ANALISA STATISTIK DISKRIPTIF

DISTRIBUSI FREKUENSI

A. Distribusi Frekuensi Katagorik

Misal : Dalam penelitian persepsi masyarakat tentang akan dibangunnya suatu kawasan industri di daerah permukiman, dng respon atau jawaban yang di tabelkan sbb :

Pendapat	Jumlah	Persen
Sangat Setuju	25	32,47
Setuju	10	12,99
Netral	-	-
Tidak Setuju	12	15,58
Sangat Tidak Setuju	30	38,96
Jumlah	77	100,00

B. *Distribusi Frekuensi Numerik*

- Disebut juga dengan distribusi frekuensi kuantitatif

. Data-2 disusun dalam tabel distribusi frekuensi yang bersumber dari raw data yang sudah dinilai kebenarannya

Contoh : Tabel Distribusi Frekuensi Nilai

Nilai	Titik Tengah	Frekuensi
10 - 20	15	7
20 - 30	25	9
30 - 40	35	13
40 - 50	45	17
50 - 60	55	9
60 - 70	65	9
70 - 80	75	6
Jumlah (n)		70

UKURAN TENDENSI PUSAT

- MODUS :

  katagori yang mempunyai frekuensi terbanyak; suatu ukuran atau nilai atau score yang paling sering terjadi dalam satu kelompok data.
- MEDIAN :

  nilai atau katagori yang berada pada posisi tengah dalam urutan data mulai dari yang terendah (atau tertinggi) kepada nilai yang tertinggi (atau terendah).
- MEAN :

  adalah ukuran tendensi pusat yang sangat umum dipakai untuk sampel dari variabel interval/ratio.

- Mean untuk data tidak dikelompokan :

$$\bar{x} = \frac{1}{n} \sum x_i$$

- Mean untuk data yang berkelompok :

$$\bar{x} = \frac{1}{n} \sum x_i f_i$$

- Kuartil, Kuantil, Desil dan Persentil

Di samping median yang menunjukkan data pada posisi tengah dari sederetan data yang terurut dari kecil ke yang terbesar dalam satu kelompok, sering juga dipakai ukuran lain untuk menunjukkan posisi tertentu :

Kuartil = posisi perempatan

Kuantil = posisi perlimaan

Desil = posisi perpuluhan

Persentil = posisi peratusan

Standart deviasi :

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

Untuk data yang berkelompok :

$$s = \sqrt{\frac{\sum (xi - x)^2 f_i}{n}}$$

s^2 = variance

Tabel penggunaan statistik ukuran tendensi pusat pd berbagai jenis variabel :

Jenis variabel / Ukuran	Nominal	Ordinal	Interval/Ratio
Modus	x	x	x
Median	-	x	x
Mean	-	-	x

x = dapat dipakai

- = tidak dapat dipakai

DISTRIBUSI PROBABILITAS

Misalkan dengan 3 kali pelemparan mata uang logam yang dilakukan secara berturut-turut, maka 8 outcome yang dapat terjadi ialah GGG, GGB, GBG, GBB, BGB, BBG, BGG dan BBB. Kemungkinan-kemungkinan tentang berapa kali B terbuka dalam ketiga pelemparan yang dilakukan secara independen itu merupakan suatu variabel x disebut sebagai variabel acak. Variabel x ini mengambil nilai : 0, 1, 2, dan 3 yaitu berapa kalinya B dapat terbuka dalam 3 kali lemparan. Nilai-nilai variabel ini berupa bilangan cacah dan disebut variabel diskrit. Peluang untuk masing-masing variabel adalah sebagai berikut :

$P(x=0) = P(G\&G\&G)$

$= P(G).P(G).P(G) = (P(G))^3 = (1/2)^3 = 0{,}125$

$P(x=1) = P(G\&G\&B)$ atau $P(G\&B\&G)$ atau $P(B\&G\&G)$

$= P(G).P(G).P(B) + P(G).P(B).P(G) + P(B).P(G).P(G)$

$= (P(G))^2 P(B) + (P(G))^2 P(B) + (P(G))^2 P(B)$

$= 3 P(B) (P(G))^2 = (3) (1/2) (1/2)^2 = 0{,}375$

$P(x=2) = P(B\&B\&G)$ atau $P(B\&G\&B)$ atau $P(G\&B\&B)$

$= P(B).P(B).P(G) + P(B).P(G).P(B) + P(G).P(B).P(B)$

$= (P(B))^2 P(G) + (P(B))^2 P(G) + (P(B))^2 P(G)$

$= 3 (P(B))^2 P(G) = (3) (1/2) (1/2)^2 = 0{,}375$

$P(x=3) = P(B\&B\&B)$

$= P(B).P(B).P(B) = (P(B))^3 = (1/2)^3 = 0{,}125$

Perhitungan ini mengikuti pola :

$P(x=a) = [P(B)]^a [P(G)]^{3-a}$ yang dalam hal ini disebut distribusi binomial.

Apabila pelemparan dilakukan N kali dan peluang untuk B = P (B) dan peluang untuk G atau B = 1 – P(B), maka fungsi peluang menjadi :

$P(X) = [P(B)]^{x}\,[P(G)]^{N-x}$ yang disebut fungsi peluang binomial. Distribusi peluang tersebut dapat disusun dalam tabel di bawah ini :

X_i	$P(X_i)$	$F(X)$
0	0,125	0,125
1	0,375	0,500
2	0,375	0,875
3	0,125	1,000
Jumlah	1,000	

F(x) pada tabel adalah peluang kumulatif yang disebut fungsi distribusi. Untuk distribusi peluang yang kontinu, salah satu distribusinya adalah distribusi normal.

DISTRIBUSI NORMAL

Distribusi normal termasuk salah satu jenis distribusi hipotetik yang kontinu. Variabel acaknya terdiri dari bilangan kontinu yang dapat mengambil nilai dari - ~ sampai + ~. Kalau variabel acaknya adalah x, maka fungsi peluang untuk x adalah :

f(x) = 1 e-½(x-m/s

s Ö2P

Π: bilangan tetap 3,1416

e : bilangan tetap 2,7183

σ: standar deviasi

m : mean

Kurve distribusi normal dng fungsi di atas memp. Bentuk lonceng.

Distribusi normal standart ialah suatu distribusi normal yang mempunyai $\mu = 0$ dan $\sigma = 1$. Untuk membedakannya dengan distribusi normal biasa, maka variabel acaknya tidak dinyatakan dalam x, tetapi dalam z. Dengan demikian fungsinya menjadi :

$$f(z) = \frac{1}{\sqrt{2}} e^{-1/2}$$

Setiap distribusi normal tertentu, yang diperoleh dari pengamatan, dapat dinyatakan dalam distribusi normal standart. Ini berarti x dari distribusi normal tertentu dinyatakan dalam distribusi z dari distribusi normal standart. Hubungan di antara kedua variabel tersebut ialah :

$$x = \mu + z\sigma \quad \text{atau } z = (x-\mu)/\sigma$$

Fungsi peluang dan fungsi distribusi dari distribusi normal standart yang digambarkan sebagai luas kurve dapat dilihat pada tabel luas kurve normal.

Fungsi peluang dan fungsi distribusi dari distribusi normal standar, yang digambarkan sebagai luas kurve normal dapat dilihat pada tabel kurve normal. Tabel berikut merupakan Fungsi distribusi normal untuk beberapa harga z sbb :

z	F(z)	Luas (%)
0,00	0,5000	50,00
0,10	0,5398	53,98
1,00	0,8413	84,13
-1,00	0,1587	15,87
1,50	0,9332	93,32
-1,50	0,0668	6,68
2,00	0,9771	97,71
-2,00	0,0229	2,29
2,50	0,9938	99,38
-2,50	0,0062	0,62
3,00	0,9987	99,87
-3,00	0,0013	0,13

Gambar berikut menunjukkan fungsi peluang dan fungsi distribusi distribusi normal untuk beberapa harga z.

Apabila suatu distribusi normal pengamatan mempunyai $\bar{x}$ = 30 dan S=5 maka frekuensi kumulatif untuk $\bar{x}$ mulai dari 20 sampai 35 adalah sbb :

Untuk x = 20, $z = \frac{20-30}{5} = -2{,}0$

Untuk x = 35, $z = \frac{35-30}{5} = 1{,}0$

Luas kurve normal di sebelah kiri z = -2,0 : 0,0229

Luas kurve normal di sebelah kiri z = 1,0 : <u>0,8413</u>

Luas kurve normal antara z = -2,0 dan z = 1,0 : 0,8184

Maka frekuensi kumulatif relatip untuk 20 < $\bar{x}$ < 35 adalah 0,8184 atau 81,84% dari seluruh populasi.

STATISTIK NON PARAMETRIK

- Tidak mengharuskan distribusi normal
- Dapat dipakai untuk level data seperti nominal dan ordinal
- Cederung lebih sederhana dan mudah dimengerti

Salah satu uji yang sering dipakai dalam praktek adalah Uji Chi-Square. Uji ini dipakai untuk menguji apakah data sebuah sampel yang diambil menunjang hipotesis yang menyatakan bahwa populasi asal sampel tsb. mengikuti distribusi yang telah ditetapkan.

Uji ini dapat juga disebut uji keselarasan (goodness of fit test), krn menguji apakah sebuah sampel selaras dengan salah satu distribusi teoritis.

Namun dalam prakteknya, uji ini tetap mengikuti prinsip dasar pengujian Chi-Square yaitu membandingkan antara frekuensi-2 harapan dengan frekuensi-2 teramati.

Contoh :

Partisipasi dlm BIMAS	Pendidikan: TT SD			Pendidikan: Tamat SD			Semua Tk. Pendidikan
	f_o	f_t	$\frac{(f_o-f_t)^2}{f_t}$	f_o	f_t	$\frac{(f_o-f_t)^2}{f_t}$	N
Ikut BIMAS	47	39,48	1,43	35	42,52	1,32	82
Tak Ikut Bimas	5	12,52	4,52	21	13,48	4,19	26
	52			56			108

$F_t (1) = \frac{52 \times 82}{108} = 39{,}48$ $F_t (1) = \frac{56 \times 82}{108} = 42{,}52$

$F_t (3) = \frac{52 \times 26}{108} = 12{,}52$ $F_t (1) = \frac{56 \times 26}{108} = 13{,}48$

$$\chi^2 = \sum \left[\frac{(f_o - f_t)^2}{f_t}\right] = 11{,}46$$

Apakah signifikan ? dk = (k-1)(b-1) = 1

Tingkat signifikansi 0,05 (3,841) & 0,01 (5,412)

Kesimpulan ada perbedaan yang signifikan pada tingkat partisipasi petani dalam BIMAS antara petani yang TT SD dan Tamat SD.

Untuk melihat hubungan keeratannya dihitung :

Koefisien kontingensi :

$$K = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

$$= 0{,}31$$

Makin besar K makin erat (K berkisar 0-1).