

Materi Kuliah Statistika

Dipersiapkan oleh Dr. Ir. Widodo, MP

Konsep dasar statistika

PSG vs Chelsea versi bola.com

Ball possession 65 : 35

Usaha Tembakan ke gawang 16 : 7

Tembakan mengarah gawang 8 : 5

Skor 2 : 1

pengertian

- Asal kata “status” artinya negara → melayani keperluan administrasi negara
- Statistik : sekumpulan data, melukiskan sesuatu
- Statistika: metode, pengumpulan dan analisis data, informasi yang bermanfaat

Statistika

- Statistika deskriptif: statistika yang membicarakan mengenai cara pengumpulan dan menyederhanakan data
- Statistika induktif/inferensia: statistika yang membicarakan semua aturan dan cara menganalisa data dan menarik kesimpulan
 - Parametrik
 - non parametrik

Statistika inferensi

- Statistika parametrik: S.I yang mempertimbangkan nilai dari satu atau lebih parameter populasi, dan membutuhkan data dengan skala pengukuran minimal interval, prosedur dan penetapan teori berbijak pada asumsi data berdistribusi normal
- Statistika non parametrik: S.I yang tidak memerlukan parameter populasi, semua skala pengukuran (lebih tepat untuk nominal dan ordinal), tidak ada asumsi berdistribusi normal

Sampel vs Populasi

- Populasi: keseluruhan elemen yang menjadi perhatian dalam suatu penelitian
 - Populasi berhingga: banyaknya elemen diketahui
 - Populasi tak berhingga: banyaknya elemen tidak (dapat) diketahui
 - Besaran yang dihasilkan dinamakan parameter
- Sampel: bagian dari keseluruhan elemen yang diambil dengan metode tertentu, yang dapat mewakili dan menggambarkan karakteristik populasinya
 - Besaran yang dihasilkan dinamakan statistik

Data

- Data (bentuk jamak dari datum) merupakan informasi yang diperoleh dari pengamatan
- Sifatnya : kualitatif vs kuantitatif
- Sumbernya: internal vs eksternal
- Cara memperoleh: primer vs sekunder
- Waktu pengumpulan: time series vs cross setion

Variabel

- Variabel: sifat yang dimiliki oleh individu contoh yang berbeda antara satu individu (kelompok individu) dengan individu (kelompok individu) lain
- Variabel diskrit vs variabel kontinu

Skala Pengukuran

- Skala Nominal: membedakan/ mengelompokkan dalam kelompok/kategori
- Skala Ordinal: mengelompokkan dalam kelompok/ kategori dan memberi peringkat
- Skala Interval: mengelompokkan dalam kelompok/ kategori, memberi peringkat, dan mempunyai jarak yang sama diantara peringkat
- Skala Rasio: mengelompokkan dalam kelompok/ kategori, memberi peringkat, mempunyai jarak yang sama diantara peringkat, dan mempunyai nilai absolut

Juta $A = 1$ $B = 2$ $C = 2,5$
 Ratusan ribu $A = 10$ $B = 20$ $C = 25$

PENDAPATAN Alif Beni dan Cezy adalah masing2 adl

- 3 ; 4 ; 5 (diukur dlm juta rupiah)
- 3.000.000 ; 4.000.000 ; 5.000.000 (diukur rupiah)
- $3 : 4 : 5 = 3000000 : 4000000 : 5000000$

TEMPERATUR badan Alif Beni dan Cezy adalah masing2 adl

- 3 ; 4 ; 5 (derajat celcius)
- 19,4 ; 21,2 ; 23 (derajat fahrenheit)
- $3 : 4 : 5 \neq 19,4 : 21,2 : 23$

Permasalahan Statistika Deskriptif

- ▶ Bgmn menyajikan data agar berguna
- ▶ Bgmn menyatakan pola yg tersembunyi di balik data
- ▶ Bgmn menyarikan data bentuk pokok data

Data Mentah

Adlh hasil pengumpulan data yg blm diatur

contoh

Data hasil pengukuran panjang 40 helai daun salam yang dicatat sampai milimeter terdekat :

138	164	150	132	144	125	149	157
146	158	140	147	136	148	152	144
168	126	138	176	163	119	154	165
146	173	142	147	135	153	140	135
151	145	135	142	150	156	145	128

Array

Adlh data yg telah disajikan secara teratur

119	135	138	144	146	150	156	164
125	135	138	144	147	150	157	165
126	135	140	145	147	152	158	168
128	136	140	145	148	153	161	173
132	138	142	146	149	154	163	176

Lebih bermanfaat?

Pola : kisaran datanya 119 - 176

panjang 140-an mm lebih banyak

Distribusi/Tabel Frekuensi

- ▶ Hasil penyarian/peringkasan data
- ▶ Data didistribusikan ke dalam kelas/katagori tertentu
- ▶ Dist Frek yg pembagian kelasnya berupa angka = Dist Frek Kuantitatif
- ✓ Angka tunggal
- ✓ Berkelompok
- ▶ Dist Frek yg pembagian kelasnya bukan berupa angka = Dist Frek Kualitatif

Membuat Dist Frek

1. BANYAKNYA selang kelas

- Data yg banyak selang kelasnya juga banyak, demikian juga sebaliknya
 - Jangan terlalu banyak atau sedikit, 5 - 20
 - Patokan Sturges
- $$k = 1 + 3,322 \log n$$
- $$k = 1 + 3,322 \log 40$$
- $$= 1 + 5,322 = 6,322 \approx 7 \text{ (bulatkan ke atas)}$$

2. INTERVAL KELAS/selang kelas

$$i = \frac{\text{Rentang}}{\text{Banyak selang kelas}} = \frac{H - L}{k}$$

$$= \frac{176 - 119}{7} = 8,1428 \approx 9 \text{ (bulatkan ke atas)}$$

119 - 127 118 - 126 117 - 125
 128 - 136 127 - 135 126 - 134
 137 - 145
 172 - 181 172 - 180 171 - 179

Membuat Dist Frek (Ljt)

3. Membuat selang kelas dan memasukkan data

Panjang daun salam (mm)	Skor	Frekuensi
117 - 125	II	2
126 - 134	III	3
135 - 143	IIII IIII	10
144 - 152	IIII IIII III	13
153 - 161	IIII I	6
162 - 170	IIII	4
171 - 179	II	2
J u m l a h		40

Membuat Dist Frek (Ljt)

Sajikan dengan diberikan judul tabel dan sumber data

Distribusi Frekuensi panjang 40 daun salam (mm)

Panjang daun salam (mm)	Frekuensi
117 – 125	2
126 – 134	3
135 – 143	10
144 – 152	13
153 – 161	6
162 – 170	4
171 – 179	2
J u m l a h	40

Sumber: Analisis data primer

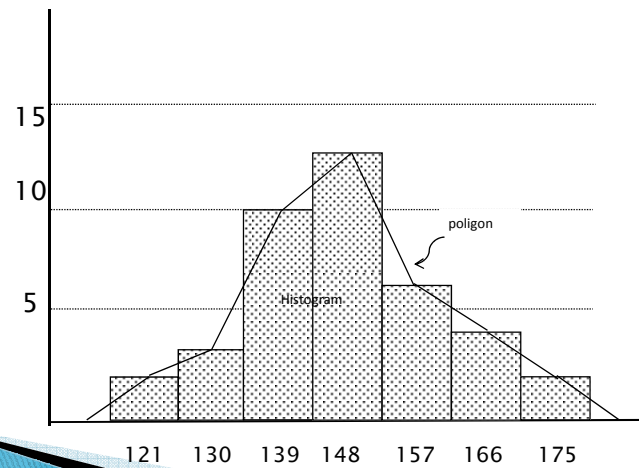
Histogram dan poligon

Histogram adalah himpunan segiempat, yang mempunyai:

- ▶ Alas pada sumbu mendatar (sumbu x) dengan pusat pada markah kelas dan lebarnya sama dengan ukuran selang kelas
- ▶ Luas segiempat sebanding dengan frekuensi kelas

Poligon adalah grafik garis yang frekuensi yang dirajah terhadap markah kelas.

Histogram dan poligon



Distribusi Frekuensi Relatif

Adlh Dist Frek yg dinyatakan dlm angka relatif

Distribusi Frekuensi Relatif panjang 40 daun salam (mm)

Panjang daun salam (mm)	Frekuensi Relatif (%)
117 – 125	$(2:40) \times 100 = 5$
126 – 134	7,5
135 – 143	25
144 – 152	32,5
153 – 161	15
162 – 170	10
171 – 179	5
J u m l a h	100

Sumber: Analisis data primer

Distribusi Frekuensi Kumulatif

Adlh total frekuensi semua nilai yang lebih kecil atau lebih besar dari batas suatu selang kelas tertentu

contoh

Distribusi frekuensi kumulatif "*kurang dari*" panjang 40 daun salam

Panjang daun salam (mm)	Frekuensi Kumulatif	Frekuensi Kumulatif relatif (%)
Kurang dari 117	0	0,0
Kurang dari 126	2	5,0
Kurang dari 135	5	12,5
Kurang dari 144	15	37,5
Kurang dari 153	28	70,0
Kurang dari 162	34	85,0
Kurang dari 171	38	95,0
Kurang dari 180	40	100,0

Sumber: Analisis data primer

Distribusi Frekuensi Kumulatif

contoh

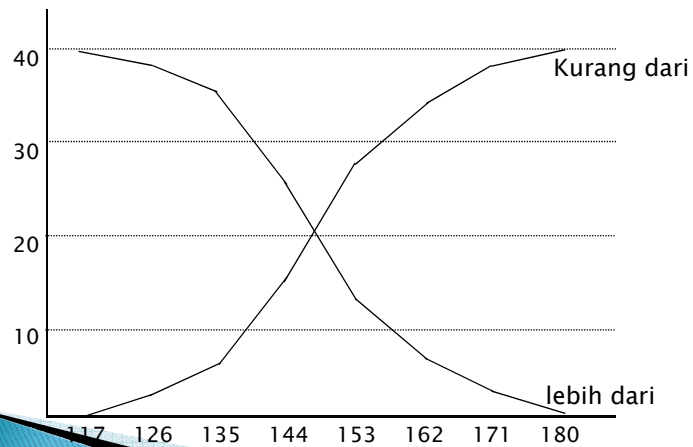
Distribusi frekuensi kumulatif "*atau lebih*" panjang 40 daun salam

Panjang daun salam (mm)	Frekuensi Kumulatif	Frekuensi Kumulatif relatif (%)
117 atau lebih	40	100,0
126 atau lebih	38	95,0
135 atau lebih	35	87,5
144 atau lebih	25	62,5
153 atau lebih	12	30,0
162 atau lebih	6	15,0
171 atau lebih	2	5,0
180 atau lebih	0	0,0

Sumber: Analisis data primer

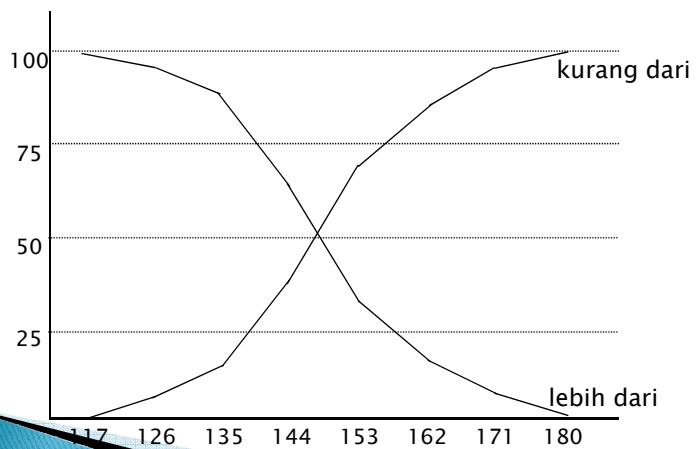
Ogife

Ogife adalah penyajian grafis dari distribusi frekuensi kumulatif



Ogife

Ogife adalah penyajian grafis dari distribusi frekuensi kumulatif





1. Nilai tengah (*arithmetic mean*)

= **Rata-rata hitung**, sering juga disebut dg **rata-rata, mean**

Nilai tengah dari suatu himpunan data $X_1, X_2, X_3, \dots, X_N$ adalah

$$\bar{X} = \frac{\sum X}{N}$$

700; 500; 500; 700; 640; 550; 660; 750

Maka nilai tengahnya adalah

$$\bar{X} = \frac{700 + 500 + 500 + 700 + 640 + 550 + 660 + 750}{8} = 625$$

$X_1, X_2, X_3, \dots, X_N$ masing-masing sebanyak $f_1, f_2, f_3, \dots, f_N$

$$\bar{X} = \frac{\sum fX}{\sum f}$$

Panjang daun salam(mm)	Markah kelas (X)	Frekuensi (f)	f.X
117 - 125	121	2	242
126 - 134	130	3	390
135 - 143	139	10	1390
144 - 152	148	13	1924
153 - 161	157	6	942
162 - 170	166	4	664
171 - 179	175	2	350
J u m l a h		40	5902

$$\bar{X} = \frac{5902}{40} = 147,55$$

$$\bar{X} = A + \frac{\sum fd}{\sum f}$$

A adalah sebarang nilai markah kelas, $d=X-A$

Panjang daun salam(mm)	Markah kelas (X)	d	Frekuensi (f)	f.d
117 - 125	121	-27	2	- 54
126 - 134	130	-18	3	- 54
135 - 143	139	-9	10	- 90
144 - 152	148 = A	0	13	0
153 - 161	157	9	6	54
162 - 170	166	18	4	72
171 - 179	175	27	2	54
J u m l a h			40	-18

$$\bar{X} = 148 + \frac{-18}{40} = 147,55$$

ukuran selang = c , dan $d = c.u \rightarrow u = d/c$

$$\bar{X} = A + \left(\frac{\sum f.u}{\sum f} \right) c$$

Panjang daun salam(mm)	Markah kelas (X)	u	Frekuensi (f)	$f.u$
117 - 125	121	-3	2	-6
126 - 134	130	-2	3	-6
135 - 143	139	-1	10	-10
144 - 152	148 = A	0	13	0
153 - 161	157	1	6	6
162 - 170	166	2	4	8
171 - 179	175	3	2	6
Jumlah			40	-2

$$\bar{X} = 148 + \left(\frac{-2}{40} \right) 9 = 147,55$$

2. Nilai tengah berbobot

rata-rata berbobot, memberikan bobot tertentu (w) pada masing-masing nilai amatan

$$\bar{X} = \frac{\sum wX}{\sum w}$$

Petani	Luas Sawah (Ha) = w	Panen (Ku/Ha) = X
A	3	40
B	4	30
C	5	45
D	2	50
E	3	70

Nilai tengah Panen per Ha

$$\bar{X} = \frac{3(40) + 4(30) + 5(45) + 2(50) + 3(70)}{3 + 4 + 5 + 2 + 3} = 45,59 \text{ Ku/Ha}$$

3. Nilai tengah geometrik

digunakan untuk mengetahui nilai tengah tingkat pertumbuhan setiap periode

Contoh: Tabungan awal = Rp. 1.000.000

Tk bunga th 1 = 7% → fk pertumbuhan = 1,07 → 1,07 X 1 jt

Tk bunga th 2 = 8% → fk pertumbuhan = 1,08 → 1,07 x 1,08 x 1 jt

Tk bunga th 3 = 10% → fk pertumbuhan = 1,10

Tk bunga th 4 = 12% → fk pertumbuhan = 1,12

Tk bunga th 5 = 18% → fk pertumbuhan = 1,18 → 1,07 x 1,08 x 1,10 x 1,12 x 1,18 x 1 jt

fk pertumbuhan = 1 + tk bunga

Nilai tengah fk pertumbuhan

$$\bar{x} = \sqrt[5]{1,07)(1,08)(1,10)(1,12)(1,18)} = 1,10933$$

Jadi nilai tengah tingkat bunga = 1,10933 - 1 = 0,10933
= 10,933%

4. Median

Median adalah nilai yang terletak di tengah dari suatu pengamatan, yaitu yg terletak pada posisi ke

$$(n+1):2 \text{ untuk data sebanyak } n$$

Contoh: skor tingkat penerapan teknologi 7 orang petani adalah

1, 1, 2, 3, 4, 4, 6

Posisi median = $(7+1):2 = 4$ → median adalah data ke-4

Median = 3

Untuk data yg telah tersusun dalam dist frek, median didekati dg rumus

$$Med = L_1 + \left(\frac{n/2 - \sum f_1}{f_{mez}} \right) \cdot c$$

	Panjang daun salam(mm)	Frekuensi
	117 - 125	2
	126 - 134	3
	135 - 143	10
Posisi median →	144 - 152	13
	153 - 161	6
	162 - 170	4
	171 - 179	2
	J u m l a h	40

Median adalah data ke $(40+1):2 = 20,5$ (antara 20 dan 21)
 Data ini terletak pada selang kelas ke 4

$$Med = 143,5 + \left(\frac{40/2 - 15}{13} \right) \cdot 9 = 146,96$$

5. Modus

Adalah data yang paling sering muncul

Suatu data yang belum dikelompokkan boleh jadi:

- tidak mempunyai modus, misalnya 1, 2, 3, 5, 8
- mono modus, misalnya 1, 2, 3, 5, 5, 8 (modusnya adalah 5)
- bimodus, misalnya 1, 2, 2, 5, 5, 8 (modusnya adalah 2 dan 5)

Data yang telah tersusun dalam dist frek, modus terletak pada selang kelas yg memiliki frekuensi terbesar

Modus didekati dengan rumus

$$M_o = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) \cdot c$$

	Panjang daun salam(mm)	Frekuensi
	117 – 125	2
	126 – 134	3
	135 – 143	10
Selang kelas modus ⇨	144 – 152	13
	153 – 161	6
	162 – 170	4
	171 – 179	2
	J u m l a h	40

modus berada pada selang kelas 144 – 152

$$Mo = 143,5 + \left(\frac{3}{3 + 7} \right) \cdot 9 = 146,2$$

- 1.000.000 → 1.070.000
- 1.070.000 → 1.155.600 (1,07 jt X 0.08) + 1,07 jt

UKURAN VARIASI

Macam ukuran variasi

- Rentang
- Nilai tengah simpangan
- Simpangan baku
- Varians
- Koefisien variasi, dan
- Skor baku

1. Rentang

- Rentang (R) adalah selisih antara nilai terbesar (H) dengan nilai terkecil (L) dalam suatu gugus data.

$$R = H - L$$

- Rentang gugus data A adalah $16 - 2 = 14$,
gugus B adalah $10 - 5 = 5$

Untuk data yang telah dikelompokkan dalam distribusi frekuensi nilai rentang (R) didekati dengan

$R = \text{markah tertinggi} - \text{markah terendah}$

Panjang daun salam(mm)	Markah kelas (X)
117 – 125	121
126 – 134	130
135 – 143	139
144 – 152	148
153 – 161	157
162 – 170	166
171 – 179	175

$$\begin{aligned} R &= 175 - 121 \\ &= 54 \end{aligned}$$

2. Nilai tengah simpangan (mean deviation – MD)

Nilai tengah simpangan (MD) adalah **jumlah harga mutlak penyimpangan** setiap data pengamatan terhadap nilai tengahnya, dan dibagi dengan banyaknya pengamatan

$$MD = \frac{\sum |X - \bar{X}|}{n}$$

Contoh produksi per hektar kacang tanah di Kabupaten Sragen tahun 2004 – 2008 adalah 1,19; 1,16; 1,13; 1,56; 1,29

Nilai tengah $(1,19 + 1,16 + 1,13 + 1,56 + 1,29) : 5 = 1,266$

$$MD = \frac{|1,19 - 1,266| + |1,16 - 1,266| + |1,13 - 1,266| + |1,56 - 1,266| + |1,29 - 1,266|}{5} = 0,1272$$

Untuk data yang telah dikelompokkan dalam distribusi frekuensi, didekati:

$$MD = \frac{\sum f |X - \bar{X}|}{n}$$

Panjang daun salam(mm)	Markah kelas (X)	$ X - \bar{X} $	Frekuensi (f)	$f \cdot X - \bar{X} $
117 – 125	121	26,55	2	53,1
126 – 134	130	17,55	3	52,65
135 – 143	139	8,55	10	85,5
144 – 152	148	0,45	13	5,85
153 – 161	157	9,45	6	56,7
162 – 170	166	18,45	4	73,8
171 – 179	175	27,45	2	54,9
Jumlah			40	382,5

$$MD = \frac{382,5}{40} = 9,5625$$

Varians

- Adalah **jumlah kuadrat** dari selisih data amatan dengan nilai tengahnya dibagi dengan banyaknya data
- Varians dilambangkan dengan σ^2 utk varians populasi atau s^2 utk varians sampel

$$\sigma^2 = \frac{\sum(X - \bar{X})^2}{n} = \frac{\sum x^2}{n} \quad x = (X - \bar{X})$$

$$s^2 = \frac{\sum(X - \bar{X})^2}{n - 1} = \frac{\sum x^2}{n - 1}$$

a b c d e s t o r

$\alpha \beta \chi \delta \varepsilon \sigma \tau \rho$

A B C D E S T O R

A B X Δ E Σ T O P

- 1,19; 1,16; 1,13; 1,56; 1,29
- Nilai tengah (\bar{x}) adalah $(1,19 + 1,16 + 1,13 + 1,56 + 1,29) : 5 = 1,266$

$$\sigma^2 = \frac{(1,19 - 1,266)^2 + (1,16 - 1,266)^2 + (1,13 - 1,266)^2 + (1,56 - 1,266)^2 + (1,29 - 1,266)^2}{5}$$

$$= 0,024504$$

- Varian bagi data yang telah tersaji dalam dist frek
Untuk varians populasi

$$\sigma^2 = \frac{\sum f(X - \bar{X})^2}{n} = \frac{\sum fx^2}{n} \quad x = (X - \bar{X})$$

Untuk varians sampel

$$s^2 = \frac{\sum f(X - \bar{X})^2}{n - 1} = \frac{\sum fx^2}{n - 1}$$

Panjang daun	Markah kelas (X)	$(X - \bar{X})$	$(X - \bar{X})^2$	Frek. (f)	$f(X - \bar{X})^2$
117 - 125	121	- 26,55	704,9025	2	1409,8050
126 - 134	130	- 17,55	308,0025	3	924,0075
135 - 143	139	- 8,55	73,1025	10	731,0250
144 - 152	148	0,45	0,2025	13	2,6325
153 - 161	157	9,45	89,3025	6	535,8150
162 - 170	166	18,45	340,4025	4	1361,6100
171 - 179	175	27,45	753,5025	2	1507,0050
J m l h				40	6471,9000

Rata-rata = 147,55

$$\sigma^2 = \frac{6471,9}{40} = 161,7975$$

Simpangan Baku

- Adalah akar dari varians

Simpangan baku populasi

$$\sigma = \sqrt{\frac{\sum(X - \bar{X})^2}{n}} = \sqrt{\frac{\sum x^2}{n}} \quad x = (X - \bar{X})$$

- Simpangan baku sampel

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}} = \sqrt{\frac{\sum x^2}{n - 1}}$$

Contoh:

Simpangan baku produksi per hektar kacang tanah di Kabupaten Sragen tahun 2004 – 2008

$$\sigma = \sqrt{\frac{0,12252}{5}} = 0,1565$$

- Untuk data yg telah tersaji dalam dist frek
Untuk simpangan baku populasi

$$\sigma = \sqrt{\frac{\sum f(X - \bar{X})^2}{n}} = \sqrt{\frac{\sum fx^2}{n}} \quad x = (X - \bar{X})$$

- Untuk simpangan baku sampel

$$s = \sqrt{\frac{\sum f(X - \bar{X})^2}{n - 1}} = \sqrt{\frac{\sum fx^2}{n - 1}}$$

- Dari contoh dist frek sblmnya

$$\sigma = \sqrt{\frac{6471,9}{40}} = 12,72$$

Koefisien Variasi

- Bandingkan
- Amatan dalam satuan ton/ha
1,13; 1,16; 1,19; 1,29; 1,56 → rentang = 0,43
- Amatan yang sama dalam satuan ku/ha
11,3; 11,6; 11,9; 12,9 ; 15,6 → rentang = 4,3
- UKURAN VARIASI TERGANTUNG SATUAN PENGUKURAN

- Koef Variasi merupakan ukuran variasi yang tdk tergantung satuan pengukuran amatan, sehingga koef variasi dapat digunakan sebagai perbandingan variasi

$$kv = \frac{\sigma}{\bar{X}} \times 100\%$$

Contoh

produksi per hektar kacang tanah (ton/ha) di Kabupaten Sragen tahun 2004 – 2008;

nilai tengah = 1,266; simpangan baku = 0,1565

$$kv_{\text{kacang}} = \frac{0,1565}{1,266} \times 100\% = 12,36\%$$

produksi kedelai;

nilai tengah = 1,6; simpangan baku = 0,2235

$$kv_{kdl} = \frac{0,2235}{1,6} \times 100\% = 13,97\%$$

Jika diperbandingkan, maka dapat disimpulkan bahwa
VARIASI PRODUKSI KEDELAI LEBIH TINGGI DIBANDINGKAN
VARIASI PRODUKSI KACANG TANAH

Skor Baku

$$z = \frac{X - \bar{X}}{\sigma}$$

- Distribusi data baru, yaitu $z_1, z_2, z_3, \dots, z_n$ mempunyai nilai tengah=0 dan simpangan baku=1
- tahun 2013 $\rightarrow X = 436; \bar{X}_0 = 412; \sigma_0 = 24$

$$z_{2013} = \frac{436 - 412}{24} = 1,00$$

- Tahun 2012 $\rightarrow X = 350; \bar{X}_0 = 335; \sigma_0 = 12$

$$z_{2012} = \frac{350 - 335}{12} = 1,25$$

Skor Baku

- Dalam penggunaannya, skor baku (z) sering dirubah menjadi distribusi data baru, yang mempunyai nilai tengah = \bar{X}_0 dan simpangan baku = σ_0

$$z = \bar{X}_0 + \sigma_0 \left(\frac{X - \bar{X}}{\sigma} \right)$$

Jika data diatas dinyatakan dalam skor baku, yg nilai tengah = 100 dan simpangan baku = 10

$$z_{2013} = 100 + 10 \left(\frac{436 - 412}{24} \right) = 110$$

$$z_{2012} = 100 + 10 \left(\frac{350 - 335}{12} \right) = 112,5$$

Contoh produksi per hektar kacang tanah di Kabupaten Sragen tahun 2004 – 2008 adalah 1,19; 1,16; 1,13; 1,56; 1,29

Nilai tengah = 1,266

Simp baku = 0,1565

$$Z_{1,19} = \frac{1,19 - 1,266}{0,1565}$$

$$Z_{1,16} = \frac{1,16 - 1,266}{0,1565}$$

- Suatu sampel terdiri dari 40 daun salam. Hasil pengukuran panjang ke 40 sampel daun salam tersebut disajikan dalam distribusi frekuensi sebagai berikut

DISTRIBUSI PROBABILITAS BINOMIAL, POISSON, DAN NORMAL

Pengertian probabilitas

$$Pr(E) = p = \frac{h}{n}$$

$$Pr(\text{bukan } E) = q = \frac{n - h}{n}$$

$$q = 1 - p$$

Contoh: pelemparan mata uang logam

$$Pr(\text{angka}) = p = \frac{1}{2} \quad Pr(\text{gambar}) = 1 - \frac{1}{2} = \frac{1}{2}$$

Pendekatan Peluang

- Pendekatan klasik, menyatakan bahwa terjadinya suatu peristiwa, E adalah rasio antara peristiwa E tersebut dengan seluruh peristiwa yang mungkin terjadi

$$Pr(E) = p = \frac{h}{n} \quad Pr(angka) = p = \frac{1}{2}$$

- Pendekatan frekuensi relatif, adalah pendekatan yang menggunakan perhitungan frekuensi relatif

Contoh:

dari pengamatan 40 daun salam, ternyata banyaknya daun yang mempunyai luas antara 117 – 126 mm² adalah sebanyak 2 lembar

Sehingga peluang satu lembar daun salam mempunyai luas antara 117 – 126 mm² adalah sebesar $2:40 = 0,05$

- Pendekatan subyektif, adalah pendekatan yang didasarkan pada tingkat kepercayaan individu yang membuat dugaan terhadap suatu peluang.

Misalnya

Pedagang sayur mayur berkeyakinan bahwa kemungkinan besar permintaan sayur mayur akan mengalami peningkatan

Para ahli berpendapat peluang Jokowi untuk menang pilpres sangat besar

Distribusi Binomial

- Adalah distribusi probabilitas yang hanya mempunyai 2 peristiwa yang mungkin terjadi

Contoh:

Pelemparan mata uang, peristiwa yang mungkin terjadi adalah munculnya sisi gambar atau sisi angka

Pendataan jenis kelamin (laki-laki ; perempuan)

Inseminasi buatan terhadap ternak sapi, maka peristiwa yang mungkin terjadi adalah berhasil ataukah gagal

- p adalah probabilitas suatu peristiwa;
- probabilitas akan terjadi sebanyak X peristiwa dari percobaan sebanyak N

$$Pr(X) = {}_N C_X \cdot p^X (1-p)^{N-X} = \frac{N!}{X!(N-X)!} p^X (1-p)^{N-X}$$

- $X = 0, 1, 2, 3, \dots, N$
- $N! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot N$ misalnya $4! = 1.2.3.4 = 24$
 $0! = 1$

Sifat-sifat distribusi binomial adalah

- Nilai tengah suatu distribusi binomial adalah

$$\mu = Np$$

- Simpangan baku suatu distribusi binomial adalah

$$\sigma = \sqrt{Np(1-p)}$$

- Varians distribusi binomial adalah $\sigma^2 = Np(1-p)$

- Contoh: dalam pelanturan mata uang (N) 6 kali; Probabilitas memperoleh "sisi angka" (X) sebanyak 2

$$\begin{aligned} Pr(X = 2) &= {}_6C_2 \cdot 0,5^2(1 - 0,5)^{6-2} \\ &= \frac{6!}{2!(6-2)!} 0,5^2(1 - 0,5)^{6-2} \\ &= \frac{720}{2(24)} (0,25)(0,0625) = 0,234375 \end{aligned}$$

- 20% dari buah durian yang dihasilkan adalah cacat $\rightarrow p = 0,2$
 - dari 4 buah durian
- a. didapati buah durian cacat sebanyak 1 buah

$$\begin{aligned}
 Pr(X = 1) &= {}_4C_1 \cdot 0,2^1(1 - 0,2)^{4-1} \\
 &= \frac{4!}{1!(4-1)!} 0,2^1(1 - 0,2)^{4-1} \\
 &= 0,4096
 \end{aligned}$$

$$N = 4$$

$$p = 0,2$$

- b. paling sedikit 2 buah akan cacat

$$\begin{aligned}
 Pr(\text{paling sedikit 2 buah}) &= Pr(X = 2, 3, 4) \\
 &= Pr(X=2) + Pr(X=3) + Pr(X=4) \\
 &= 1 - (Pr(X=0) + Pr(X=1))
 \end{aligned}$$

$$Pr(X = 1) = 0,4096$$

$$Pr(X = 0) = 0,4096$$

$$\begin{aligned}
 Pr(\text{paling sedikit 2 buah}) &= 1 - (0,4096 + 0,4096) \\
 &= 0,1808
 \end{aligned}$$

- c. Paling banyak 1 buah
- d. Paling banyak 2 buah
- e. Paling banyak 4 buah

Distribusi Poisson

- turunan dari distribusi binomial
- yang secara khusus berlaku untuk kasus yang mempunyai ukuran populasi (N) yang sangat besar
- mempunyai peluang terjadinya suatu peristiwa (p) teramat kecil

$$Pr(X) = \frac{\lambda^X e^{-\lambda}}{X!}$$

- Probabilitas ditolaknya pasokan daging ayam pada suatu supermarket adalah 0,001. Probabilitas dari 4.000 potong daging ayam yang disetor ke supermarket

- a. Probabilitas daging ayam akan ditolak sebanyak 3 potong

$$\lambda = \mu = 4.000 (0,001) = 4$$

$$Pr(X = 3) = \frac{4^3 2,71828^{-4}}{3!} = 0.19537$$

- b. Probabilitas daging ayam tidak ada yang ditolak
 $\rightarrow X = 0$

$$Pr(X = 0) = \frac{4^0 2,71828^{-4}}{0!} = 0,01832$$

- c. Probabilitas daging ayam yang ditolak paling banyak 3 potong

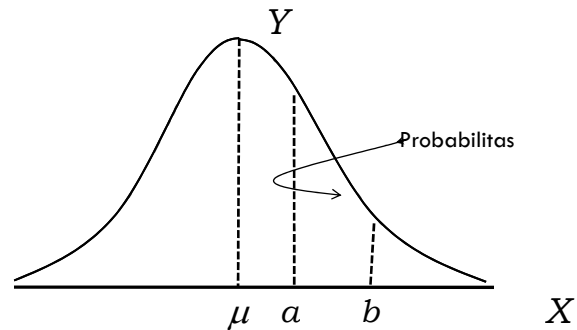
Distribusi Normal

- Distribusi normal didefinisikan sbg

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

μ = nilai tengah, σ = simpangan baku,
 π = 3,14159; e = 2,71828

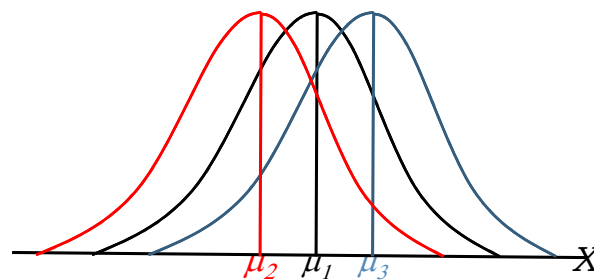
Kurva Normal



- Luas seluruh daerah bawah kurva = 1 → probabilitas
- Luas daerah bawah kurva antara $X = a$ dan $X = b$ → probabilitas X mempunyai nilai antara a dan b

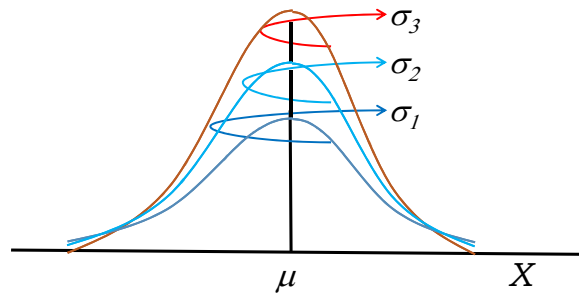
Bentuk kurva normal

- Jika μ berbeda namun σ sama



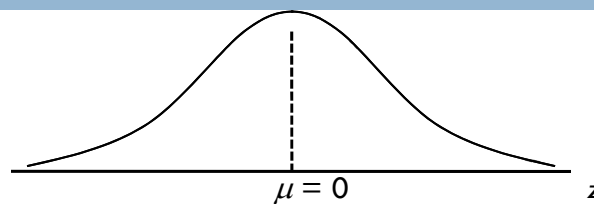
Bentuk sama namun terjadi pergeseran letak kurva norma

Jika μ sama namun σ berbeda



σ yang lebih kecil \rightarrow bentuk kurva yang lebih tinggi dan penyebaran data X lebih rendah

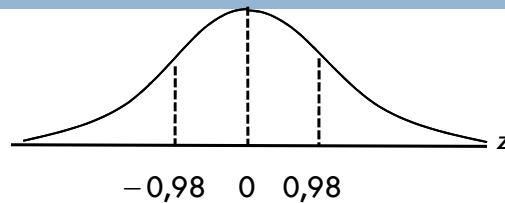
Kurva normal baku



- Distribusi data skor baku
- skor baku (z) merupakan $(X-\mu)/\sigma$
- Nilai tengah $z \rightarrow \mu = 0$;
- Simpangan baku $z \rightarrow \sigma = 1$

$$Y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

Luas daerah bawah kurva normal



- Tabel Luas Kurva normal antara $z=0$ sampai $z=z_1$

Z	0	1	2		8	9
0,0	0,0000	0,0040	0,0080			
0,1						
0,9					0,3365	

Penerapan distribusi normal baku

- Distribusi normal baku digunakan untuk menghitung banyaknya nilai X antara $X=a$ sampai $X=b$ dalam suatu distribusi data yang mempunyai nilai tengah sebesar μ dan simpangan baku sebesar σ
- suatu data tentang hasil panen padi yang dihasilkan **300** orang petani menyebar mengikuti distribusi normal dengan nilai tengah $\mu = 50$ (ku/ha) dan simpangan baku $\sigma = 20$ (ku/ha)

Contoh: $n=300$; $\mu=50$; $\sigma=20$

a. Peluang petani menghasilkan panen padi antara 50 sampai 60 (ku/ha)

□ Rubahlah 50 dan 60 dalam bentuk skor baku, z

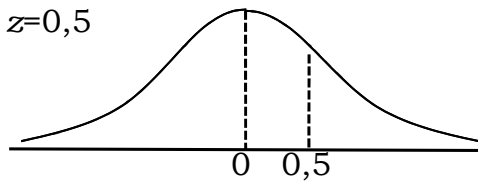
$$z = \frac{X - \mu}{\sigma}$$

$$z_{50} = \frac{50 - 50}{20} = 0$$

$$z_{60} = \frac{60 - 50}{20} = 0,5$$

Contoh: $n=300$; $\mu=50$; $\sigma=20$

□ Luas daerah di bawah kurva normal antara $z=0$ dan $z=0,5$



Z	0	1	2	8	9
0,0					
0,1					
0,2					
0,3					
0,4					
0,5	0,1915				

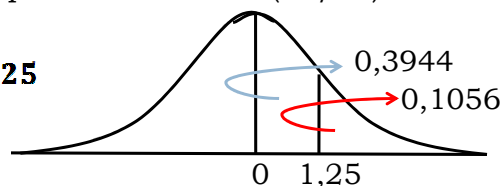
Contoh: $n=300$; $\mu=50$; $\sigma=30$

- Jadi peluang petani menghasilkan panen padi antara 50 sampai 60 (ku/ha) adalah sebesar 0,1915
- b. Banyaknya petani yang mempunyai hasil panen padi antara 50 sampai 60 (ku/ha)
 - = peluang x banyaknya petani padi
 - = $0,1915 \times 300 = 57,45$.
- Jadi banyaknya petani yang mempunyai hasil panen padi antara 50 sampai 60 (ku/ha) adalah 57 orang.

Contoh: $n=300$; $\mu=50$; $\sigma=20$

- c. Peluang hasil panen padi lebih dari 75 (ku/ha)

$$z_{75} = \frac{75 - 50}{20} = 1,25$$



- Luas daerah bawah kurva normal antara $z=0$ sampai $z=1,25$ adalah 0,3944
- lebih dari dari 75 (ku/ha) → Luas daerah bawah kurva normal antara $z=1,25$ sampai $z=\infty$ yaitu $0,5 - 0,3944 = 0,1056$
- Jadi peluang hasil panen padi lebih dari dari 75 (ku/ha) adalah 0,1056

Contoh: $n=300$; $\mu=50$; $\sigma=20$

- d. Banyaknya petani yang mempunyai hasil panen lebih dari 75 (ku/ha) adalah
 = peluang x banyaknya petani padi
 = $0,1056 \times 300 = 31,68$
- Jadi banyaknya petani yang mempunyai hasil panen lebih dari 75 (ku/ha) adalah $0,1056 \times 300 = 31,68 \approx 31$ orang

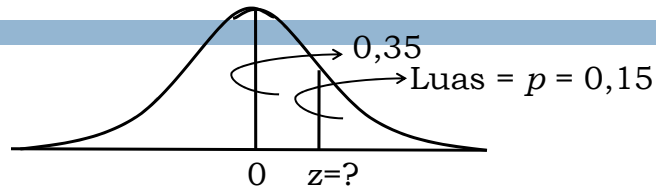
Contoh: $n=300$; $\mu=50$; $\sigma=20$

- e. Jika 45 orang petani yang panennya tertinggi akan diberikan penghargaan studi banding, tentukan batas terendah panen padi petani yang akan studi banding
- Proporsi petani yang akan studi banding

$$p = \frac{n_x}{n}$$

$$p = \frac{45}{300} = 0,15$$

Contoh: $n=300$; $\mu=50$; $\sigma=30$



- Luas daerah antara $z=0$ sampai $z=?$ adalah $0,5 - 0,15 = 0,35$
- Tabel luas daerah kurva normal baku $\rightarrow z=1,04$

Z	0	1	...	4		
0,0						
0,1						
1,0				0,3508		

Contoh: $n=300$; $\mu=50$; $\sigma=20$

- Berdasarkan tabel diperoleh $z = 1,04$
- Dimasukkan dalam rumus skor baku untuk memperoleh nilai X, yaitu batas panen terrendah

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

$$1,04 = \frac{X - 50}{20} \rightarrow 20,8 = X - 50 \rightarrow X = 70,8$$

- Jadi, batas terrendah dari 45 orang yang panennya terbaik adalah 70,8(kw/ha)

Contoh: $n=300$; $\mu = 50$; $\sigma = 20$

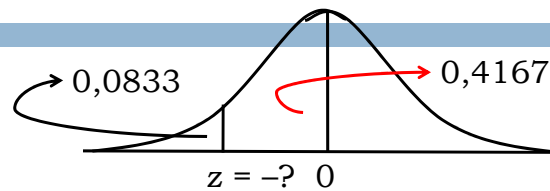
f. Jika 25 orang yang panennya terburuk akan mendapat pembinaan, tentukan batas tertinggi panen petani yang akan mendapat pembinaan!

- Proporsi petani yang akan mendapat pembinaan

$$p = \frac{n_x}{n}$$

$$p = \frac{25}{300} = 0,0833$$

Contoh: $n=300$; $\mu = 50$; $\sigma = 20$



- Panen terburuk \rightarrow lebih rendah dari nilai tengah \rightarrow nilai z adalah negatif (nilai z berada di kiri 0)
- Luas daerah kurva normal antara $z=0$ sampai $z=-?$ adalah $0,5 - 0,0833 = 0,4167$

Z	0	1	...	8	
0,0					
					↑
1,3				0,4162	

Contoh: $n=300$; $\mu=50$; $\sigma=20$

- Dari tabel diperoleh $z = -1,38$
- Dimasukkan dalam rumus skor baku untuk memperoleh nilai X , yaitu batas panen tertinggi

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

$$-1,38 = \frac{X - 50}{20} \rightarrow -27,6 = X - 50$$

$$X = 50 - 27,6 = 22,4$$

- Jadi, batas tertinggi 25 orang yang panennya terburuk adalah 22,4 (kw/ha)

Contoh: $n=300$; $\mu=50$; $\sigma=20$

- Berapa orang petani yang panennya antara 40 sampai 60 (kw/ha)?
- Berapa orang petani yang panennya antara 60 sampai 75 (kw/ha)?
- 80% petani yang panennya terbaik diundang ke Istana. Berapa batas terendah petani yang diundang ke Istana?