

STATISTIKA DESKRIPTIF DAN STATISTIKA INFERENSIAL

I. STATISTIKA DESKRIPTIF

Setelah perkuliahan ini diharapkan dapat:

- Membedakan data kuantitatif dan kategorial dan memberikan contohnya
 - Menjelaskan perbedaan antara statistika dan parameter
 - Menjelaskan pengertian “distribusi normal” dan “kurva normal”
 - Menjelaskan pengertian Mean (rerata), Median, dan Modus untuk distribusi frekuensi data
 - Menjelaskan pengertian tentang standar deviasi suatu distribusi.
 - Menjelaskan tentang Z score
 - Menginterpretasikan distribusi normal
 - Menjelaskan tentang koefisien korelasi
 - Menghitung koefisien korelasi Pearson
-

Setelah peneliti membuat instrumen untuk menjangkau data, kemudian data di nilai dan ditabulasikan, maka langkah pertama dalam analisis data adalah mendeskripsikan dengan menggunakan statistika deskriptif. Pada jenis penelitian yang menggunakan kuesioner dan sejenisnya, proses analisis data menggunakan perhitungan statistika yang kemudian harus diinterpretasikan maknanya. Pada pertemuan 6 dan 7 kita telah mempelajari tentang 2 tipe dasar data numerik .

Dalam suatu penelitian, informasi dapat dikumpulkan melalui berbagai cara, tetapi hanya dapat dilaporkan melalui dua cara, yaitu dengan kata-kata (narasi) atau dengan menyajikan hasil perhitungan. Dalam penelitian etnografi, atau studi kasus, peneliti umumnya menjelaskan temuannya secara naratif. Mereka tidak mereduksi informasi menjadi bentuk numerik tetapi menyajikannya secara deskriptif (kualitatif).

Dalam penelitian pendidikan informasi yang dilaporkan umumnya berupa angka seperti: skor tes, persentase, IP, *rating* (skala sikap), frekuensi dan sejenisnya. Alasannya sederhana, yaitu

untuk menyederhanakan informasi. Informasi numerik biasanya biasanya dinyatakan sebagai data, yang diklasifikasikan menjadi data kuantitatif dan data kategorial.

1. PERBEDAAN ANTARA DATA KUANTITATIF dan DATA KATEGORIAL

Data kuantitatif	Data kategorial
1) Diperoleh bila variabel yang diteliti, diukur dengan skala yang menyatakan “berapa” jumlah yang dicapai oleh variabel tersebut	Variabel ditentukan berdasarkan kategori.
2) Dilaporkan dalam bentuk angka/skor. Skor tinggi menunjukkan variabel yang diukur (mis. Tinggi badan, berat badan, kemampuan akademik, rasa percaya diri, minat dalam matematika dll) adalah lebih tinggi dibandingkan dengan skor yang rendah.	Dilaporkan dalam bentuk persentase atau frekuensi dari masing-masing kategori. Data kategorial menunjukkan jumlah objek, individu, atau peristiwa yang dipilah berdasarkan kategori, kemudian dinyatakan dengan persentase.
3) Contoh: <ul style="list-style-type: none"> Jumlah dana yang dikeluarkan oleh berbagai sekolah untuk membeli alat-alat laboratorium (variabel: jumlah dana yang dikeluarkan) Skor Tes Potensi Akademik (variabel: kemampuan akademik) Hasil pengukuran suhu badan penderita malaria selama 24 jam. (variabel: suhu) Tingkat kecemasan semua mahasiswa baru di UPI tahun 2005 (variabel: kecemasan) 	Contoh: <ul style="list-style-type: none"> Variabel kelompok etnis di suatu sekolah: pribumi 55%, etnis Tionghoa 20%, etnis Arab 15%, etnis Eropa 10%. (variabel: etnis) Jumlah mahasiswa putri yang belajar ilmu kimia di Jurusan Pendidikan kimia. (variabel: gender) Jumlah guru di suatu sekolah yang menggunakan (1) metode ceramah dan (2) metode diskusi (variabel: metode mengajar) Jumlah masing-masing tipe alat yang ada di lab. IPA (variabel: tipe alat)

2. STATISTIKA DAN PARAMETER

Manfaat statistika deskriptif dalam penelitian adalah peneliti dapat mendeskripsikan informasi yang diperoleh melalui berbagai cara penskoran seperti: Mean dan median dalam suatu waktu tertentu. Bila perhitungan diperoleh dari sampel yang diambil dari populasi, maka hasil tersebut disebut statistika. Sementara bila perhitungan ditarik dari suatu seluruh populasi, maka hasilnya disebut parameter.

Pengukuran	Statistika Sampel	Parameter Populasi
Rerata (mean)	\bar{X}	μ
Simpangan baku	S (SD)	
Variansi	S^2	σ^2

3. TEKNIK MENYIMPULKAN DATA KUANTITATIF

Data kuantitatif dapat disimpulkan melalui berbagai cara yang meliputi:

- Pengelompokan data dengan membuat tabel distribusi frekuensi
- Menyajikan data dengan menggambarkan histogram dan poligon
- Menghitung kemiringan poligon
- Menghitung mean (\bar{X}), simpangan baku (S), variansi (S^2), modus, median
- Menguji normalitas distribusi data : penting untuk menentukan langkah pengujian selanjutnya, apakah menggunakan statistika parametrik atau nonparametrik.

3.1. Skor Baku dan Kurva Normal

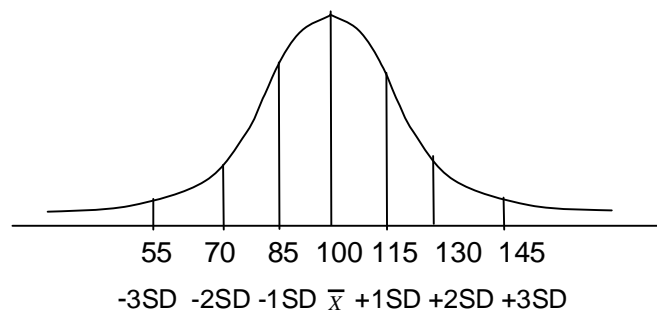
Peneliti seringkali tertarik untuk melihat bagaimana skor seseorang dibandingkan dengan yang lain. Untuk ini digunakan skor baku (*standard score*). Skor baku juga menggunakan skala untuk mengetahui bagaimana kedudukan individu dalam kelompoknya. Skor ini sangat membantu dalam membandingkan

posisi relatif individual. Kedua skor baku yang seringkali digunakan dalam penelitian pendidikan adalah *Zscore* dan *T score*.

a. Z score

Merupakan bentuk skor baku yang paling sederhana yang mengekspresikan jarak skor mentah dari mean dalam unit simpangan baku. Jarak skor mentah dari mean berdasarkan Zscore adalah: +1SD, +2SD, +3SD di atas nilai mean (ke kanan) atau : - 1SD, -2SD, - 3 SD di bawah nilai mean (ke kiri)

SKOR IQ



RUMUS Zscore:
$$z = \frac{X - \text{mean}}{SD}$$

Contoh: bila skor mentah =80; mean=65; SD=12, maka Zscore adalah:

$$Z = \frac{80 - 65}{12} = 1,25$$

Bila seorang siswa memperoleh nilai biologi 60 dan kimia 80; seorang pengamat yang naif akan segera mengatakan bahwa anak tersebut lebih cakap dalam ilmu kimia dibandingkan biologi. Benarkah demikian?

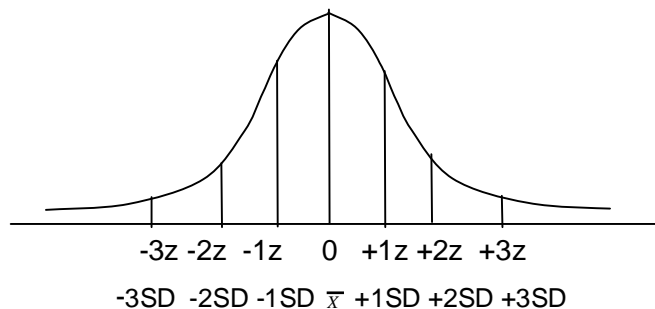
Mean biologi = 50; SD=5; Mean kimia=90;SD=10

maka Zscore siswa tersebut :

$$Z_{\text{BIOLOGI}} = \frac{60 - 50}{5} = +2 \quad Z_{\text{KIMIA}} = \frac{80 - 90}{10} = -1$$

Jadi bagaimanakah kedudukan siswa tersebut dalam kedua mata pelajaran di atas? Ternyata capaian nilai biologi siswa tersebut lebih baik daripada ilmu kimia

Z score seringkali diasosiasikan dengan kurva normal:



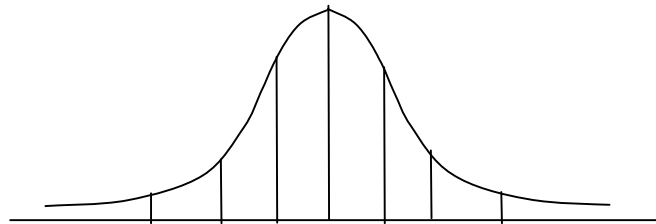
b. T score

Skor mentah di bawah mean adalah negatif; untuk menghilangkan tanda negatif, maka Z score dikonversi ke T score. Untuk mengubah Z score menjadi T score, maka Z score dikalikan 10 lalu ditambah 50. Dengan demikian:

Z score = +1 setara dengan T score: $(1 \times 10) + 50 = 60$

Z score = -2 setara dengan T score: $(-2 \times 10) + 50 = 30$

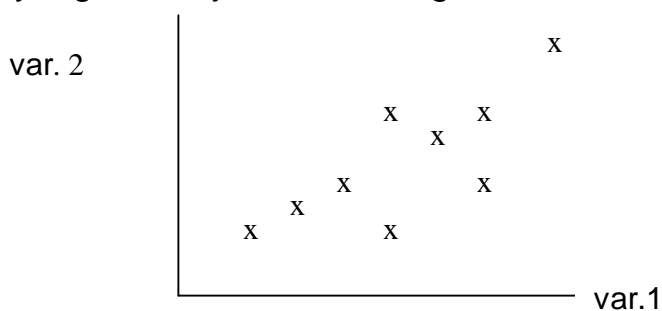
Ada pula sistem penilaian yang mirip dengan T score yang dibedakan pada pilihan nilai mean dan SD. Dua diantaranya adalah yang digunakan dalam *Graduate Record Examination/ GRE*: $\bar{x} = 500$; SD 100, dan *Weschler Intelligence Scale*: $\bar{x} = 100$, SD=15.



Simpangan Baku	-3z	-2z	-z	0	+1z	+2z	+3z
Zscore	-3	-2	-1	0	+1	+2	+3
Tscore	20	30	40	50	60	70	80
GREscore	200	300	400	500	600	700	800
IQscore	55	70	85	100	115	130	145

3.2. Korelasi

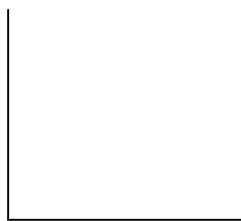
Dalam mempelajari adanya hubungan, maka peneliti dapat membandingkan dua variabel, baik variabel kategorial (misalnya; metode A dan metode B) maupun variabel kuantitatif (misalnya; usia dan berat badan; usia dengan kemampuan membaca atau berhitung dll). Kadang-kadang hubungan ini sangat berguna untuk memprediksi. Langkah awal untuk mengetahui korelasi adalah membuat *scatterplot diagram*, yaitu gambaran berupa titik-titik yang menunjukkan hubungan antara dua variabel kuantitatif.



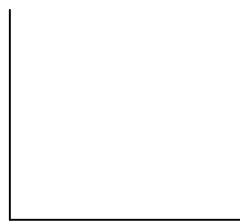
Bagaimana menginterpretasikan *Scatterplot*?

Titik-titik yang digambarkan dalam scatterplot (diagram pencar) menunjukkan arah hubungan antara variabel 1 dan variabel 2 apakah menunjukkan hubungan yang positif atau negatif. Koefisien korelasi (r) berkisar antara -1,00 hingga +1,00. yang menunjukkan hubungan positif sempurna (+1,00) atau negatif sempurna (-1,00).

Beberapa contoh diagram pencar/*scatterplot*



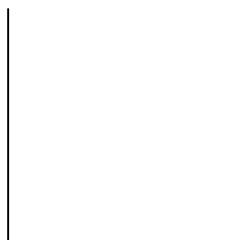
$r = 0.90$



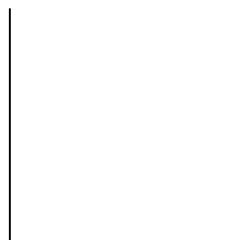
$r = -0.90$



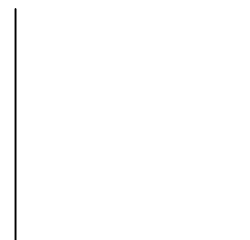
$r = 0.65$



$r = -0.50$



$r = -0.00$



$r = -0.10$

Koeffisien Product-Moment PEARSON

Bila data dari kedua variabel adalah data kuantitatif, maka peneliti dapat menggunakan rumus Product Momen Pearson

$$r = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2] [(n \sum Y^2 - (\sum Y)^2)]}}$$

contoh penggunaannya:

Siswa	Var X	Var Y
A	20	20
B	18	16
C	28	20
D	15	12
E	10	10

Siswa	X	Y	X ²	Y ²	XY
A	20	20	400	400	400
B	18	16	324	256	288
C	28	20	324	400	360
D	15	12	225	144	180
E	10	10	100	100	100
	81	78	1373	1300	1328

maka :

$$r = \frac{5(1328) - (81)(78)}{\sqrt{[5(1373) - 81^2][5(1300) - 78^2]}} = 0,90$$

4. TEKNIK MENYIMPULKAN DATA KATEGORIAL

4.1. Tabel Distribusi Frekuensi

Seandainya peneliti menggunakan kuesioner untuk mengumpulkan data secara random dari 50 orang guru di suatu kota. Kuesioner ini meliputi beberapa variabel yang berkaitan dengan minat dan kegiatan mengajar. Salah satu variabelnya adalah “ Kegiatan pembelajaran yang paling sering digunakan guru di kelas”, maka bentuk tabel frekuensinya harus menunjukkan adanya beberapa kegiatan pembelajaran, yang menunjukkan kategori dari kegiatan pembelajaran di kelas. Peneliti kemudian menghitung dengan rigi (tally)

Contoh:

Respon guru	Tally	Frekuensi	Persentase (%)
Ceramah		15	30
Diskusi		10	20
Laporan lisan		4	8
Demonstrasi		6	12
Praktikum		8	16
Studi pustaka		7	14
		n=50	100

Data di atas dapat disajikan dalam bentuk

- (1) Grafik batang
- (2) Diagram pie

4.2. Tabel Kontingensi

Tabel kontingensi merupakan tabel yang menggambarkan hubungan antara 2 variabel kategorial. Tabel kontingensi yang paling sederhana adalah tabel kontingensi 2x2

Contoh:

Tabel 1
Data hipotetik jenjang guru dan gender guru

Jenjang	Laki-laki	Perempuan	Jumlah
Guru SMP	40	60	100
Guru SMA	60	40	100
	100	100	100

Data di atas merupakan data hipotetik hasil pengamatan (observasi), namun dengan asumsi yang sesuai kita berharap bahwa guru laki-laki dan wanita sama jumlahnya, Maka peneliti dapat menambahkan data yang diharapkan (*expected*) selain data pengamatan (*observed*)

Tabel 2
Data hipotetik jenjang guru dan gender guru
dengan frekuensi harapan (dalam kurung)

Jenjang	Laki-laki	Perempuan	Jumlah
Guru SMP	40 (50)	60 (50)	100
Guru SMA	60 (50)	40 (50)	100
	100	100	100

Jumlah yang diharapkan (*expected*) adalah 50 laki-laki dan 50 perempuan, karena diasumsikan jumlah laki-laki dan perempuan seimbang. Contoh lainnya sebagai pembandingan, di mana datanya tak dapat diasumsikan seperti halnya 'gender'

Tabel 3
Jabatan dan Etnis pimpinan sekolah

	Administrator	Guru	Total
Pribumi	70 (62)	300(308)	370
Non pribumi	30 (38)	200(192)	230
Total	100	500	600

Angka yang ada di dalam kurung adalah angka 'expected'. Tabel ini menunjukkan hubungan antara jabatan dan etnis. Maka untuk menghitung nilai 'expected' adalah:

Jumlah administrator dibandingkan total jumlah adalah $100/600$

Jumlah administrator pribumi = $1/6 \times 370 = 62$

Jumlah administrator non pribumi = $1/6 \times 230 = 38$

Jumlah guru dibandingkan total jumlah adalah $500/600$

Jumlah guru pribumi = $5/6 \times 370 = 308$

Jumlah guru non pribumi = $5/6 \times 230 = 192$

Data kategorial ini dapat digabungkan dengan kategori 'gender' seperti contoh berikut ini:

Tabel 4
Jabatan dan Gender pimpinan sekolah

	Administrator	Guru	Total
Laki-laki	70 (50)	230 (250)	300
Perempuan	30 (50)	270 (250)	300
Total	100	500	600

Berdasarkan contoh di atas, perbandingan antar gender adalah sama (1:1) sehingga angka 'expected' 50, 50; 250, 250

Tabel Jabatan dan Etnis serta Tabel Jabatan dan Gender di atas dapat pula diubah menjadi Tabel Gender dan Etnis berikut:

Tabel 5
Gender dan Etnis pimpinan sekolah

	Pribumi	Nonpribumi	Total
Laki-laki	200 (185)	100 (115)	300
Perempuan	170 (185)	130 (115)	300
Total	370	230	600

Berdasarkan tabel-tabel di atas kita melihat ada diskrepansi antar variabel gender, etnis, peran/jabatan. Tetapi kita dapat menyimpulkan bahwa ada kecenderungan lebih banyak pribumi daripada nonpribumi (tabel 3); lebih banyak laki-laki daripada perempuan (Tabel 5). Namun terdapat hubungan kuat bahwa lebih banyak laki-laki yang menjadi administrator dan lebih banyak perempuan yang menjadi guru (Tabel 4). Mari kita bandingkan diskrepansi antara frekuensi *expected* dan *observed*.

Tabel 6
Diskrepansi total antara frekuensi *observed* dan *expected*

Tabel 3		Tabel 4		Tabel 5	
(70 dan 62)	= 8	(70 dan 50)	=20	(200 dan 185)	= 15
(30 dan 38)	= 8	(30 dan 50)	=20	(170 dan 185)	= 15
(300 dan 308)	= 8	(230 dan 250)	=20	(100 dan 115)	= 15
(200 dan 192)	= 8	(270 dan 250)	=20	(130 dan 115)	= 15
Total	32	Total	80	Total	60

LATIHAN 8

STATISTIKA DESKRIPTIF

1. Pertanyaan penelitian atau hipotesis penelitian untuk tesis saya adalah:
2. Variabelnya adalah (1) _____
(2) _____
3. Variabel (1) adalah: kuantitatif _____ atau kategorial _____
4. Variabel (2) adalah: kuantitatif _____ atau kategorial _____
5. Masing-masing variabel akan disimpulkan dengan menggunakan: ...(beri tanda "V")

	Variabel 1	Variabel 2	Lain-lain
Poligon frekuensi			
Mean			
Median			
Simpangan baku			
Tabel distribusi frekuensi			
Grafik batang			
Diagram 'pie'			

6. Hubungan antara variabel 1 dan 2 digambarkan dalam (beri tanda "V")
 - a. perbandingan rerata
 - b. tabel kontingensi
 - c. koefisien korelasi