



VALIDITAS DAN RELIABILITAS

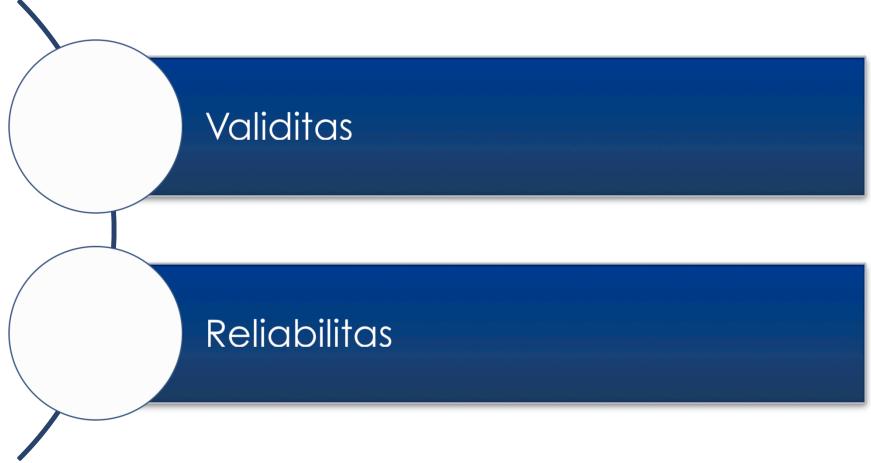
Debrina Puspita Andriani

www.debrina.lecture.ub.ac.id

E-mail: debrina@ub.ac.id



Outline



www.debrina.lecture.ub.ac.id

Transformasi Data Menjadi Informasi

Data yang baik

- Data harus akurat atau valid
- Data harus relevan
- Data harus up to date

Kaidah Pengukuran yang Baik

Objektivitas

Apakah pengukuran dilakukan apa adanya? Atau tidak bias?

YA, berarti pengukuran objektif

Validitas

Apa pengukuran benar – benar mengukur apa yang dikehendaki atau mengukur dengan akurat?

Adakah keterkaitan metode dan alat ukur dengan objek yang dikur?

YA, berarti pengukuran valid

Reliabilitas

Apakah hasil pengukuran mencerminkan nilai sesungguhnya suatu variabel?

Apakah hasil pengukuran konsisten?

YA, berarti pengukuran reliabel

Uji Kualitas data

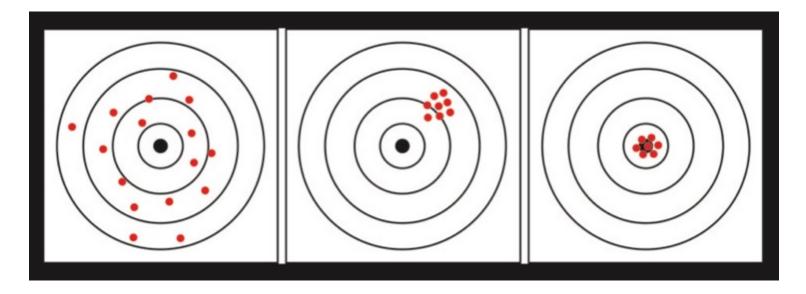
Dua Konsep untuk mengukur kualitas data:



Realibilitas

- Kesimpulan akan bias jika data tidak reliabel dan tidak valid
- Kualitas data penelitian ditentukan oleh kualitas instrumen untuk pengumpulan data

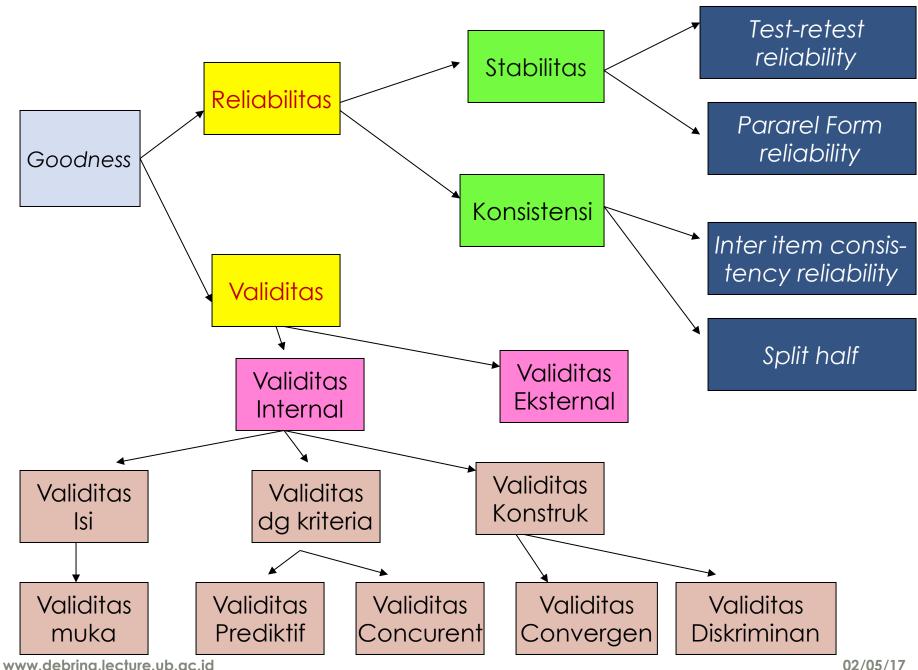
VALIDITY vs. RELIABILITY



Low Reliability Low Validity

High Reliability Low Validity

High Reliability High Validity



02/05/17

VALIDITAS

Validitas mengukur ketepatan (akurasi).

VALIDITAS

- Menunjukkan sejauh mana suatu alat (instrumen) mengukur apa yang seharusnya diukur (Ghiselli, 1981).
- Sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya (Azwar, 2000).
- Validitas menunjukkan tingkat dari kemampuan test untuk mencapai sasarannya (Isaac dan Michael, 1981).

VALIDITAS

- Alat pengumpul data dapat dikatakan valid atau sahih apabila alat ukur tersebut mampu mengukur apa yang seharusnya diukur / diinginkan.
- Bila skala pengukuran tidak valid maka ia tidak bermanfaat bagi peneliti karena tidak mengukur atau melakukan apa yang seharusnya dilakukan.



A tape measure is used to measure length or distance

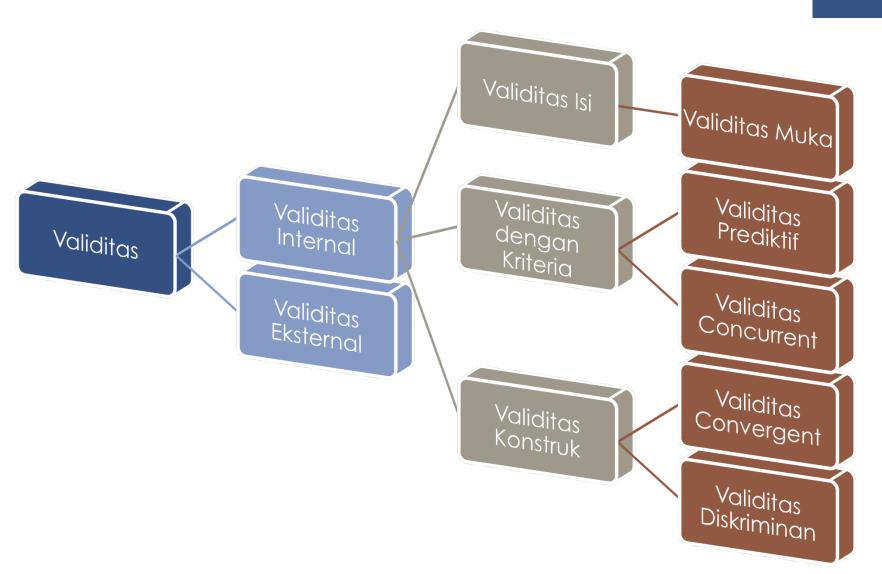


A thermometer is used to measure temperature



A stopwatch is used to measure time

Different tools are used things for measuring different things



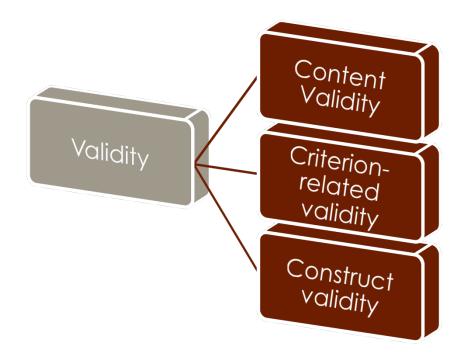
Validitas Eksternal

- Dapat atau tidaknya hasil penelitian digeneralisasikan pada populasi tempat sampel tersebut diambil.
- Bila sampel penelitian representatif, instrumen penelitian valid dan reliabel, cara mengumpulkan dan menganalisis data benar, penelitian akan memiliki validitas eksternal yang tinggi.

Contoh:

- Dilakukan penelitian yang bertujuan mengetahui pola belajar mahasiswa di kampus area malang.
- Hasil penelitian diperoleh kesimpulan bahwa dari sampel diketahui sebagian besar mahasiswa lebih sering belajar SKS (sistem kebut semalam) dan lebih intens mengerjakan praktikum.
- Jika hasil penelitian ini mewakili pola populasi mahasiswa di malang maka validitas eksternal dikatakan tinggi.

- Digunakan untuk menjawab pertanyaan apakah penelitian sudah menggunakan konsep yang seharusnya (actually).
- Validitas internal biasanya membantu mengatasi kelemahan validitas eksternal.



Validitas Internal Content Validity

 Validitas isi memastikan bahwa sejumlah item yang representatif telah diperhitungkan dalam menyusun sebuah konsep.

 Validitas isi merupakan sebuah fungsi yang menunjukkan seberapa baik dimensi dan elemen sebuah konsep digambarkan.

Contoh: Content Validity

 Seorang peneliti tertarik untuk meneliti pengaruh program matematika realistik terhadap kemampuan matematika siswa kelas 5. Peneliti memberikan tes berupa 15 soal matematika yang mengandung 15 masalah berbeda. Kinerja siswa pada tes ini sangat penting untuk melihat tingkat kemampuan siswa dalam memecahkan masalah. Untuk itu soal tes yang diberikan harus memiliki validitas untuk mengukur kemampuan siswa yang diharapkan oleh program.

Agar mempunyai content validity maka:

- Soal tes terdiri dari soal yang mudah dan sukar saja agar soal tes representatif dan memberi informasi yang valid.
- √ Sampel cukup.
- ✓ Format instrumen, yang meliputi kejelasan cetakan, ukuran huruf, ruang yang cukup untuk menjawab dan keterbacaan soal.
- ✓ Untuk itu maka soal yang disusun berdasarkan kisi-kisi, perlu diuji coba kepada sejumlah sampel di luar sampel penelitian, kemudian dianalisis.

Criterion-related Validity

- Untuk mendapatkan pembuktian validitas kriteria, peneliti biasanya membandingkan bentuk instrumen yang satu dengan instrumen lainnya.
- Pengukuran indeks validitas kriteria tersebut dengan koefisien korelasi (r) yang menunjukkan derajat hubungan antara skor individual yang diperoleh melalui dua instrumen tersebut.
- Koefisien korelasi berada pada kisaran angka +1,00 dan -1,00; r =0 menunjukkan tidak ada hubungan.
- Koefisien korelasi (r) disebut pula koefisien validitas.
- Misalnya suatu instrumen dirancang untuk mengukur kemampuan akademis, maka skor siswa berdasarkan instrumen tersebut harus dibandingkan dengan IPK mereka (kriteria eksternal).
- Jika instrumen benar-benar mengukur kemampuan akademis, maka siswa yang mendapat nilai tinggi untuk tes tersebut, seharusnya juga tinggi IPK-nya.

Criterion-related Validity

Concurrent Validity

- Apabila data instrumen dan data kriteria telah terkumpul pada waktu yang hampir bersamaan, lalu dibandingkan hasilnya.
- Contohnya: Bila seorang peneliti melakukan tes IQ pada sekelompok siswa kelas 8 dan kemudian membandingkannya dengan hasil penilaian 'self esteem' yang dilakukan oleh gurunya, di mana kedua penilaian ini dilakukan pada waktu yang hampir bersamaan.

Predictive validity

- Peneliti harus menyediakan selang waktu antara pencatatan instrumen dan perolehan skor kriteria.
- Contoh: seorang peneliti ingin mebuktikan hasil tes keterampilan sains siswa SMA yang diberikan pada awal semester untuk melakukan prediksi. Kemudian skor keterampilan sains ini dibandingkan dengan hasil tes akhir semester untuk mata pelajaran sains.

Construct Validity

- Validitas konstruk membuktikan seberapa bagus hasil yang diperoleh dari penggunaan ukuran sesuai dengan teori dimana pengujian dirancang.
- Hal ini dinilai dengan convergent validity (instrument yang memiliki korelasi tinggi) dan discriminant validity (variabel yang tidak berkorelasi).

- Pembuktian validitas konstruk merupakan yang terluas penggunaannya di antara tiga kategori validitas. Ada tiga tahap untuk memperoleh pembuktian validitas konstruk:
 - variabel yang diukur didefinisikan dengan jelas
 - 2. hipotesis, berdasarkan teori yang melandasi variabel
 - 3. hipotesis diuji secara logis dan empiris

Contoh Construct Validity

- Peneliti berniat mengembangkan tes tertulis untuk mengukur "kejujuran" dan bermaksud menggunakan pembuktian validitas konstruk.
- Maka yang harus dilakukannya adalah:
 - ✓ Mendefinisikan tentang "kejujuran"
 - ✓ Merumuskan teori tentang kejujuran seseorang dibandingkan dengan yang tidak jujur, sehingga jelas perbedaan antara jujur dan tidak jujur.
 - ✓ Berdasarkan teori ini peneliti menyusun hipotesis, bahwa individu yang mendapatkan skor tinggi dalam 'tes kejujuran' akan lebih tergerak untuk mengembalikan barang milik orang lain yang ditemukannya ibandingkan dengan individu yang mndapatkan skor rendah
 - ✓ Selanjutnya peneliti mencatat hasil tes 'kejujuran', dan memisahkan siswa yang mendapat skor tinggi dan skor rendah
 - ✓ Kemudian ia melakukan tes lainnya, misalnya meninggalkan sebuah dompet berisi uang Rp 50.000,- di kelas yang di dalamnya dompet terdapat kartu nama pemiliknya.
 - ✓ Jika hipotesis penelitiannya benar, maka individu mengembalikan dompet adalah individu yang mendapat skor tinggi untuk tes 'kejujuran'.

RELIABILITAS

sejauh mana hasil suatu pengukuran dapat dipercaya. Kepercayaan itu dalam bentuk keandalan instrumen yaitu **konsistensi** hasil dari waktu kewaktu jika suatu instrumen digunakan pada subjek.

RELIABILITAS

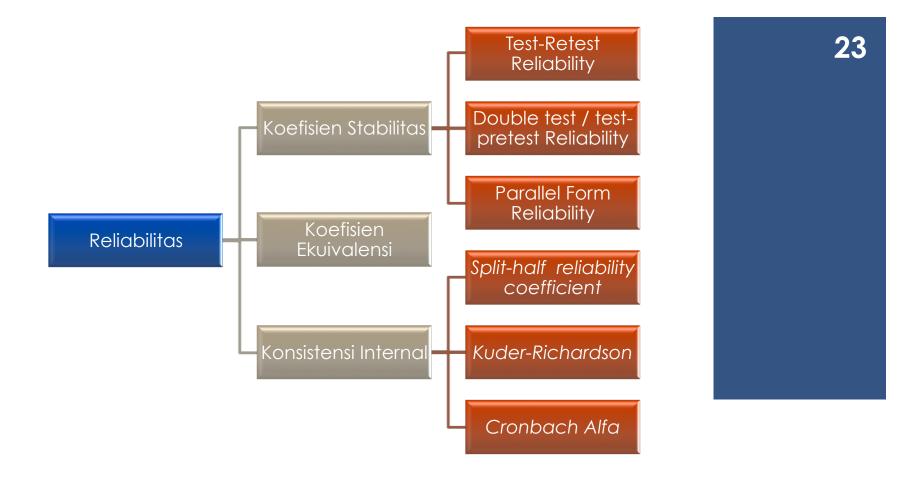
- Seberapa besar variasi tidak sistematik dari suatu individu jika individu yang sama diukur berkali kali (Ghiseli, 1981).
- Ukuran yang menunjukkan stabilitas dan konsistensi suatu instrumen yang mengukur suatu konsep dan berguna untuk mengukur kebaikan (goodness) dari suatu pengukur (Sekaran, 2003).
- Reliabilitas berkaitan dengan keajegan atau konsistensi dari skor yang diperoleh, yaitu bagaimana konsistensinya antara setiap individu yang dites oleh instrumen tersebut

Contoh:

Suatu tes dirancang untuk mengukur logika berpikir. Bila tes ini reliable, maka dapat diharapkan bila siswa mendapat skor tinggi pada tes pertama, ia pun akan mendapatkan skor tinggi pula pada kesempatan lain bila ia mengambil tes tersebut.

RELIABILITAS vs. VALIDITAS

- Skor yang diperoleh dari suatu instrumen <u>dapat reliabel, tetapi belum</u> tentu valid.
- Contoh.
 - Seorang peneliti memberikan dua macam tes (A dan B) kepada sekelompok siswa kelas untuk mengukur pemahaman mereka tentang "Statistika" lalu mendapatkan bahwa skor tes mereka konsisten: artinya siswa yang mendapat skor tinggi pada Tes A juga mendapat skor tingi pada Tes B; dan siswa yang mendapat skor rendah pada tes A juga mendapat skor rendah pada tes B, Maka dapat dikatakan bahwa skor tersebut reliabel.
- Tetapi bila peneliti menggunakan skor tes "statistika" yang sama untuk melakukan prediksi keberhasilan siswa dalam pelajaran OMI, maka skor tes tsb tidak valid



Konsep Reliabilitas

Reliabilitas dapat diukur dengan 3 pendekatan

Koefisien Stabilitas Test – Retest Reliability

Menggunakan instrumen yang sama terhadap subjek penelitian tertentu sebanyak 2 kali pada saat yang berbeda

Bermaksud untuk menguji stabilitas jawaban responden dari suatu waktu ke waku berikutnya dengan cara menghitung koefisien korelasi dan skor jawaban responden yang diukur dengan instrumen yang sama pada saat yang berbeda

- Yang biasa digunakan → Pearson Correlation
- Kelemahan:
 - Kemungkinan adanya perubahan kondisi subyek sejalan dengan perbedaan waktu.
 - Sulitnya mencari kembali responden yang sama pada periode yang berbeda.
 - Sulitnya menentukan tenggang waktu yang pas

1. Koefisien Stabilitas

Contoh: Test – Retest Reliability

Question:

If infrastructure in the country is not good, your company will cancel your decision to do Foreign Direct Investment (FDI) in that following country.

Answer

(now): strongly disagree

■ (30 days later): agree

➤ Conclusion: Low Reliability

Koefisien Stabilitas Double test / test-pretest

Pada teknik ini kita membagi kuesioner yang sama pada waktu yang berbeda.

Misalnya:

- Pada minggu I ditanyakan: "Bagaimana tanggapan saudara terhadap kualitas dosen di Universitas X?"
- Pada minggu III ditanyakan: Ditanyakan lagi pada responden yang sama dengan pertanyaan yang sama.

1. Koefisien Stabilitas

Parallel Form

Pada teknik ini kita membagi kuesioner kepada responden yang intinya sama akan tetapi menggunakan kalimat yang berbeda.

Misalnya:

- Apakah menurut saudara harga tiket di kereta ini tidak mahal?
- Apakah harga di kereta ini telah sesuai dengan pelayanan yang saudara terima?

Contoh lain:

- Apakah saudara betah tinggal di perumahan ini?
- Apakah saudara ingin pindah dari perumahan ini?

1. Koefisien Stabilitas

Contoh: Parallel Form

- Do you think that Susi is beautiful?
 - Answer: YES
- Do you think that Tukul's wife is beautiful?
 - Answer: NO
- Fact: Susi is Tukul's wife.
 - ➤ Conclusion: Low Reliability

2. Koefisien Ekuivalensi

- Pengukuran menggunakan instrumen yang berbeda untuk subjek penelitian yang sama pada saat yang sama.
- Menguji korelasi skor jawaban responden dengan instrumen yang berbeda.
- Dua buah eksperimen disusun berdasarkan satu kisi-kisi. Kedua buah instrumen diuji cobakan semua. Sesudah uji coba terlaksana, maka hasil kedua instrumen tsb dihitung korelasinya dengan menggunakan rumus product moment (korelasi pearson).
 - Korelasi antara skor skor yang didapatkan pada kedua bentuk itu merupakan koefisien reabilitas tes.

- Konsistensi antara butir-butir pertanyaan atau pernyataan dalam suatu instrumen
- Tingkat keterkaitan antara butir pertanyaan atau pernyataan menunjukkan reliabilitas konsistensi internal instrumen yang bersangkutan
- Ada 3 macam instrumen untuk mengukur konsistensi internal yaitu:
 - Split-half reliability coefficient
 - Kuder-Richardson
 - Cronbach Alfa

Split-half reliability coefficient

- Peneliti boleh hanya memiliki seperangkat instrumen saja dan hanya diuji cobakan satu kali, kemudian hasilnya dianalisis, yaitu dengan cara membelah seluruh instrumen menjadi 2 sama besar.
- Langkah-langkah:
 - 1. Membuat tabel analisis butir.
 - 2. Mengelompokan skor menjadi dua bagian soal.
 - a. Belahan Ganjil-Genap
 - b. Belahan Awal-Akhir
 - 3. Korelasikan skor belahan pertama dengan skor belahan kedua dan diperoleh r_{xv}.

 Hitung nilai reliabilitas internal dengan persamaan sebagai berikut:

$$r = \frac{2 \times r \ half \ test}{1 + r \ half \ test}$$

$$r_{11} = \frac{2.r_{1/2.1/2}}{(1 + r_{1/2.1/2})}$$

Keterangan:

r₁₁ : Reliabilitas instrumen

 $r_{1/2.1/2}:r_{xy}$ indeks korelasi antara dua

belahan instrumen.

Kriteria:

Instrumen dikatakan reliabel:

Jika
$$r_{11} > r_{tabel}$$
 (df: α , n-2)

Contoh: Split-half reliability coefficient

Resp.	X1_1	X1_2	X1_3	X1_4	Total Ganjil	Total Genap
1	5	4	5	4	10	8
2	4	5	4	4	8	9
3	5	4	5	5	10	9
4	5	5	5	3	10	8
5	4	5	5	5	9	10
6	3	4	4	3	7	7
7	4	4	4	5	8	9
8	4	4	5	4	9	8
9	4	4	4	4	8	8
10	4	4	5	4	9	8
11	5	4	5	5	10	9
12	5	4	4	4	9	8
13	4	4	5	4	9	8
14	5	4	4	4	9	8
15	3	4	5	4	8	8
16	4	4	5	5	9	9
17	5	4	4	3	9	7
18	5	4	5	4	10	8
19	3	3	3	3	6	6
20	5	5	5	4	10	9
21	4	4	5	4	9	8
22	5	4	5	4	10	8
23	4	3	4	4	8	7
24	4	4	4	4	8	8
25	4	5	5	5	9	10
26	4	4	3	3	7	7
27	5	4	5	5	10	9
28	4	4	4	4	8	8
29	4	4	4	5	8	9
30	3	4	4	4	7	8
					F _{Gan jil. Genap}	0.493

$$r_{11} = \frac{2.r_{1/2.1/2}}{(1 + r_{1/2.1/2})}$$

$$r_{11} = \frac{2.(0,493)}{(1+0,493)} = 0,661$$

Kriteria:

 r_{11} =0,661 > r_{tabel} =0,374 maka instrumen dinyatakan <u>reliabel</u>.

3. Konsistensi Internal Kuder-Richardson

- Cara ini diberlakukan bila instrumen digunakan untuk mengukur satu gejala psikologis atau perilaku yang sama, artinya alat ukur tersebut dapat dikatakan reliabel bila terbukti ada konsistensi jawaban antar item yang satu dengan item yang lain.
- Didasarkan pada konsistensi respons terhadap semua butir soal dalam tes.

$$r = \frac{n}{n-1} \left(\frac{s_t^2 - npq}{s_t^2} \right)$$

KR 20

$$r = \frac{n}{n-1} \left(\frac{s_t^2 - pq}{s_t^2} \right)$$

Keterangan:

n = jumlah item dalam tes

St = simpangan baku seluruh tes

p = mean dibagi jumlah item

q = 1-p

Contoh: Kuder-Richardson

No.	Jawaban Angket						
Resp	1	2	3	4	5	Total	Total Kuadrat
1	4	4	3	4	4	19	361 ^(a)
2	3	3	4	4	4	18	324
3	2	2	2	2	4	12	144
4	3	4	4	4	4	19	361
5	3	4	4	4	5	20	400
6	4	3	3	4	4	18	324
7	2	3	3	4	5	17	289
8	4	4	4	2	4	18	324
9	4	4	4	2	4	18	324
10	4	4	4	4	4	20	400
Jumlah	33	35	35	34	42	179	3251 ^(c)
	115 ^{(b}						
Jumlah Kuadrat		127	127	124	178	115	

Keterangan:

 $361^{(a)} \rightarrow 19^2$

 $115^{(b)} \rightarrow 4^2+3^2+2^2+3^2+3^2+4^2+2^2+4^2+4^2$

 $3251^{(c)} \rightarrow 361+324+144+361+400+324+289+324+324+400$

www.debrina.lecture.ub.ac.id

35

Contoh: Kuder-Richardson

1. Menghitung Total Varians Butir $(\sum \sigma_b^2)$ 2. Menghitung Total Varians (σ_t^2)

Contoh menghitung varians Butir (σ_b) pertama

$$\sigma_b^2 = \frac{33^2}{10}$$

$$= 0,61$$

Varians butir ke-2 sampai ke-5 dapat dihitung dengan cara yang sama seperti menghitung varians butir I. Dengan demikian total varians butir:

$$\sum \sigma_b^2 = 0.61 + 0.45 + 0.45 + 0.84 + 0.16$$
$$= 2.51$$

3. Menghitung Koefisien Cronbach Alpha

$$r = \left[\frac{k}{(k-1)}\right]\left[1 - \frac{\sum \sigma_b^2}{\sigma_t^2}\right]$$

$$r = \left[\frac{5}{(5-1)}\right] \left[1 - \frac{2.51}{4,69}\right]$$

= 0.581023 (dibulatkan menjadi 0,58)

Cronbach Alfa

- Perbedaan pokok dengan model Kuder Richardson adalah bahwa teknik ini tidak hanya untuk instrumen dengan dua pilihan tetapi tidak terikat pada dua pilihan saja, sehingga penerapannya lebih luas. Misalnya untuk menguji reliabilitas skala pengukuran dengan 3, 5 atau 7 pilihan.
- Langkah-langkah metode ini:
 - 1. Membuat tabel analisis butir
 - 2. Menghitung nilai total item pertanyaan
 - 3. Menghitung nilai varian butir dan varian total
 - 4. Menjumlahkan nilai varian butir

 Hitung nilai reliabilitas internal dengan persamaan sebagai berikut:

$$\alpha = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum \sigma b^2}{(\sigma^2 t)}\right)$$

Keterangan:

α : Koefisien Alpha Cronbach

k : Jumlah butir pertanyaan

 $\Sigma \sigma_{\rm b}^2$: Jumlah varian butir

 σ_{t^2} : Varian total

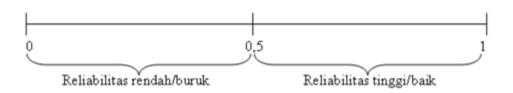
Kriteria:

Instrumen dikatakan reliabel:

Jika $\alpha > r_{tabel}$ (df: α , n-2)

3. Konsistensi Internal Cronbach Alfa

- Tinggi/rendahnya reliabilitas secara empirik ditunjukkan oleh suatu angka yang disebut nilai koefisien reliabilitas.
- Reliabilitas yang tinggi ditunjukkan dengan nilai 1,00. Reliabilitas yang dianggap cukup memuaskan atau tinggi adalah ≥ 0,70.



No	Interval	Kriteria		
1.	< 0,200	Sangat rendah		
2.	0,200 – 0,399	Rendah		
3.	0,400 – 0,599	Cukup		
4.	0,600 – 0,799	Tinggi		
5.	0,800 – 1,000	Sangat Tinggi		

Contoh: Cronbach Alfa

Resp.	X1_1	X1_2	X1_3	X1_4	Total
1	5	4	5	4	18
2	4	5	4	4	17
3	5	4	5	5	19
4	5	5	5	3	18
5	4	5	5	5	19
6	3	4	4	3	14
7	4	4	4	5	17
8	4	4	5	4	17
9	4	4	4	4	16
10	4	4	5	4	17
11	5	4	5	5	19
12	5	4	4	4	17
13	4	4	5	4	17
14	5	4	4	4	17
15	3	4	5	4	16
16	4	4	5	5	18
17	5	4	4	3	16
18	5	4	5	4	18
19	3	3	3	3	12
20	5	5	5	4	19
21	4	4	5	4	17
22	5	4	5	4	18
23	4	3	4	4	15
24	4	4	4	4	16
25	4	5	5	5	19
26	4	4	3	3	14
27	5	4	5	5	19
28	4	4	4	4	16
29	4	4	4	5	17
30	3	4	4	4	15
Varian	0.461	0.231	0.395	0.438	2.921
Jlh.∀ar.Butir				1.525	

$$\alpha = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum \sigma b^2}{(\sigma^2 t)}\right)$$

$$\alpha = (\frac{4}{4-1})(1 - \frac{1,525}{2,921}) = 0,637$$

Kriteria:

 α = 0,637 > r_{tabel} = 0,374 maka instrumen dinyatakan <u>reliabel</u>.

PENDEKATAN	DEFINISI	KELEMAHAN	39
Test-retest	Menyajikan tes dua kali pada satu kelompok subyek dengan tenggang waktu diantara kedua penyajian tersebut.	 Kemungkinan adanya perubahan kondisi sub-yek sejalan dengan per-bedaan waktu Adanya efek bawaan (<i>carry-over effect</i>) pada subyek Sulitnya menentukan tenggang waktu yang pas 	
Parallel-form	Menggunakan dua tes yang sama tujuan ukurnya dan setara isi item-nya	- Sulitnya menyusun dua tes yang paralel yang setara	
Internal Consistency	Menggunakan satu bentuk tes yang dikenakan sekali pada sekelompok subyek dan dibagi menjadi beberapa belahan	- Diperlukan bantuan teknik komputasi yang cermat	

Summary