

Variabel dan Konstanta

Ada dua istilah yang sangat penting dalam dunia ilmiah yaitu variabel dan konstanta. Variabel berasal dari kata vary dan able, vary artinya berbeda sedangkan able mampu. Jadi arti dari variabel adalah sesuatu yang memiliki Seperti namanya, variabel adalah

KONSEP STATISTIK

Ketika angka-angka atau grafik digunakan untuk menggambarkan, meringkas, atau menampilkan data maka ini masuk statistik deskriptif. Namun jika data digunakan untuk mengestimasi nilai populasi yang berdasarkan nilai sampel atau menguji hipotesis, maka statistik inferensial yaitu rangkaian prosedur statistik yang berdasarkan kepada teori probabilitas, digunakan.

Tabel 1: Data Mentah: 60 Skor Tes

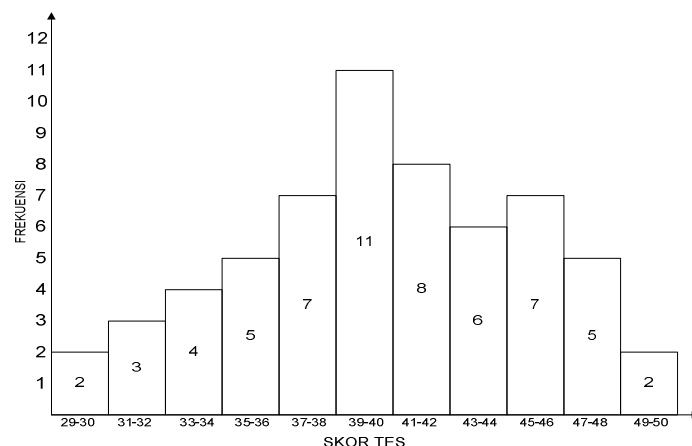
41 50 39 40 40 31 42 29 37 36 34 37 38 40 39 39 36 32 48 39
35 45 44 49 38 34 35 32 41 41 33 42 44 48 47 40 33 46 46 40
39 47 30 45 43 47 35 46 42 41 44 37 45 43 39 42 37 45 43 38

Distribusi Frekuensi

Sebelum menggunakan rumus statistik, adalah hal yang baik jika data mentah itu diatur sedemikian rupa sehingga bisa dipahami. Biasanya cara yang digunakan adalah dengan menggunakan distribusi frekuensi.

Tabel 2 Distribusi Frekuensi Data Mentah Tabel 2.2

Setelah ditata menurut distribusi frekuensi, sebuah data bisa juga ditampilkan kedalam sebuah grafik. Grafik batang (bar graphic) dan lingkaran (pie graphic) adalah untuk data yang berupa data diskrit atau kategori, sedangkan grafik histogram atau polygon untuk data kontinum atau metrik.



Grafik

Disamping menyajikan frekuensi yang kemudian disajikan kedalam grafik, statistik deskriptif juga menyediakan deskripsi numerik data, yaitu tendensi sentral dan variabilitas.

Tendensi sentral

Modus: adalah skor yang paling sering muncul dalam sebuah distribusi. Sebagian besar data memiliki satu modus, yang disebut dengan *unimodal*, namun tidak jarang sebuah data memiliki dua angka atau lebih yang memiliki frekuensi tertinggi. Jika memiliki dua modus disebut *bimodal*, jika lebih disebut data *multimodal*. Semakin data itu memiliki banyak modus maka data itu semakin tidak variatif. Untuk data diatas, modulusnya adalah 39 dan termasuk data yang unimodal.

Median: adalah angka yang membagi sebuah distribusi yang berurutan menjadi dua bagian, atau menjadi bagian atas dan bawah. Jika frekuensi sebuah data adalah ganjil maka tinggal menunjuk angka diantara dua bagian, namun jika frekuensi genap maka yang menjadi median adalah nilai tengah dari dua angka tengah. Untuk data diatas, mediannya adalah 40. Data diatas jumlah skornya adalah genap sehingga harus membagi dua angka yang tepat berada ditengah-tengah. Kebetulan dua angka yang berad ditengah kedua-duanya adalah 40, maka median data diatas adalah 40.

Mean atau rata-rata arithmatik (μ untuk mean populasi dan M untuk mean sampel) diperoleh dengan menjumlahkan semua nilai angka dan dibagi dengan jumlah angka dalam sebuah distribusi. Rata-rata ini sangat dipengaruhi oleh data ekstrim. Jika ada kasus atau angka ekstrim yang tinggi maka rata-ratanya akan secara drastis menjadi tinggi, namun jika angka ekstrimnya rendah rata-ratanya akan turun drastis. Rumus rata-rata adalah $\frac{\sum X}{N}$ atau $\frac{\sum fX}{N}$.

Tabel 3: Data untuk Penghitungan Mean, Kuartil, Persentil, Deviasi Standar, dan Varian.

SKOR (X)	f				fX	$\frac{x}{(X - \bar{X})}$	fx	x^2	fx^2
29	1	1	0.017	0.017	29	-11.13	-11.1	124	123.951
30	1	2	0.017	0.033	30	-10.13	-10.1	102.7	102.684
31	1	3	0.017	0.05	31	-9.13	-9.13	83.42	83.4178
32	2	5	0.033	0.083	64	-8.13	-16.3	66.15	132.302
33	2	7	0.033	0.117	66	-7.13	-14.3	50.88	101.769
34	2	9	0.033	0.15	68	-6.13	-12.3	37.62	75.2356
35	3	12	0.05	0.2	105	-5.13	-15.4	26.35	79.0533
36	2	14	0.033	0.233	72	-4.13	-8.27	17.08	34.1689
37	4	18	0.067	0.3	148	-3.13	-12.5	9.818	39.2711
38	3	21	0.05	0.35	114	-2.13	-6.4	4.551	13.6533
39	6	27	0.1	0.45	234	-1.13	-6.8	1.284	7.70667
40	5	32	0.083	0.533	200	-0.13	-0.67	0.018	0.08889
41	4	36	0.067	0.6	164	0.87	3.467	0.751	3.00444
42	4	40	0.067	0.667	168	1.87	7.467	3.484	13.9378
43	3	43	0.05	0.717	129	2.87	8.6	8.218	24.6533
44	3	46	0.05	0.767	132	3.87	11.6	14.95	44.8533
45	4	50	0.067	0.833	180	4.87	19.47	23.68	94.7378
46	3	53	0.05	0.883	138	5.87	17.6	34.42	103.253
47	3	56	0.05	0.933	141	6.87	20.6	47.15	141.453
48	2	58	0.033	0.967	96	7.87	15.73	61.88	123.769
49	1	59	0.017	0.983	49	8.87	8.867	78.62	78.6178

50	1	60	0.017	1	50	9.87	9.867	97.35	97.3511
	N = 60				$\sum fX = 2408$				

Rata-rata: $\bar{X} = 2408 : 60 = 40.13333$

Variabilitas

Statistik yang mengukur variabilitas adalah statistik untuk menunjukkan sejauhmana sebaran (*dispersion atau scatter*) sebuah data. Ukuran sebaran atau variabilitas dalam tes psikologi yang paling banyak digunakan adalah *range*, rentang interkuartil, varian, dan deviasi standar.

Range adalah jarak antara dua angka ekstrim (terendah dan tertinggi) dalam sebuah distribusi. Meskipun sangat lemah, maka kita bisa mengatakan bahwa semakin panjang jaraknya maka dikatakan data itu semakin variatif. Untuk data diatas skor minimal atau terendah adalah 29 dan skor maksimal atau tertinggi adalah 50. Jadi range-nya adalah $50 - 29 = 21$.

Rentang interkuartil adalah jarak antara angka yang menjadi Q_1 atau C_{25} (titik 25% data dari urutan terkecil) dan Q_3 atau C_{75} (titik 75% data dari urutan yang terkecil). Jadi rentang interkuartil memiliki rentang 50% data di tengah-tengah. Untuk menunjukkan variasi, rentang interkuartil lebih baik daripada range. Bisa dikatakan bahwa semakin besar interkuartil maka semakin variatif sebuah data.

$$Q_1 = bb + \left(\frac{0,25 - \sum p_b}{p_i} \right) i$$

$$Q_2 = bb + \left(\frac{0,50 - \sum p_b}{p_i} \right) i$$

$$Q_3 = bb + \left(\frac{0,75 - \sum p_b}{p_i} \right) i$$

Q_1 = Kuartil Satu atau Persentil 25

Q_2 = Kuartil Dua atau Median atau Persentil 50

Q_3 = Kuartil Tiga atau Persentil 75

bb = Batas bawah interval dimana Q_1, Q_2, Q_3 berada

$\sum p_b$ = Jumlah proporsi dibawah interval dimana Q_1, Q_2, Q_3 berada

p_i = Proporsi interval dimana Q_1, Q_2, Q_3 berada

i = Rentang interval

Dengan data pada tabel 1 maka hasil dari Q_1, Q_2, Q_3 adalah

$$Q_1 = 36,5 + \left(\frac{0,25 - 0,23}{0,07} \right) 1 = 36,8$$

$$Q_2 = 39,5 + \left(\frac{0,50 - 0,45}{0,083} \right) 1 = 40,1 \text{ dibulatkan menjadi } 40.$$

$$Q_3 = 43,5 + \left(\frac{0,75 - 0,717}{0,05} \right) 1 = 44,16$$

$$\text{Rentang Interkuartil} = 44,16 - 36,8 = 7,36$$

Varian adalah jumlah kwadrat skor perbedaan atau deviasi antara setiap angka (X) dalam sebuah distribusi dengan mean sebuah distribusi (M) yang dibagi dengan N . Atau varian adalah rata-rata dari jumlah kwadrat (Sum of Square: SS). Deviasi harus dikwadratkan dulu untuk menghilangkan nilai negatif, karena jika langsung dijumlahkan hasilnya akan nol.

$$s^2 = \frac{\sum fx^2}{n-1} = \frac{1518,93}{59} = 25,7446$$

Deviasi standar adalah akar kwadrat dari varian. Deviasi adalah ukuran yang menunjukkan jarak perbedaan satu skor dengan skor yang lain. Misalnya jika deviasi standar sebuah distribusi data adalah 5, maka sebuah skor dengan skor lain jika memiliki perbedaan satu unit deviasi maka perbedaan keduanya adalah 5. Jika perbedaannya 0,3 unit maka perbedaannya adalah $0,3 \times 5 = 1,5$. Jika memiliki perbedaan 1,5 unit maka $1,5 \times 5 = 7,5$.

Penghitungan deviasi standar sampel dengan data tabel 1.

$$\sum fx^2 = 1518,93$$

$$s = \sqrt{\frac{\sum fx^2}{n-1}} = \sqrt{\frac{1518,93}{59}} = \sqrt{25,7446} = 5,07392$$

Dalam penghitungan varian dan deviasi standar harus diperhatikan apakah penghitungannya itu untuk populasi atau sampel. Jika untuk populasi maka rumus varian yang digunakan adalah $s = \sqrt{\frac{\sum fx^2}{n}}$, jika untuk sampel rumusnya adalah $s = \sqrt{\frac{\sum fx^2}{n-1}}$.

Manfaat Variabilitas

Variabilitas atau perbedaan adalah anugrah dari Tuhan, tanpa ada perbedaan maka hidup ini akan hampa. Tanpa perbedaan tidak akan ada pengetesan. Tes digunakan adalah untuk mengetahui perbedaan satu orang dengan orang yang lain dalam suatu atribut perilaku.

Distribusi Normal

Model kurva normal adalah model matematis yang merupakan sebuah distribusi yang jika digambarkan seperti sebuah bel. Pada sumbu X kurva ini menampilkan unit deviasi standar. Meskipun kurva normal ini adalah sebuah model, seringkali didasarkan kepada sebuah distribusi data yang sebenarnya. Diperkirakan jika peristiwa atau variabel

di dunia ini jika didistribusikan akan menghasilkan sebuah distribusi yang mirip dengan kurva normal.

Karakteristik model kurva normal yaitu:

1. Berbentuk seperti lonceng.
2. Dianggap simetris, artinya setengah bagiannya adalah sama (jika kita lipat kurva itu kedua bagiannya memiliki 50% bagian kurva).
3. Unimodal, memiliki satu nilai yang memiliki frekuensi tertinggi.
4. Mean, median, dan modus bersama-sama berada di tengah distribusi karena titik ini merupakan titik keseimbangan antara dua bagian kurva.

Tests of Normality

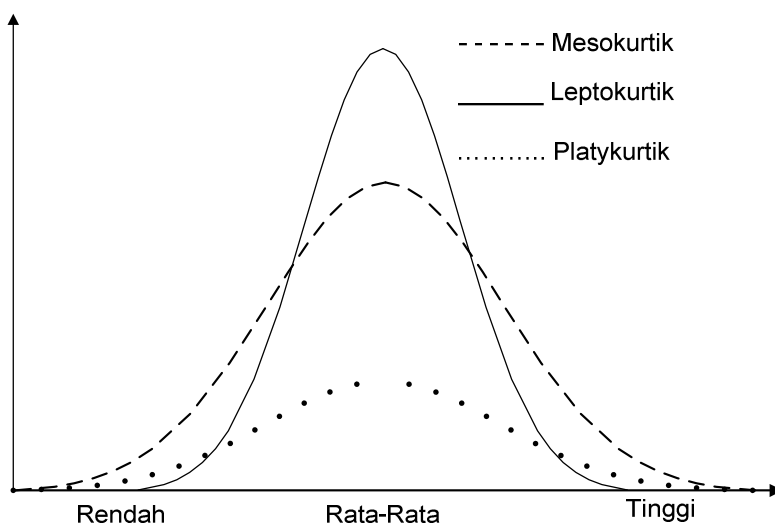
	Kolmogorov-Smirnov		
	Statistic	df	Sig.
Skor Tes Psikologi Umum	.074	22	.200

Dari data di atas didapat uji normalitas distribusi data dengan metode Kolmogorof-smirnof bahwa data tersebut distribusinya normal. Hal ini bisa dilihat bahwa probabilitasnya lebih besar dari 0,05. Jika kita bandingkan dengan nilai mean, median, dan modus yaitu 40,133, 40,00 dan 39 maka bisa kita lihat bahwa perbedaan antara ketiga nilai tersebut tidak jauh berbeda.

Distribusi Tidak Normal

Distribusi adakalanya tidak mengikuti pola hipotesis distribusi normal atau disebut tidak normal. Ada dua statistik untuk mengukur ketidaknormalan distribusi yaitu kurtosis dan skewness.

Kurtosis



Model kurva di atas adalah distribusi skor inteligensi laki-laki dan perempuan, dimana distribusi skor inteligensi perempuan bersifat *leptokurtik*, dan distribusi skor inteligensi laki-laki bersifat normal atau *mesokurtik*.

Kurtosis berkaitan secara langsung dengan tingkat penyebaran dalam sebuah distribusi. Distribusi *platykurtic* memiliki penyebaran yang paling tinggi, distribusi *leptokurtic* adalah distribusi yang tingkat penyebarannya paling rendah. Distribusi normal atau *mesokurtic* memiliki tingkat penyebaran sedang.

Jika α_4 adalah negative, maka bentuk distribusi adalah Platikurtik

Jika α_4 adalah positif, maka bentuk distribusi adalah Leptokurtik.

Jika α_4 adalah nol, maka bentuk distribusi adalah Mesokurtik atau normal.

Skewness

Skewness (*Sk*) adalah kemencengan sebuah distribusi atau cenderung tidak simetris atau normal. Jika sebuah data simetris secara sempurna atau distribusinya normal maka nilai skewnessnya adalah 0. Jika sebagian besar skor sebuah data itu berada pada level atas sebuah skala dan ekor panjangnya menuju ke nilai rendah, maka distribusinya *menceng negatif* ($Sk < 0$), di sisi lain, jika sebagian besar skor sebuah data berada pada level bawah dan ekornya memanjang ke skor atas maka data itu *menceng positif* ($Sk > 0$). Data yang menceng negatif nilai mediannya lebih tinggi daripada meannya sehingga $mean - median = \text{negatif}$. Data yang menceng positif, mediannya lebih kecil daripada meannya, sehingga $mean - median = \text{positif}$.

Jika diterapkan pada sebuah tes, maka kemencengan data bisa memberi informasi yang bermanfaat. Jika sebuah data menceng negatif maka tes itu level kesulitannya mudah, karena sebagian besar testee memiliki skor tinggi, jika distribusi skor tesnya menceng positif maka tes itu level kesulitannya sulit karena sebagian besar testee skornya rendah.

Tabel 4: Statistik Deskriptif Data pada Tabel 2.

N		60
Mean		40.1333
Std. Error of Mean		.65504
Median		40.0000
Mode		39.00
Std. Deviation		5.07392
Variance		25.74463
Skewness		-.163
Std. Error of Skewness		.309
Kurtosis		-.626
Std. Error of Kurtosis		.608
Range		21.00
Minimum		29.00
Maximum		50.00
Sum		2408.00

Percentiles	25	37.0000
	50	40.0000
	75	44.0000
	100	50.0000

Korelasi

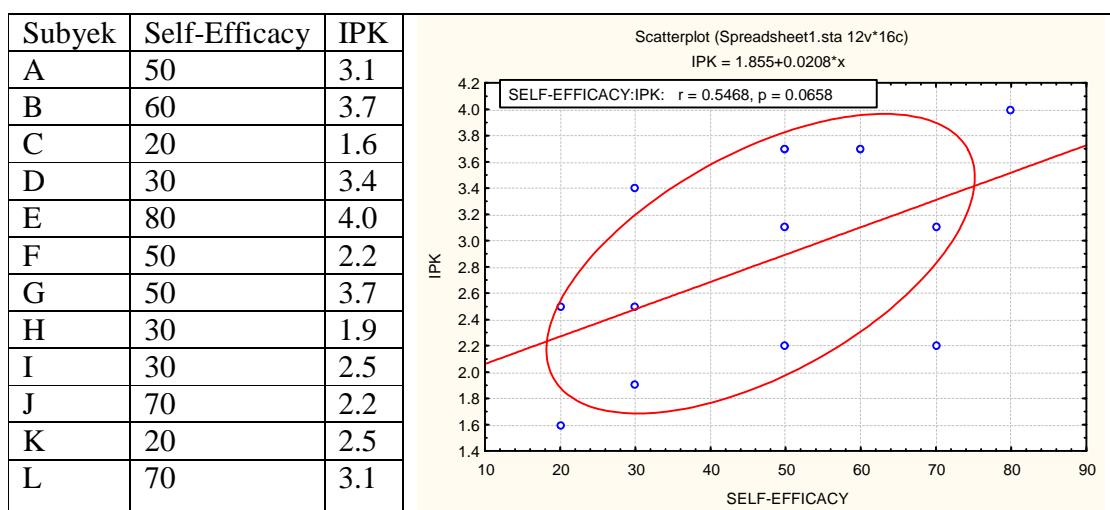
Cara yang paling sederhana untuk menggambarkan hubungan antar skor adalah menggunakan scatterplot. Scatterplot adalah grafik yang menunjukkan posisi sebuah kelompok orang di titik temu skor dua variabel. Sumbu X pada grafik digunakan untuk satu variabel dan sumbu Y digunakan untuk variabel kedua, dan skor setiap orang pada dua variabel itu ditampilkan dalam satu titik.

Dalam korelasi ada banyak jenis korelasi jika berkaitan dengan tes, yaitu (a) korelasi skor antar tes yang berbeda, (b) korelasi skor tes dengan variabel non-tes, (c) antara skor item tes dengan skor total tes, (d) antara skor item dengan variabel non-tes, (e) antar skor item dalam satu tes, dan (f) antar skor item beda tes.

Perlu juga diperhatikan adalah jenis variabel yang dikorelasikan. Ada dua jenis variabel dalam statistik yaitu variabel diskrit (nominal atau ordinal) dan variabel kontinyus (interval atau rasio).

Variable Y\X	Kontinyus X	Ordinal X	Nominal X
Kontinyus Y	Pearson r		
Ordinal Y	Biserial r_b	Spearman rho/Tetrachoric r_{tet}	
Nominal Y	Point Biserial r_{pb}	Rank Biserial r_{rb}	Kontingensi

Tabel 2.1: Skor Self-Efficacy dan IPK



Gambar 2.1: Hubungan antara Self-Efficacy dan IPK

Misalnya, data dalam Tabel 2.1 skor dua variabel yaitu Self-Efficacy dan IPK. Data dalam tabel itu menunjukkan skor 12 mahasiswa pada Self-Efficacy dan Indeks prestasi Kumulatif. Dalam bentuk tabel, hubungan antara dua variabel tidak terlihat. Gambar 2.1 memperlihatkan data yang sama dalam bentuk scatterplot. Seperti yang bisa dilihat dalam scatterplot, tampak bahwa IPK cenderung naik jika skor Self-efficacy naik. Namun, hubungan ini jauh dari sempurna, yaitu jelas bahwa setiap kenaikan atau tambahan skor tidak dibarengi oleh gradasi kenaikan skor.

Hubungan yang ditunjukkan dalam gambar 2.1 disebut korelasi positif karena peningkatan di satu variabel diikuti oleh peningkatan skor di variabel yang lain. Yaitu, semakin tinggi skor Self-efficacy semakin tinggi skor IPK. Meskipun sebagian besar korelasi yang dilakukan dalam banyak penelitian adalah positif, tapi ada korelasi negatif yang menunjukkan arah yang berlawanan. Misalnya korelasi usia dan kemampuan mengingat, semakin tua orang maka kemampuan mengingatnya semakin turun.

Alat korelasi yang sederhana dan paling umum adalah koefisien korelasi yang disimbolkan dengan r . Untuk menghitung r , kalikan setiap skor z subyek pada variabel X (self-Efficacy) dengan skor z IPK-nya dan hitunglah rata-rata hasil perkalian itu. Dengan kata lain:

$$r = \frac{\sum(z_x \times z_y)}{N}$$

Sekali lagi, statistik ini adalah rata-rata sederhana. Pertanyaannya mengapa rata-rata z_x dikalikan z_y menjadi indeks korelasi antara X dan Y .

Ketika dua variabel tidak berkorelasi, pola perubahan skor dua variabel tersebut adalah sebagai berikut: ada skor X tinggi skor Y ikut tinggi ($z_x \times z_y$, hasilnya tinggi dan positif), ada skor X tinggi skor Y sedang-sedang saja (z_y mendekati nol, jadi $z_x \times z_y$ hasilnya hampir nol), ada skor X tinggi skor Y rendah ($z_x \times z_y$ hasilnya tinggi dan negatif). Jadi ketika dirata-ratakan hasilnya 0.

Prediksi

Salah satu kegunaan korelasi adalah prediksi. Jika skor pada satu variabel, seperti kinerja, dikorelasikan dengan skor lainnya seperti tes pemahaman mekanik, implikasinya kita dapat memprediksi kinerja seseorang dari skor-skor tes. Banyak keputusan yang kita buat atas seseorang berdasarkan kepada (setidaknya secara implisit) kepada prediksi-prediksi. Perguruan tinggi menerima siswa sebagian besar diinginkan berhasil. Dunia industri menerima orang yang diprediksi bekerja dengan baik dan tidak memilih mereka yang diprediksi gagal. Psikolog klinis sering mendasarkan treatment mereka kepada hasil-hasil prediktif dari berbagai terapi. Oleh karena itu, topik prediktif adalah salah satu hal praktis yang sangat penting untuk dibahas.

Persoalan prediksi dipenuhi dengan kenyataan bahwa prediktor (X) dan variabel yang ingin diprediksi (Y) sering memiliki skala yang berbeda. Misalnya, Tes Potensi Akademik, yang rentang skornya 200 sampai 800, sering digunakan untuk memprediksi IPK, yang memiliki rentang 0,0 sampai 4,0. teknik yang digunakan untuk membuat prediksi oleh karena itu, harus menghitung dua skala itu X dan Y dan sejauhmana X dan Y berkorelasi.

Skor yang diprediksi yaitu Y, berdasarkan kepada skor orang pada X, rumusnya adalah:

$$\hat{Y} = a + bX$$

Dimana:

\hat{Y} : Skor yang diprediksikan pada Y

a : Intercept

b : Slope atau koefisien regresi

X : Skor pada prediktor

Dengan kata lain, skor prediksi Y adalah sama dengan skor pada X yang dikalikan dengan b , plus konstanta, a . Konstanta (intercept) membuat kita mampu menyesuaikan untuk berbagai skala pengukuran untuk X dan Y. Koefisien regresi, b , menunjukkan perubahan yang diharapkan dalam Y per unit perubahan dalam X. Misalnya, persamaan yang ditunjukkan di bawah ini dan digambarkan di Gambar 2.2:

IPK prediktif = $(0,5) + 0,005 \times \text{TPA-verbal}$.

Seorang mahasiswa dengan skor TPA 550 diprediksi akan memiliki IPK 3,25. Seorang mahasiswa dengan skor 650 diprediksi akan memperoleh IPK 3,75, yaitu setengah unit ($100 \times 0,5$) lebih tinggi dari skor sebelumnya.

Persamaan regresi adalah persamaan linier yang memprediksi bahwa setiap peningkatan satu unit pada X akan menghasilkan peningkatan b unit pada Y.

Memaknai Korelasi: Analisis Faktor

Tes mental sering menunjukkan interkorelasi sedang sampai tinggi. Analisis faktor mewakili aplikasi metode statistik lanjut untuk masalah menjelaskan mengapa dua tes dikorelasikan. Misalkan, dua tes yang didesain untuk mengukur pemahaman bacaan dan vokabulari yang menunjukkan korelasi 0,53. Jika seseorang bertanya mengapa tes ini dikorelasikan, jawabannya mungkin adalah bahwa korelasi keduanya dijeaskan oleh fakta bahwa mereka mengukur banyak hal yang sama – yaitu kemampuan verbal. Penjelasan ini mewakili aplikasi penalaran induktif: mengidentifikasi variabel yang umum, abstrak, dari kemampuan verbal dengan memberi gambaran bahwa dua tes itu memiliki kesamaan. Analisis faktor adalah metode statistik untuk memperoleh hasil yang sama. Analisis faktor berusaha, mengetahui variabel dasar yang dihitung dengan cara mengorelasikan skor tes yang ada. Variabel dasar ini disebut *faktor*, dua tes bisa memiliki faktor yang sama secara umum, dan mungkin saja beberapa faktor akan muncul dalam sekelompok tes melalui interkorelasi.

Tabel 2.2 menunjukkan interkorelasi antar empat tes, reading comprehension test (RC), vocabulary test (Vocab), figure rotation test (FR), dan exploded figure test (EF). Dua tes yang awal memiliki korelasi yang tinggi karena memang keduanya mengukur kemampuan yang sama yaitu kemampuan verbal, sedangkan dua tes yang terakhir juga berkorelasi kuat karena mengukur kemampuan spasial.

Tabel 2.2 Korelasi antara Empat Tes

	RC	VOCAB	FR	EF
Reading Comprehension (RC)	1,0			
Vocabulary (Vocab)	0,62	1,0		
Figure Rotation (FR)	0,12	0,09	1,0	
Exploded Figure (EF)	0,04	0,11	0,76	1,0

Jadi ketika dianalisis faktor akan memunculkan dua faktor yaitu faktor verbal dan faktor spasial. Hasilnya bisa dilihat ditabel 2.3

	Faktor I	Faktor II
Reading Comprehension (RC)	0,88	0,09
Vocabulary (Vocab)	0,76	0,15
Figure Rotation (FR)	0,04	0,72
Exploded Figure (EF)	0,20	0,78

Rumus-rumus korelasi:

Korelasi poin biserial adalah korelasi antara data nominal dan interval. Korelasi ini biasa digunakan untuk analisis item dengan jawaban dikotomis, ya/idak, benar/salah. Selain korelasi point biserial, korelasi biserial bisa juga digunakan untuk analisis item. Korelasi biserial pada dasarnya sama dengan korelasi point-biserial tapi hasilnya korelasi biserial selalu lebih daripada korelas point-biserial. Berikut adalah rumus-rumus point-biserial dan biserial.

$$\text{Korelasi point-biserial: } r_{pb} = \frac{M_1 - M_0}{s_x} \sqrt{\frac{pq}{n(n-1)}}$$

$$\text{Korelasi biserial: } r_b = (X_1 - X_0) \left(\frac{pq}{\sigma_X} \right)$$

Korelasi juga bisa dilakukan pada variabel yang skalanya nominal. Misalnya korelasi antara jenis kelamin dan focus coping. Disini jenis kelamin terdiri dari laki-laki dan perempuan dan focus coping terdiri dari emosional dan problem solving. Disini hanya akan dilakukan penghitungan berapa orang yang masuk jenis kelamin laki-laki dan perempuan. Diantara yang perempuan berapa orang yang masuk focus coping emosional

dan problem solving dan diantara laki-laki berapa yang masuk focus coping emosional dan problem solving.

Data korelasi ini harus dimasukkan kedalam tabel kontingensi, dan dengan penghitungan

phi: berikut rumus $\phi = (BC-AD) / \sqrt{((A+B) \times (C+D) \times (A+C) \times (B+D))}$

Y\X	0	1	Totals
1	A	B	A + B
0	C	D	C + D
Totals:	A + C	B + D	N

FC\JK	PEREMPUAN (0)	LAKI-LAKI (1)	Total
EMOSIONAL	10	5	15
PROBLEM SOLVING	5	10	15
Total:	15	15	30

Hasilnya: $\phi = (25-100) / \sqrt{(15 \times 15 \times 15 \times 15)} = -75/225 = -0.33$