

## **BAB II**

### **MATEMATIKA DASAR DAN PENGUKURAN**

**Prof. Dr. Patta Bundu, M.Ed.**

*“One does not need to be a statistical wizard to grasp the basic mathematical concepts needed to understand major measurement issues”*  
(Reynolds, Livingstone, Willson, 2006: h.33)

#### **A. SKALA PENGUKURAN**

##### **1. Matematika dan Pengukuran**

Setiap semester dalam mata kuliah Tes dan Pengukuran sering kita mendengar mahasiswa takut dengan istilah statistik yang rumit. Padahal seperti pada kutipan pembuka tadi “seeorang tidak harus menjadi pakar statistik untuk menangkap makna konsep dasar matematika yang diperlukan untuk memahami isu utama pengukuran”. Kubiszyn dan Borich (2003) memperkirakan bahwa kurang dari 1% siswa dalam program pengujian dan penilaian mereka tampil buruk sepenuhnya karena keterampilan matematika tidak cukup. Namun demikian, satu hal yang tidak bisa dipungkiri bahwa semua pengukuran dalam pendidikan dan psikologi memiliki sifat matematika. Termasuk penggunaan tes dan penilaian lain, baik buatan guru atau pun prosedur standar perlu memiliki pemahaman tentang konsep-konsep dasar matematika dan statistik.

Pengukuran tidak akan lepas dari konsep matematika. Oleh sebab itu, kenalkan secara umum tekanan akan difokuskan pada pengembangan pemahaman konseptual masalah daripada berfokus pada perhitungan matematika. Dalam beberapa kasus akan disajikan rumus dan formula tertentu dalam aplikasinya. Pada prinsipnya para guru sangat diharapkan untuk dapat mengembangkan, mengelola, dan menafsirkan hasil asesmen pembelajaran di kelasnya. Apakah guru harus menjadi ahli pengembangan tes? Tentu saja tidak, karena guru secara rutin mengembangkan, menggunakan, dan menginterpretasikan penilaian, mereka harus kompeten dalam penggunaannya.

Pengukuran adalah seperangkat aturan untuk menetapkan angka yang mewakili benda, sifat, atribut, atau perilaku. Tes pendidikan atau psikologis adalah alat pengukur, dan karena itu melibatkan aturan (misalnya, item tertentu, administrasi, instruksi dan scoring) untuk menetapkan posisi individu melalui penafsiran tertentu. Sebagai contoh, jumlah pertanyaan matematika dimana siswa menjawab dengan benar pada kuis matematika tertentu

***Pengukuran adalah seperangkat aturan untuk menetapkan angka pada benda, sifat, atribut, atau perilaku.***

dapat ditafsirkan sebagai pencemiran pemahaman mereka tentang dua-digit perkalian. Contoh lain, adalah bahwa tanggapan pertanyaan tentang seberapa sering merasa cemas tentang aspek kehidupan, dapat ditafsirkan sebagai ungkapan relatif tingkat kecemasan. Ketika kita mengukur sesuatu, unit pengukuran memiliki perangkat

matematika yang disebut *skala pengukuran*. Skala adalah suatu sistem atau skema untuk memberikan nilai atau skor tentang karakteristik yang sedang diukur (Sattler, 1992). Ada empat skala pengukuran, dan perbedaan skala ini memiliki sifat yang berbeda dan menyampaikan jenis yang informasi unik. Keempat skala pengukuran tersebut adalah nominal, ordinal, interval, dan rasio. Skala membentuk hirarki, dan seperti yang kita kemajuan dari skala nominal ke skala rasio kita dapat melakukan pengukuran yang semakin canggih yang menangkap informasi lebih rinci.

## 2. Skala nominal

Skala nominal adalah skala yang paling sederhana dari empat skala. Skala nominal menyediakan sistem kualitatif untuk mengategorikan orang atau benda ke dalam kategori, kelas, atau set. Dalam kebanyakan situasi, kategori-kategori ini sangat eksklusif. Jenis kelamin adalah contoh dari skala nominal yang menunjukkan individu untuk kategori sangat eksklusif. Contoh

***Skala Nominal mengklasifikasikan orang atau benda ke dalam kategori, kelas, atau kelompok.***

lain, memposisikan orang dalam kategori berdasarkan jurusan akademik perguruan tinggi mereka (misalnya, Pendidikan, Psikologi, Kimia). Dalam situasi ini penetapan angka ke skala nominal hanya untuk mengidentifikasi atau memberi label kategori dan tidak diperuntukan untuk sesuatu angka yang berarti. Pada contoh tadi, dapat digunakan nomor 1 (satu) untuk jurusan Pendidikan, nomor 2 (dua) untuk Psikologi, dan nomor 3 (tiga) untuk Kimia.. Kategori ini tidak bermakna

bahwa tiga tidak lebih besar dari dua, dan dua tidak lebih besar dari satu. Apalagi jika menyimpulkan bahwa Kimia 3 kali bidang Pendidikan. Penetapan angka hanya untuk mudah menyebut apakah merah, biru, hijau, dan sebagainya. Bilangan ditetapkan dalam skala nominal tidak boleh dijumlah, dikurangi, diranking, atau dimanipulasi.

### **3. Skala Ordinal**

Pengukuran skala ordinal digunakan untuk memberi peringkat pada orang atau benda sesuai dengan jumlah atau kuantitas dari karakteristik yang dimiliki. Skala ordinal memungkinkan mengidentifikasi variabel dalam ujian dan memberikan informasi secara lebih substansif dari skala nominal. Misalnya, peringkat anak-anak di kelas sesuai dengan ketinggian dari yang tertinggi ke yang terpendek. Anak tertinggi dikelas akan menerima peringkat 1, tertinggi berikutnya peringkat 2, dan seterusnya. Meskipun pengukuran skala ordinal memberikan

***Skala ordinal menentukan peringkat orang atau benda sesuai dengan jumlah dari karakteristik yang ditampilkan.***

informasi kuantitatif, tetapi tidak menjamin interval antara jajaran secara konsisten.

Artinya, perbedaan ketinggian antara anak-anak peringkat 1 dan 2 mungkin tiga inci sedangkan perbedaan antara peringkat 3 dan 4 mungkin hanya satu inci. Skala ordinal menunjukkan posisi peringkat antara individu-individu atau objek, tetapi itu tidak menunjukkan sejauh mana mereka berbeda. Skala ordinal agak terbatas baik dalam informasi pengukuran yang diberikan dan prosedur statistik yang

dapat diterapkan. Namun demikian, penggunaan skala ini cukup umum dalam lingkungan pendidikan. Peringkat persentil, setara usia, dan kelas setara merupakan contoh skala ordinal.

#### 4. Skala Interval

Skala Interval memberikan informasi lebih baik dari pada skala nominal atau ordinal. Pengukuran skala interval memungkinkan untuk memberi peringkat orang atau benda seperti skala ordinal, tetapi pada skala yang sama. Dengan skala unit yang sama, dapat diartikan perbedaan antara unit yang berdekatan pada skala setara. Perbedaan antara skor dari 70 dan 71 adalah sama dengan perbedaan antara skor 50 dan 51 (atau 92 dan 93, 37 dan 38;dll). Banyak tes pendidikan dan psikologis dirancang untuk menghasilkan skor tingkat interval. Contoh skor untuk tiga orang pada tes bakat. Asumsikan individu A memperoleh skor 100, individu B memperoleh skor 110, dan individu C memperoleh skor 120. Pertama, diketahui bahwa C memperoleh nilai tertinggi diikuti oleh B selanjutnya A. Kedua, mengingat bahwa skor berada pada skala interval, kita juga tahu

***Skala interval merangking orang atau objek seperti pada skala ordinal, tetapi pada skala dengan unit yang sama.***

bahwa perbedaan antara individu A dan B (yaitu 10 poin) setara dengan perbedaan antara B dan C (yaitu 10 poin). Akhirnya, kita tahu perbedaan antara individu A

dan C (yaitu 20 poin) dua kali lebih besar dari perbedaan antara individu A dan B (yaitu 10 poin). Data tingkat Interval dapat dimanipulasi dengan menggunakan operasi matematika umum

seperti, penambahan, pengurangan, perkalian, dan pembagian. Keuntungannya adalah bahwa prosedur statistik dapat digunakan dengan data berskala interval.

Skala interval mewakili peningkatan yang substansial dari skala ordinal dan/memberikan informasi yang cukup. Keterbatasannya adalah bahwa skala interval tidak memiliki titik nol benar. Artinya, pada skala interval nilai nol tidak mencerminkan ketiadaan total atribut. Sebagai contoh, jika seseorang tidak mampu menjawab pertanyaan dengan benar pada tes kecerdasan dan mencetak nol, itu tidak akan menunjukkan “tidak ada” kecerdasan. Demikian juga, meskipun IQ 100 dua kali lebih besar dari IQ 50, itu tidak berarti bahwa orang dengan IQ 100 adalah dua kali lebih cerdas dari orang dengan IQ 50. Dalam pendidikan, skor skala interval yang paling sering terlihat dalam bentuk nilai standar.

## **5. Skala Rasio**

Skala rasio memiliki sifat interval ditambah titik nol sejati yang mencerminkan tidak adanya kelengkapan karakteristik yang sedang diukur. Mil per jam, panjang, dan berat badan merupakan contoh skala rasio. Seperti namanya, dengan skala ini kita dapat menafsirkan rasio antara skor. Sebagai contoh, 60 mil per jam adalah dua kali lebih cepat dari pada 30 mil per jam, 20 meter dua kali lebih panjang dari pada 10 kaki, dan 60 pon adalah tiga kali lebih banyak dari pada 20 kilogram. Rasio tidak berarti atau ditafsirkan dengan skala interval. Seperti yang kita ketahui, anak dengan intelligence quotient (IQ) 100 tidak dua kali secerdas

**Skala Rasio memiliki sifat skala interval ditambah titik nol benar.**

dengan IQ 50 anak dengan skor tes prestasi matematika standar 100 tidak tahu dua kali lebih banyak sebagai salah satu dengan skor 50.

Untungnya, sebagian besar masalah pengukuran dalam pendidikan dapat menggunakan skala interval.

TABEL 2.1. Skala Nominal, Ordinal, Interval, dan Rasio

Skala	Contoh	Contoh Skor
Nominal	Jenis Kelamin peserta	Wanita =1 Laki-laki =2
	Etnis	Makassar = 1 Bugis = 2 Jawa = 3 Minang = 4
Ordinal	Pilihan untuk kegiatan	1 = paling disukai 2. = disukai 3. = kurang disukai 4 = tidak disukai
	Peringkat Kelulusan	1 = pujian 2 = sangat memuaskan 3 = memuaskan
Interval	Skor Intelligensi	IQ 100
	Tes Personaliti	Skor kecemasan 75
Rasio	Tinggi dalam cm	Tinggi 60 cm
	Berat dalam ons	100 kg

Dalam kebanyakan situasi, pengukuran akan menggunakan skala yang memberikan informasi yang paling banyak.

## **B. DESKRIPSI SKOR TES**

Skor tes individu jika berdiri sendiri memberikan sangat sedikit informasi. Jika skor individu pada tes prestasi membaca adalah 79, maka hanya sedikit informasi yang diketahui tentang kemampuan orang tersebut dalam membaca. Untuk penafsiran yang lebih bermakna diperlukan kerangka acuan. Misalnya, jika dalam kelas terdapat 25 anak, skor 79 adalah nilai tertinggi yang dicapai, itu akan mencerminkan kinerja di atas rata-rata (atau mungkin superior). Sebaliknya, jika 79 adalah skor terendah, itu akan mencerminkan kinerja di bawah rata-rata. Bagian berikut memberikan informasi tentang distribusi skor dan statistik yang digunakan untuk menggambarkan dan menginterpretasikan skor tes.

### **1. Distribusi Skor**

Distribusi secara sederhana adalah *satu set skori*. Skor ini dapat diperoleh dari tes membaca, skor pada tes kecerdasan, atau nilai pada pengukuran lainnya. Distribusi skor dapat mencerminkan karakteristik fisik seperti berat, tinggi, atau kekuatan. Distribusi skor dapat direpresentasikan dalam berbagai cara, termasuk tabel dan grafik. Tabel 2.2 menyajikan skor untuk 20 siswa pada tugas pekerjaan rumah yang mirip dengan apa yang dapat terekam dalam buku kelas guru. Tabel 2.3 menyajikan distribusi frekuensi yang tidak bergroup dengan skornya sama. Perhatikan bahwa dalam contoh ini hanya ada kemungkinan tujuh kategori pengukuran atau nilai (yaitu, 4, 5, 6, 7, 8, 9, dan 10).

**TABEL. 2.2. Distribusi Skor untuk 20 Siswa**

Nama	Skor
Patta	9
Rahman	8
Rahim	7
Gafur	10
Rahmah	5
Erna	7
Irma	7
Latiha	8
Ghaniya	9
Khaliqa	8
Adhwa	6
Rosdia	6
Mustari	7
Hasan	7
Husain	7
Yaser	4
Rini	6
Lia	6
Emi	8
Jamal	7

**TABEL 2.3. Ditribusi Frekuensi tdk Berkelompok**

Skor	Frekuensi
10	1
9	2
8	4
7	7
6	4
5	1
4	1

Ada banyak kemungkinan skor dimana hal tersebut tidak praktis untuk membuat daftar skor masing-masing individual. Dalam situasi ini lazim untuk menggunakan distribusi frekuensi

berkelompok. Skor "digabung" atau "dikelompokkan" ke dalam interval kelas tertentu. Tabel 2.4 menyajikan distribusi frekuensi berkelompok dari 100 skor hipotetis dikelompokkan ke dalam interval kelas yang mencakup lima skor nilai.

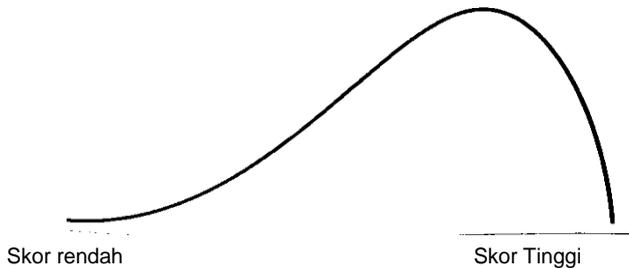
**TABEL. 2.4. Distribusi Frekuensi Berkelompok**

Kelas Interval	Frekuensi
81 – 100	6
61 – 80	38
41 – 60	42
21 – 40	12
0 – 20	2

Grafik frekuensi seperti Gambar 2.1, juga populer dan memberikan representasi visual dari sebuah distribusi. Ketika membaca grafik frekuensi, skor secara tradisional tercantum pada sumbu horisontal dan frekuensi skor tercantum pada sumbu vertikal. Dalam memeriksa skor ini terlihat bahwa hanya ada satu skor 10 (mencerminkan kinerja yang sempurna) dan hanya ada satu skor 4 (mencerminkan respon yang benar hanya empat pertanyaan).

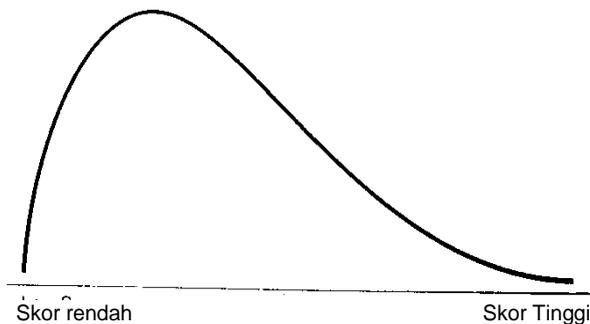


Tidak semua distribusi simetris karena ada yang mempunyai kemiringan yang bisa condong positif atau negatif. Distribusi condong negatif (Gambar 2.3), besar kemungkinan bahwa tes ini terlalu mudah sehingga ada banyak skor yang tinggi dan relatif sedikit skor rendah.



Gbr. 2.3. Distribusi Condong Negatif

Sebaliknya, distribusi condong positif jika skor terlalu sulit sehingga banyak skor nilai rendah dan skor yang tinggi relatif sedikit (Gambar, 2.4)..



Gbr. 2.4. Distribusi Condong Positif

## 2. Pengukuran Tendensi Sentral

Skor pada banyak distribusi cenderung berkonsentrasi pada sekitar pusat atau tengah yang sering dikenal dengan istilah tendensi sentral. Ada tiga deskripsi kecenderungan tendensi sentral yakni *mean*, *median*, dan *modus*.

### a. Mean.

Secara sederhana means berarti skor rata-rata. Dalam kehidupan sehari-hari sangat sering digunakan konsep rata-rata. Sangat banyak contoh seperti suhu rata-rata, gaji buruh rata-rata perjam, rata-rata berat badan, dan skor rata-rata. Secara formal rata-rata satu set skor dapat didefinisikan dengan persamaan sederhana berikut:

$$\text{Mean} = \text{Jumlah Skor} / \text{Banyaknya Skor}$$

Mean dari skor pekerjaan yang tercantum dalam Tabel 2.2 dihitung dengan menjumlahkan nilai 20 dalam distribusi dan membaginya dengan 20. Hasilnya =  $142 : 20 = 7,1$ . Meskipun tidak ada yang memperoleh skor 7,1, tetapi skor angka tersebut memberikan makna kecenderungan sentral dari kelompok skor. Beberapa karakteristik matematika penting dari mean membuatnya berguna sebagai ukuran tendensi sentral. Pertama, rata-rata dapat dihitung dengan data interval dan rasio, tetapi tidak dengan data ordinal dan nominal. Kedua, rata-rata sampel adalah perkiraan yang baik dari rata-rata untuk populasi dari mana sampel tersebut diambil. Hal ini berguna ketika

mengembangkan tes standar di mana sampel standardisasi diuji dan distribusi yang dihasilkan diyakini mencerminkan karakteristik seluruh populasi. Ciri positif lain dari mean adalah sangat penting untuk definisi dan perhitungan statistik deskriptif dalam konteks pengukuran.

Kelemahan penggunaan means adalah sangat sensitif terhadap nilai ekstrim yang tidak seimbang. Artinya, jika ada skor yang relatif sangat tinggi atau sangat rendah terhadap seluruh nilai dalam distribusi. Contoh pada skor ekstrim berikut: 1, 2, 2, 3, 4, 4, 4, dan 36 maka rata-rata adalah 7,0. Rata-rata = 7,0, jauh jaraknya dari setiap nilai yang benar-benar ada dalam distribusi. Biasanya nilai rata-rata ekstrim akan berkurang jika distribusi skor semakin banyak.

#### **b. Median.**

Median adalah nilai atau skor potensi yang membagi distribusi menjadi dua kelompok. Misalnya, satu set skor yang terdiri atas: 9, 8, 7, 6, 5, maka median adalah 7 karena dua nilai tersebut berada di atas dan dua jatuh di bawahnya. Median dapat dihitung untuk distribusi data rasio, interval, atau skor ordinal, tetapi tidak sesuai untuk skor data nominal. Median adalah ukuran yang serba guna pada tendensi sentral.

*Median adalah nilai atau skor potensi yang membagi distribusi menjadi dua.*

### c. Modus.

Modus distribusi adalah skor yang paling sering muncul. Mengacu kembali ke Tabel 2.3, yang menyajikan distribusi frekuensi groupnya dari 20 siswa akan terlihat bahwa skor yang paling banyaak muncul adalah 7. Keuntungan dari modus adalah dapat digunakan dengan data nominal (misalnya, perguruan tinggi yang paling sering banyak dipilih oleh siswa) serta data ordinal, interval, dan rasio (Hays , 1994). Namun, modus memiliki keterbatasan sebagai ukuran tendensi sentral. Pertama, beberapa distribusi memiliki dua skor yang sama dalam frekuensi dan lebih tinggi dari skor lainnya. Hal ini disebut sebagai “bimodal distribusi” dan modus tidak efektif sebagai ukuran tendensi sentral. Kedua, mode bukanlah ukuran yang stabil untuk tendensi sentral, terutama dengan sampel kecil.

*Modus adalah skor yang paling sering muncul dalam distribusi*

Manakah antara Mean, Median, dan Modus sebagai ukuran tendensi sentral yang paling berguna atau tepat? Jawabannya bergantung pada sejumlah faktor. Pertama , means memiliki penggunaan yang cukup sebagai ukuran tendensi sentral. Namun, untuk tujuan murni deskriptif median sering menjadi ukuran yang paling fleksibel dan bermanfaat dalam tendensi sentral. Ketika distribusi miring, pengaruh nilai ekstrim tidak seimbang pada rata-rata cenderung merusak kegunaannya.

### 3. Pengukuran Variabilitas

Dua distribusi dapat memiliki rata-rata, median, dan modus yang sama, namun berbeda dalam cara skor didistribusikan sekitar ukuran pemusatan. Oleh karena itu, tidak cukup untuk menandai satu set nilai hanya dengan ukuran pemusatan. Gambar 2.7. menyajikan grafik tiga distribusi dengan cara yang sama tetapi berbeda derajat variabilitas. Sebuah ukuran dispersi, penyebaran, atau variabilitas dari satu set nilai akan membantu menjelaskan distribusi lebih lengkap. Ada tiga pengukuran variabilitas yang umum digunakan untuk menggambarkan distribusi: *range*, *deviasi standar*, dan *varians*.

#### a. Range

Range (rentang) adalah jarak antara skor terkecil dan terbesar dalam distribusi. Range ini (rentang) dihitung:

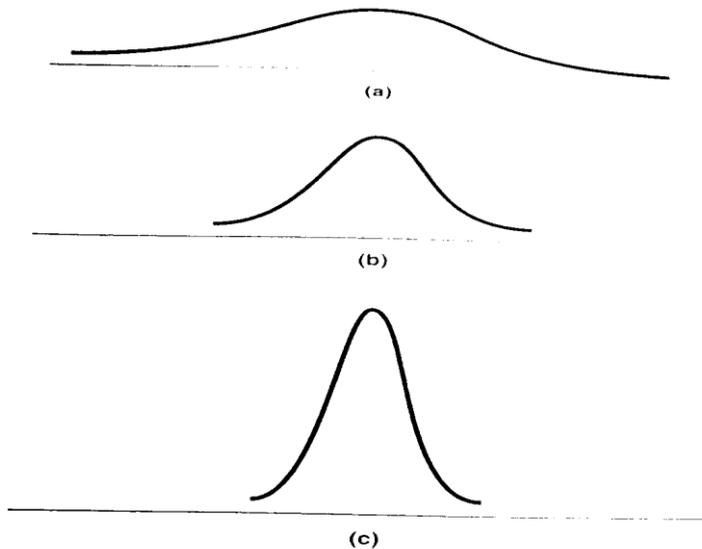
$$\text{Range} = \text{Skor Tertinggi} - \text{Skor Terendah}$$

Misalnya, dalam merujuk kembali ke distribusi skor yang tercantum dalam Tabel 2.3, terlihat bahwa skor terbesar adalah

***Range adalah jarak antara terkecil dan terbesar skor dalam distribusi***

10 dan skor terkecil adalah 4. Dengan mudah dilihat bahwa  $\text{range} = 10 - 4 = 6$ . (Catatan: Beberapa penulis mendefinisikan rentang

sebagai skor tertinggi dikurangi skor terendah, *ditambah satu*, ini dikenal sebagai rentang inklusif).



GAMBAR 2.7 Tiga Distribusi dengan Perbedaan Derajat Variabilitas

*Sumber: Robert J. Gregory, Psychological Testing : Sejarah, Prinsip dan Aplikasi. Boston: Allyn & Bacon, 2000.*

Range ini hanya mempertimbangkan dua nilai paling ekstrim dalam distribusi, namun, tidak memberikan informasi tentang bagaimana nilai sisanya tersebar atau terdispersi dalam batas-batas distribusi. Untuk itu perlu statistik deskriptif lainnya, yang menggambarkan distribusi lebih lengkap. Secara umum yang sering digunakan adalah deviasi standar (standar deviasi) dan varians, untuk memberikan informasi tentang penyebaran atau dispersi skor.

## b. Deviasi Standar.

Mean dan deviasi standar adalah ukuran jarak rata-rata statistik yang paling banyak digunakan dalam pengujian

***Deviasi standar adalah ukuran dari jarak rata-rata yang bervariasi dari nilai rata-rata dari distribusi***

pendidikan dan psikologis dengan nilai bervariasi dari rata-rata pada penelitian ilmu-ilmu sosial dan perilaku. Deviasi

standar dihitung dengan langkah-langkah sebagai berikut:

*Langkah 1.* Hitung mean dari distribusi.

*Langkah 2.* Kurangi setiap skor distribusi dengan mean.

Hal ini akan menghasilkan beberapa angka negatif, dan jika semua selisih dijumlah akan menjadi nol. Solusinya jika masing-masing dipangkat dua setiap angka selalu positif.

*Langkah 3.* Jumlahkan semua skor yang sudah dipangkat dua.

*Langkah 4.* Bagilah jumlah ini dengan jumlah skor untuk menurunkan rata-rata kuadrat penyimpangan dari rata-rata. Nilai ini adalah varians dengan simbol  $\sigma^2$ .

*Langkah 5.* Deviasi standar ( $\sigma$ ) adalah akar kuadrat positif dari varians ( $\sigma^2$ ). Itu adalah akar kuadrat karena sudah mengkuadratkan semua skore sebelum menjumlahkannya.

Langkah-langkah ini diilustrasikan dalam Tabel 2.5 dengan menggunakan nilai yang tercantum pada Tabel 2.2. Contoh ini menggambarkan perhitungan deviasi standar populasi, yang ditunjuk dengan ( $\sigma$ ). Deviasi standar yang biasa disimbolan SD atau s.

Deviasi standar adalah ukuran dari jarak rata-rata nilai yang bervariasi dalam satu distribusi. Semakin besar deviasi standar, semakin banyak skor berbeda berarti semakin besar variabilitas ada dalam distribusi. Jika skor secara luas tersebar atau menyebar ke seluruh mean, deviasi standar akan besar. Sebaliknya, jika dispersi atau penyebaran nilai sekitar mean, deviasi standar akan menjadi kecil.

### **c. Varians.**

Dalam menghitung deviasi standar, lebih dahulu menghitung varians ( $\sigma^2$ ). Seperti digambarkan pada Tabel 2.5, standar deviation sebenarnya akar kuadrat positif dari varians.

***Varians adalah ukuran variabilitas yang memiliki arti khusus sebagai konsep teoritis dalam teori pengukuran dan statistik***

Oleh karena itu, varians juga adalah ukuran variabilitas skor. Alasan deviasi standar lebih sering digunakan ketika menafsirkan nilai individu adalah bahwa varians dalam satuan kuadrat pengukuran, yang mempersulit interpretasi. Sebagai contoh, kita dapat dengan mudah menafsirkan berat dalam pound, tetapi lebih sulit untuk menafsirkan dan menggunakan berat badan yang dilaporkan dalam pound kuadrat. Deviasi standar (akar kuadrat dari varians) adalah dalam satuan yang sama sehingga lebih mudah dipahami. Meskipun varians adalah sulit untuk diterapkan ketika menggambarkan nilai individu, tetapi memiliki arti khusus sebagai konsep teoritis dalam teori pengukuran dan statistik. Varians adalah ukuran tingkat variabilitas dalam skor.

**TABEL 2.5. Menghitung Deviasi Standar dan Varians**

Skor Siswa	Perbedaan (Nilai - Mean)	Perbedaan (Nilai - Mean) <sup>2</sup>
7	$(7 - 7,3) = -0,3$	0,09
8	$(8 - 7,3) = 0,7$	0,49
9	$(9 - 7,3) = 1,7$	2,89
6	$(6 - 7,3) = -1,3$	1,69
7	$(7 - 7,3) = -0,3$	0,09
6	$(6 - 7,3) = 1,3$	1,69
10	$(10 - 7,3) = 2,7$	7,29
8	$(8 - 7,3) = 0,7$	0,49
5	$(5 - 7,3) = -2,3$	5,29
9	$(9 - 7,3) = 1,7$	2,89
9	$(9 - 7,3) = 1,7$	2,89
9	$(9 - 7,3) = 1,7$	2,89
8	$(8 - 7,3) = 0,7$	0,49
4	$(4 - 7,3) = -3,3$	10,89
5	$(5 - 7,3) = -2,3$	5,29
6	$(6 - 7,3) = -1,3$	1,69
7	$(7 - 7,3) = -0,3$	0,09
8	$(8 - 7,3) = 0,7$	0,49
8	$(8 - 7,3) = 0,7$	0,49
7	$(7 - 7,3) = -0,3$	0,09
Jumlah = 146 n = 20 Mean = 7.3	Jumlah = 48.2 Varian = $48.2 / (n = 48.2 / 20)$ = 2.41 Deviasi standar = $\sqrt{\text{Varians}}$ = $\sqrt{2.41}$ = 1.55	

Seperti kita ketahui, Range menyampaikan informasi tentang batas-batas distribusi, tetapi tidak memberitahu kita bagaimana skor tersebar dalam batas-batas didistribusi. Deviasi standar menunjukkan jarak yang bervariasi dari nilai rata-rata dari distribusi. Semakin besar deviasi standar, semakin banyak variabilitas yang ada dalam distribusi. Deviasi standar ini sangat berguna dalam menggambarkan distribusi dan akan menjadi penting ketika kita memperhatikan pembahasan pada bab berikutnya. Varians adalah satu langkah penting dan berguna dalam variabilitas. Karena varians dinyatakan dalam unit pengukuran kuadrat, hal tersebut tidak berguna dalam menafsirkan nilai individu seperti pada deviasi standar. Namun, varians penting sebagai konsep teoritis, terutama saat membahas reliabilitas dan validitas dalam kajian selanjutnya.

## **C. KONSEP KORELASI**

### **1. Koefisien Korelasi**

Kebanyakan siswa terbiasa dengan konsep korelasi. Ketika orang berbicara tentang korelasi, mereka mengacu pada hubungan antara dua variabel. Variabel dapat bersifat fisik seperti berat dan tinggi atau psikologis seperti kecerdasan dan prestasi akademik. Sebagai contoh, hal tersebut masuk akal untuk melihat korelasi (hubungan) antara tinggi badan dengan berat badan. Individu – individu lebih tinggi cenderung lebih berat dari orang yang lebih pendek. Hubungan ini tidak sempurna karena ada beberapa orang pendek yang beratnya lebih tinggi. Mungkin juga menduga lebih banyak orang cerdas pada skor

lebih tinggi pada tes prestasi akademik dari pada orang yang kurang cerdas, dan ini adalah hasil penelitian. Sekali lagi, hubungan ini tidak sempurna, tetapi sebagai aturan umum individu lebih cerdas tampil lebih baik pada tes prestasi akademik dibanding rekan-rekan mereka yang kurang cerdas.

Sebuah koefisien korelasi adalah ukuran kuantitatif dari hubungan antara dua variabel. Koefisien korelasi dikembangkan oleh Karl Pearson (1857-1936) dan ditunjuk oleh huruf  $r$ .

***Sebuah koefisien korelasi adalah ukuran kuantitatif dari hubungan antara dua variabel.***

koefisien korelasi dapat berkisar dari  $-1.0$  ke  $+1.0$ . Ketika menafsirkan koefisien korelasi, ada dua parameter untuk dipertimbangkan.

Parameter pertama adalah tanda dari koefisien. Sebuah koefisien korelasi positif menunjukkan bahwa peningkatan pada satu variabel dikaitkan dengan peningkatan variabel lain. Sebaliknya, sebuah koefisien korelasi negatif menunjukkan bahwa peningkatan pada satu variabel dikaitkan dengan penurunan pada variabel lain.

Parameter kedua yang perlu dipertimbangkan ketika menafsirkan koefisien korelasi adalah besarnya atau ukuran absolut dari koefisien. Besarnya koefisien menunjukkan kekuatan hubungan antara dua variabel. Nilai 0 menunjukkan tidak adanya hubungan antara variabel. Sebuah koefisien yang mendekati nilai 1,0, menunjukkan kekuatan hubungan yang meningkat. Sebuah koefisien 1,0 (baik positif atau negatif) menunjukkan korelasi yang sempurna, dimana perubahan dalam satu variabel disertai

dengan perubahan yang sesuai dan proporsional dalam variabel lain. Koefisien korelasi yang sempurna jarang terjadi dalam pengukuran psikologis dan pendidikan.

Ada banyak cara kualitatif dan kuantitatif yang menggambarkan koefisien korelasi. Bagaimana menggambarkan koefisien korelasi apakah lemah, sedang, atau kuat? Meskipun tidak ada standar yang diterima secara universal untuk menggambarkan kekuatan korelasi, patokan berikut dapat digunakan:  $< 0.30$ , lemah;  $0.30 - 0.70$ , sedang; dan  $> 0.70$ , yang kuat (Reynolds, Livingstone, Willson, 2006.). Dalam konteks lain mungkin lebih penting untuk menentukan apakah korelasi "signifikan secara statistik." Signifikan secara statistik ditentukan oleh ukuran koefisien korelasi dan ukuran sampel.

***Koefisien determinasi diinterpretasikan sebagai jumlah varians bersama oleh dua variabel.***

Cara lain untuk menggambarkan koefisien korelasi adalah dengan mengkuadratkannya untuk memperoleh koefisien determinasi (yaitu,  $r^2$ ). Koefisien determinasi diinterpretasikan sebagai jumlah varians bersama oleh dua variabel. Dengan kata lain, koefisien determinasi mencerminkan jumlah varians dalam satu variabel yang diprediksi dari variabel lain, dan sebaliknya. Misalnya, h. Asumsikan korelasi antara tes kecerdasan dan tes prestasi 0,60. ( $r = 0.60$ ). Dengan mengkuadratkan nilai ini diperoleh koefisien determinasi adalah 0,36 ( $r^2 = 0,36$ ). Hal ini menunjukkan bahwa 36% dari varians dalam satu variabel yang diprediksi dari variabel lain.

## 2. Scatter Plot

Penelitian scatter plots dapat meningkatkan pemahaman kita tentang hubungan antara variabel. Sebuah scatter plots adalah sebuah grafik yang secara visual menampilkan hubungan antara dua variabel. Untuk membuat sebuah scatter plots harus

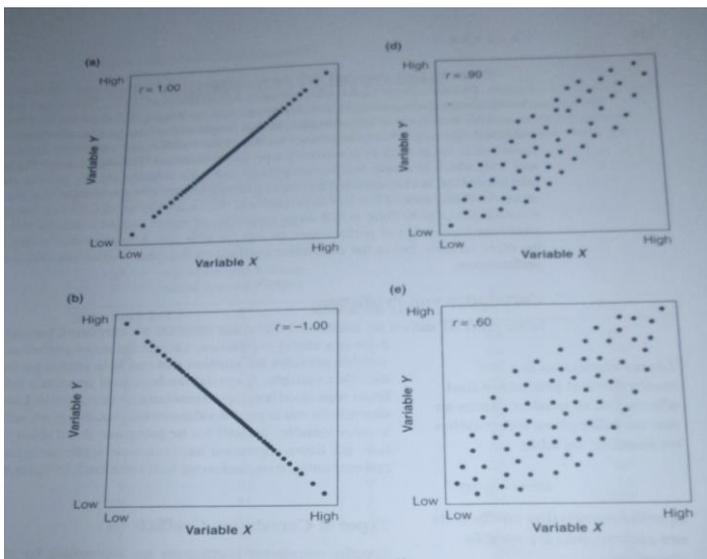
***Sebuah scatter plot adalah grafik yang secara visual menampilkan hubungan antara dua variabel.***

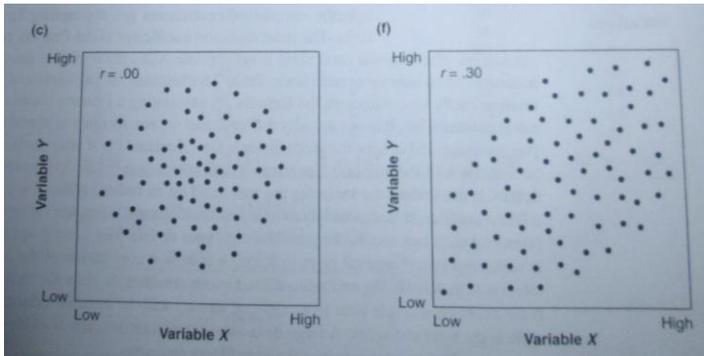
memiliki dua nilai untuk setiap individu. Sebagai contoh, korelasi antara grafik berat badan dan tinggi masing-masing individu.

Dalam konteks pengujian pendidikan, harus ada skor untuk siswa dalam satu kelas pada dua pekerjaan rumah yang berbeda. Dalam scatter plots sumbu X merupakan salah satu variabel dan sumbu Y variabel lainnya. Setiap tanda scatter plots mewakili dua nilai individu pada variabel X dan variabel Y.

Gambar 2.8 menunjukkan scatter plot untuk berbagai macam nilai korelasi. Pertama, lihat Gambar 2.8a, yang menunjukkan sebuah hipotesis korelasi positif yang sempurna (+1.0). Perhatikan bahwa dengan korelasi sempurna semua tanda-tanda akan jatuh pada garis lurus. Karena ini adalah **korelasi positif** peningkatan pada satu variabel dikaitkan dengan peningkatan yang sesuai pada variabel lain. Karena merupakan korelasi sempurna. Skor individu pada satu variabel yang dapat memprediksi skor pada variabel lain dengan presisi sempurna. Berikutnya lihat Gambar 2.8b, yang menggambarkan korelasi **negatif sempurna** (-1.0). Menjadi korelasi sempurna

semua tanda jatuh pada garis lurus, tetapi karena itu adalah korelasi negatif meningkat pada satu variabel dikaitkan dengan penurunan yang sesuai pada variabel lain. Pemberian skor pada satu variabel, Anda masih dapat memprediksi kinerja individu pada variabel lain dengan presisi sempurna. Sekarang lihat Gambar 2.8c, yang menggambarkan korelasi 0,0 (tidak ada hubungan antara variabel). Dalam situasi ini, pengetahuan tentang kinerja pada satu variabel tidak memberikan informasi apapun tentang kinerja individu pada variabel lain atau meningkatkan prediksi.





GBR 2.8. Scatter plots dari Koefisien Korelasi Berbeda

*Sumber:* D.Hopkins, Measurement in Psychology and Education. Boston, MA: Allyn & Bacon, 1998.

Pada gambar 2.8d, yang menunjukkan korelasi +0,90. Tanda korelasi jelas berkelompok sepanjang garis lurus. Hal ini memberi pemahaman bahwa situasi kinerja pada satu variabel membantu kita memprediksi kinerja pada variabel lain, tetapi kemampuan kita untuk memprediksi tidak sempurna. Akhirnya, gambar 2,8e dan 2.8f, yang menggambarkan masing-masing koefisien 0,60 dan 0,30. Hubungan 0,60 ditandai dengan tanda yang masih berkelompok sepanjang garis lurus, tapi ada lebih variabilitas sekitar garis ini daripada dengan korelasi 0,90. Demikian pula dengan korelasi 0,30 masih ada variabilitas lebih dari tanda di sekitar garis lurus. Dalam kondisi seperti ini, situasi pengetahuan kinerja pada satu variabel akan membantu kita memprediksi kinerja pada variabel lain, tetapi sebagai koefisien korelasi berkurang begitu pula kemampuan kita untuk memprediksi kinerja.

### 3. Korelasi dan Prediksi

Pada bagian sebelumnya telah disebutkan bahwa ketika variabel yang berkorelasi, terutama ketika ada korelasi yang kuat, pengetahuan tentang penampilan pada satu variabel menyediakan informasi yang dapat membantu memprediksi kinerja pada variabel lain. Sebuah prosedur matematika khusus disebut sebagai regresi linier dirancang tepat untuk tujuan ini. Regresi linier memungkinkan Anda untuk memprediksi nilai-nilai pada satu variabel jika diberi informasi tentang variabel lain. Kami tidak akan pergi ke detail tentang perhitungan, tetapi regresi linier memiliki berbagai aplikasi dalam mengembangkan dan mengevaluasi tes, dan jadi kita akan datang kembali untuk regresi linier di bab berikutnya.

### 4. Jenis Koefisien Korelasi

Koefisien korelasi spesifik sesuai untuk situasi tertentu. Koefisien yang paling umum adalah *Pearson Product - Moment Correlation* (Koefisien Korelasi Pearson). Formula ini sesuai ketika variabel yang berkorelasi diukur pada interval atau skala rasio. Tabel 2.6 menggambarkan perhitungan koefisien korelasi Pearson. Meskipun rumus untuk menghitung korelasi Pearson mungkin tampak agak rumit, itu tidak benar-benar sulit, dan kami mendorong Anda untuk meninjau bagian ini jika Anda tertarik pada bagaimana koefisien ini dihitung ( atau jika dosen Anda ingin Anda menjadi akrab dengan proses).

Koefisien korelasi lain yang populer digunakan adalah koefisien korelasi Spearman. Digunakan ketika variabel diukur

pada skala ordinal. Ada juga *Koefisien korelasi point- biserial* juga banyak digunakan dalam pengembangan tes ketika salah satu variabel dikotomi (yang berarti hanya dua nilai yang mungkin, misalnya, lulus atau gagal, 0 atau 1, dll) dan variabel lain yang diukur pada interval atau rasio skala. Sebuah aplikasi umum dari korelasi point- biserial adalah dalam menghitung nilai ujian korelasi item-total.

### **TABEL 2.6 Menghitung Koefisien Korelasi Pearson**

Ada dua formula/rumus yang digunakan untuk menghitung Koefisien Korelasi Pearson, dan berikut ditampilkan yang lebih sederhana. dan kami akan menggambarkan salah satu yang lebih sederhana. Untuk ilustrasi ini kita akan menggunakan nilai tugas pekerjaan rumah sebagai variabel X dan satu set dari 20 skor hipotesis sebagai variabel Y. Rumusnya adalah:

$$r_{xy} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{N\sum X^2 - (\sum X)^2} \sqrt{N\sum Y^2 - (\sum Y)^2}}$$

$\sum XY$  = Jumlah produk XY

$\sum X$  = Jumlah skor X

$\sum Y$  = Jumlah skor Y

$\sum X^2$  = Jumlah skor X kuadrat

$\sum Y^2$  = Jumlah skor Y kuadrat

<b>Pekerjaan Rumah 1 (X)</b>	<b>X<sup>2</sup></b>	<b>Pekerjaan Rumah 2 (Y)</b>	<b>Y<sup>2</sup></b>	<b>(X) (Y)</b>
7	49	8	64	56
8	64	7	49	56
9	81	10	100	90
6	36	5	25	30
7	49	7	49	49
6	36	6	36	36
10	199	9	81	90
8	64	8	64	64
5	25	5	25	25
9	81	9	81	81
9	81	8	64	72
9	81	7	49	63
8	64	7	49	56
4	16	4	16	16
5	25	6	36	30
6	36	7	49	42
7	49	7	49	49
8	64	9	81	72
8	64	8	64	64
7	49	6	36	42
<b>X = 146</b>	<b>X<sup>2</sup> = 1,114</b>	<b>Y = 143</b>	<b>Y<sup>2</sup> = 1,067</b>	<b>XY = 1,083</b>

$$\begin{aligned}
 r_{xy} &= \frac{20(1,083)-(140)(143)}{\sqrt{20(1,114)-(146)^2}\sqrt{20(1,067)-(143)^2}} \\
 &= \frac{20(1,083)-(140)(143)}{\sqrt{20(1,114)-(146)^2}\sqrt{20(1,067)-(143)^2}} \\
 &= \frac{21,660-20,878}{\sqrt{22,280-21,316}\sqrt{21,340-20,449}} \\
 &= \frac{782}{\sqrt{964}\sqrt{891}} \\
 &= \frac{782}{(31.048)(29.849)} = 0.843
 \end{aligned}$$

## 5. Korelasi Versus Kausalitas

Pada waktu variabel berkorelasi maka akan meningkatkan kemampuan prediksi kinerja pada variabel lain. Telah diketahui bahwa dengan mengkuadratkan koefisien korelasi untuk mendapatkan koefisien deteminasi, kita dapat membuat pernyataan tentang jumlah varians bersama oleh dua variabel.

***Analisis korelasi tidak memungkinkan seseorang untuk menetapkan kausalitas***

Pemahaman umum sering terjadi yakni percaya bahwa jika dua variabel yang berkorelasi maka satu menyebabkan yang lain. Ada kemungkinan bahwa variabel kausal berkaitan, tetapi juga mungkin bahwa variabel ketiga menjelaskan hubungan tersebut. Mari kita lihat sebuah contoh. Anggaphlah kita menemukan korelasi antara jumlah es krim yang dikonsumsi di New York dan jumlah kematian karena tenggelam

di Texas. Jika menafsirkan hubungan ini sebagai menyimpulkan sebab-akibat, maka orang yang makan es krim di New York membuat orang tenggelam di Texas atau bahwa orang-orang yang tenggelam di Texas menyebabkan orang untuk makan es krim di New York. Jelas tidak akan benar! Bagaimana menjelaskan hubungan ini? Jawabannya adalah bahwa perubahan musiman suhu membentuk hubungan tersebut. Pada akhir musim semi dan musim panas orang di New York mengkonsumsi lebih banyak es krim dan orang-orang di Texas terlibat dalam lebih banyak kegiatan yang berhubungan dengan air (misalnya, berenang, ski, berperahu, dll) dan akibatnya lebih sering ada yang tenggelam.. Kasus ini cukup jelas dari variabel ketiga menjelaskan hubungan. Namun, mengidentifikasi variabel ketiga tidak selalu begitu mudah. Pada saat mendengar di televisi atau membaca di koran bahwa para peneliti menemukan korelasi antara variabel A dan variabel B, tidak boleh langsung disimpulkan bahwa korelasi ini berarti bahwa A menyebabkan B, atau sebaliknya. Meskipun analisis korelasi tidak memungkinkan kita untuk membangun hubungan sebab akibat, prosedur statistik tertentu secara khusus dirancang untuk memungkinkan kita untuk menyimpulkan kausalitas. Prosedur ini disebut statistik inferensial sebagai dan akan dibahas dalam mata kuliah statistik.

\* \* \*