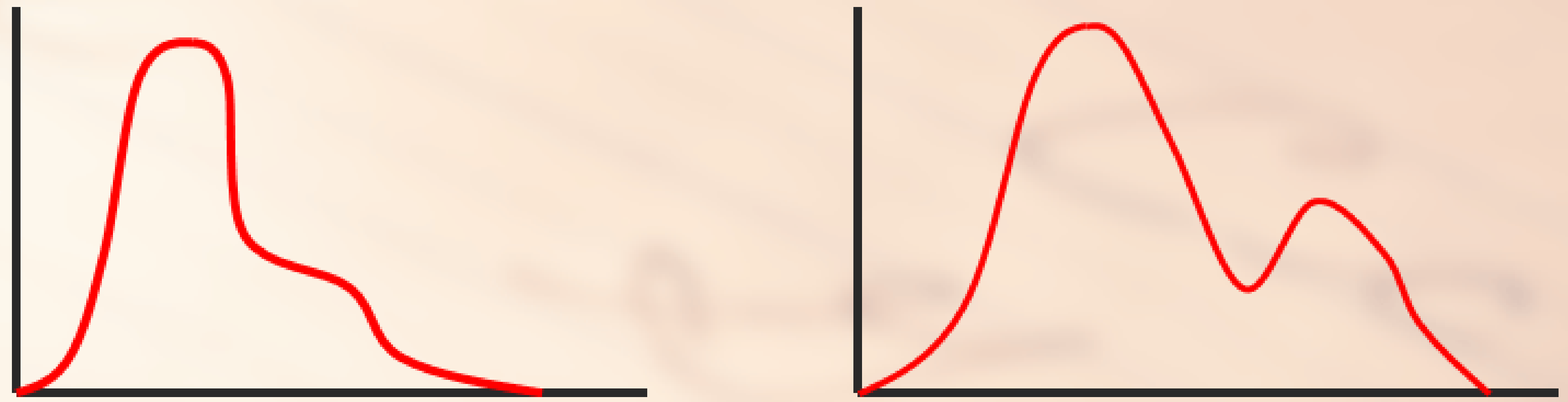
Sehingga $\bar{x} = \int_{-\infty}^{+\infty} x\, f(x)dx$, dimana $\bar{x}$ merupakan perkiraan terbaik dari nilai benar. Jika N $\rightarrow \infty$ maka $\bar{x} = x$ benar.

Lebar dan sempitnya kurva merupakan kriteria ketelitian dari pengukuran.



Gambar 16: dua grafik distribusi, pertama untuk presisi yang tinggi dan yang kedua untuk presisi yang rendah.

Suatu fungsi distribusi *f(x)* dari hasil pengukuran belum tentu meiliki bentuk umum seperti bell, tetapi dapat juga berbentuk non simetri bergantung dari hasil pengukuran yang diperoleh.



Selain itu terdapat suatu nilai dalam statistik yang digunakan untuk menunjukkan perbedaan karakteristik suatu populasi dibanding populasi yang lain. Nilai itu adalah nilai z, dimana

$$z = \frac{x - \mu}{\sigma}$$

Model Deskripsi Matematis dari fungsi Distribusi Data

Suatu persamaan analitis yang dapat memprediksikan distribusi data dari suatu hasil ekspermen, sangat bermanfaat untuk menginterpretasikan data dan untuk memandu suatu design eksperimen.

Parameter yang digunakan adalah cacah eksperimen yang dilakukan (*N*) dan kebolehjadian hasil yang diharapkan (*p*)

Berikut dijabarkan model matematis untuk fungsi distribusi data yang sering digunakan dalam eksperimen fisika, yaitu: distribusi normal (gaussian), distribusi binomial dan distribusi poisson.

Distribusi Normal (Gaussian)

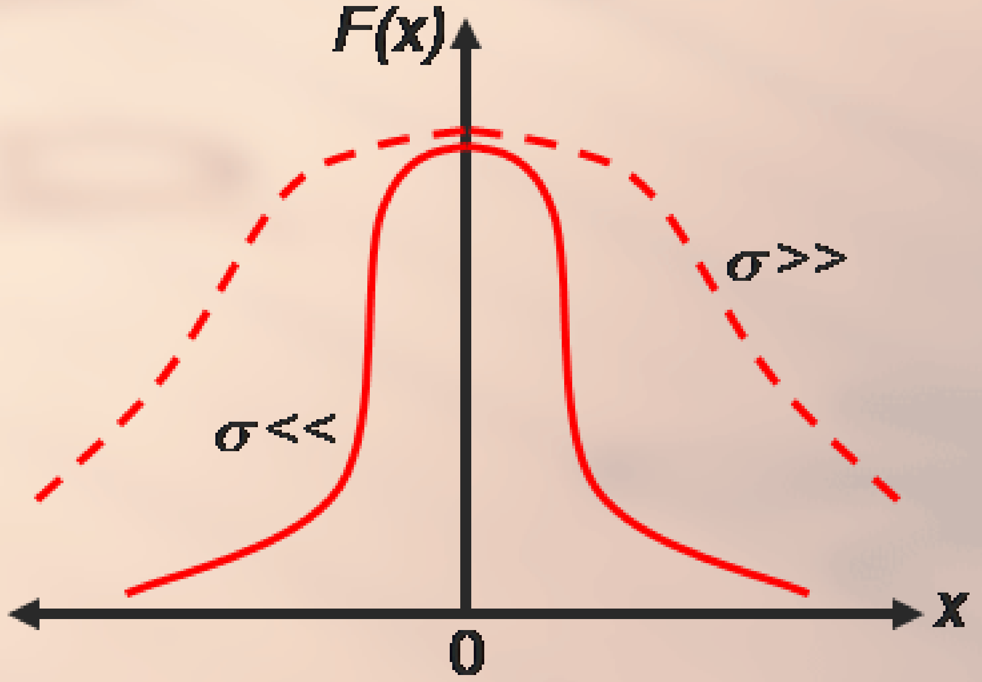
Fungsi distribusi yang paling penting untuk eksperimen fisika dan data eksperimen secara umum, adalah distribusi gaussian atau distribusi normal atau *bell curve distribution*.

Distribusi normal merupakan fungsi distribusi kebolehjadian yang bersifat kontinyu, yang memiliki makna bahwa setiap nilai dari variabel x yang bebas adalah mungkin.

Diketahui fungsi gauss

$$f(x) = e^{-x^2/2\sigma^2}$$

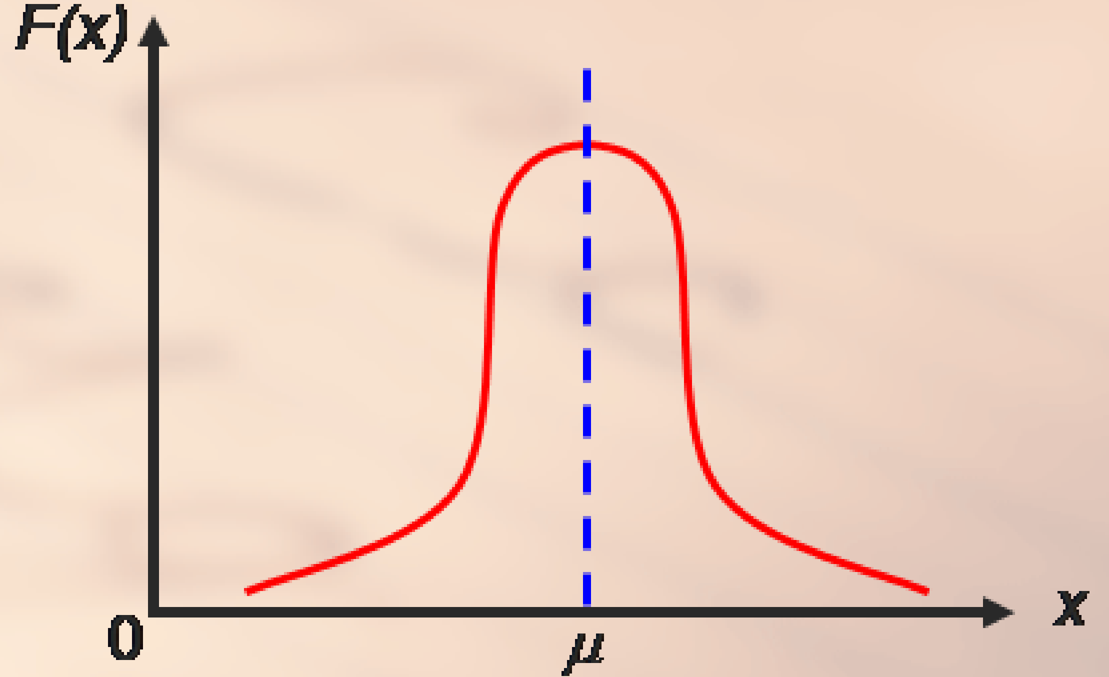
σ Melukiskan lebar sempitnya kurva, karena untuk σ kecil maka$(x^2/2\sigma^2)$ besar. Sehingga fungsi $f(x)$ cepat berubah



Secara umum dapat diturunkan persamaan gauss sbb:

Diperoleh fungsi gauss

$$f(x) = e^{-(x-\mu)^2 / 2\sigma^2}$$



Fungsi distribusi harus ternormalisir, yaitu

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

dimana $f(x) = N\,e^{-(x-\mu)^2 / 2\sigma^2}$

Maka diperoleh

$$N \int_{-\infty}^{\infty} e^{-(x-\mu)^2 / 2\sigma^2}\,dx = 1$$

Penyelesaian dengan subtitusi $x - \mu = y$ dan $dx = dy$

$$N\int_{-\infty}^{\infty} e^{-y^2/2\sigma^2}\, dy = 1$$

Diambil subtitusi lagi yaitu $z = y/\sigma$ dan $dy = \sigma\, dz$, diperoleh

$$N\sigma\int_{-\infty}^{\infty} e^{-z^2/2}\, dz = 1$$

$$N\sigma\sqrt{2\pi} = 1 \quad \rightarrow \quad N = 1/\sigma\sqrt{2\pi}$$

Dari penyelesaian tersebut diperoleh fungsi distribusi normal

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}}$$

Dengan μ = nilai benar (hanya diperoleh jika cacah data = ∞).

Bagaimana dengan $\bar{x}$? Nilai ini merupakan pendekatan terbaik dari nilai benar μ

$$\bar{x} = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

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

Diambil substitusi $x-\mu = y$; dx = dy, diperoleh

$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}} \left\{ \underbrace{\int_{-\infty}^{\infty} y e^{-y^2/2\sigma^2}\, dy}_{\text{fungsi gasal} = 0} + \mu \underbrace{\int_{-\infty}^{\infty} e^{-y^2/2\sigma^2}\, dy}_{\sigma\sqrt{2\pi}} \right\}$$

Sehingga

$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}}\mu\ \sigma\sqrt{2\pi} \rightarrow \bar{x} = \mu$$

Dengan kata lain dapat dinyatakan bahwa dalam distribusi normal nilai $\bar{x}$ = nilai benar.

Lantas bagaimana dengan nilai standar deviasi σ_x? Apakah sama dengan σ pada distribusi induk? Diketahui $\sigma_x = x - \bar{x}$

Kemudian hal ini dapat dibuktikan dengan menyelesaikan

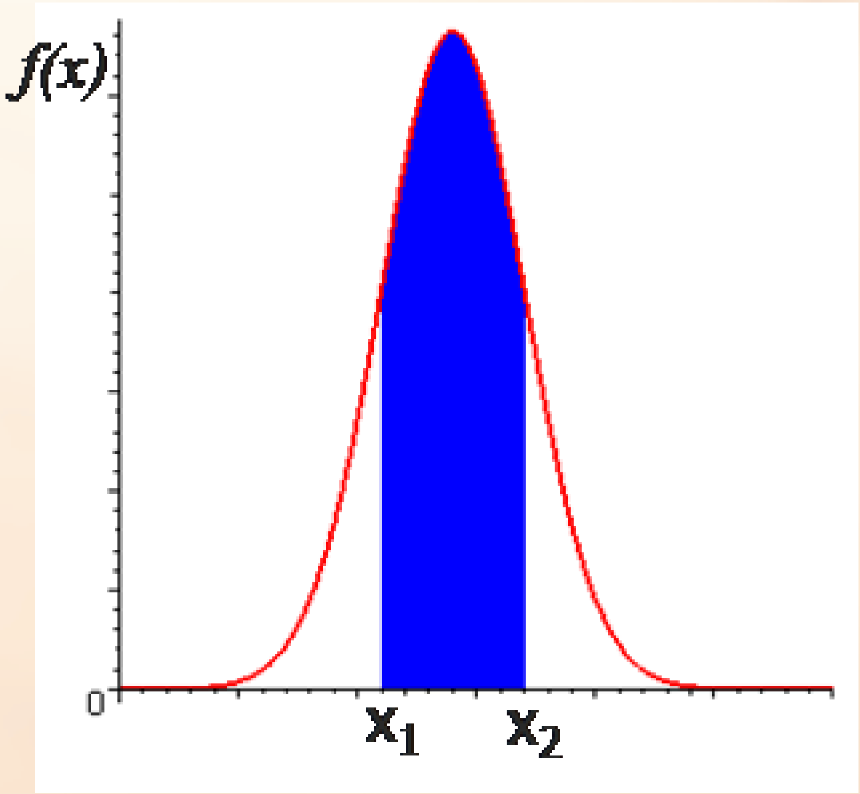
$$\sigma_x{}^2 = \int_{-\infty}^{\infty} (x-\bar{x})^2 f(x)\,dx$$

$$= \int_{-\infty}^{\infty} (x-\bar{x})^2 e^{-(x-\mu)^2/2\sigma^2}\,dx$$

dengan subtitusi $x - \sigma = y$ dan $dx = dy$, akan diperoleh $\sigma_x = \sigma$.

Integral kebolehjadian pada distribusi normal

Secara umum fungsi distribusi mengandung makna berapa kebolehjadian suatu hasil ukur dapat diperoleh. Nilainya dapat diketahui dengan melakukan Integral tertentu dari fungsi distribusi tersebut. Hal ini juga berlaku pada distribusi normal, dimana



$$P(x) = \int_{x_1}^{x_2} f(x)\,dx$$

Yang menyatakan kebolehjadian suatu hasil ukur dapat ditemukan di antara x_1 dan x_2.
Tentunya dengan $f(x)$ adalah fungsi distribusi normal.

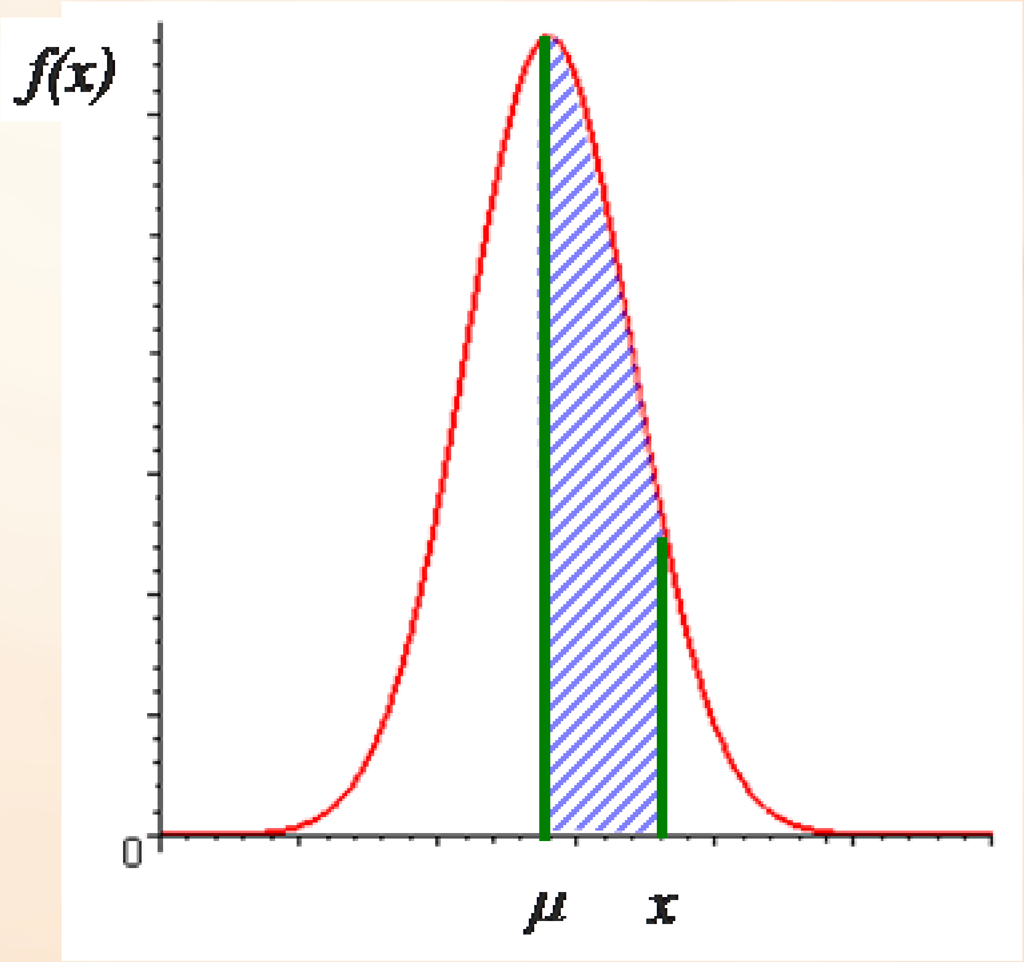
Pendahuluan

Analisa data secara statistik merupakan langkah analisis lebih lanjut jika kita melakukan pengukuran berulang (*repeated measurement*).

Pada analisa ini data yang diperoleh akan dianggap sesuai dengan model fungsi distribusi statistik tertentu dan kemudian dicari aspek-aspek pengukuran (peluang, nilai rata-rata, nilai ketakpastian/standar deviasi pengukuran, dll) berdasarkan model distribusi tersebut.

Umumnya yang paling ingin untuk diketahui adalah berapa kebolehjadian suatu hasil ukur akan berbeda sebesar Δx dari nilai rata-rata μ, sehingga

$$P(\mu \leq x) = \int_{\mu}^{x} f(x)\,dx$$



Dengan $\Delta x = |x - \mu|$ dan

$$\frac{|x-\mu|}{\sigma} = z$$

maka $\Delta x = z\,\sigma$

Sehingga kebolehjadian hasil ukur ditemukan antara $\mu - \Delta x \leq x \leq \mu + \Delta x$, adalah

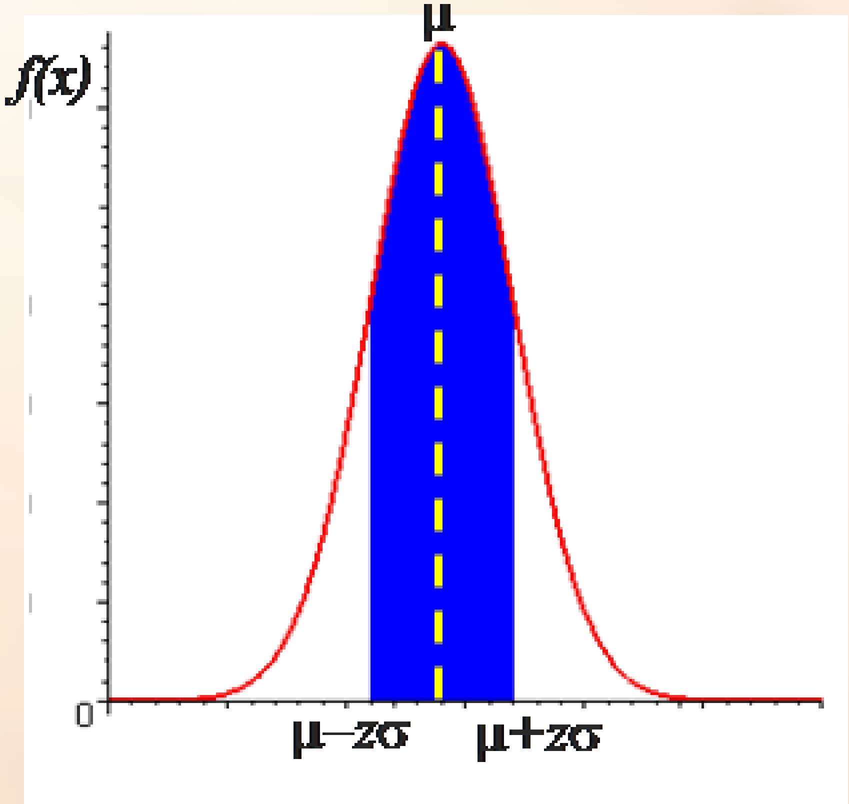
$$P(\mu \pm \Delta x) = \int_{\mu - z\sigma}^{\mu + z\sigma} f(x)\,dx$$

Nilai dari

$$P(\mu \pm \Delta x) = \int_{\mu - z\sigma}^{\mu + z\sigma} f(x)\, dx$$

Untuk

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

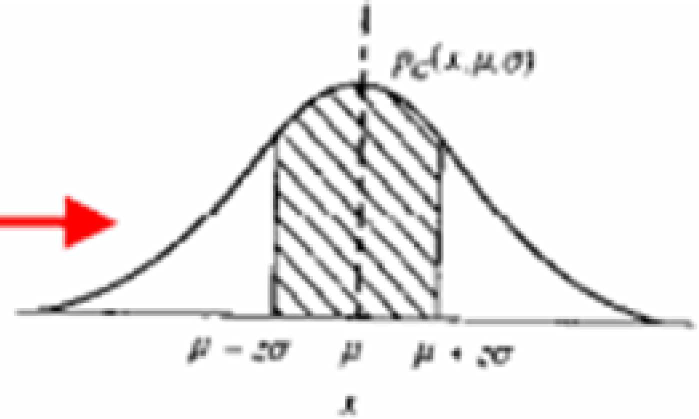


telah tersedia dalam tabel integral distribusi normal.

Nilai kebolehjadian *P* dapat diketahui jika nilai z diketahui

Tabel nilai integral distribusi normal

Perhatikan gambar kurva, karena nilai *P* akan berbeda jika daerah yang diarsir berbeda, bergantung jenis tabel yang anda gunakan



Contoh:
Jika nilai z = 0,74 maka nilai P diperoleh sebesar 0,54070 atau sebesar 54,070%

TABLE C.2
Integral of Gaussian distribution. The integral of the Gaussian probability density distribution. $P_G(x; \mu, \sigma)$ versus $z = |x - \mu|/\sigma$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0	0.00798	0.01596	0.02393	[illegible]	0.03988	0.04784	0.05581	0.06376	0.07171
0.1	0.07966	0.08759	0.09552	0.10343	[illegible]	0.11924	0.12712	0.13499	0.14285	0.15069
0.2	0.15852	0.16633	0.17413	0.18191	[illegible]	0.19741	0.20514	0.21284	0.22052	0.22818
0.3	0.23582	0.24344	0.25103	0.25860	[illegible]	0.27366	0.28115	0.28862	0.29605	0.30346
0.4	0.31084	0.31819	0.32551	0.33280	[illegible]	0.34729	0.35448	0.36164	0.36877	0.37587
0.5	0.38292	0.38995	0.39694	0.40389	[illegible]	0.41768	0.42452	0.43132	0.43809	0.44481
0.6	0.45149	0.45814	0.46474	0.47131	[illegible]	0.48431	0.49075	0.49714	0.50350	0.50981
0.7	[illegible]	[illegible]	[illegible]	[illegible]	0.54070	0.54674	0.55274	0.55870	0.56461	0.57047
0.8	0.57629	0.58206	0.58778	0.59346	0.59909	0.60467	0.61021	0.61570	0.62114	0.62653
0.9	0.63188	0.63718	0.64243	0.64763	0.65278	0.65789	0.66294	0.66795	0.67291	0.67783
1.0	0.68269	0.68750	0.69227	0.69699	0.70166	0.70628	0.71085	0.71538	0.71985	0.72428
1.1	0.72866	0.73300	0.73728	0.74152	0.74571	0.74985	0.75395	0.75799	0.76199	0.76595
1.2	0.76985	0.77371	0.77753	0.78130	0.78502	0.78869	0.79232	0.79591	0.79945	0.80294
1.3	0.80639	0.80980	0.81316	0.81647	0.81975	0.82298	0.82616	0.82930	0.83240	0.83546
1.4	0.83848	0.84145	0.84438	0.84727	0.85012	0.85293	0.85570	0.85843	0.86112	0.86377
1.5	0.86638	0.86895	0.87148	0.87397	0.87643	0.87885	0.88123	0.88355	0.88588	0.88816
1.6	0.89039	0.89259	0.89476	0.89689	0.89898	0.90105	0.90308	0.90507	0.90703	0.90896
1.7	0.91086	[illegible]	0.91456	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

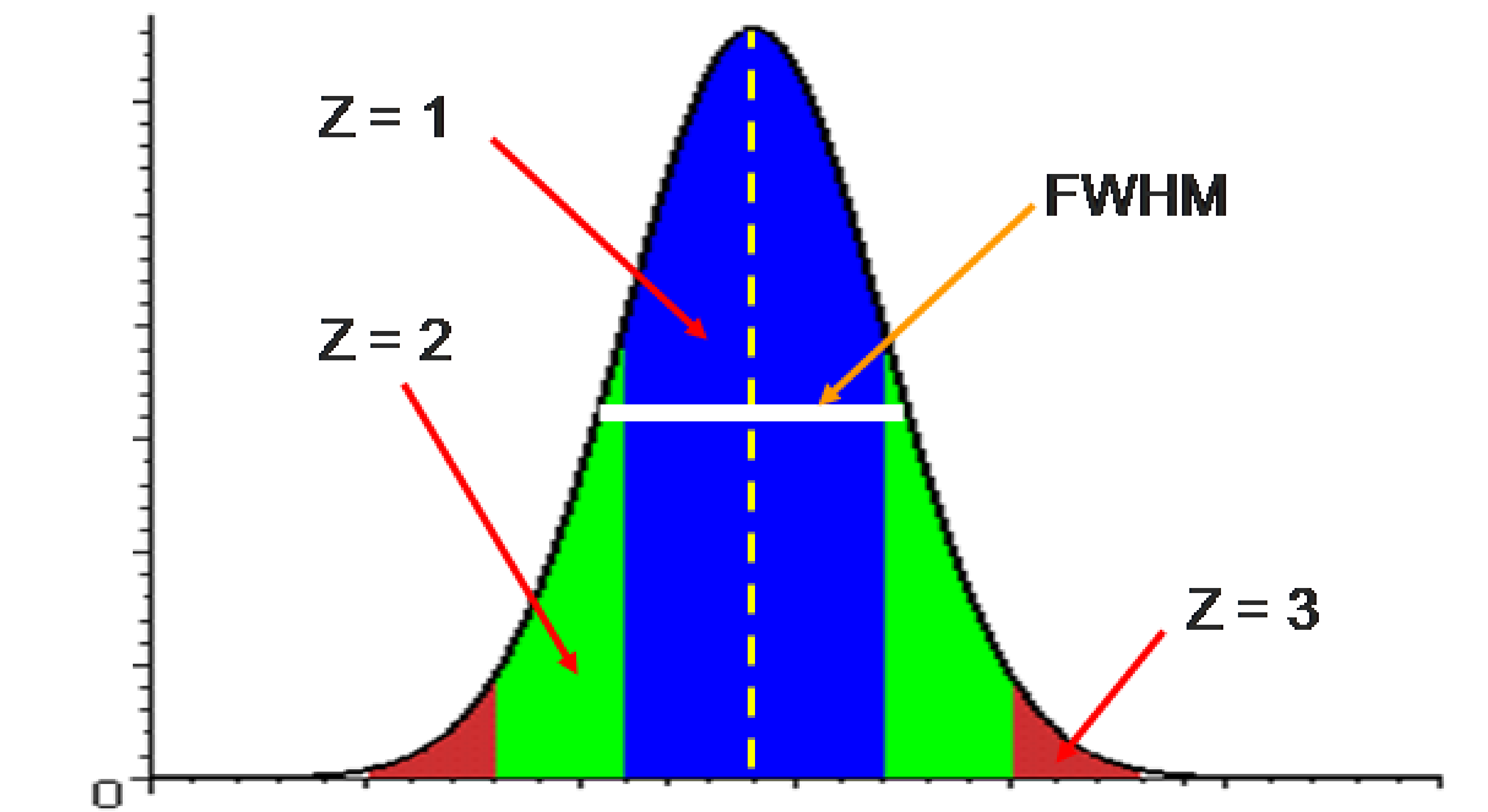
Dari tabel tersebut dapat diketahui bahwa

For $\Delta x = \sigma$, $z=1$ $P(z=1) = 0.683 \approx 68\%$

For $\Delta x = 2\sigma$, $z=2$ $P(z=2) = 0.954 \approx 95\%$

For $\Delta x = 3\sigma$, $z=3$ $P(z=3) = 0.997 \approx 99{,}7\%$

For $\Delta x = 4\sigma$, $z=4$ $P(z=4) = 0.99994 \approx 99{,}9\%$

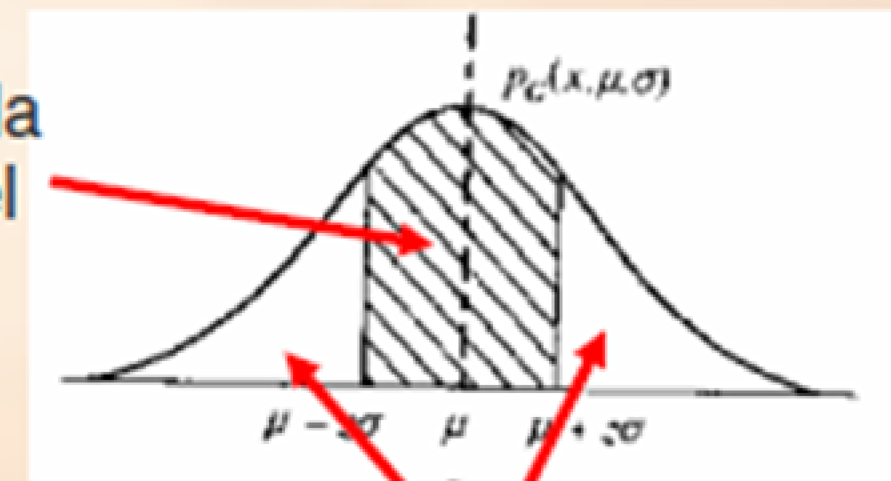


FWHM = *Full Width at Half Maximum* = $2{,}35\sigma$

Uji Kelayakan nilai ukur

- Jika dimiliki hasil pengukuran yang terpusat di nilai rata-rata μ dengan sebaran standar deviasi *S*, maka dapat dilakukan uji apakah sebuah nilai ukur *x* dapat dikatakan sesuai dengan kecenderungan atau tidak.
- Uji dilakukan dengan melihat apakah perbedaan antara *x* dengan μ dapat dikatakan signifikan atau tidak. Hal ini dilakukan dengan cara menghitung kebolehjadian *P* perbedaan hasil ukur berada diluar.

Kebolehjadian berada di dalam atau P tabel



Kebolehjadian berada di luar atau 1 - P tabel

Uji Kelayakan nilai ukur (lanjutan)

- Jika kebolehjadian *P* perbedaan hasil ukur berada diluar telah diketahui, maka untuk melihat apakah perbedaan tersebut signifikan atau tidak, nilai *P* tersebut dibandingkan dengan tingkat signifikansi (TS) yang disepakati misal 1% atau 5%.
- Kesimpulan diambil dengan melihat perbandingan hasil sbb:
  - Jika *P* diluar < TS maka perbedaan signifikan
  - Jika *P* diluar ≥ TS maka perbedaan tidak signifikan
- Perbedaan signifikan berarti nilai *x* tidak sesuai dengan kecenderungan pengukuran.
- Perbedaan tidak signifikan berarti nilai *x* sesuai dengan kecenderungan pengukuran.

Contoh soal :

1. Masa pakai, dinyatakan dengan X, untuk semacam alat dapat dilukiskan oleh fungsi densitas eksponensial dengan persamaan :

$f(x) = \frac{1}{2} e^{-\frac{1}{2} x}$, $x \geq 0$, dalam bulan dan $e = 2{,}7183$.

Tentukan peluang sebuah alat demikian yang dapat dipakai selama :

- antara 3 dan 3½ bulan,
- lebih dari 3 bulan,
- tentukan pula rata-rata masa pakainya.

Berdasarkan : Peluang X = x antara a dan b:

$$P\,(a< X < b) = \int_a^b f(x)dx.$$

Maka :

$$P\,(3 < X < 3\tfrac{1}{2}) = \int_3^{3\frac{1}{2}} \tfrac{1}{2}e^{-\frac{1}{2}x}\,dx = -e^{-\frac{1}{2}x}\Big|_{x=3}^{x=3\frac{1}{2}}$$

$$= -e^{-,75} + e^{-1,5}$$

$$= -\,0{,}1738 + 0{,}2231$$

$$= 0{,}0493.$$

Peluang masa pakai alat antara 3 dan 3½ bulan ialah 0,0493 atau 4,93%

$$P(3 < X < \infty) = \int_{3}^{\infty} \tfrac{1}{2} e^{-\frac{1}{2}x}\, dx = -e^{-\frac{1}{2}x} \Big|_{x=3}^{x=\infty}$$

$$= -0 + e^{-1,5} = 0,2231$$

Peluang masa pakai alat lebih dari 3 bulan adalah 0,2231 atau 22,31%

- Untuk x ≥ 0, maka :

$$E(X) = \int_{0}^{\infty} x \, \frac{1}{2} \, e^{-\frac{1}{2}x} \, dx$$

rata-rata masa pakainya adalah 2 bulan

Hal ini dapat dilakukan karena setiap pengukuran pasti memiliki ralat. Jika ralat sistematis diabaikan maka diharapkan hasil pengukuran yang diperoleh akan terdistribusi secara acak di sekitar “nilai benar”.

Jika tidak terdapat gangguan pengukuran, tentu hasil berikutnya akan diharapkan mempunyai kebolehjadian lebih besar untuk mendekati nilai benar.

Jika hal ini terpenuhi maka hasil pengukuran akan mempunyai distribusi kebolehjadian yang mengerucut, seperti bentuk bel (*bell shaped*), di sekitar nilai benar.

2. Dari penelitian terhadap 150 orang laki-laki yang berumur 40–60 tahun didapatkan rata-rata kadar kolesterol (μ) mereka 215 mg % dan simpangan baku σ = 45 mg %. Hitunglah peluang kita mendapatkan seorang laki-laki yang kadar kolesterolnya:

a. < 250 mg %

b. > 200 mg %

c. antara 200 –275 mg %

Latihan Soal:

1. Telah dilakukan pengukuran suatu besaran fisis dengan hasil akhir adalah $x = 46 \pm 5$. Jika salah satu hasil ukur bernilai 58, berapa kebolehjadian mendapatkan hasil ukur lebih besar daripada 58?

2. Dari suatu penelitian di Lab fisika, diperoleh data bahwa rata-rata cacah radiasi latar yang terukur adalah 20 cacah/menit, dengan simpangan baku 2 cacah/menit. Jika cacah radiasi latar berdistribusi normal, maka tentukan ada:
   - Berapa banyak data yang bernilai kurang dari 15 cacah/menit jika dilakukan 1000 kali pengukuran ?
   - Berapa banyak data yang bernilai lebih dari 21 cacah/menit jika dilakukan 1000 kali pengukuran ?
   - Berapa banyak data yang bernilai $18<x<22$?

Distribusi Binomial

Distribusi binomial merupakan salah satu dari distribusi kebolehjadian yang bersifat diskret.

Distribusi ini menggambarkan kebolehjadian suatu pengamatan akan sukses sebanyak x dari n percobaan dengan kebolehjadian sukses dari tiap percobaan adalah p.

Distribusi binomial diperoleh dari percobaan binom yang mempunyai ciri-ciri sbb:

- Percobaan terdiri atas n ulangan
- Dalam setiap ulangan, hasilnya dapat digolongkan sebagai berhasil (ya) atau gagal (tidak)
- Peluang berhasil p untuk setiap ulangan adalah sama
- Ulangan-ulangan bersifat bebas satu sama lain

Dari definisi tersebut dapat dirumuskan :

$$p_b(x,n,p) = C_x^n \, p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$p = \text{peluang berhasil}$$

$$q = \text{peluang gagal} = 1 - p$$

$$x = 0, 1, 2, 3, \ldots, n$$

$$\text{dengan nilai rata-rata distribusi} \quad \mu = np$$

$$\text{dan nilai simpangannya adalah} \quad \sigma^2 = np(1-p) = npq$$

$$\text{catatan: untuk } n >>, \text{ berlaku} \quad P_{binomial} \approx P_{gauss}$$

Contoh Soal:

Jika sebuah dadu dilempar sebanyak 5 kali, berapa:

a. Kebolehjadian memperoleh bulatan dua sebanyak 2 kali
b. Kebolehjadian memperoleh bulatan dua dan empat sebanyak 2 kali
c. Kebolehjadian memperoleh bulatan dua atau bulatan empat sebanyak 2 kali
d. Kebolehjadian memperoleh bulatan dua sebanyak lebih dari 2 kali
e. Kebolehjadian memperoleh bulatan dua kurang dari 3 kali
f. Kebolehjadian memperoleh bulatan dua sebanyak 2 sampai 4 kali

Ingat ! $P(A \text{ dan } B) = P(A)\,P(B)$

$P(A \text{ atau } B) = P(A) + P(B)$

Penerapan Distribusi Binomial:

Uji Hipotesa (langkah-langkahnya):

1. Pengamatan sistem (populasi)
2. Pengambilan hipotesa $\rightarrow$ dengan mengetahui p dan n, maka $\rightarrow$ $\mu_H = n\,p$
3. Pengambilan data, diperoleh $\rightarrow$ μ_E
4. Bandingkan antara μ_H dengan μ_E
5. Ambil dan uji kesimpulan, dengan asumsi:
   a. H_0 (Hipotesa Nol), yaitu hipotesa yang dirumuskan dengan harapan akan ditolak. Pada H_0 berlaku $\mu_H = \mu_E$.
   Atau
   b. H_1 (Hipotesa Alternatif), berlaku $\mu_H \neq \mu_E$ yang dapat berarti $\rightarrow$ $\mu_H > \mu_E$ atau $\mu_H < \mu_E$.

Untuk mengetahui apakah pada H_1 hipotesa dapat diterima atau ditolak, kita dapat memilih salah satu uji sebagai berikut:

- Untuk $\mu_H > \mu_E$, hitung kebolehjadian $\mu_H > \mu_E$ yaitu $P(\mu_H \geq \mu_E)$ yang hasilnya dibandingkan dengan tingkat signifikansi/TS (biasanya 1% atau 5%).

  Jika diperoleh $P(\mu_H \geq \mu_E) > TS \rightarrow$ Hipotesa ditolak

  Jika diperoleh $P(\mu_H \geq \mu_E) < TS \rightarrow$ Hipotesa diterima

- Untuk $\mu_H < \mu_E$, hitung kebolehjadian $\mu_H < \mu_E$ yaitu $P(\mu_H < \mu_E)$ yang hasilnya juga dibandingkan dengan TS.

  Jika diperoleh $P(\mu_H < \mu_E) > TS \rightarrow$ Hipotesa ditolak

  Jika diperoleh $P(\mu_H < \mu_E) < TS \rightarrow$ Hipotesa diterima

Ket: parameter TS 1% lebih signifikan daripada 5%

Contoh Soal 1:

Suatu pabrik lilin pelapis (*wax*) untuk papan selancar menyatakan bahwa "produknya bekerja dengan baik", selancar berjalan lebih cepat. Bagaimana me-nguji pernyataan tersebut? Diterima atau tidak?

Misal kita menguji 10 pasang selancar (10 menggunakan lilin dan 10 lainnya tanpa lilin), ternyata hasilnya adalah 8 selancar dengan lilin lebih cepat.

Jawab:

Untuk menguji kita ambil H_0. H_0 menyatakan tidak ada pengaruh wax, yang berarti *kebolehjadian selancar dengan lilin berjalan lebih cepat = kebolehjadian selancar tanpa lilin berjalan lebih cepat*. Sehingga p = ½, maka $\mu_H = n\,p = 10\ ½ = 5$

Karena hasil eksperimen adalah 8 dengan lilin lebih cepat maka kita uji hipotesa dengan menghitung $P(\mu_H \geq \mu_E)$, yaitu:

$$P(\mu_H \geq \mu_E) = P_8 + P_9 + P_{10}$$

$$P(\mu_H) = \frac{n!}{\mu!(n-\mu)!} p^{\mu} q^{n-\mu}$$

$$P(8) = \frac{10!}{8!(10-8)!}\left(\tfrac{1}{2}\right)^{8}\left(\tfrac{1}{2}\right)^{10-8} \approx 0{,}0439 \approx 4{,}39\%$$

$$P(9) = \frac{10!}{9!(10-9)!}\left(\tfrac{1}{2}\right)^{9}\left(\tfrac{1}{2}\right)^{10-9} \approx 0{,}0098 \approx 0{,}98\%$$

$$P(10) = \frac{10!}{10!(10-10)!}\left(\tfrac{1}{2}\right)^{10}\left(\tfrac{1}{2}\right)^{10-10} \approx 0{,}0010 \approx 0{,}10\%$$

$$P(\mu_H \geq \mu_E) \approx 5{,}47\%$$

$P(\mu_H \geq \mu_E)$ lebih besar dari TS (5% ataupun 1%) sehingga hipotesa ditolak! Hal ini berarti pernyataan pabrik adalah benar.

Contoh Soal 2:

Di sebuah negara akan diadakan pemilihan presiden, dengan dua calon yaitu Obam dan Hilar. Obam menyatakan 60% pemilih berada dipihaknya. LSI menguji pernyataan tersebut melalui polling, dengan hasil: dari 600 responden, 330 orang memilih Obam. Bagaimana hipotesa Obam Diterima atau tidak?

Jawab:

Hipotesa p = 0,6, maka $\mu_H = n\,p = 600 \times 0{,}6 = 360$

Pengujian dengan $P(\mu_H \geq \mu_E)$, yaitu mencari nilai $P(\mu_H \geq 330)$:

$$P(\mu_H \geq 330) = P_{330} + P_{331} + P_{332} + \ldots + P_{600}$$

$$P(\mu_H) = \frac{n!}{\mu!(n-\mu)!} p^{\mu} q^{n-\mu}$$

dimana $n = 600;\ p = 0{,}6;\ q = 0{,}4\ dan\ \mu = 360$

Histogram

Misal diperoleh hasil pengukuran sbb: 39, 38, 40, 43, 37, 39, 42, 41, 39, 38, 40, 41, 39, 42 dan 41.

Data tersebut akan lebih "mudah dilihat" jika diurutkan menjadi: 37, 38, 38, 39, 39, 39, 39, 40, 40, 41, 41, 41, 42, 42, 43.

Tampilan lebih baik akan terlihat dalam bentuk tabel:

X_i	37	38	39	40	41	42	43
n_i	1	2	4	2	3	2	1

Karena tidak dapat dihitung secara langsung, maka diselesaikan dengan penyelesaian distribusi gaussian.

$$z = \frac{x - \mu}{\sigma}$$

$$\text{dengan } \sigma = \sqrt{npq} = \sqrt{600 \times 0{,}6 \times 0{,}4} = 12, \text{ maka}$$

$$z = \frac{330 - 360}{12} = -2{,}5; \quad \text{dari tabel diperoleh :}$$

$$P(x \le \mu - 2{,}5\sigma) = 0{,}6\%$$

Berarti $P(\mu_H \geq \mu_E)$ lebih kecil dari TS sehingga hipotesa diterima. Hal ini berarti pernyataan Obam adalah benar.

Contoh :

1. Peluang untuk mendapatkan 6 muka G ketika melakukan undian dengan sebuah mata uang homogin sebanyak 10 kali adalah :

$$P\ (x = 6) = C_6^{10} (\tfrac{1}{2})^6 (\tfrac{1}{2})^4 = (210) (\tfrac{1}{2})^{10} = 0{,}2050$$

Dengan x= jumlah muka G yang nampak

2. Undian dengan menggunakan 10 buah dadu homogin sekaligus. Peluang nampaknya mata 6 sebanyak 8 buah, yaitu:

P (mata 6) = 1/6 dan N = 10, x = 8 dimana x berarti muka bermata 6 nampak disebelah atas, maka :

$$P\ (x=8) = C_8^{10} (1/6)^8 (5/6)^2 = 0{,}000015$$

Berarti undian dengan 10 dadu akan diperoleh mata 6 sebanyak 8 kali, terjadi kira-kira 15 dari tiap sejuta.

SOAL LATIHAN

Debit puncak banjir sungai Citarum-Nanjung periode T=5 tahun adalah 359m^3/det. Tentukan dalam waktu 10 tahun peluang debit banjir tersebut:

- Tidak terjadi
- Terjadi satu kali
- Terjadi dua kali
- Terjadi tiga kali
- Rata-rata dan deviasi standarnya ?

DISTRIBUSI POISSON

SEJARAH DISTRIBUSI POISSON

- Distribusi poisson disebut juga distribusi peristiwa yang jarang terjadi, ditemukan oleh S.D. Poisson (1781–1841), seorang ahli matematika berkebangsaan Perancis. Distribusi Poisson termasuk distribusi teoritis yang memakai variabel random diskrit.

- Menurut Walpole (1995), distribusi poisson adalah distribusi peluang acak poisson X, yang menyatakan banyaknya sukses yang terjadi dalam suatu selang waktu atau daerah tertentu.

Distribusi Poisson

Distribusi poisson juga merupakan salah satu dari distribusi kebolehjadian yang bersifat diskret.

Distribusi ini menggambarkan kebolehjadian banyaknya hasil percobaan yang terjadi selama selang waktu tertentu atau di suatu daerah tertentu.

Distribusi poisson diperoleh dari percobaan poisson yang mempunyai ciri-ciri sbb:

- Banyaknya hasil percobaan yang terjadi dalam selang waktu atau daerah tertentu tidak bergantung pada banyaknya hasil percobaan yang terjadi pada selang waktu atau daerah lain.
- Peluang terjadinya suatu hasil percobaan selama selang waktu yang singkat sekali atau daerah yang sempit sekali, sebanding dengan panjang selang waktu atau besarnya daerah tersebut.

- Peluang bahwa lebih dari satu hasil percobaan akan terjadi dalam selang waktu yang singkat atau daerah yang sempit tersebut, dapat diabaikan.

- Distribusi Poisson dapat pula dianggap sebagai pendekatan kepada distribusi binomial :
  n cukup besar dan P(A), sangat dekat kepada nol sehingga
  $\mu = n\,p$ tetap,
  ➔ distribusi binomial menjadi distribusi Poisson, dilakukan pendekatan $n \geq 50$ sedangkan $np < 5$

- Sehingga dapat dikatakan bahwa distribusi poisson adalah turunan dari distribusi binomial untuk *n* yang sangat besar dan μ yang sangat kecil. Hal ini dapat dirumuskan

$$P_p(x,\mu) = \frac{\mu^x e^{-\mu}}{x!}$$

dimana $\mu = n\,p$

dan $\sigma^2 = \mu$ maka $\sigma = \sqrt{\mu}$

Sehingga dalam distribusi poisson dapat dirumuskan:

$$\text{rata - rata} = \mu$$

$$\text{ralat mutlak} = \sigma = \sqrt{\mu}$$

$$\text{ralat relatif} = \sigma / \mu$$

Hal ini mempunyai implikasi pada desain eksperimen, yaitu: semakin besar μ maka ketakpastian relatifnya akan semakin kecil. Sebagai contoh dalam eksperimen fisika nuklir, untuk memperoleh ralat yang kecil maka waktu pengamatan diperpanjang.

Distribusi poisson banyak digunakan dalam hal berikut : Menghitung probabilitas terjadinya peristiwa menurut satuan waktu, ruang atau isi, luas, panjang tertentu, seperti menghitung probabilitas dari:

a. Banyaknya penggunaan telepon per menit
b. Banyaknya mobil yang lewat selama 5 menit di suatu ruas jalan,
c. Banyaknya bakteri dalam satu 1 liter air,
d. Banyaknya kesalahan ketik per halaman sebuah buku
e. Banyaknya kecelakaan mobil di jalan tol selama minggu pertama bulan Oktober.

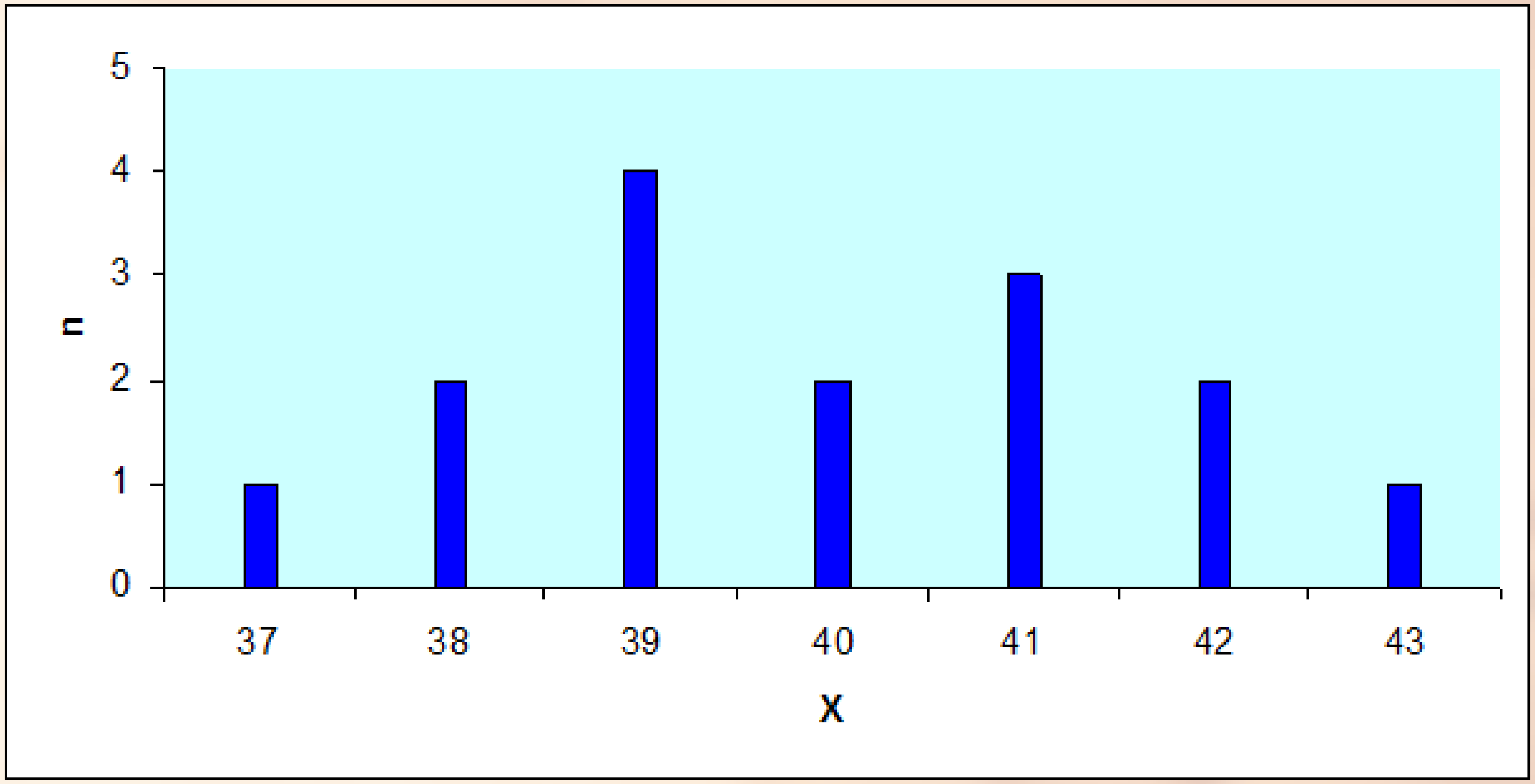
Contoh :

1. Banyak orang yang lewat melalui depan pasar setiap hari, tetapi sangat jarang terjadi seseorang menemukan barang hilang dan mengembalikannya kepada si pemilik atau melaporkannya kepada polisi.

2. Dalam tempo setiap 5 menit, operator telepon banyak menerima permintaan nomor untuk disambungkan, diharapkan jarang sekali terjadi salah sambung.

3. Misalkan rata-rata ada 1,4 orang buta huruf untuk setiap 100 orang. Berapa kemungkinan tidak ada orang buta huruf dari sebuah sampel berjumlah 200 orang

Data dapat disajikan dalam bentuk histogram



Jika hasil ukur bukan bilangan bulat, misal:

X_i	37 – 38	38 – 39	39 – 40	40 – 41	41 – 42	42 – 43
n_i	….	….	….	….	….	….

Contoh aplikasi

1. Peluang seseorang akan mendapat reaksi buruk setelah disuntik = 0,0005. Dari 4000 orang yang disuntik, tentukan peluang yang mendapat reaksi buruk:

a. tidak ada

b. ada 2 orang

c. lebih dari 2 orang, dan

d. Rata-rata ada berapa orang akan mendapat reaksi buruk.

2. Dalam suatu DPS dibangun dam pengendali banjir dengan umur bangunan 100 tahun. Berapa peluang terjadinya banjir 550 m^3/det dengan priode ulang 200 tahun selama priode umur dam tersebut, apabila ditentukan dengan Distribusi Poisson ?

1. a. Dengan menggunakan pendekatan distribusi Poisson kepada distribusi binomial, maka = n p = 4000 X 0,0005 = 2.
R = banyak orang yang mendapat reaksi buruk akibat suntikan

P(R=0) = $\dfrac{e^{-2}2^{0}}{0!} = 0{,}1353.$

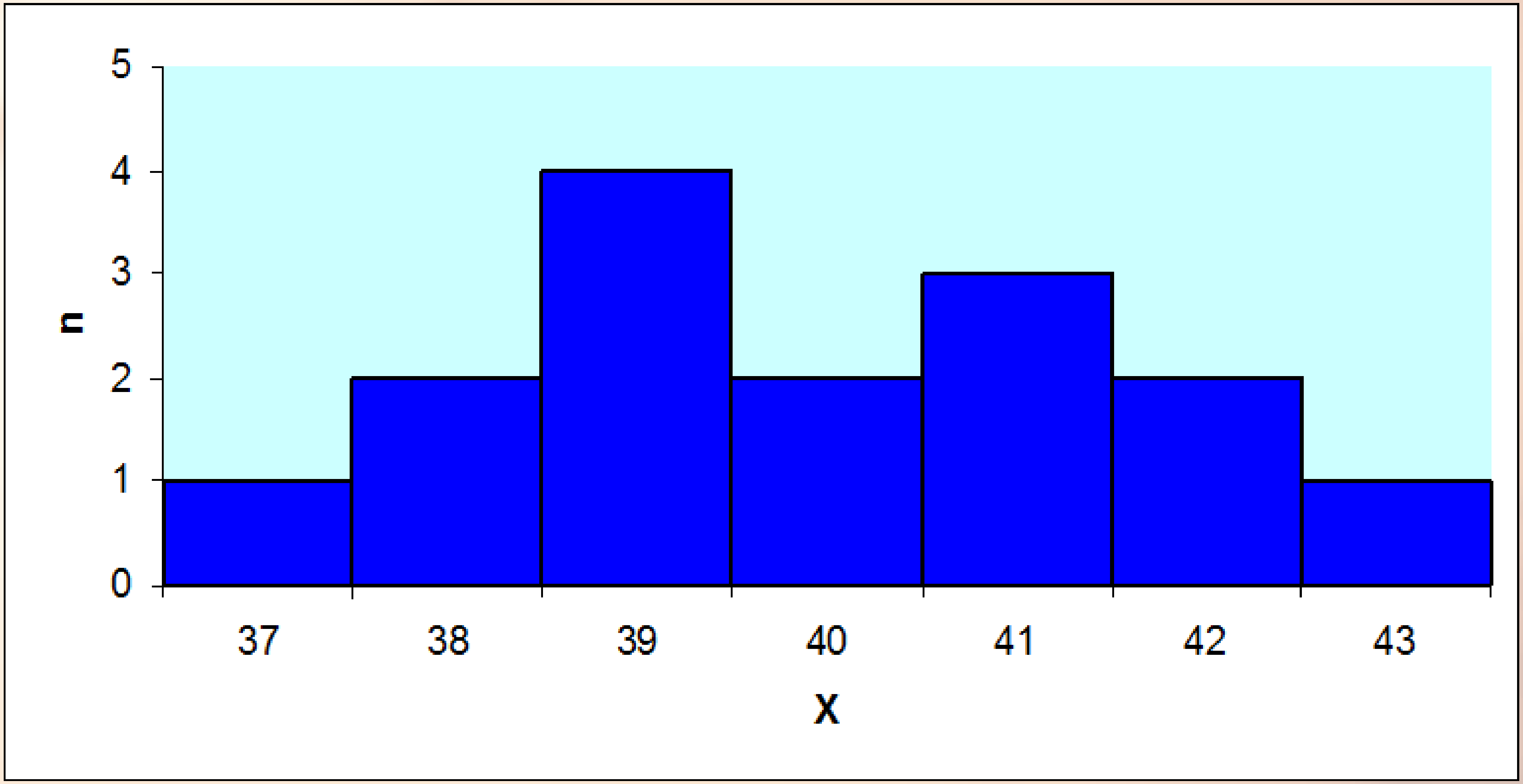
Peluang tidak ada orang yg mendapat reaksi buruk adalah **0,1353**

b. Dalam hal ini X = 2, sehingga :

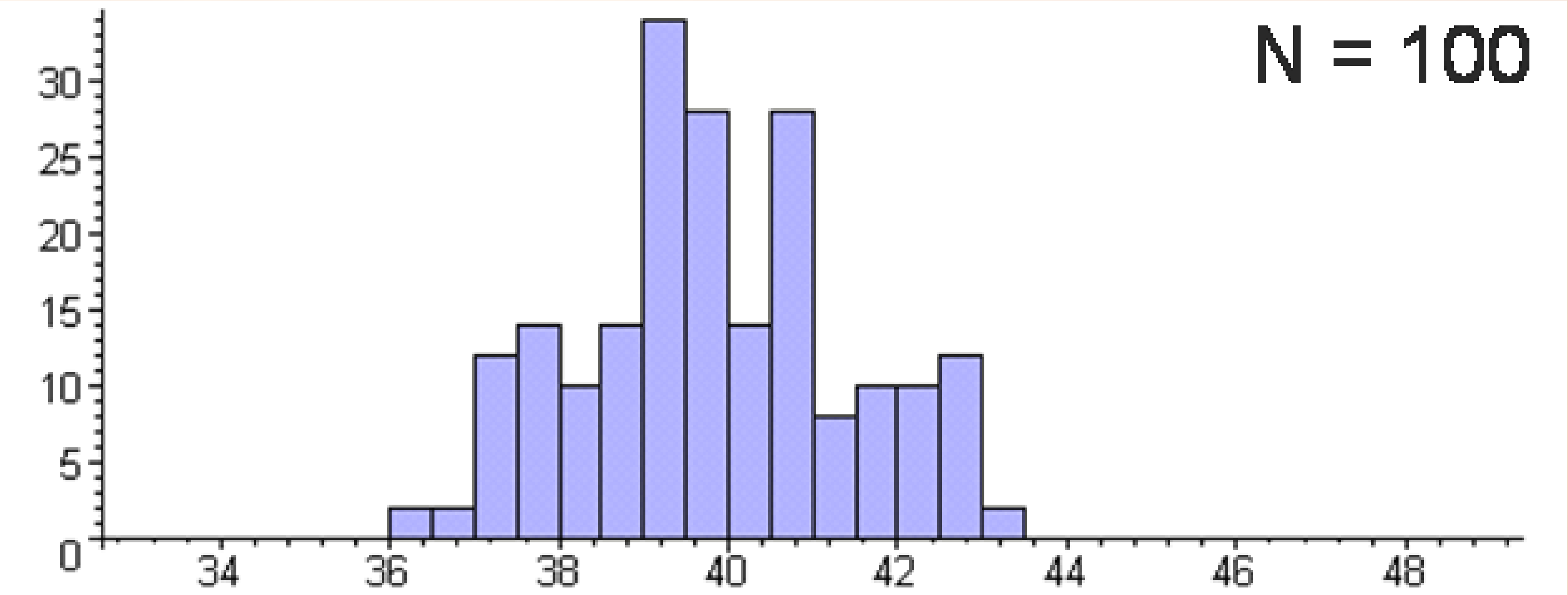
P(R=2) = $\dfrac{e^{-2}2^{2}}{2!} = 0{,}2706.$

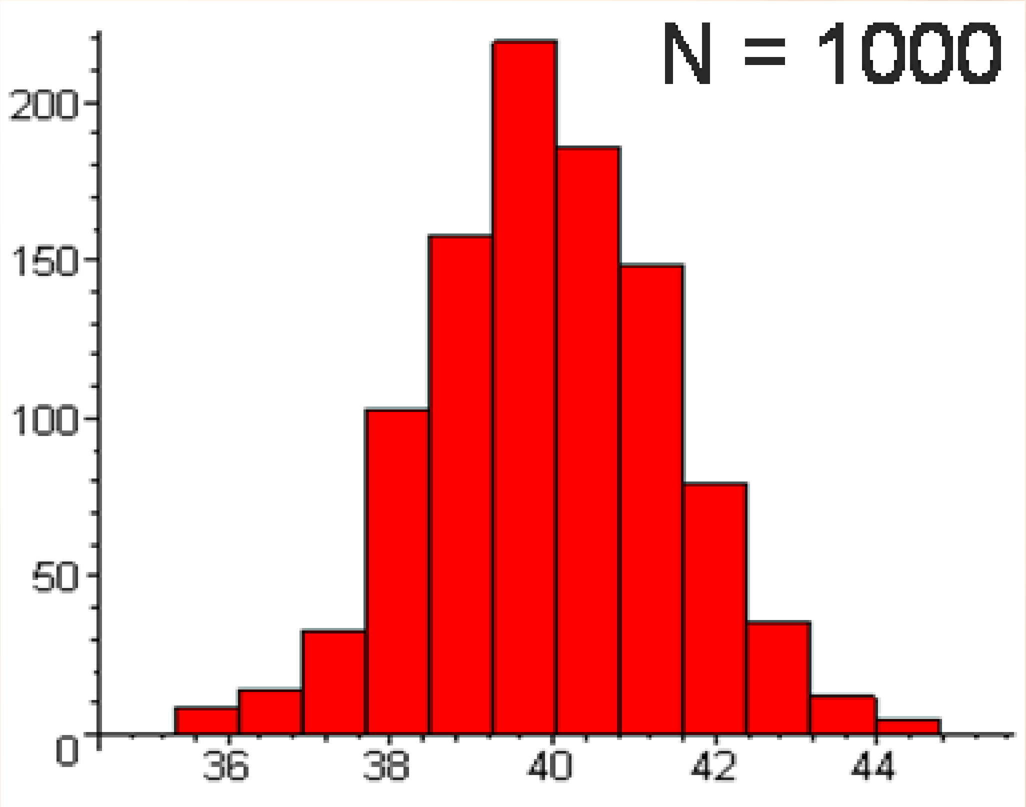
Peluang ada 2 orang mendapat reaksi buruk ialah 0,2706.

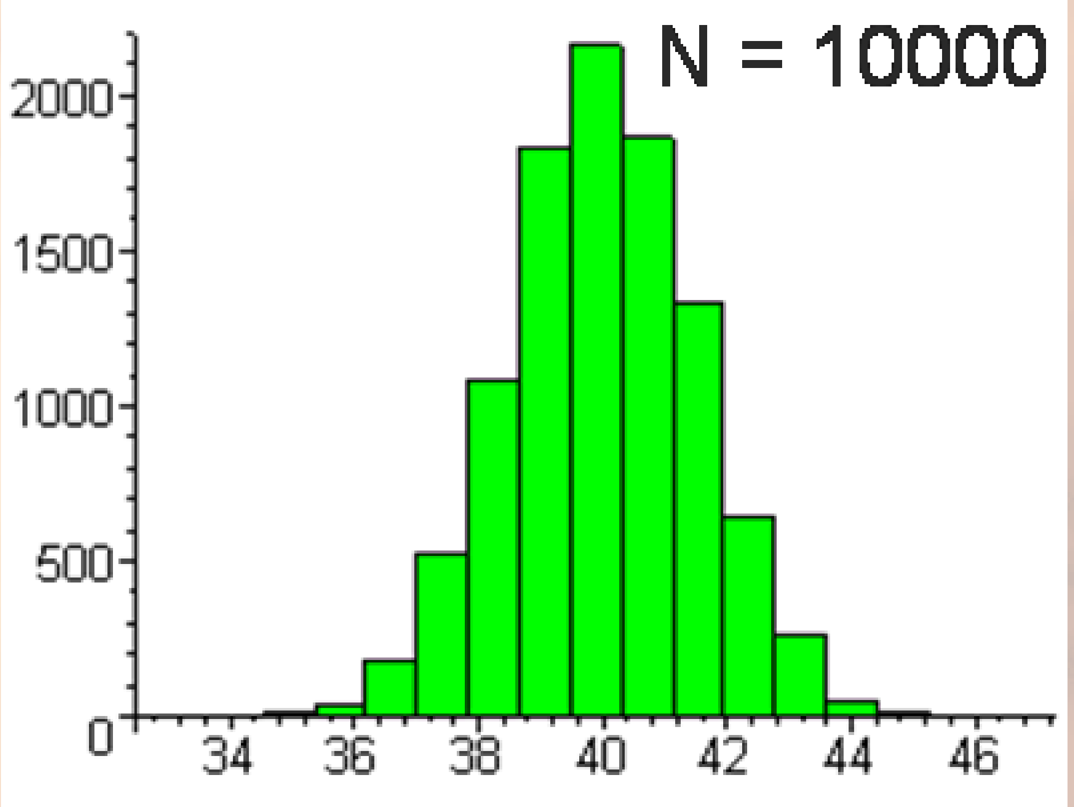
Data tersebut dapat disajikan dalam bentuk histogram sbb:



Jika cacah pengukuran diperbanyak maka histogram akan semakin berbentuk umum (*bell shaped*).

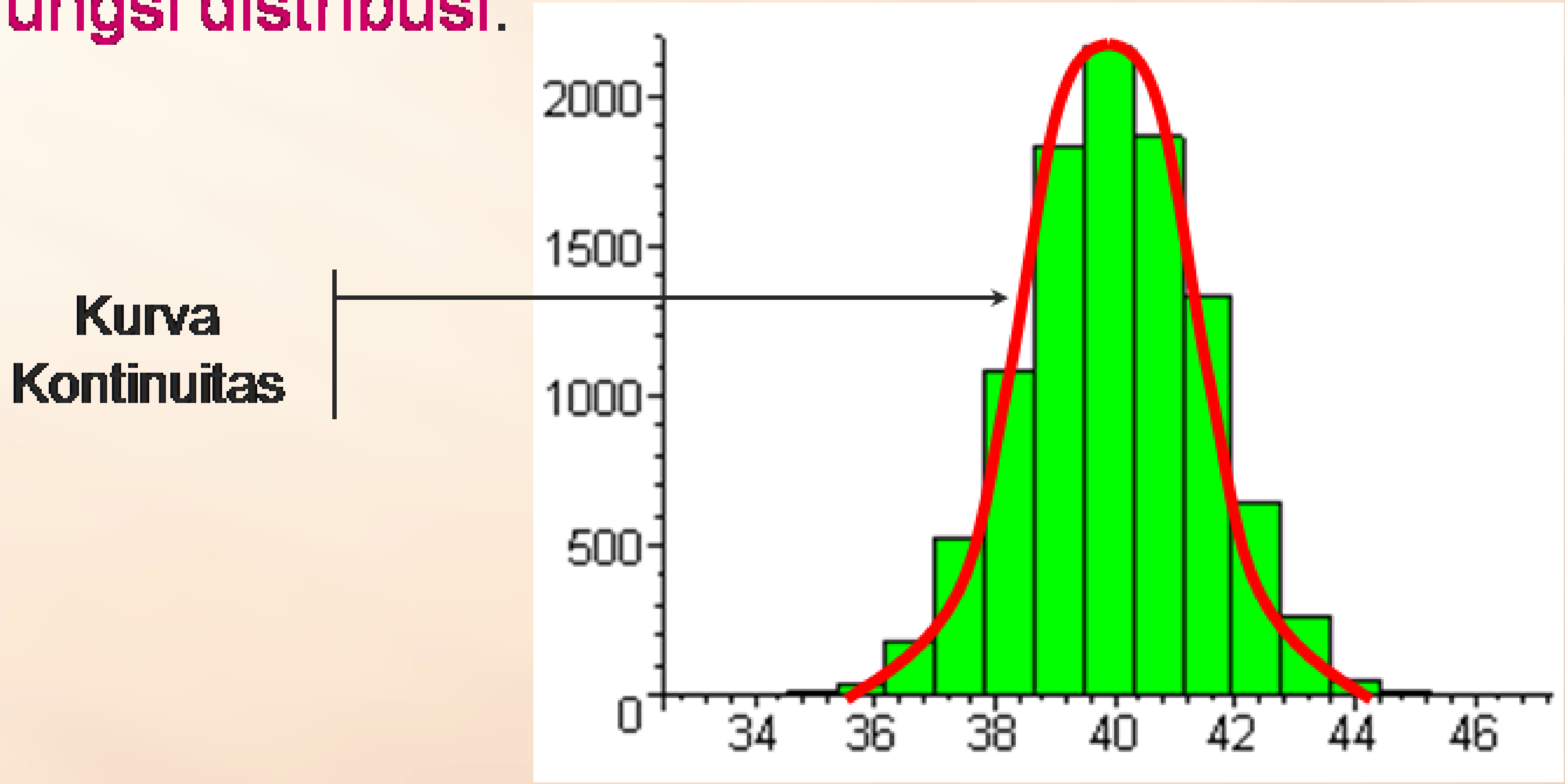

N = 100
0
5
10
15
20
25
30
34
36
38
40
42
44
46
48

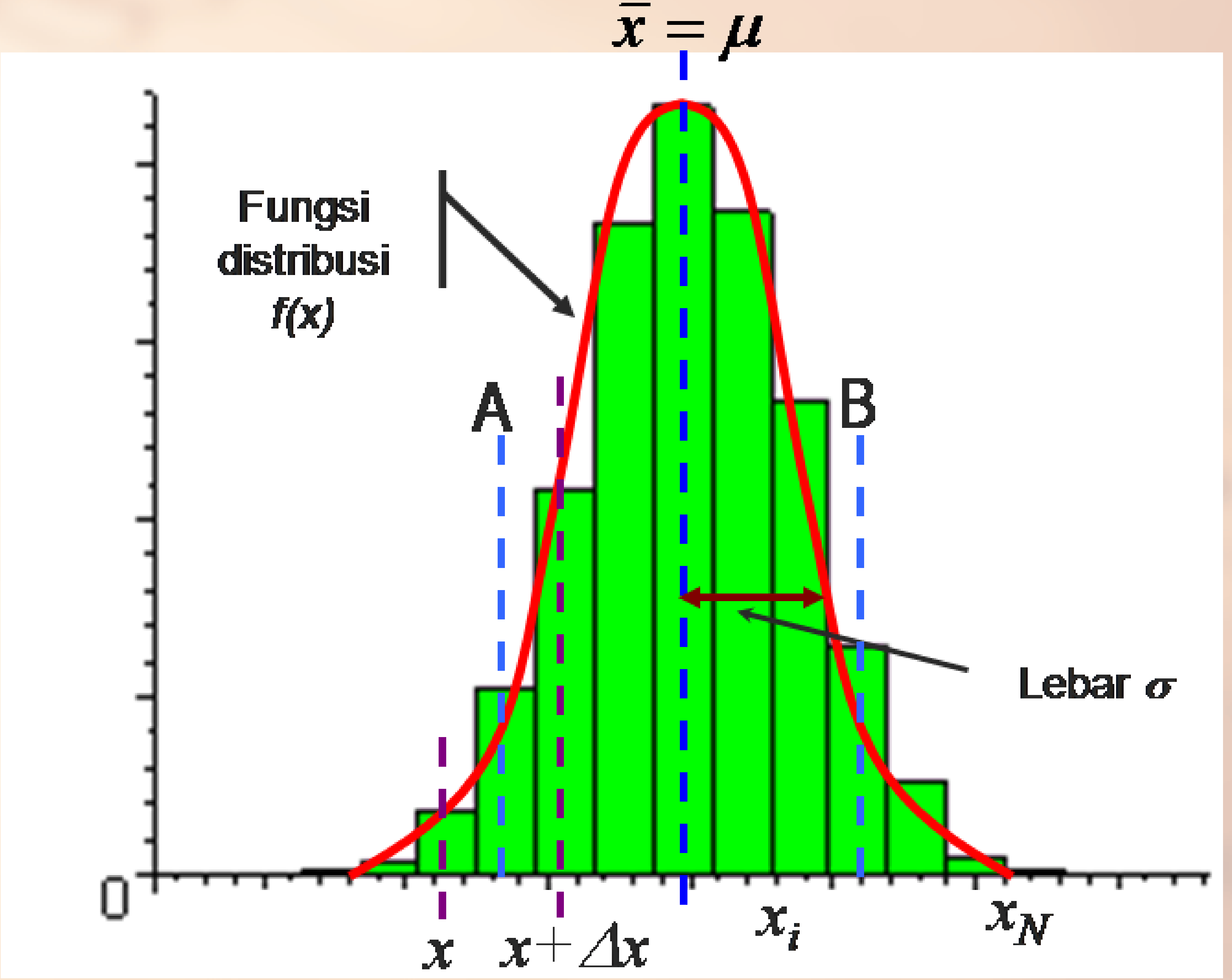

N = 1000
0
50
100
150
200
36
38
40
42
44


N = 10000
0
500
1000
1500
2000
34
36
38
40
42
44
46

Fungsi Distribusi

Jika cacah pengukuran ditambah sampai N → ∞ (distribusi induk) dan rentang antar satuan ukur diperkecil sampai Dx → 0, maka akan diperoleh fungsi kontinyu yang membentuk kurva kontinuitas. Fungsi kontinyu inilah yang dikenal sebagai fungsi distribusi.





$\bar{x} = \mu$
Fungsi distribusi $f(x)$
A
B
Lebar σ
0
x
$x + \Delta x$
x_i
x_N