



STiPRAM
Sekolah Tinggi Pariwisata Ambarukmo
Y O G Y A K A R T A

TOURISM STATISTIC PART 2 (PERTEMUAN 9-15)

Oleh: Dyah Wahyuning Tyas, S.E., M.Ec.Dev.

Sekolah Tinggi Pariwisata Ambarukmo
Yogyakarta

ANALISIS REGRESI (SEDERHANA DAN BERGANDA)

Konsep Regresi

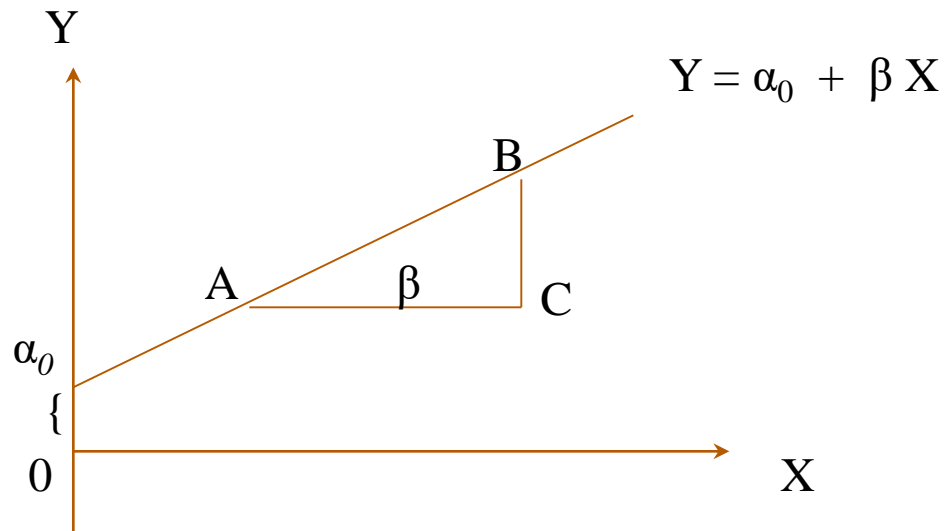
- ❑ Regresi adalah suatu metode yang digunakan untuk mengetahui atau mengukur pola hubungan (hubungan fungsional) antarvariabel.
- ❑ Persamaan umum untuk garis regresi linear sederhana dapat ditulis sebagai berikut:

$$\bar{Y} = \alpha_0 + \beta_1 X_t + \varepsilon_t$$

di mana: \bar{Y} = variabel dependent (terikat)
 X_t = variabel independent (bebas)
 α_0 dan β_1 = koefisien regresi
 ε_t = *error term*.

Garis regresi linear dibedakan menjadi dua, yaitu garis linear sederhana dan garis linear berganda. Yang membedakan terletak pada jumlah variabel X (variabel bebas/independen) yang digunakan dalam persamaan regresi. Jika dalam persamaan regresi hanya terdiri dari satu variabel dependen/terikat dan satu variabel independen/bebas, maka disebut persamaan regresi sederhana. Sedangkan jika dalam persamaan regresi terdiri dari satu variabel dependen/terikat dan lebih dari satu variabel independen/bebas, maka disebut persamaan regresi berganda (*multiple regression*).

Grafik persamaan linear digambarkan sebagai berikut:



Nilai α_0 sama dengan ordinat titik potong antara garis regresi dengan sumbu Y, atau dapat disebut juga *intercept*. Sedangkan nilai β sama besarnya dengan sudut yang dibentuk oleh garis regresi dengan sumbu horisontal (sumbu X), atau merupakan kemiringan (*slope*) dari suatu garis regresi. Dengan demikian sama besarnya dengan *tangens* dari sudut yang dibentuk oleh garis regresi dengan sumbu mendatar (sumbu X), di mana harga dari :

$$\text{tg } \alpha = \frac{BC}{AC}$$

Rumus untuk menentukan nilai α_0 (intercept/constant) dan β (slope) dengan metode kuadrat terkecil atau minimum (*least square methods*)

Metode *least square* menggunakan dasar perhitungan BLUE (*Best Line Unbiased Estimate*). Dengan menggunakan metode *least square*, maka harga-harga koefisien regresi α dan β dari persamaan regresi $Y = \alpha + \beta X$ dapat dihitung melalui persamaan berikut ini:

$$\alpha_0 = \frac{\sum x^2 \cdot \sum y - \sum x \cdot \sum xy}{n \sum x^2 - (\sum x)^2}$$

$$\beta = \frac{n \sum xy - \sum x \cdot \sum y}{n \sum x^2 - (\sum x)^2}$$

Metode kuadrat terkecil (*least square*) adalah penentuan persamaan regresi dengan meminimumkan jumlah kuadrat jarak vertikal nilai Y pengamatan dan nilai Y prediksi (Kusmayadi, 2004).

Contoh soal regresi linear sederhana

No.	Volume Penjualan Paket Wisata (ribuan unit)	Laba Bersih (puluhan juta rupiah)
1.	2,5	6,0
2.	2,6	5,9
3.	2,8	6,2
4.	2,9	6,4
5.	3,0	6,9
6.	3,5	7,2
7.	3,4	7,0
8.	4,0	7,5
9.	3,4	6,8
10.	3,2	6,9
11.	3,3	6,7
12.	3,8	6,9

Pertanyaan:

1. Tentukan persamaan regresi linearnya!
2. Jika volume penjualan paket wisata di Biro Perjalanan Wisata tersebut mencapai 4.200 unit, berapakah perkiraan laba bersih yang diterima Biro Perjalanan Wisata tersebut?

Penyelesaian:

$$\alpha = \frac{\sum x^2 \cdot \sum y - \sum x \cdot \sum xy}{n \sum x^2 - (\sum x)^2} = \frac{125,20 \cdot (80,40) - 38,40 \cdot (259,49)}{12 \cdot (125,20) - 1474,56}$$

$$\alpha = \frac{10066,08 - 9964,416}{1502,4 - 1474,56} = \frac{101,664}{27,84} = 3,65172 = 3,65$$

$$\beta = \frac{n \sum xy \cdot \sum y - \sum x \cdot \sum y^2}{n \sum x^2 - (\sum x)^2} = \frac{12 \cdot (259,49) - (38,40) \cdot (80,40)}{12 \cdot (125,20) - 1474,56}$$

$$\beta = \frac{3113,88 - 3087,36}{1502,4 - 1474,56} = \frac{26,52}{27,84} = 0,9526 = 0,95$$

1. Persamaan regresi linear sederhana antara volume penjualan paket wisata terhadap laba bersih adalah:

$$Y = \alpha + \beta X$$

$$Y = 3,65 + 0,95 X$$

2. Jika volume penjualan mencapai 4.200 unit, maka laba bersih yang diterima adalah:

$$Y = 3,65 + 0,95 (4.200)$$

$$Y = 3,65 + 3990$$

$$Y = 3.993,65$$

Jadi, laba bersih yang diterima sebesar 3.993,65 puluhan juta rupiah.

Interpretasi Analisis Regresi

Setelah model persamaan regresi (baik regresi sederhana maupun berganda) di estimasi menggunakan beberapa perangkat alat analisis statistik, seperti Ms. Excel dan SPSS (untuk mempercepat perhitungan) maka langkah selanjutnya menginterpretasikan hasil estimasi regresi tersebut dengan 3 analisis pengujian statistik yang meliputi:

- Uji *t-statistic*
- Uji *F-statistic*
- Analisis Koefisien Determinasi (*R-square*)

Ketiga analisis statistik tersebut digunakan untuk menguji tingkat signifikansi dari data empiris yang telah diestimasi, sehingga pada akhirnya akan diperoleh pengambilan keputusan atas dasar hipotesis (dugaan sementara), baik secara statistik dan penelitian.

Kriteria pengambilan keputusan untuk analisis uji statistik ditentukan sesuai dengan hipotesis statistik.

Uji Statistik

1. Uji *t*-statistic

t-statistic digunakan untuk menguji hubungan atau pengaruh antara satu variabel independen terhadap variabel dependen secara parsial (individu).

2. Uji *F*-statistic

F-statistic digunakan untuk menguji hubungan atau pengaruh semua variabel independen secara bersama-sama terhadap variabel dependen.

3. Analisis Koefisien Determinasi (*R*-square)

Koefisien determinasi digunakan untuk melihat seberapa besar kontribusi variabel independen/bebas (*X*) dalam menentukan perubahan variabel tidak bebas (*Y*)

KRITERIA PENGUJIAN DALAM PENGAMBILAN KEPUTUSAN

Statistik Parametrik

1. Uji *t-Statistic* (distribusi *t-student*)

(dalam hipotesis asosiatif untuk menentukan korelasi ganda)

untuk populasi berdistribusi normal, maka nilai t distribusi student dengan $df^* = (n - 1)$.

*Keterangan: df = degree of freedom (derajat kebebasan)

Kriteria pengujian (uji signifikansi) berdasarkan nilai kritis (*critical value*) atau tabel distribusi t adalah sebagai berikut:

Tipe Hipotesis	H_0 : hipotesis nol	H_a : hipotesis alternatif	Keputusan : menolak H_0 , jika
dua sisi (<i>two-tail</i>)	$\beta_2 = \beta_2^*$	$\beta_2 \neq \beta_2^*$	$ t > t_{1/2\alpha, df}$
sisi kanan (<i>right-tail</i>)	$\beta_2 \leq \beta_2^*$	$\beta_2 > \beta_2^*$	$t > t_{\alpha, df}$
sisi kiri (<i>left-tail</i>)	$\beta_2 \geq \beta_2^*$	$\beta_2 < \beta_2^*$	$t < -t_{\alpha, df}$

Note: β_2 = nilai koefisien parameter masing-masing variabel independen; β_2^* = nilai koefisien parameter yang dihipotesiskan masing-masing variabel independen; β = rata-rata hitung parameter atau dinotasikan dengan μ ; t = nilai t -hitung atau *t-statistic*; t_{α} atau $t_{1/2\alpha}$ = nilai kritis/tabel distribusi *t-student*./ $|t|$ = nilai absolut dari *t-statistic* (nilai negatif diabaikan, dianggap positif).

Uji *t*-Statistic (distribusi *t*-student)...(2)

a. Uji Dua Pihak (*two tail*)

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$

- Jika nilai *t*-tabel (*t-critical value*; $\frac{1}{2} \alpha$) < **|nilai *t*-statistic|** , maka H_0 ditolak atau H_a diterima.
- Jika nilai *t*-tabel (*t-critical value*; $\frac{1}{2} \alpha$) > **|nilai *t*-statistic|** , maka H_0 diterima atau H_a ditolak.

Keterangan:

|nilai *t*-statistic| merupakan nilai absolut dari nilai *t*-hitung atau *t*-statistic (artinya mengabaikan

tanda negatif, di mana jika nilai *t*-statistic negatif maka dianggap bernilai positif)

b. Uji satu pihak atau *one-tail* (pihak kiri)

$$H_0: \mu \geq \mu_0$$

$$H_a: \mu < \mu_0$$

Kriteria pengujian berdasarkan nilai kritis (*critical value*) atau tabel distribusi *t* adalah sebagai berikut:

- Jika $t_{\text{statistic}} < -t_{\text{kritis}}$ dengan harga *t* diperoleh dari tabel distribusi *t* student pada taraf nyata/signifikansi tertentu (α) dan $df = n-1$, maka **H_0 ditolak**.
- Jika $t_{\text{statistic}} > -t_{\text{kritis}}$ dengan harga *t* diperoleh dari tabel distribusi *t* student pada taraf nyata/signifikansi tertentu (α) dan $df = n-1$, maka **H_0 diterima**.

Uji *t*-Statistic (distribusi *t*-student)...(3)

c. Uji satu pihak atau *one-tail* (pihak kanan)

$$H_0: \mu \leq \mu_0$$

$$H_a: \mu > \mu_0$$

Kriteria pengujian berdasarkan nilai kritis (*critical value*) atau tabel distribusi *t* adalah sebagai berikut:

- Jika $t_{\text{statistic}} > t_{\text{kritis}}$ dengan harga *t* diperoleh dari tabel distribusi *t* student pada taraf nyata/signifikansi tertentu (α) dan $df = n-1$, maka **H_0 ditolak**.
- Jika $t_{\text{statistic}} < t_{\text{kritis}}$ dengan harga *t* diperoleh dari tabel distribusi *t* student pada taraf nyata/signifikansi tertentu (α) dan $df = n-1$, maka **H_0 diterima**.

Uji *t-Statistic*

(kriteria pengujian berdasar nilai probabilitas atau sig. t-stat)

Selain dengan menggunakan kriteria pengujian perbandingan antara nilai *t-statistic* dengan nilai *distribusi t-student*, maka dapat juga dilakukan dengan perbandingan nilai probabilitas (sig. t-stat.) dengan taraf signifikansi (baik 1%, 5%, maupun 10% yang disesuaikan dengan pilihan pengujian *two-tail* atau *one-tail*). Ketentuannya adalah sebagai berikut:

- Jika nilai probabilitas *t-statistic* < taraf signifikansi tertentu, maka H_0 ditolak atau H_a diterima dan berarti signifikan.
- Jika nilai probabilitas *t-statistic* > taraf signifikansi tertentu, maka H_0 diterima atau H_a ditolak dan berarti tidak signifikan.

Contoh:

Nilai sig.t-stat = 0,006 pada Confidence level = 95%, taraf signifikan 5 % = 0,05
Hal ini menunjukkan nilai sig.t-stat = 0,006 < 0,05 (taraf signifikan 5%), artinya signifikan dan H_a diterima (sesuai dengan teori atau hipotesis penelitian)

2. Uji *F-Statistic*

(dalam hipotesis asosiatif untuk menentukan korelasi ganda)

Untuk uji signifikansi korelasi ganda dapat dilakukan melalui uji F, dengan persamaan berikut:

$$F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)}$$

Cara untuk menentukan nilai kritis F dalam tabel distribusi F :

Df numerator (pembilang) = k

di mana k = variabel independen (bebas)

Df denominator (penyebut) = (n-k-1)

Kriteria pengujiannya berdasarkan nilai kritis F adalah :

- Jika $F_{\text{statistic}} > F_{\text{kritis}}$ pada taraf signifikansi tertentu, maka diputuskan H_0 ditolak atau H_a diterima.
- Jika $F_{\text{statistic}} < F_{\text{kritis}}$ pada taraf signifikansi tertentu, maka diputuskan H_0 diterima atau H_a ditolak.

Kriteria pengujiannya berdasarkan nilai probabilitas *F-statistic* atau *significane F* adalah :

- Jika nilai probabilitas $F_{\text{statistic}} <$ taraf signifikansi tertentu (pada umumnya 5%), maka diputuskan H_0 ditolak atau H_a diterima dan berarti signifikan ada pengaruh.
- Jika nilai probabilitas $F_{\text{statistic}} >$ taraf signifikansi tertentu (pada umumnya 5%), maka diputuskan H_0 diterima atau H_a ditolak dan berarti tidak signifikan atau tidak ada pengaruh.

3. Analisis Koefisien Determinasi

Koefisien determinasi (R^2) bagian dari keragaman total variabel tidak bebas Y yang dapat diterangkan atau diperhitungkan keragaman variabel bebas X (Kusmayadi, 2004)

$$\text{Koefisien determinasi} = r^2 \times 100\%$$

Koefisien determinasi dinyatakan dalam persentase yang menunjukkan besarnya pengaruh perubahan variabel bebas X terhadap perubahan variabel tidak bebas Y.

SOAL DATA 1

Tahun	Jumlah Wisman	Tingkat Penghunian Kamar (TPK)	
		Hotel Bintang	Hotel Non Bintang
1985	749351	46.40	28.40
1986	825035	47.30	28.30
1987	1060347	48.40	29.60
1988	1301049	53.80	31.60
1989	1625965	55.70	31.20
1990	2177566	54.97	34.13
1991	2569870	54.20	33.17
1992	3064161	51.39	33.58
1993	3403138	51.15	30.33
1994	4006312	50.52	33.66
1995	4324229	47.98	31.80
1996	5034472	49.06	31.47
1997	5185243	47.02	30.86
1998	4606416	38.13	29.03
1999	4727520	42.22	30.98
2000	5064217	43.23	31.84
2001	5153620	44.79	31.01
2002	5033400	44.28	30.57
2003	4467021	45.03	29.88
2004	5321165	44.98	28.33
2005	5002101	45.03	28.86
2006	4871351	46.19	29.80
2007	5505759	46.89	32.44
2008	6234497	48.06	34.65
2009	6323730	48.31	35.56
2010	7002944	48.86	35.98
2011	7649731	51.25	38.74
2012	8044462	51.55	38.22
2013	8802129	52.22	37.34
2014	9435411	52.56	35.87

Source: BPS

Pertanyaan:

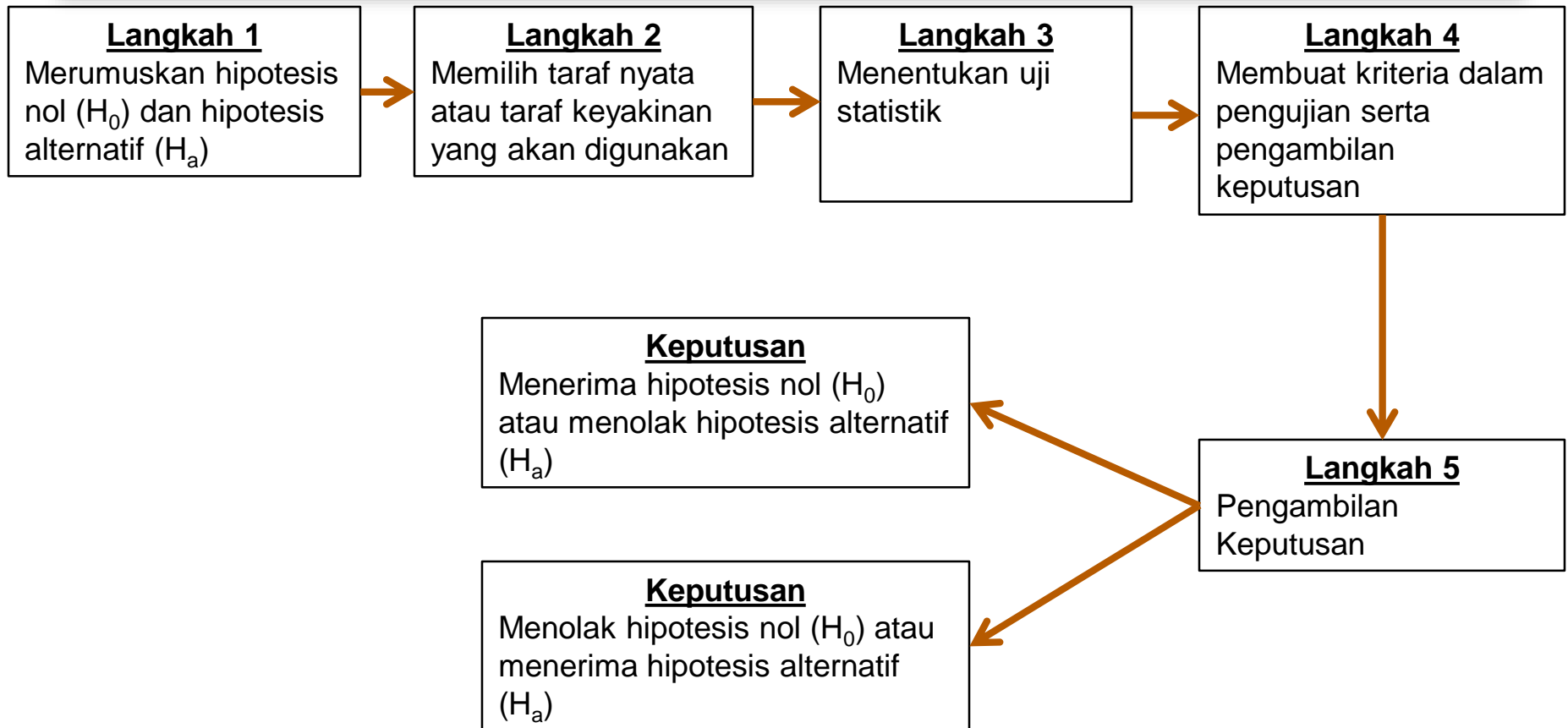
1. Berapa nilai koefisien parameter α (alpha) dan β (beta) pada persamaan regresi linear berikut:
 - a. $\text{TPK Hotel Bintang} = \alpha + \beta \text{ Jumlah wisatawan}$
 - b. $\text{TPK Hotel Non Bintang} = \alpha + \beta \text{ Jumlah wisatawan}$
2. Tentukan nilai korelasi (r) dari masing-masing persamaan regresi linear tersebut di atas!

PENGUJIAN HIPOTESIS

Definisi Hipotesis

- Hipotesis (dalam statistika) diartikan sebagai suatu pernyataan mengenai harga suatu parameter populasi yang dikembangkan untuk maksud pengujian.
- Hipotesis(dalam penelitian) diartikan sebagai dugaan sementara yang kebenarannya perlu dibuktikan berdasarkan data empiris.

Langkah-langkah Pengujian Hipotesis



Taraf nyata atau taraf signifikansi adalah probabilitas atau peluang untuk menolak hipotesis nol apabila hipotesis nol tersebut adalah benar. Taraf nyata yang dapat digunakan adalah 0,10 (taraf nyata 10%); 0,05 (taraf nyata 5%); 0,01 (taraf nyata 1%). Taraf nyata ini disebut sebagai tingkat risiko, yang menggambarkan besaran risiko yang ditanggung untuk menolak hipotesis nol padahal hipotesis nol tersebut benar.

Tipe Kekeliruan yang Mungkin Terjadi Dalam Membuat Kesimpulan Tentang Hipotesis

KESIMPULAN	KEADAAN SEBENARNYA	
	HIPOTESIS BENAR	HIPOTESIS SALAH
Tidak Ditolak Hipotesis	BENAR	KELIRU (Kekeliruan Tipe II)
Tolak Hipotesis	KELIRU (Kekeliruan Tipe I)	BENAR

Pengujian hipotesis, ada dua macam kekeliruan yang dapat terjadi yaitu:

1. Kekeliruan tipe I : yaitu kita menolak yang seharusnya diterima
2. Kekeliruan tipe II : yaitu kita menerima hipotesis yang seharusnya ditolak.

Perbedaan Statistik Parametrik dan Statistik Nonparametrik

STATISTIK PARAMETRIK

- Pengujian data dihadapkan pada populasi yang berdistribusi normal atau diasumsikan distribusi normal
- Jenis datanya: data interval dan data ratio

STATISTIK NONPARAMETRIK

- Pengujian data tidak mengikuti distribusi normal
- Jenis datanya : data nominal dan data ordinal

Beberapa Bentuk Rumusan Suatu Hipotesis

1. Hipotesis Deskriptif
2. Hipotesis Hubungan (Asosiatif)
3. Hipotesis Komparatif

1. Hipotesis Deskriptif

Hipotesis ini merupakan hipotesis atau dugaan mengenai harga atau nilai suatu variabel mandiri, tidak membuat perbandingan atau hubungan.

Tabel Uji Statistik yang Digunakan Dalam Menguji Hipotesis Deskriptif (satu sampel)

No.	Jenis Data	Statistik Uji
1.	Nominal	1. Tes Binomial 2. Chi kuadrat (satu sampel)
2.	Ordinal	Run Test
3.	Interval/Ratio	T-test (satu sampel)

Sumber: Sugiyono, Statistik untuk Penelitian, 1999

2. Hipotesis Hubungan (Asosiatif)

Adalah suatu pertanyaan yang menunjukkan dugaan tentang adanya hubungan antara dua variabel atau lebih.

- Hipotesis nol (H_0): tidak ada **hubungan** antara motivasi kerja karyawan bagian tiketing dengan kualitas pelayanan
- Hipotesis alternatif (H_a): ada **hubungan** antara motivasi kerja karyawan bagian tiketing dengan kualitas pelayanan

Hipotesis statistik:

- H_0 : $r = 0$
- H_a : $r \neq 0$

3. Hipotesis Komparatif

Adalah pernyataan yang menggambarkan dugaan terhadap suatu harga dalam satu variabel atau lebih pada sampel yang berbeda.

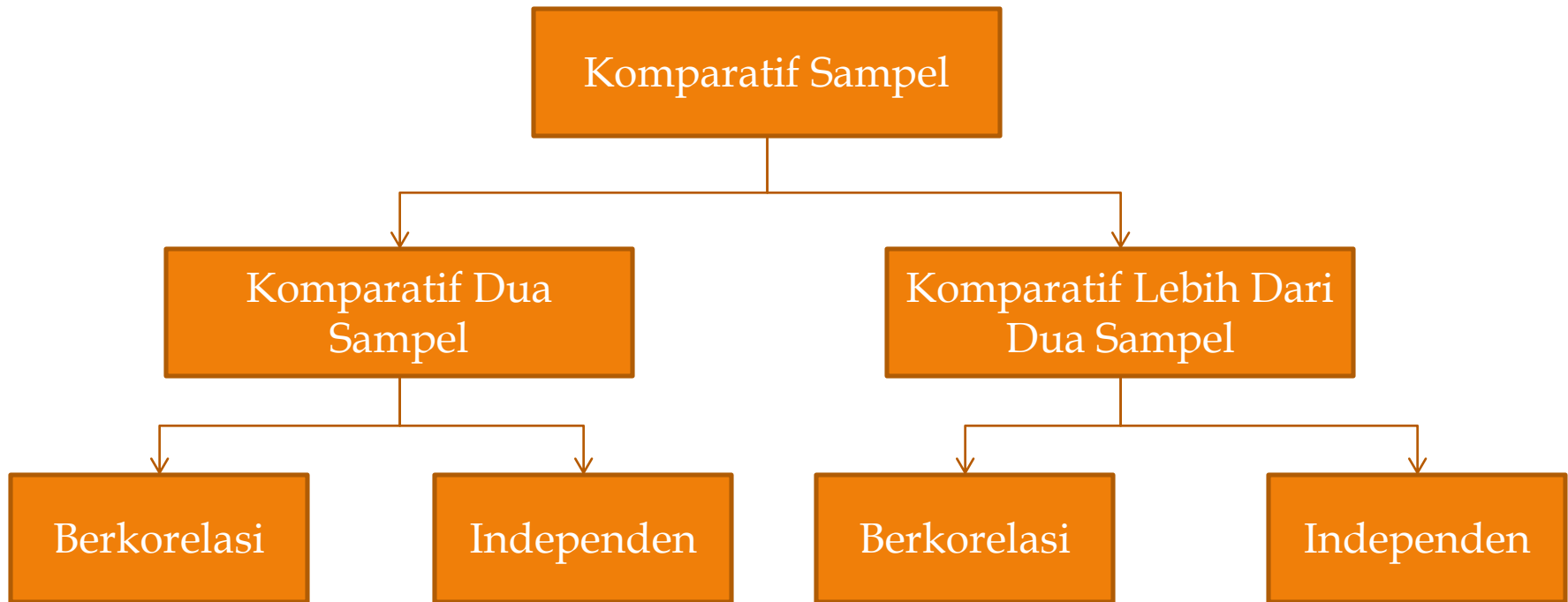
Contoh rumusan masalah: adakah perbedaan antara motivasi kerja karyawan di bagian tiketing dan di bagian tour?

- Hipotesis nol (H_0): tidak terdapat **perbedaan** motivasi kerja karyawan bagian tiketing dan bagian tour
- Hipotesis alternatif (H_a): ada **perbedaan** antara motivasi kerja karyawan bagian tiketing dan bagian tour

Hipotesis statistik:

Uji Dua Pihak	Uji Pihak Kiri	Uji Pihak Kanan
$H_0: \mu_1 = \mu_2$	$H_0: \mu_1 \geq \mu_2$	$H_0: \mu_1 \leq \mu_2$
$H_a: \mu_1 \neq \mu_2$	$H_a: \mu_1 < \mu_2$	$H_a: \mu_1 > \mu_2$

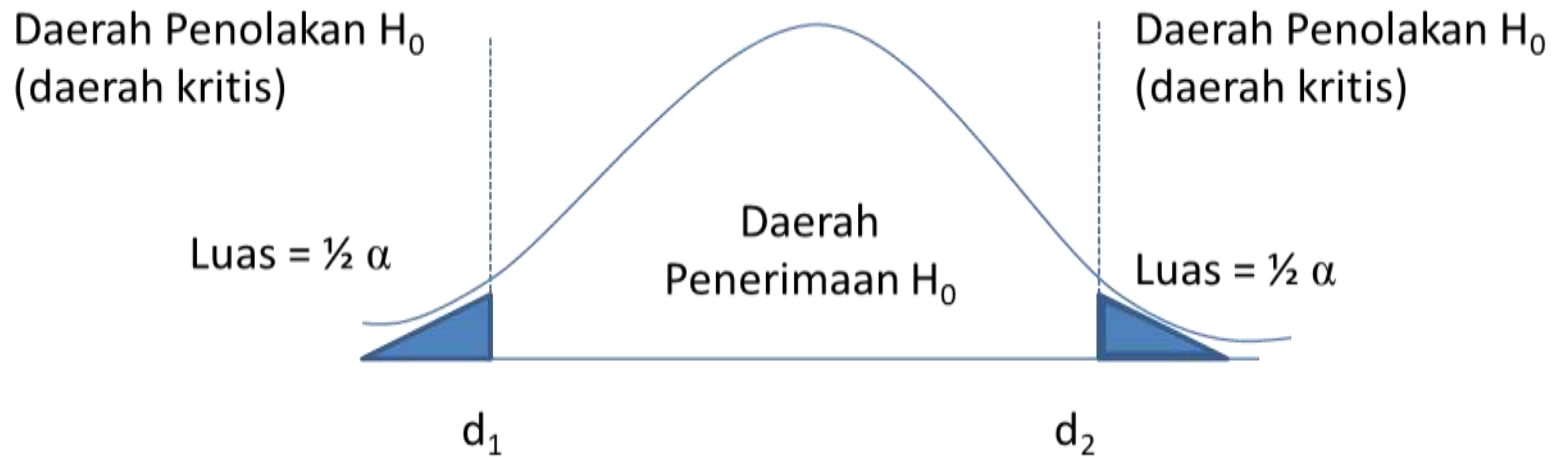
Model Komparatif Sampel



- Komparatif sampel yang berkorelasi lebih sering digunakan dalam penelitian-penelitian eksperimen. Contohnya: perbandingan kemampuan karyawan sebelum dan setelah mengikuti pelatihan (*treatment*) dengan membandingkan nilai pre-test dan pasca test.
- Komparatif sampel yang independen, contohnya perbandingan kualitas pelayanan dari dua buah restoran, dua buah hotel, dan sebagainya.

Peran Hipotesis Alternatif atau Hipotesis Kerja dalam penentuan daerah kritis

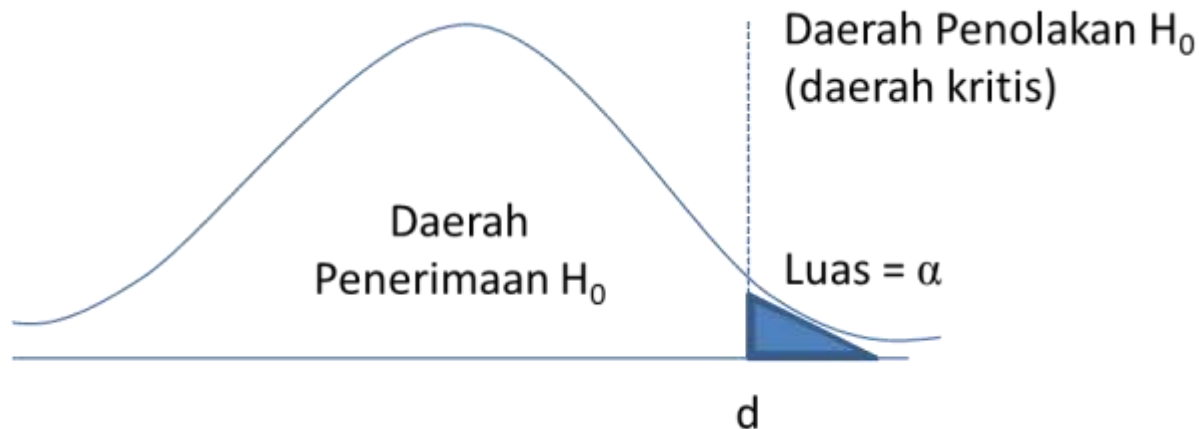
1. Apabila hipotesis alternatif (H_a) mempunyai perumusan *tidak sama* (\neq), maka dalam distribusi statistik yang digunakan, normal untuk angka Z , *Student* untuk t di dapat dua daerah kritis masing-masing pada ujung-ujung distribusi. Luas daerah kritis atau daerah penolakan hipotesis pada setiap ujung adalah $\frac{1}{2} \alpha$, oleh karena adanya dua daerah penolakan ini maka pengujian hipotesis dinamakan uji dua pihak.



Kriteria yang didapat adalah: terima H_0 jika harga statistik yang dihitung berada di daerah antara d_1 dan d_2 , dalam hal lainnya H_0 ditolak.

Peran Hipotesis Alternatif atau Hipotesis Kerja dalam penentuan daerah kritis (2)

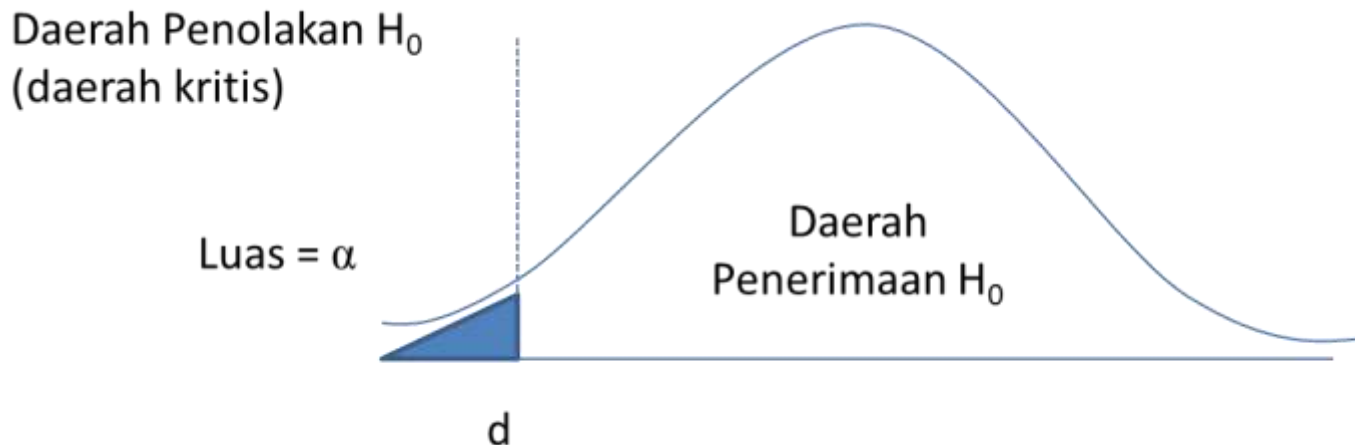
2. Untuk hipotesis alternatif (H_a) mempunyai perumusan lebih besar ($>$), maka dalam distribusi statistik yang digunakan, diperoleh sebuah daerah kritis yang letaknya ujung sebelah kanan. Luas daerah kritis atau daerah penolakan ini adalah sama dengan α .



Kriteria yang didapat adalah: tolak H_0 jika harga statistik yang dihitung berdasarkan sampel yang diambil tidak kurang dari "d", dalam hal lainnya terima H_0 . Pengujian ini disebut **uji satu pihak** atau **uji pihak kanan**

Peran Hipotesis Alternatif atau Hipotesis Kerja dalam penentuan daerah kritis (3)

3. Untuk hipotesis alternatif (H_a) mempunyai perumusan lebih kecil ($<$), maka daerah kritis atau daerah penolakan H_0 ada di ujung sebelah kiri dari distribusi yang digunakan. Luas daerah ini = α . Yang menjadi batas daerah penerimaan H_0 oleh "d" yang nilainya diperoleh dari daftar distribusi yang bersangkutan. Peluang untuk mendapatkan "d" ditentukan oleh taraf nyata α .



Kriteria yang digunakan: terima H_0 jika harga statistik yang dihitung berdasarkan penelitian lebih besar dari "d", sedangkan dalam hal lainnya H_0 ditolak. Pengujian ini disebut **uji satu pihak** atau **uji pihak kiri**.

ESTIMASI PARAMETER

- **Fungsi Estimasi Statistik Inferensial**

Statistik inferensial digunakan untuk menaksir populasi berdasarkan sampel, misalnya berupa data deskriptif seperti rata-rata hitung (mean), simpangan baku, modus, median, dan lain-lain. Data deskriptif yang dihitung dari sampel biasa disebut data statistik (atau statistik), sedangkan data deskriptif yang dihitung dari populasi biasa disebut sebagai parameter. Keakuratan penaksiran dalam statistik inferensial sangat dipengaruhi oleh ketepatan pengambilan sampel. Pada umumnya yang dihitung dalam estimasi statistik inferensial adalah besarnya rata-rata hitung dan kesalahan sampling (selisih antara rata-rata hitung populasi dengan rata-rata hitung sampel). Teknik statistik inferensial digunakan untuk menguji hipotesis.

- **Interval Kepercayaan (*confidence interval*)**

Interval kepercayaan digunakan untuk menaksirkan letak rata-rata hitung parametrik, yang mana rata-rata hitung tersebut diperoleh dari hasil perhitungan kesalahan sampling atau simpangan baku rata-rata hitung. Interval kepercayaan merupakan rentangan bilangan dari angka dan ke angka tertentu yang di dalamnya diperkirakan terletak kemungkinan bilangan statistik parametrik berada. Penaksiran letak rata-rata hitung pada umumnya menggunakan taraf signifikansi 5% dan 1%, bahkan ada yang mencapai 10%. Penaksiran dengan taraf signifikansi 5% sama artinya dengan menggunakan nilai Z-skor 1,96, yang mana nilai ini memiliki daerah probabilitas sebanyak 95% atau 0,95. Jadi, daerah penerimaan probabilitas rata-rata hitung adalah seluas 95% dan kemungkinan penolakannya atau salahnya adalah sebesar 5%. Di sisi lain, penaksiran dengan taraf signifikansi 1% sama artinya dengan menggunakan nilai z-skor 2,58, yang mana memiliki daerah probabilitas sebanyak 99% atau 0,99. Jadi daerah penerimaan seluas 99% dengankemungkinan penolakan atau salah sebesar 1 %.

	104,135 – 107,005	
2,5%	95%	2,5%
Daerah Penerimaan Probabilitas Letak Rata-rata Hitung		

	103,68 – 107,46	
0,5%	99%	0,5%
Daerah Penerimaan Probabilitas Letak Rata-rata Hitung		

- **Distribusi Statistik**

Distribusi statistik dibuat dari rata-rata hitung yang biasanya disebut sebagai distribusi sampling rata-rata hitung. Distribusi ini mengikuti asumsi distribusi normal sehingga memiliki karakteristik kurva normal.

1. Distribusi Normal Baku

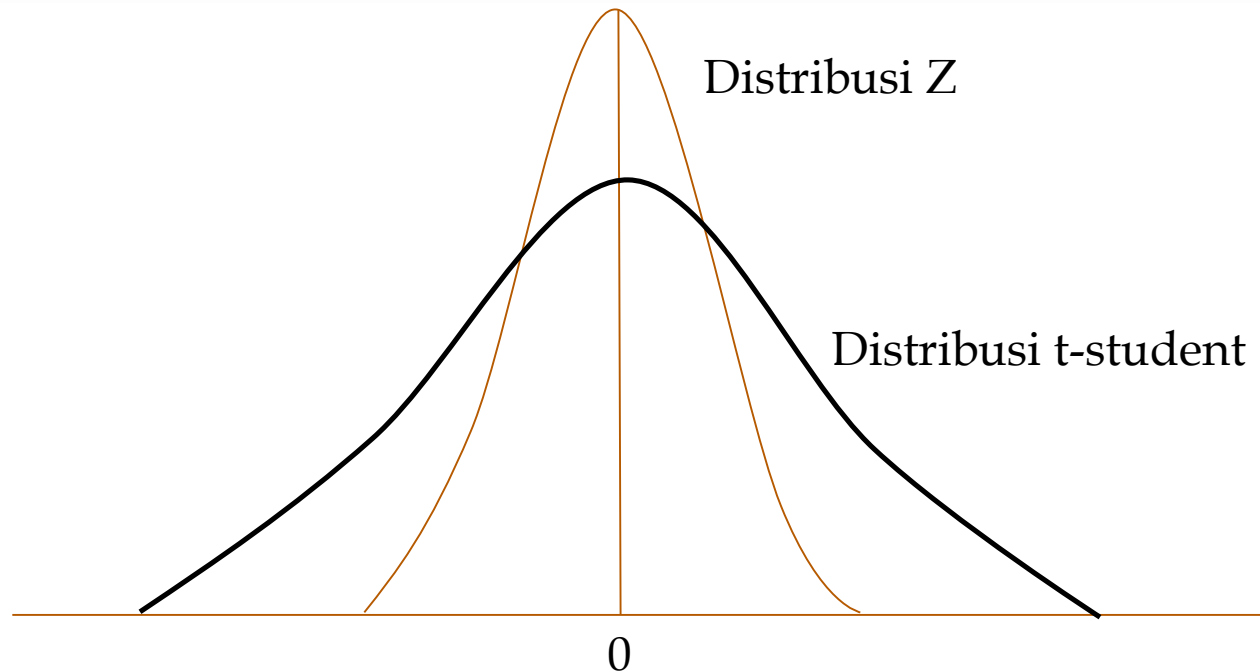
- Distribusi normal baku (Distribusi Gauss) adalah salah satu distribusi peluang (sebuah daftar dari keseluruhan hasil suatu percobaan yang disertai peluang dari masing-masing hasil yang bersangkutan) yang penting dan banyak digunakan sebagai pendekatan dalam menguji suatu hipotesis.
- Distribusi normal ini digunakan untuk variabel yang bersifat acak kontinyu. Rumusnya:

$$Z = \frac{x - \mu}{\sigma}$$

di mana: x = nilai observasi, μ = nilai rata-rata populasi,
dan

σ = simpangan baku populasi

2. Distribusi *t*-student



Kurva distribusi *t*-student lebih mendatar dan melebar di tengah bila dibandingkan dengan kurva distribusi normal baku. Namun dengan semakin besar ukuran sampel, maka kurva yang menggambarkan distribusi *t*-student mendekati kurva distribusi normal baku - Z.

Kriteria Pengujian

(uji signifikansi nilai *t-statistic* dibandingkan nilai distribusi *t-student*)

Tipe Hipotesis	H_0 : hipotesis nol	H_a : hipotesis alternatif	Keputusan : menolak H_0 , jika
dua sisi (<i>two-tail</i>)	$\beta_2 = \beta_2^*$	$\beta_2 \neq \beta_2^*$	$ t > t_{1/2\alpha, df}$
sisi kanan (<i>right-tail</i>)	$\beta_2 \leq \beta_2^*$	$\beta_2 > \beta_2^*$	$t > t_{\alpha, df}$
sisi kiri (<i>left-tail</i>)	$\beta_2 \geq \beta_2^*$	$\beta_2 < \beta_2^*$	$t < -t_{\alpha, df}$

Note: β_2 = nilai koefisien parameter masing-masing variabel independen; β_2^* = nilai koefisien parameter yang dihipotesiskan masing-masing variabel independen; β = rata-rata hitung parameter atau dinotasikan dengan μ ; t = nilai t-hitung atau *t-statistic*; t_α atau $t_{1/2\alpha}$ = nilai kritis/tabel distribusi *t-student*./ $|t|$ = nilai absolut dari *t-statistic* (nilai negatif diabaikan, dianggap positif).

1. Uji Satu Pihak (*one -tail*) Kanan
 - Jika nilai t-hitung(*t-statistic*) > nilai t-tabel (*t-critical value*; α), berarti signifikan atau H_0 ditolak atau H_a diterima.
 - Jika nilai t-hitung(*t-statistic*) < nilai t-tabel (*t-critical value*; α), berarti tidak signifikan atau H_0 diterima atau H_a ditolak.
2. Uji Satu Pihak (*one -tail*) Kiri
 - Jika nilai t-hitung(*t-statistic*) < nilai - t-tabel (*t-critical value*; α), berarti signifikan atau H_0 ditolak atau H_a diterima.
 - Jika nilai t-hitung(*t-statistic*) > nilai - t-tabel (*t-critical value*; α), berarti tidak signifikan atau H_0 diterima atau H_a ditolak.

Kriteria pengujian distribusi *t-student* (2)

3. Uji Dua Pihak (*two-tail*)
- Jika nilai *t*-tabel (*t-critical value*; $\frac{1}{2} \alpha$) < $|nilai\ t-statistic|$, maka H_0 ditolak atau H_a diterima.
 - Jika nilai *t*-tabel (*t-critical value*; $\frac{1}{2} \alpha$) > $|nilai\ t-statistic|$, maka H_0 diterima atau H_a ditolak.

Keterangan:

$|nilai\ t-statistic|$ merupakan nilai absolut dari nilai *t*-hitung atau *t-statistic* (artinya mengabaikan tanda negatif, di mana jika nilai *t-statistic* negatif maka dianggap bernilai positif)

Uji *t-Statistic*

(kriteria pengujian berdasarkan nilai probabilitas atau sig. t-stat)

Selain dengan menggunakan kriteria pengujian perbandingan antara nilai *t-statistic* dengan nilai *distribusi t-student*, maka dapat juga dilakukan dengan perbandingan nilai probabilitas (sig. t-stat.) dengan taraf signifikansi (baik 1%, 5%, maupun 10% yang disesuaikan dengan pilihan pengujian *two-tail* atau *one-tail*). Ketentuannya adalah sebagai berikut:

- Jika nilai probabilitas *t-statistic* $<$ taraf signifikansi tertentu, maka H_0 ditolak atau H_a diterima dan berarti signifikan.
- Jika nilai probabilitas *t-statistic* $>$ taraf signifikansi tertentu, maka H_0 diterima atau H_a diterima dan berarti tidak signifikan.

3. Distribusi F

Hipotesis pengambilan keputusan uji F menyatakan bahwa:

$$H_0 = \beta_i = 0$$

Artinya hipotesis nol uji F adalah sama dengan menguji hipotesis koefisien determinasi (R^2) dalam populasi = 0. Rumus untuk menentukan uji signifikansi melalui uji F-statistic atau F -hitung adalah berikut.

$$F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)}$$

Cara untuk menentukan nilai kritis F dalam tabel distribusi F adalah sebagai berikut:

Df_{numerator (pembilang)} = k, di mana k = variabel independen (bebas)

Df_{denominator (penyebut)} = (n-k-1)

Kriteria pengujiannya berdasarkan nilai kritis F adalah sebagai berikut:

Jika $F_{\text{statistic}} > F_{\text{kritis}}$ pada taraf signifikansi tertentu, maka diputuskan H_0 ditolak atau H_a diterima.

Jika $F_{\text{statistic}} < F_{\text{kritis}}$ pada taraf signifikansi tertentu, maka diputuskan H_0 diterima atau H_a ditolak.

4. Koefisien Determinasi

- Koefisien determinasi adalah kuadrat dari koefisien korelasi, yaitu $(r)^2$.
- Koefisien determinasi dapat menjelaskan sejauhmana penyebaran harga-harga variabel X (variabel bebas) dapat menjelaskan harga variabel terikat (variabel Y).

Contoh: harga koefisien korelasi $r = 0,9$, maka harga koefisien determinasi adalah $(0,9)^2 = 0,81 = 81\%$.

Artinya sebesar 81 % variabel bebas dapat mempengaruhi variabel terikat atau masalah (variabel Y), sedangkan yang 19% dipengaruhi oleh variabel lain di luar model persamaan estimasinya.

Koefisien determinasi...(2)

Harga koefisien korelasi berada pada interval -1 sampai +1, atau dapat dituliskan: $-1 \leq r \leq +1$.

	r			
-1	0			+1
←	makin kuat		makin kuat	→

Koefisien korelasi ganda untuk statistik parametrik diperoleh dengan menggunakan rumus korelasi *product moment* Pearson. Rumus perhitungan diperoleh sebagai berikut:

$$r_{xy} = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \cdot \sqrt{n \sum y^2 - (\sum y)^2}}$$

UJI BEDA RATA-RATA

Konsep Uji Beda

- Untuk menguji keadaan atau “sesuatu” yang terdapat dalam suatu kelompok dengan kelompok (-kelompok) yang lain.
- Teknik statistik dimaksudkan untuk menguji hipotesis komparatif, baik berupa hipotesis nol maupun hipotesis alternatif (hipotesis kerja).
- Teknik statistik yang tepat untuk menguji apakah terdapat perbedaan yang signifikan di antara kelompok-kelompok yang diuji adalah **uji beda**.

Metode untuk Uji Beda

- **T-tes** (jika sampel yang ingin diuji perbedaan rata-rata hitungnya “hanya” terdiri dari dua kelompok)
- **Analisis Varians (Anova)** (untuk menguji perbedaan rata-rata hitung dua kelompok sampel, dan t-tes untuk kelompok sampel yang lebih dari dua kelompok)
- **Chi-kudrat atau Chi-square (X^2)** (jika berupa data berskala nominal, uji beda dilakukan untuk menguji perbedaan frekuensi pemunculan).

1. Pengujian T-TEST (Hipotesis Komparatif)

1. DISTRIBUSI SAMPEL BEBAS (*INDEPENDENT SAMPLES*)

- Sampel-sampelnya berasal dari dua populasi yang berbeda, atau kelompok subyeknya berbeda.
- Contoh: membandingkan kualitas pelayanan dari dua buah restoran, dua buah hotel, dan sebagainya.

2. DISTRIBUSI SAMPEL BERHUBUNGAN ATAU BERPASANGAN (*CORRELATED SAMPLES* ATAU *PAIRED SAMPLES*)

- Berasal dari sampel yang sama, atau kelompok subyek yang sama.
- Sering dijumpai dalam penelitian eksperimental, contohnya membuat perbandingan kemampuan karyawan sebelum dan setelah mengikuti pelatihan (*treatment*) yaitu dengan membandingkan nilai pre-test dan pasca-test.

Hipotesis Komparatif Dua Sampel (*t-test*)

Uji hipotesis komparatif dua sampel dengan menggunakan uji statistik baik untuk dua sampel yang berkorelasi maupun untuk dua sampel yang bersifat independen. Tergantung rumusan hipotesis komparatif dua sampel yang digunakan.

- Hipotesis komparatif dirumuskan untuk menguji parameter populasi dalam bentuk perbandingan dengan cara membandingkan statistik sampel.
- Pada umumnya, teknik statistik parametrik yang biasa digunakan untuk menguji hipotesis komparatif dilakukan dengan cara menguji perbedaan rata-rata hitung dari dua kelompok sampel dengan metode *t-test*.

Perumusan Uji Hipotesis Komparatif

t-Test Dua Sampel

(interpretasi analisis *correlated /paired samples t-test* mengikuti tabel distribusi *t-student*)

a) **Uji dua pihak**

H_0 = tidak ada perbedaan dalam variabel yang diamati pada sampel 1 dan sampel 2.

H_a = ada perbedaan dalam variabel yang diamati pada sampel 1 dan sampel 2.

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

b) **Uji satu pihak (pihak kanan)**

H_0 = variabel yang diamati pada sampel 1 lebih kecil atau sama dengan variabel yang diamati pada sampel 2.

H_a = variabel yang diamati pada sampel 1 lebih besar dari sampel 2.

$$H_0 : \mu_1 \leq \mu_2$$

$$H_a : \mu_1 > \mu_2$$

b) **Uji satu pihak (pihak kiri)**

H_0 = variabel yang diamati pada sampel 1 lebih besar atau sama dengan variabel yang diamati pada sampel 2.

H_a = variabel yang diamati pada sampel 1 lebih kecil dari sampel 2.

$$H_0 : \mu_1 \geq \mu_2$$

$$H_a : \mu_1 < \mu_2$$

Komparatif Dua Sampel yang Independen

Sampel Independen : statistik yang digunakan untuk menguji hipotesis komparatif rata-rata hitung dua sampel yang bersifat independen atau tidak berkorelasi.

Contohnya : jika akan membandingkan produktivitas kerja antara pramusaji dengan pramugraha; kualitas pelayanan bagian ticketing dan bagian tour; kinerja karyawan dapur dan karyawan kantor depan; dan lain-lain jika datanya berbentuk interval atau rasio maka dapat diuji dengan menggunakan *t-student*.

Bentuk persamaannya sebagai berikut:

☐ Separated varians:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

□ Polled varians:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

di mana:

\overline{X}_1 = rata-rata sampel 1

\overline{X}_2 = rata-rata sampel 2

S_1^2 = variansi sampel 1

S_2^2 = variansi sampel 2

S_1 = simpangan baku sampel 1

S_2 = simpangan baku sampel 2

n_1 = jumlah sampel 1

n_2 = jumlah sampel 2

Cara menghitung variansi (S^2) dan simpangan baku (S)

Rumus menghitung variansi:

$$S^2 = \frac{n \sum X^2 - (\sum X)^2}{n(n-1)}$$

Rumus menghitung simpangan baku:

$$S = \sqrt{\frac{n \sum X^2 - (\sum X)^2}{n(n-1)}}$$

Untuk mengetahui derajat kebebasan (dk) atau *degree of freedom* (df) dalam penentuan nilai t-tabel (*t-critical value* atau distribusi *t-student*) adalah **dk = (n₁ + n₂ - 2)/2**

Komparatif Dua Sampel yang Berkorelasi

Sampel Berkorelasi : statistik uji yang digunakan untuk menguji hipotesis komparatif rata-rata hitung dua sampel yang berkorelasi jika datanya terbentuk interval atau rasio adalah *t-student*. Persamaannya ditulis sebagai berikut:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} - 2r \left(\frac{S_1}{\sqrt{n_1}} \right) \left(\frac{S_2}{\sqrt{n_2}} \right)}}$$

di mana:

\overline{X}_1	= rata-rata sampel 1
\overline{X}_2	= rata-rata sampel 2
S_1^2	= variansi sampel 1
S_2^2	= variansi sampel 2
S_1	= simpangan baku sampel 1
S_2	= simpangan baku sampel 2
n_1	= jumlah sampel 1
n_2	= jumlah sampel 2

Untuk menentukan nilai r (jika terdiri dua sampel) maka digunakan rumus korelasi *product moment* Pearson sebagai berikut:

$$r = \frac{n \sum X_1 X_2 - (\sum X_1) \cdot (\sum X_2)}{\sqrt{n \sum X_1^2 - (\sum X_1)^2} \cdot \sqrt{n \sum X_2^2 - (\sum X_2)^2}}$$

Rumus Paired Samples t-test lain yang juga dapat digunakan untuk perhitungan nilai t-test

Rumus *t*-test =

$$t = \frac{\bar{X}_D - \mu_0}{S_D / \sqrt{n}}$$

di mana: $\bar{X}_d = \frac{\sum D}{n}$

$$S = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

Keterangan:

D = selisih antara sampel X_1 dan X_2 atau $(X_1 - X_2)$

n = banyaknya data

\bar{X}_D = rata-rata hitung

S_D = Standar deviasi dari D

2. Pengujian Chi-Square

Chi-square merupakan metode statistik nonparametrik yang digunakan untuk mengestimasi *correlated/paired samples*. Pada umumnya digunakan untuk mengukur korelasi antara data kualitatif dengan data kuantitatif atau dapat juga merubah data kualitatif menjadi data kuantitatif ,dan untuk menguji perbedaan frekuensi pada data berskala nominal. Rumus untuk menentukan nilai chi-square hitung adalah sebagai berikut.

$$X^2 = \sum \frac{(f_0 - f_e)^2}{f_e}$$

Keterangan:

X^2 = *Chi-square*

f_0 = frekuensi yang diobservasi

f_e = frekuensi yang diharapkan

Interpretasi Analisis Uji *Chi-square*

Hipotesis *chi-square*:

- H_0 menyatakan bahwa tidak ada perbedaan yang signifikan antara frekuensi yang diobservasi terhadap frekuensi yang diharapkan .
- H_a menyatakan bahwa ada perbedaan yang signifikan antara frekuensi yang diobservasi terhadap frekuensi yang diharapkan.

Kriteria pengujian hipotesis:

- Jika nilai X^2 hitung $\leq X^2$ tabel/kritis, maka menerima H_0 atau menolak H_a .
- Jika nilai X^2 hitung $> X^2$ tabel/kritis, maka menerima H_a atau menolak H_0 .

Untuk menentukan nilai kritis/tabel dari X^2 pada tabel distribusi *chi-square* perlu diketahui terlebih dahulu posisi dengan cara *degree of freedom* (df) = (baris-1) (kolom-1)

LATIHAN PENYELESAIAN UJI BEDA

Soal Uji Beda – *Correlated/ Paired Samples T-test* (Dua Sampel dalam Statistik Parametrik)

No. Urut Pramugraha	Nilai Produktivitas Kerja		X_1	X_2	X_1X_2
	Sebelum Pelatihan (X_1)	Setelah Pelatihan (X_2)			
1	79	87	6241	7569	6873
2	76	80	5776	6400	6080
3	80	85	6400	7225	6800
4	91	92	8281	8464	8372
5	88	88	7744	7744	7744
6	79	75	6241	5625	5925
7	75	80	5625	6400	6000
8	90	90	8100	8100	8100
9	67	85	4489	7225	5695
10	65	72	4225	5184	4680
11	66	69	4356	4761	4554
12	80	84	6400	7056	6720
13	80	86	6400	7396	6880
14	75	80	5625	6400	6000
15	77	79	5929	6241	6083
Jumlah	1168	1232	91832	101790	96506

Berdasarkan data di samping, apakah ada perbedaan produktivitas kinerja pramugraha antara sebelum pelatihan dan setelah pelatihan? Gunakan metode analisis statistik untuk menyelesaikan soal tersebut.

Sumber: contoh soal data dikutip dari tugas kelompok mahasiswa S1-V Hospitality

Cara 1: penyelesaian dengan perhitungan manual

$$\bar{X}_1 = \frac{\sum X_1}{n} = \frac{1168}{15} = 77,86$$

$$\bar{X}_2 = \frac{\sum X_2}{n} = \frac{1232}{15} = 82,13$$

$$S_1^2 = \frac{n \sum X_1^2 - (\sum X_1)^2}{n(n-1)}$$

$$S_2^2 = \frac{n \sum X_2^2 - (\sum X_2)^2}{n(n-1)}$$

$$S_1^2 = \frac{15 \times 91832 - (1168)^2}{15(15-1)}$$

$$S_2^2 = \frac{15 \times 101790 - (1232)^2}{15(15-1)}$$

$$S_1^2 = \frac{13256}{210} = 63,12$$

$$S_2^2 = \frac{9026}{210} = 42,98$$

$$S_1 = \sqrt{63,12} = 7,94$$

$$S_2 = \sqrt{42,98} = 6,55$$

$$r = \frac{n \sum X_1 X_2 - (\sum X_1) \cdot (\sum X_2)}{\sqrt{n \sum X_1^2 - (\sum X_1)^2} \cdot \sqrt{n \sum X_2^2 - (\sum X_2)^2}}$$

$$r = \frac{15 \times 96506 - (1168)(1232)}{\sqrt{15 \times 91832 - (1364224)} \cdot \sqrt{15 \times 101790 - (1517824)}}$$

$$r = \frac{1447590 - 1438976}{\sqrt{13256} \cdot \sqrt{9026}} = \frac{8614}{115,1347 \times 95,0052} = \frac{8614}{10938,39} = 0,7875$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} - 2r \left(\frac{S_1}{\sqrt{n_1}} \right) \left(\frac{S_2}{\sqrt{n_2}} \right)}}$$

$$t = \frac{77,86 - 82,13}{\sqrt{\frac{63,12}{15} + \frac{42,98}{15} - 2 \times 0,7875 \left(\frac{7,94}{\sqrt{15}} \right) \left(\frac{6,55}{\sqrt{15}} \right)}}$$

$$t = \frac{-4,27}{\sqrt{(4,208 + 2,865) - 1,575(2,051)(1,692)}}$$

$$t = \frac{-4,27}{\sqrt{7,073 - 5,465}} = \frac{-4,27}{1,268} = -3.36$$

Cara 2 : penyelesaian dengan perhitungan manual

No. Urut Pramugraha	Nilai Produktivitas Kerja		D = X ₁ -X ₂	D ²
	Sebelum Pelatihan (X ₁)	Setelah Pelatihan (X ₂)		
1	79	87	-8	64
2	76	80	-4	16
3	80	85	-5	25
4	91	92	-1	1
5	88	88	0	0
6	79	75	4	16
7	75	80	-5	25
8	90	90	0	0
9	67	85	-18	324
10	65	72	-7	49
11	66	69	-3	9
12	80	84	-4	16
13	80	86	-6	36
14	75	80	-5	25
15	77	79	-2	4
Jumlah	1168	1232	-64	610

$$S = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

$$S = \sqrt{\frac{1}{15-1} \left\{ 610 - \frac{(-64)^2}{15} \right\}}$$

$$S = \sqrt{24,0671} = 4,9058$$

$$t = \frac{\left(\frac{-64}{15} \right) - 0}{4,9058 / \sqrt{15}}$$

$$t = \frac{-4,26}{1,2676} = -3,3606$$

Penyelesaian dengan SPSS

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Sebelum_X1	77.8667	15	7.94505	2.05140
	Setelah_X2	82.1333	15	6.55599	1.69275

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Sebelum_X1 & Setelah_X2	15	.788	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Sebelum_X1 - Setelah_X2	-4.26667	4.90578	1.26667	-6.98340	-1.54994	-3.368	14	.005

Penyelesaian dengan Ms. Excel

t-Test: Paired Two Sample for Means		
	<i>Variable 1</i>	<i>Variable 2</i>
Mean	77.86666667	82.13333333
Variance	63.12380952	42.98095238
Observations	15	15
Pearson Correlation	0.787500712	
Hypothesized Mean Difference	0	
df	14	
t Stat	-3.36842105	
P(T<=t) one-tail	0.002296263	
t Critical one-tail	1.761310115	
P(T<=t) two-tail	0.004592527	
t Critical two-tail	2.144786681	

Keterangan:

Mean = nilai rata-rata hitung (\bar{x})

Variance = variansi sampel (S^2)

Pearson Correlation = nilai koefisien korelasi (r)

Observation = banyaknya data (n)

Interpretasi Analisis *Correlated/ Paired* Samples T-Test

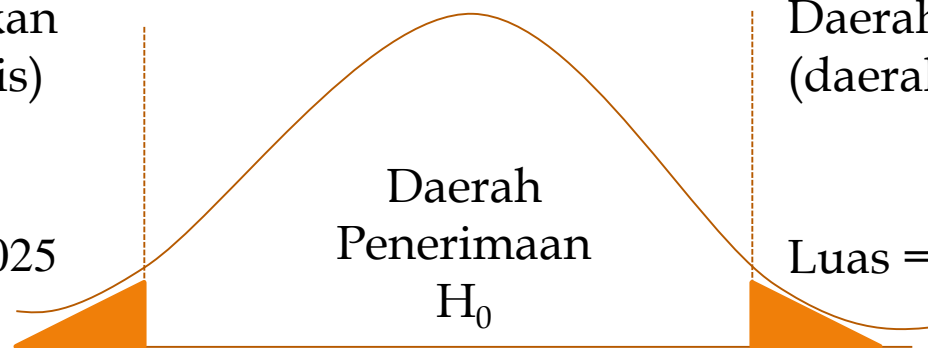
- Hipotesis statistik komparatif
 $H_0 : \mu_1 = \mu_2$
 $H_a : \mu_1 \neq \mu_2$
 H_0 menyatakan bahwa tidak ada perbedaan nilai produktivitas kerja antara sebelum pelatihan (X_1) dan sesudah pelatihan (X_2).
 H_a menyatakan bahwa ada perbedaan antara nilai produktivitas kerja antara sebelum pelatihan (X_1) dan sesudah pelatihan (X_2).
- Nilai t-tabel/kritis $\{df = (n_1+n_2-2)/2 = (15+15-2)/2 = 14; 1/2 \alpha = 0,025\} = 2,144$
- Nilai *t-statistic* = - 3,36
- Jadi, nilai *t-statistic* > nilai t-kritis atau $|3,36| > 2,144$, maka H_a diterima atau H_0 ditolak.
- Kesimpulannya: Ada perbedaan nilai produktivitas kerja antara sebelum pelatihan dan sesudah pelatihan.

Daerah Penolakan
 H_0 (daerah kritis)

Daerah Penolakan H_0
(daerah kritis)

Luas = $1/2 \alpha = 0,025$

Luas = $1/2 \alpha = 0,025$



-2,144

2,144

$|3,36|$

Soal Uji Beda dengan Metode *Chi-square* (Dua Sampel dalam Statistik Nonparametrik)

Fasilitas dan Aksesibilitas	Tanggapan Pengunjung			TOTAL
	Sangat	Cukup	Kurang	
	Luas/Baik/Mudah	Luas/Baik/Mudah	Luas/Baik/Mudah	
Area parkir	12	23	35	70
Variasi produk	26	36	8	70
Kualitas produk	15	30	25	70
Jarak dari bandara, stasiun	30	30	10	70
Kemudahan pencapaian	25	40	5	70
TOTAL	108	159	83	350

Sumber: contoh soal data dikutip dari tugas kelompok mahasiswa S1-V Hospitality

Berdasarkan data di atas, apakah ada perbedaan persepsi kepuasan pengunjung pada setiap fasilitas dan aksesibilitas di destinasi wisata belanja Malioboro Yogyakarta? Gunakan metode analisis statistik untuk menyelesaikan soal tersebut.

Penyelesaian dengan perhitungan manual

Fasilitas dan Aksesibilitas	Tanggapan Pengunjung			TOTAL
	Sangat	Cukup	Kurang	
	Luas/Baik/Mudah	Luas/Baik/Mudah	Luas/Baik/Mudah	
Area parkir	12/21,6	23/31,8	35/16,6	70
Variasi produk	26/21,6	36/31,8	8/16,6	70
Kualitas produk	15/21,6	30/31,8	25/16,6	70
Jarak dari bandara, stasiun	30/21,6	30/31,8	10/16,6	70
Kemudahan pencapaian	25/21,6	40/31,8	5/16,6	70
TOTAL	108	159	83	350

- Harga f_e pada kolom I dapat dihitung sebagai berikut:

$$f_{e1} = \frac{70}{350} \times 108 = 21,6$$

$$f_{e2} = \frac{70}{350} \times 159 = 31,8$$

$$f_{e3} = \frac{70}{350} \times 83 = 16,6$$

Hasil perhitungan chi-square (X^2)

$$\begin{aligned} X^2 &= \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(12 - 21,6)^2}{21,6} + \frac{(26 - 21,6)^2}{21,6} + \frac{(15 - 21,6)^2}{21,6} + \frac{(30 - 21,6)^2}{21,6} + \frac{(25 - 21,6)^2}{21,6} + \\ &\frac{(23 - 31,8)^2}{31,8} + \frac{(36 - 31,8)^2}{31,8} + \frac{(30 - 31,8)^2}{31,8} + \frac{(30 - 31,8)^2}{31,8} + \frac{(40 - 31,8)^2}{31,8} + \\ &\frac{(35 - 16,6)^2}{16,6} + \frac{(8 - 16,6)^2}{16,6} + \frac{(25 - 16,6)^2}{16,6} + \frac{(10 - 16,6)^2}{16,6} + \frac{(5 - 16,6)^2}{16,6} \\ X^2 &= \frac{92,16}{21,6} + \frac{19,36}{21,6} + \frac{43,56}{21,6} + \frac{70,56}{21,6} + \frac{11,56}{21,6} + \frac{77,44}{31,8} + \frac{17,64}{31,8} + \\ &\frac{3,24}{31,8} + \frac{3,24}{31,8} + \frac{67,24}{31,8} + \frac{338,56}{16,6} + \frac{73,96}{16,6} + \frac{70,56}{16,6} + \frac{43,56}{16,6} + \frac{134,56}{16,6} \\ X^2 &= \frac{237,2}{21,6} + \frac{168,8}{31,8} + \frac{661,2}{16,6} = 10,99 + 5,31 + 39,83 \\ X^2 &= 56,13 \end{aligned}$$

Interpretasi *Chi-square*

- Hipotesis *chi-square*:
 - H_0 menyatakan bahwa tidak ada perbedaan persepsi/tanggapan kepuasan pengunjung antara kualitas produk, fasilitas, dan aksesibilitas objek wisata di kota Bandung.
 - H_a menyatakan bahwa ada perbedaan persepsi/tanggapan kepuasan pengunjung antara kualitas produk, fasilitas, dan aksesibilitas objek wisata di kota Bandung.
- Kriteria pengujian hipotesis:
 - Jika nilai X^2 hitung $\leq X^2$ tabel/kritis, maka menerima H_0 atau menolak H_a .
 - Jika nilai X^2 hitung $> X^2$ tabel/kritis, maka menerima H_a atau menolak H_0 .
- *Degree of freedom* (df) = (baris-1) (kolom-1) = (5 - 1).(4 - 1) = 12
- X^2 kritis/tabel pada $\alpha = 5\%$ sebesar 21,026
- Berdasarkan hasil perhitungan, diperoleh kriteria pengujian bahwa nilai X^2 hitung sebesar 56,13 $> 21,026$ (X^2 tabel) maka diputuskan untuk menerima H_a dan menolak H_0 .
- Dengan demikian dapat disimpulkan dan diputuskan bahwa ada perbedaan persepsi kepuasan pengunjung antara kualitas produk, fasilitas, dan aksesibilitas objek wisata di kota Bandung pada taraf signifikan 5%.

Latihan Soal Dua Sampel Statistik Nonparametrik dengan Metode Chi-square

Fasilitas dan Pelayanan	Tanggapan Konsumen/Tamu Hotel			TOTAL
	Sangat Baik/Nyaman/Puas	Cukup Baik/Nyaman/Puas	Kurang Baik/Nyaman/Puas	
Bar/Ruang Makan	35	25	20	80
Kamar	45	20	15	80
Kolam Renang	40	25	15	80
Area Parkir	35	35	10	80
TOTAL	155	105	60	320

Pertanyaan:

Berdasarkan data di atas, apakah ada perbedaan persepsi kepuasan tamu hotel pada setiap fasilitas dan pelayanan Hotel Holiday? Gunakan metode analisis statistik untuk menyelesaikan soal tersebut.

UJI KORELASI

Konsep Korelasi

Korelasi adalah suatu metode yang dapat digunakan untuk mengukur tingkat atau derajat hubungan (asosiasi) antarvariabel (Sulaiman dan Kusherdyana, 2013)

- ❖ Untuk variabel pengamatan yang terdiri dari dua variabel maka disebut korelasi sederhana dan persamaan regresi sederhana
- ❖ Untuk variabel pengamatan yang lebih dari dua variabel disebut korelasi berganda dan persamaan regresi berganda

a. Koefisien korelasi *product moment* Pearson

Suatu parameter untuk mengukur tingkat atau derajat hubungan yang sering disebut koefisien korelasi (r). Rumusnya adalah berikut:

$$r = \frac{n \sum xy - (\sum x) \cdot (\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \cdot \sqrt{n \sum y^2 - (\sum y)^2}}$$

b. Koefisien korelasi Rank Spearman

Untuk data dalam bentuk ordinal (berjenjang) (termasuk dalam statistik nonparametrik), misalnya pada penggunaan Skala Likert, Skala Rating, Skala Semantik diferensial dalam melakukan analisis menggunakan koefisien korelasi *product moment* Pearson tidak tepat, sehingga persamaannya adalah berikut:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

di mana:

r_s = koefisien korelasi Rank Spearman

d = beda ranking antara variabel bebas (X) dan variabel terikat (Y)

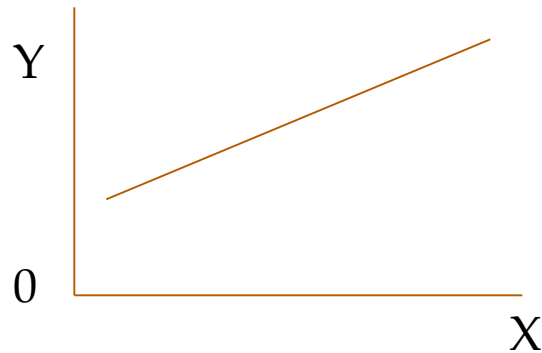
n = jumlah sampel/ pengamatan

c. Koefisien Determinasi

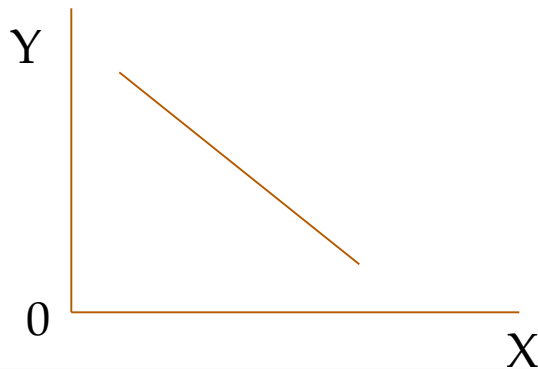
Koefisien determinasi adalah kuadrat dari koefisien korelasi, yaitu $(r)^2$. Koefisien determinasi dapat menjelaskan sejauhmana penyebaran harga-harga variabel X (variabel bebas) dapat menjelaskan harga variabel terikat (variabel Y).

Nilai koefisien korelasi	Sifat Korelasi (hubungan)
$r < 0,2$	Hubungan sangat lemah
$0,2 < 0,4$	Hubungan lemah
$0,4 < 0,7$	Hubungan cukup kuat
$0,7 < 0,9$	Hubungan kuat
$0,9 \leq 1,0$	Hubungan sangat kuat

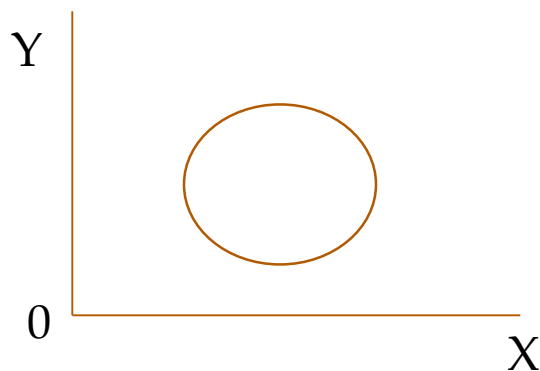
Plot diagram pencar yang menggambarkan titik-titik koordinat x dan y (x, y)



$r = +1 \rightarrow$ korelasi kuat, positif
Artinya makin besar nilai x, makin besar pula nilai y, atau sebaliknya.
Korelasi atau hubungan yang semacam ini dapat disebut pula hubungan atau korelasi langsung.



$r = -1 \rightarrow$ korelasi kuat, negatif
Artinya makin besar nilai x, makin kecil nilai y, atau sebaliknya.
Korelasi atau hubungan yang semacam ini dapat disebut pula hubungan atau korelasi terbalik.



$r = 0 \rightarrow$ tidak ada korelasi antar variabel
Artinya tidak ada korelasi antara nilai-nilai x dan nilai-nilai y.

Contoh 1: Soal Korelasi *Product Moment Pearson*

No.	Volume Penjualan Paket Wisata (ribuan unit)	Laba Bersih (puluhan juta rupiah)
1.	2,5	6,0
2.	2,6	5,9
3.	2,8	6,2
4.	2,9	6,4
5.	3,0	6,9
6.	3,5	7,2
7.	3,4	7,0
8.	4,0	7,5
9.	3,4	6,8
10.	3,2	6,9
11.	3,3	6,7
12.	3,8	6,9

Pertanyaan:

1. Hitunglah harga koefisien korelasinya!
2. Bagaimana interpretasi hasil perhitungan korelasi tersebut?

Penyelesaian:

No.	Volume Penjualan (X)	Laba Bersih (Y)	XY	X ²	Y ²
1	2.5	6.0	15.00	6.25	36.00
2	2.6	5.9	15.34	6.76	34.81
3	2.8	6.2	17.36	7.84	38.44
4	2.9	6.4	18.56	8.41	40.96
5	3.0	6.9	20.70	9.00	47.61
6	3.5	7.2	25.20	12.25	51.84
7	3.4	7.0	23.80	11.56	49.00
8	4.0	7.5	30.00	16.00	56.25
9	3.4	6.8	23.12	11.56	46.24
10	3.2	6.9	22.08	10.24	47.61
11	3.3	6.7	22.11	10.89	44.89
12	3.8	6.9	26.22	14.44	47.61
TOTAL	38.40	80.40	259.49	125.20	541.26

$$(\sum X)^2 = (38,40)^2 = 1474,56$$

$$(\sum Y)^2 = (80,40)^2 = 6464,16$$

$$r = \frac{n \sum xy - (\sum x) \cdot (\sum y)}{\sqrt{n \sum x^2 - (\sum x)^2} \cdot \sqrt{n \sum y^2 - (\sum y)^2}}$$

$$r = \frac{12.(259,49) - (38,40).(80,40)}{\sqrt{12.(125,20) - 1474,56} \sqrt{12.(541,26) - 6464,16}}$$

$$r = \frac{3113,88 - 3087,36}{\sqrt{1502,40 - 1474,56} \sqrt{6495,12 - 6464,16}}$$

$$r = \frac{26,52}{\sqrt{27,84} \cdot \sqrt{30,96}} = \frac{26,52}{(5,28).(5,56)}$$

$$r = \frac{26,52}{29,36} = 0,90326 = 0,90$$

Interpretasi hasil perhitungan :

1. Harga koefisien korelasi antara volume penjualan paket wisata dan laba bersih sebesar 0,90.
2. Dengan r sebesar 0,90 dapat diinterpretasikan bahwa ada hubungan yang sangat kuat dan positif antara volume penjualan paket wisata dan laba bersih di Biro Perjalanan Wisata tersebut, dengan meningkatnya volume penjualan paket wisata akan memberikan kecenderungan meningkatnya laba bersih.

Contoh 2: Soal Korelasi *Product Moment Pearson*

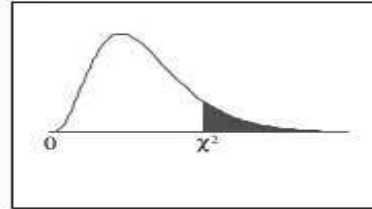
Tingkat Huni Kamar dan *Rooms Revenue* di Hotel Alamack
Periode Agustus 2010 – Juli 2011

No.	Bulan	THK (%)	<i>Rooms Revenue</i> (ratusan juta rupiah)
1.	Agustus 2010	74,0	8,9
2.	September 2010	69,0	8,5
3.	Oktober 2010	72,0	8,5
4.	November 2010	76,0	9,1
5.	Desember 2010	85,0	10,8
6.	Januari 2011	87,0	11,0
7.	Februari 2011	68,0	8,8
8.	Maret 2011	64,0	8,7
9.	April 2011	64,0	8,4
10.	Mei 2011	65,0	8,5
11.	Juni 2011	66,0	8,7
12.	Juli 2011	70,0	8,7

Pertanyaan:

1. Hitunglah nilai korelasi dan koefisien determinasinya!
2. Bagaimanakah interpretasi hasil perhitungan tersebut?

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_{\alpha}$.

<i>df</i>	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

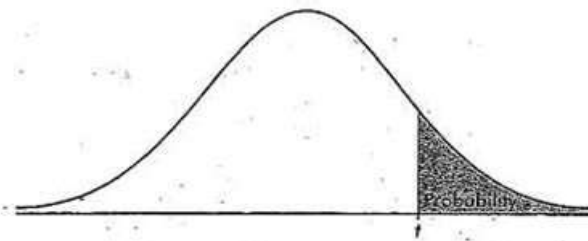


TABLE B: t-DISTRIBUTION CRITICAL VALUES

df	Tail probability <i>p</i>											
	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level <i>C</i>											

Sumber Pustaka

- Kusmayadi, (2004), *Statistika Pariwisata Deskriptif*. Bandung: Tarsito.
- Sulaiman dan Kusherdyana, (2013), *Pengantar Statistika Pariwisata*, aplikasinya dalam bidang pariwisata, usaha perjalanan, dan perhotelan. Bandung: Penerbit Alfabeta.
- Nurgiyantoro, B., Gunawan dan Marzuki, (2012), *Statistik Terapan: untuk penelitian ilmu-ilmu sosial*, Yogyakarta: Gadjah Mada University Press.
- Gujarati, D. N, (2003), *Basic Econometrics*, Fourth Edition. McGraw-Hill, Singapore.

Good Luck and Bring to success