PROBABILITAS dan STATISTIKA

Oleh :
Nova Rijati, Yuniarsi Rahayu, Achmad Wahid Kurniawan

Penerbit
UDINUS Press
Semarang

KATA PENGANTAR

Puji syukur bagi Tuhan Yang Maha Kasih atas segala karuniaNya yang telah dilimpahkan kepada penulis sehingga buku Probabilitas dan Statistika ini dapat terselesaikan dengan baik.

Probabilitas dan Statistika merupakan mata kuliah wajib untuk mahasiswa Fakultas Ilmu Komputer. Penyusunan buku ini dimaksudkan untuk memberikan kemudahan kepada mahasiswa dalam memahami mata kuliah Probabilitas dan Statistika bagi mahasiswa Fakultas Ilmu Komputer Universitas Dian Nuswantoro Semarang khususnya mahasiswa program studi Teknik Informatika.

Materi yang dibahas dalam buku ini meliputi : Pengantar Probabilitas, Distribusi Probabilitas, Analisis Regresi, Analisis Korelasi, Distribusi Sampling, Penaksiran Parameter dan Pengujian Hipotesa

Untuk lebih membantu mahasiswa dalam memahami buku ini, pada tiap-tiap bab disertai pula dengan contoh dan latihan soal. Meskipun begitu untuk lebih memahami materi-materi dalam buku ini, diharapkan mahasiswa mempelajari pula buku-buku textbook seperti tercantum dalam Daftar Pustaka atau buku-buku yang disarankan dosen sebagai buku pegangan.

Kami menyadari, penyusunan buku ini belum dapat dikatakan sempurna, sehingga kritik dan saran dari pengguna buku ini sangat kami harapkan. Meskipun begitu kami berharap buku ini dapat memberikan manfaat dan membantu mahasiswa dalam mata kuliah Probabilitas dan Statistika

Penulis

DAFTAR ISI

BAB 1

PENGANTAR PROBABILITAS

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa mampu memahami konsep dan permasalahan probabilitas serta dapat menentukan peluang suatu kejadian.

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa dapat memahami pengertian probabilitas, mengetahui rumus-rumus probabilitas, menghitung probabilitas suatu peristiwa dan menghitung kombinasi dan permutasi suatu peristiwa.

1.1 Konsep Dasar Probabilitas

1.1.1 Pengantar

Dalam percakapan sehari–hari seringkali kita menggunakan kata *kemungkinan*, dalam bahasa Inggris *chance*, dalam bahasa Belanda *kans*. Misalnya dalam pertandingan sepakbola, kesebelasan A mempunyai kemungkinan (kans) menang lebih besar daripada kesebelasan B. Dalam politik Luar Negeri : Apakah sekarang ini ada kemungkinan pecah Perang Dunia ketiga ? Dalam dunia kedokteran : Apakah pasien ini mungkin diserang penyakit *Typhus* ? Dan masih banyak lagi.

Lebih banyak lagi kata *kemungkinan* dipakai dalam bidang permainan (Teori Permainan, *Theory of Games*)., yaitu, jika kita bermain kartu, misalnya bridge dan lain – lain. Memang lahirnya teori kemungkinan tidak pada tempat yang seharusnya, yaitu di meja perjudian, tetapi beberapa sarjana besar yang taat beragama seperti Pascal, Laplace, de Moivre, dan lain – lain, mulai memikirkan teori kemungkinan (*Theory of Probability)*: tidak untuk memakainya dalam judi tetapi dalam ilmu (*science).* Cabang ilmu yang pada masa sekarang menjadi penting dan tumbuh pesat, yaitu ilmu *statistika* , yang berdasarkan pada teori kemungkinan ini. Tidak terbatas pada ilmu statistika saja, tetapi ilmu lain juga memakainya.

Jika kita melempar sebuah mata uang, misalnya mata uang 500 rupiah, pada sebuah lantai yang licin, kita semua tahu bahwa lemparan tadi dapat menghasilkan sisi Gambar (bunga melati) atau sisi Angka. Karena mata uang tadi simetris, dalam arti tidak cekung atau cembung, dan tidak berdiri pada pinggirannya setelah dilemparkan, kita katakan nilai kemungkinan untuk mendapatkan sisi Gambar, sama dengan nilai kemungkinan untuk mendapatkan sisi Angka, yaitu masing – masing $\frac{1}{2}$. Kita singkat dengan : P(gambar) = P(angka) = $\frac{1}{2}$. P adalah singkatan dari *peluang, probabilitas* atau *nilai kemungkinan.*

1.1.2 Pengertian Probabilitas

Definisi Klasik

Misalkan kejadian A dapat terjadi dalam *p* cara dari seluruh n cara yang mungkin, dan *n* cara ini *mempunyai kemungkinan sama* (*equally likely)*, maka *probabilitas* kejadian A = P(A) = $\frac{p}{n}$.

Contoh 1.1

Sebuah dadu dilemparkan. Ada 6 cara yang mungkin, yaitu hasil lemparan adalah sisi 1, 2, 3, 4, 5 atau 6, sehingga P(1)=P(2)=P(3)=P(4)=P(5)=P(6) = $\frac{1}{6}$.

Jika A adalah kejadian jatuh pada sisi genap, maka kejadian tersebut dapat terjadi dengan 3 cara, yaitu 2, 4, atau 6; jadi P(A) = $\frac{3}{6} = \frac{1}{2}$. Begitu juga dengan B yang merupakan kejadian jatuh hasil ≤ 2, dapat terjadi dengan 2 cara yaitu 1, 2 sehingga P(B) = $\frac{2}{6} = \frac{1}{3}$

Contoh 1.2

Pandanglah 9 angka 1, 2, 3, 4, 5, 6, 7, 8, dan 9. Suatu bilangan diambil secara *acak* atau *random* , yaitu setiap angka mempunyai kemungkinan terambil yang sama. Berapa nilai kemungkinannya mendapat angka yang habis dibagi 3 ?

Jawab :

Ada 9 cara yang mungkin untuk mendapat satu angka, yaitu 1, atau 2, atau 3, dst sampai dengan 9. Dari 9 cara yang mungkin ini hanya ada 3 cara untuk mendapatkan bilangan yang habis dibagi 3.

Jadi P (bilangan yang habis dibagi 3) = $\frac{3}{9} = \frac{1}{3}$

1.2. Hukum-Hukum Probabilitas

Pandang sebuah eksperimen lemparan 2 mata uang logam , Rp. 500 dan Rp. 1000. Semua hal yang mungkin terjadi (*possible cases)* adalah (A,A), (A,G), (G,A), (G,G). Dikatakan semua hal yang mungkin ini membentuk *ruang sampel* atau *ruang kemungkinan*. Dalam statistika, ini disebut *populasi.* Satu hal dari ruang sampel diatas disebut *titik sampel* , sehingga ruang sampel ini mempunyai 4 titik sampel.

Pada teori himpunan, ruang sampel disebut *himpunan semesta.* Biasanya ruang sampel dinyatakan dengan huruf besar S, misalnya,

S = { (A,A), (A,G), (G,A), (G,G) }

Satu titik sampel adalah satu unsur himpunan semesta. Satu *kejadian (event)* merupakan satu himpunan bagian (*subset)* dari ruang sampel. Jadi, kejadian A (hasil tos sama) adalah {(A, A), (G, G)}; A = {(A, A), (G, G)}; dan kejadian B (hasil tos berbeda) adalah {(A,G), (G,A)}, B = {(A, G), (G, A)};

Nilai kemungkinan mempunyai interval $0 \leq P(A) \leq 1$

Contoh 1.3

Eksperimen lemparan 2 buah dadu, yang satu disebut dadu putih, yang kedua disebut dadu merah, menghasilkan ruang sampel S yang terdiri atas 6 x 6 = 36 titik.

	1	2	3	4	5	6
1	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
2	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
3	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
4	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

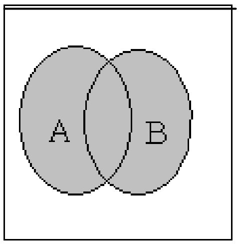
Semua titik mempunyai nilai kemungkinan yang sama, karena kita misalkan tiap – tiap dadu bentuknya simetris dan sudah tentu hasil yang putih tidak mempengaruhi hasil dadu merah; Jadi, P(tiap titik) = $\frac{1}{36}$. Misalkan kejadian A adalah hasil jumlah 7, maka A menyatakan himpunan bagian A = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)} dan P(A) = jumlah semua ukuran titik di A = 6 x $\frac{1}{36} = \frac{1}{6}$.

Kejadian B menyatakan hasil jumlah 8, maka P(B) = $\frac{5}{36}$

1.2.1 Gabungan dan Irisan Kejadian

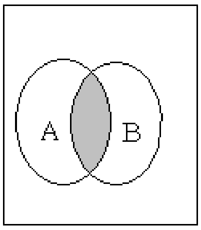
Misalkan diketahui 2 kejadian A dan B yang kita lukiskan dengan diagram Venn.

Untuk kejadian “A atau B” dilambangkan dengan simbol $A \cup B$. Arti “atau” disini adalah A atau B atau keduanya. Jadi $A \cup B$ adalah kumpulan semua titik yang berada di A atau pada B atau pada keduanya (Gambar 1.1)



Gambar 1.1 $A \cup B$

Untuk kejadian majemuk “A dan B”, berarti A terjadi dan juga bersamaan B terjadi, yang dinyatakan dengan simbol $A \cap B$. $A \cap B$ adalah kumpulan semua titik yang ada pada A dan juga pada B (Gambar 1.2)



Gambar 1.2. A $\cap$ B

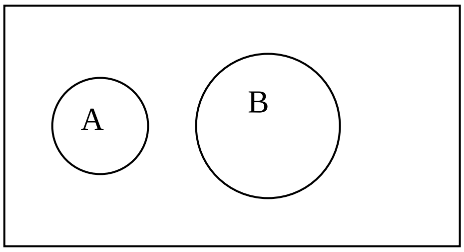
A $\cup$B juga disebut *A gabungan B* atau *A union B*, sedangkan A $\cap$B disebut *A irisan B* atau *A intersection B.*

Jika kita akan menghitung P(A $\cup$B), dari diagram Venn dapat dilihat bahwa P(A) + P(B) – P(A $\cap$B) = P(A $\cup$B)

$$\boxed{\mathbf{P(A \cup B) = P(A) + P(B) - P(A \cap B)}}$$

1.2.2. Kejadian yang Saling Bertentangan

Dua kejadian yang tidak dapat terjadi bersamaan disebut *saling bertentangan.* Dalam teori himpunan A dan B disebut *saling lepas (disjoint).* Pada diagram Venn 2 kejadian yang saling bertentangan atau 2 kejadian yang saling lepas dapat dilukiskan seperti Gambar. 1.3 dibawah ini



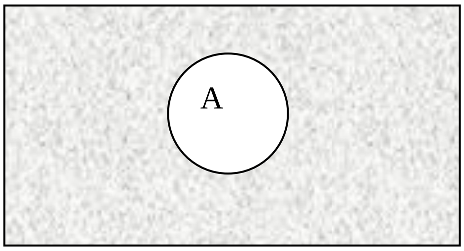
Gambar 1.3 A $\supset\subset$ B

Dengan begitu P(A $\cap$B) = Ǿ (himpunan kosong)

Untuk dua kejadian yang saling lepas berlaku : **P(A$\cup$B) = P(A) + P(B)**

1.2.3. Kejadian Komplementer

Kejadian A dan kejadian Ã (bukan A), yang terdiri dari semua titik di S yang tidak di A, disebut kejadian komplementer



Gambar 1.4 $\overline{A}$

A dan $\overline{A}$ tidak mempunyai persekutuan, sehingga A dan $\overline{A}$ adalah dua himpunan yang saling lepas dan kejadian A dan kejadian $\overline{A}$ saling bertentangan.

Maka $A \cup \overline{A} = S$.

$P(S) = P(A \cup \overline{A}) = P(A) + P(\overline{A})$, dan $P(S) = 1$, sehingga

$$\boxed{\mathbf{P(\overline{A}) = 1 - P(A)}}$$

1.2.4. Kejadian Yang Saling Bebas

Kejadian yang saling bebas adalah kejadian yang tidak bersangkut paut.

Definisi

Kejadian A dan B disebut bebas, jika dan hanya jika, $P(A \cap B) = P(A)\ \mathbf{.}\ P(B)$

Jika syarat ini tidak dipenuhi, A dan B disebut *saling bergantung.*

Contoh 1.4

Dalam eksperimen lemparan dua buah dadu, apakah kejadian C (merah + putih = 11) dan kejadian D (merah $\neq$ 5) saling bebas ?

Jawab :

$C = \{(5, 6), (6, 5)\}$, $\quad P(C) = \dfrac{1}{18}$

$\bar{D} = \{(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5)\}$, $P(\bar{D}) = \dfrac{1}{6}$

Jadi $P(D) = 1 - P(\bar{D}) = \dfrac{5}{6}$

$C \cap D = \{(5,6)\}$, $\quad P(C \cap D) = \dfrac{1}{36}$

Ternyata $P(C \cap D) \neq P(C) . P(D)$

Jadi C dan D saling bergantung.

1.2.5 Probabilitas Bersyarat

Definisi

Nilai kemungkinan bersyarat kejadian A jika kejadian B diketahui, ditulis dengan P(A|B) dan ditentukan oleh

$$\boxed{\mathbf{P(A|B) =} \frac{P(A\cap B)}{P(B)}} \quad , \text{dimana } P(B) \neq 0$$

Jika kedua ruas dikalikan dengan P(B), maka $\boxed{\mathbf{P(A\cap B) = P(A|B) .P(B)}}$

Karena $A\cap B = B\cap A$, maka **$P(A\cap B) = P(B|A) \,.\, P(A)$**

Rumus tersebut dapat diperluas untuk irisan 3 kejadian, sehingga

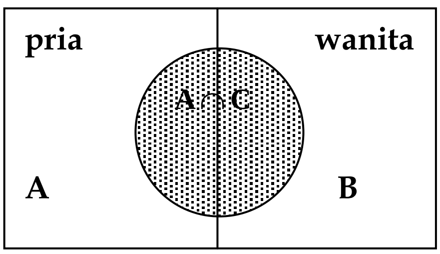
$$\boxed{\mathbf{P(A\cap B\cap C) = P(A|B\cap C) \,.\, P(B|C \,.\, P(C)}}$$

Contoh 1.5

Misalkan 5% pria dan 1% wanita buta warna, dan masing – masing merupakan 50% dari populasi. Seorang peneliti mengambil seorang buta warna secara acak. Berapa nilai kemungkinan ia (a) pria, (b) wanita ?

Jawab :

Misalkan diagram Venn untuk kasus di atas adalah sebagai berikut :



Gambar 1.5

	C	N	
A	0.025	0.475	0.500
B	0.005	0.495	0.500
	0.030	0.970	1.000

Dapat dilihat bahwa P(A)=P(B) = $\frac{1}{2}$, P(C) = 0.030 , dan P(N) = 0.970

Nilai kemungkinan untuk pria, jika diketahui buta warna, ditulis sebagai P{pria|buta warna} = P(A|C) = $\frac{0.025}{0.030} = \frac{5}{6}$

P(A|C) disebut nilai kemungkinan bersyarat A, jika C diketahui.

Dapat dilihat bahwa P(A|C) = $\frac{0.025}{0.030} = \frac{P(A \cap C)}{P(C)}$

Demikian P(B|C) = $\frac{0.005}{0.030} = \frac{P(B \cap C)}{P(C)}$

Nilai kemungkinan bersyarat untuk kejadian yang saling bebas :

Jika A dan B adalah 2 kejadian yang bebas dengan nilai kemungkinan bukan nol, maka

P(A|B) = P(A) dan P(B|A) = P(B)

1.3 Formulasi Bayes

Formulasi Bayes adalah pengembangan dari probabilitas bersyarat (*conditional probability*) dan aturan umum hukum perkalian. Andaikan terdapat sekelompok peristiwa B_1, B_2, …, B_n yang *mutually exlusive* dan *exhaustive*, artinya masing-masing peristiwa tidak memiliki keluaran yang sama dan secara bersama-sama memuat keseluruhan keluaran didalam ruang sampel. Secara matematis, peristiwa B_i yang *mutually exclusive* bersifat *exhaustive* jika dan hanya jika :

$$\sum_{i=1}^{n} P(B_i) = 1$$

Sebagai tambahan, terdapat sebuah peristiwa lain A yang didefinisikan pada ruang sampel yang sama. Karena peristiwa-peristiwa B_i bersifat exhaustive maka peristiwa A pasti beririsan dengan satu atau lebih peristiwa B_i. Oleh karena itu, satu cara untuk mendapatkan probabilitas peristiwa A adalah dengan menjumlahkan probabilitas P ($A \cap B_i$) untuk seluruh harga i. Maka :

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$

Sekarang jika persoalan diubah dan diasumsikan bahwa peristiwa A telah terjadi. Bagaimana menentukan probabilitas masing-masing peristiwa B_i juga terjadi ?

$$P(B_i \mid A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(B_i).P(A|B_i)}{\sum_{j=1}^{n} P(B_j \cap A)} = \frac{P(B_i).P(A|B_i)}{\sum_{j=1}^{n} P(B_i).P(A|B_i)}$$

Contoh 1.6

Supplier I, II, III, dan IV menyediakan seluruh keperluan telur bagi Toko Roti Swiss sebanyak masing-masing 25%, 35%, 10 % dan 30%. Dari pengalaman selama ini diketahui bahwa supplier I, II, III, dan IV masing-masing mengirimkan 20%, 5%, 30%, dan 10% telur yang rusak. Maka probabilitas bahwa sebutir telur yang dipilih secara acak merupakan telur yang rusak dapat dihitung sebagai berikut :

Misalkan A adalah peristiwa ditemukannya sebutir telur yang rusak, dan B_1, B_2, B_2 dan B_4 adalah peristiwa pemilihan telur yang dikirim oleh supplier I, II, III, dan IV. Maka :

$$P(A) = \sum_{i=1}^{4} P(A \cap B_i) = \sum_{i=1}^{4} P(A \mid B_i) \cdot P(B_i)$$

$$= P(A \mid B_1).P(B_1)+P(A \mid B_2).\ P(B_2) + P(A \mid B_3)\ .P(B_3)+P(A \mid B_4)\ .\ P(B_4)$$

$$= (0,2)\ (0,25) + (0,05)\ (0,35) + (0,3)\ (0,1) + (0,1)\ (0,3)$$

$$= 0,1275$$

Kemudian jika ditemukan ada sebutir telur yang rusak, maka probabilitas telur rusak itu berasal dari supplier III adalah :

$$P(B_i \mid A) = \frac{P(B_3 \cap A)}{P(A)} = \frac{P(B_i).P(A|B_3)}{P(A)} = \frac{(0,3)(0,1)}{0,1275} = \frac{0,03}{0,1275} = 0,2353$$

1.4 Analisa Kombinatorial

Untuk dapat menghitung nilai kemungkinan dari kejadian yang lebih majemuk, digunakan *analisa kombinatorial*

Simbol untuk *n faktorial* adalah **n!**

Definisi : *n faktorial* = **n!** = 1 . 2 . ... (n-1).(n).

Jadi : 2! = 1 . 2 = 2

3! = 1 . 2 . 3 = 6

4! = 1 . 2 . 3 . 4 = 24

5! = 1 . 2 . 3 . 4. 5 = 120

Catatan : untuk 1! = 0! = 1

1.4.1 Permutasi

Pandang 3 unsur yang berlainan, misal *a*, *b*, dan *c*. Kita dapat mengurut-kannya sebagai *abc*, *acb*, *bac*, *bca*, *cba*, dan *cab*. Tiap urutan disebut dengan permutasi 3 unsur dari ketiganya dinyatakan dengan simbol ${}_3P_3$ dan ${}_3P_3 = 6$. Jika hanya diambil 2 unsur saja, kita mendapatkan permutasi *ab*, *ba*, *ac*, *ca*, *bc*, dan *cb*. Banyaknya permutasi 3 unsur diambil dari ${}_3P_2 = 6$

Definisi :

Suatu permutasi *r* unsur, yang diambil dari *n* unsur yang berlainan, yiatu penempatan *r* unsur itu dalam satu urutan ($r \le n$)

$$\boxed{\begin{array}{l} {}_nP_r = n\,(n-1)\,(n-2)\ldots(n-r+1) = \dfrac{n!}{(n-r)!} \\ {}_nP_n = n\,! \end{array}}$$

1.4.2. Kombinasi

Pandang 3 unsur *a*, *b* dan *c* . Sekarang diambil 2 unsur tanpa mengindahkan urutannya, jadi *ab* sama dengan *ba*, *ac* sama dengan *ca*. Pilihan *ab*, *ac*, dan *bc* adalah 3 kombinasi 3 unsur diambil 2.

Definisi :

Suatu kombinasi *r* unsur yang diambil dari *n* unsur yang berlainan, adalah suatu pilihan dari *r* unsur tanpa memperhatikan urutannya ($r \le n$)

Banyaknya kombinasi 3 unsur diambil 2 dinyatakan dengan simbol ${}_3C_{2,}$ atau $\binom{3}{2} = 3$. Jika dari 3 unsur di atas, diambil 1 unsur saja, akan ada kombinasi, yaitu *a*, *b* dan *c*. Jadi ${}_3C_1 = 3$

$$\boxed{{}_nC_r = \binom{n}{r} = \frac{n\,\mathrm{Pr}}{r!} = \frac{n!}{r!(n-r)!}}$$

Contoh 1.7

$${}_3C_2 = \frac{3!}{2!\ 1!} = \frac{6}{2\ .1} = 3$$

$${}_4C_2 = \frac{4!}{2!\ 2!} = \frac{24}{2\ .2} = 6$$

Contoh Pemakaian Permutasi dan Kombinasi

1. Ada 5 buku : A= Aljabar, B= Biologi, C = Kimia, D = Jerman, E = Inggris. Jika kita mengambil 3 buku kemudian mengurutkannya pada rak buku, ada berapa permutasikah ?

   Jawab :
   Banyaknya permutasi 5 buku yang berlainan diambil 3, adalah ${}_5P_3 = \frac{5!}{2!} = 60$

2. Kita hendak mengirimkan surat per pos dan biayanya Rp 2300. Kantor pos memberikan 4 perangko yang berlainan, yaitu 100, 300, 800, 1100. Dengan berapa permutasi kita dapat menempelkan 4 perangko ini pada surat kita ?
   Jawab :
   Banyaknya permutasi 4 unsur diambil 4 unsur adalah ${}_4P_4 = 24$

3. Ada berapa cara satu panitia terdiri atas 3 orang dapat dipilih dari 4 pasangan suami istri.
   a. Jika semua orang ini dipilih
   b. Jika panitia ini harus terdiri dari 2 pria dan 1 wanita ?

   Jawab :

   a. Ada 8 orang dan diambil 3; jadi ada () = $\frac{8!}{3!\,5!} = \frac{6.7.8}{1.2.3} = \frac{336}{6} = 56$

   b. Dua pria dipilih dari 4 suami, jadi ada $\binom{4}{2} = \frac{4!}{2!\,2!} = 6$ cara

      dan seorang wanita dari 4 istri, jadi ada $\binom{4}{1} = \frac{4!}{1!\,3!} = 4$ cara

      sehingga ada 6 x 4 = 24 cara

1.5 Latihan Soal

1. Dua mata uang dilempar, berapa nilai kemungkinan untuk mendapat hasil satu gambar dan satu angka ?
2. Dua dadu dilempar, berapa nilai kemungkinan untuk mendapatkan jumlah mata dadu 7 ?
3. Satu baskom berisi 3 bola putih dan 5 bola hitam. Berapa probabilitas untuk mendapat bola putih, jika kita mengambil satu bola secara acak ?
4. Satu baskom berisi 10 bola putih dan 15 bola merah. Akan kita ambil 4 bola secara acak. Berapa probabilitas untuk mendapatkan :
   a. semua putih
   b. semua merah
   c. dua putih dan 2 merah
   d. satu putih dan 3 merah
5. Seorang pegawai perpustakaan akan mengatur 8 buku di rak buku, 3 diantaranya berwarna merah. Ada berapa carakah pegawai tersebut dapat mengatur buku itu bila :
   a. 3 buku yang berwarna merah selalu berdampingan
   b. 3 buku yang berwarna merah itu tidak boleh bersebelahan
6. Tentukanlah jumlah susunan huruf yang dapat dibentuk dengan menggunakan huruf dalam kata “PERMUTASI”, apabila :

a. semua huruf dipakai tanpa syarat
b. semua huruf dipakai tetapi semua huruf hidup berdampingan
c. semua huruf dipakai tetapi 3 huruf tertentu selalu berdampingan

7. Berapakah jumlah plat mobil dapat dibuat dengan menggunakan 2 huruf, kemudian 4 angka, dan terakhir 1 huruf ?
8. Sepasang pengantin baru mengatakan bahwa mereka ingin mempunyai 3 anak dari pernikahannya. Bila keinginannya terpenuhi, tentukanlah nilai kemungkinan bahwa anaknya :
   a. wanita semua
   b. pria semua
   c. satu pria dan 2 wanita
9. Suatu panitia terdiri dari 10 orang. Mereka mengadakan rapat dan mengelilingi meja bundar. Ada berapa carakah mereka menempati 10 kursi yang tersedia mengelilingi meja itu ?
10. Sebuah kota kecil mempunyai satu unit pemadam kebakaran dan satu ambulance yang tersedia dalam keadaan darurat. Probabilitas bahwa unit pemadam kebakaran akan siap diperlukan adalah 0,98 dan probabilitas bahwa ambulance siap bila dipanggil adalah 0,92. Dalam peristiwa terbakarnya sebuah gedung di kota itu,
    a. berapa probabilitasnya keduanya akan siap beroperasi
    b. berapa probabilitasnya bahwa yang siap beroperasi hanya salah satu
11. Jika diketahui probabilitas bahwa Amir masih hidup 20 tahun lagi adalah 0,7 dan probabilitas bahwa Badu masih hidup 20 tahun lagi adalah 0,9.
    a. berapa probabilitasnya bahwa keduanya tidak hidup 20 tahun lagi
    b. berapa probabilitasnya bahwa hanya Amir yang masih hidup 20 tahun lagi
12. Seorang mahasiswa mempunyai peluang untuk lulus dalam mata kuliah A, B, C dan D masing-masing adalah 0,6 ; 0,8 ; 0,7 dan 0,9. Jika kelulusan dalam suatu mata kuliah tidak tergantung dari kelulusan mata kuliah lainnya, hitunglah probabilitasnya bahwa mahasiswa tersebut :
    a. Lulus dalam keempat mata kuliah
    b. Tidak lulus dalam empat mata kuliah
    c. Lulus dalam dua mata kuliah
13. Probabilitas seorang dokter akan mendiagnosis dengan tepat adanya suatu penyakit tertentu adalah 0,7. Pada peristiwa seorang dokter membuat kesalahan dalam pemeriksaan suatu penyakit (diagnosis tidak tepat), probabilitas bahwa

pasien akan mengajukan tuntutan adalah 0,9. Hitunglah probabilitasnya bahwa dokter tersebut membuat diagnosis yang tidak tepat dan pasien tersebut mengajukan tuntutan.

14. Pengantin baru mengatakan bahwa mereka menginginkan 3 orang anak dari pernikahannya. Bila keinginannya terpenuhi, tentukanlah nilai kemungkinan bahwa anaknya
    a. wanita semua;
    b. satu pria dan dua wanita;
    c. pria semua
15. Suatu Biro Perjalanan Wisata menawarkan 3 armada barunya. Ternyata diketahui bahwa 25% orang memilih Bus Safari, 30% orang memilih Bus Nusantara dan sisanya memilih Bus Rahayu. Berdasarkan pengalaman diketahui bahwa dari ke-3 armada tersebut, 4% dari Bus Safari kemungkinan rusak armadanya, 6% dari Bus Nusantara kemungkinan rusak armadanya dan 8% dari Bus Rahayu kemungkinan rusak armadanya. Berdasarkan data-data diatas, tentukan :
    a. Berapa probabilitas armada yang rusak ?
    b. Jika ternyata ditemukan ada armada yang rusak, berapa probabilitas bahwa armada tersebut adalah Bus Rahayu

BAB 2

DISTRIBUSI PROBABILITAS

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa dapat menentukan jenis distribusi probabilitas yang tepat untuk suatu masalah beserta pemakaian rumus dan tabel yang sesuai.

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa dapat memahami pembentukan distribusi probabilitas, penentuan peubah acak dan pemanfaatan harapan matematis, membedakan antara distribusi peluang diskrit dan distribusi peluang kontinu, memahami masalah Binomial, membaca tabel binomial, memahami masalah Poisson, menentukan nilai variabel dan memahami manfaatnya lebih lanjut serta dapat menghitung luas kurva dengan tabel Z, dan dapat membedakan pemakaian distribusi normal dengan distribusi student serta menghitung probabilitas kejadian dengan distribusi tersebut.

2.1. Variabel Acak

Eksperimen probabilitas memiliki keluaran yang bisa berupa suatu nilai numerik (angka/bilangan), suatu cacahan/hitungan, atau suatu hasil pengukuran. Variabel Acak (*random variable*), biasa ditandai dengan sebuah simbol seperti X, adalah variabel yang memiliki sebuah nilai numerik tunggal untuk setiap keluaran dari sebuah eksperimen probabilitas. Jadi X dapat bernilai angka berapapun tergantung pada keluaran yang mungkin dihasilkan dari eksperimen. Dengan kata lain, nilai tertentu dari X dalam sebuah eksperimen adalah suatu kemungkinan keluaran yang acak.

Variabel acak diskrit adalah variabel acak yang memiliki nilai yang dapat dicacah. Sementara variabel acak kontinu memiliki nilai yang tak terhingga banyaknya sepanjang sebuah interval yang tak terputus. Variabel acak kontinu biasanya diperoleh dari hasil pengukuran

Contoh 2.1. :

Saat kita melakukan pelemparan dua buah uang logam , maka kejadian yang terjadi adalah : GG, AA, GA, AG.

Jika X menyatakan banyak sisi G yang muncul, maka X = 0, 1, 2.

Simbol X diatas disebut variable acak diskrit

2.2 Distribusi Probabilitas Diskrit

Jika pada sebuah eksperimen probabilitas didaftarkanlah seluruh keluaran yang mungkin dari variabel diskrit X, yaitu x_1, x_2, x_3, …, x_n, dan kemudian didaftarkan pula nilai probabilitas yang berkaitan dengan keluaran tersebut, yaitu $P(X=x_1)$, $P(X=x_2)$, $P(X=x_3)$, …, $P(X=x_n)$, (biasa dinotasikan juga dengan $p(x_1)$, $p(x_2)$,$p(x_3)$, …, $p(x_n)$) maka telah dibentuk suatu distribusi probababilitas diskrit dari variabel X.

Pernyataan matematis $P(X=x) = p(x)$, yang dibaca “ probabilitas X yang menyandang nilai x”, disebut fungsi probabilitas dari variabel acak X. Perlu dicatat bahwa p(x) didefinisikan bernilai nol untuk setiap nilai x yang tidak berkaitan dengan suatu keluaran yang ada dalam ruang sampel.

Terdapat dua aturan yang harus dipenuhi :

1. Nilai-nilai dari suatu fungsi probabilitas adalah angka-angka yang berada dalam interval antara 0 dan 1. Jadi nilai-nilai dari fungsi yang mungkin akan selalu berada dalam interval $0 \leq p(x) \leq 1$
2. Jumlah seluruh nilai fungsi probabilitas adalah 1. Jadi $\sum p(x) = 1$

2.2.1. Distribusi Binomial

Distribusi Binomial adalah salah satu distribusi probabilitas diskrit yang paling sering digunakan dalam analisis statistika modern. Di bidang teknik, distribusi ini erat kaitannya dengan pengendalian kualitas (*quality control*).

Pandang sebuah eksperimen yang hanya menghasilkan dua kejadian A dan $\overline{A}$, dengan P(A) = p = peluang terjadinya kejadian A.

Jika pada tiap percobaan dalam eksperimen itu p = P(A) berharga tetap, maka percobaan yang berulang-ulang dari eksperimen itu disebut *percobaan Bernoulli.*

Sekarang lakukan percobaan Bernoulli n kali, X diantaranya menghasilkan A sisanya yaitu (n-X) menghasilkan $\overline{A}$

Jika p = P(A), maka P($\overline{A}$) = 1 – p = q

Maka peluang terjadinya kejadian A sebanyak X = x kali diantara n, dihitung dengan rumus :

$$P(x) = P(X=x) = nCx\ p^{x}\ (1 - p)^{n-x}$$

dimana x = 0, 1, 2, …, n , $0 \le p \le 1$

dan

$$nCx = \frac{n!}{x!(n-x)!}$$

Distribusi binom ini mempunyai parameter diantaranya rata-rata μ dan simpangan baku σ. Rumusnya adalah :

$$\mu = n\,p$$

$$\sigma = \sqrt{n\,p\,q}$$

Contoh 2.2 :

1. Peluang untuk mendapatkan 6 sisi G ketika melakukan undian dengan sebuah mata uang sebanyak 10 kali adalah :

   $$P(x=6) = {}_{10}C_6\,(\tfrac{1}{2})^6\,(\tfrac{1}{2})^4 = \frac{10!}{6!4!}(\tfrac{1}{2})^{10} = \frac{7.8.9.10}{1.2.3.4}\,(\tfrac{1}{2})^{10}$$

   $$= 210.\,(/\tfrac{1}{2})^{10}$$

   $$= 0{,}2050$$

2. Lakukan undian dengan 10 dadu sekaligus. Berapa peluang yang muncul mata 6 sebanyak 8 buah.

   Kita tahu bahwa p = P(mata 6) = $\frac{1}{6}$, n = 10, x=8, maka

   $P(x=8) = {}_{10}C_8 \left(\frac{1}{6}\right)^8 \left(\frac{5}{6}\right)^2 = 0{,}000015$

3. 10% dari suatu benda tergolong dalam kategori A.

   Sebuah sample berukuran 30 telah diambil secara acak. Berapa peluang sample itu akan berisi benda kategori A.

   a) Semuanya
   b) Sebuah
   c) Dua buah
   d) Paling sedikit sebuah
   e) Paling banyak 2 buah
   f) Tentukan rata-rata terdapatnya kategori A

   Jawab :

   Misal x = banyak benda kategori A.

   Maka peluang p = peluang benda termasuk katergori A = 0,1

   a) Semua masuk kategori A berarti X = 30

      $P(x=30) = {}_{30}C_{30} (0{,}10)^{30} (0{,}90)^{0} = 10^{-30}$

   b) X = 1

      $P(x=1) = {}_{30}C_{1} (0{,}10)^{1} (0{,}90)^{29} = 0{,}1409$

   c) x = 2

      $P(x=2) = {}_{30}C_{2} (0{,}10)^{2} (0{,}90)^{28} = 0{,}2270$

   d) Paling sedikit 1, berarti x = 1, 2, 3, …, 30.

      Jadi perlu P(x=1) + P(x=2) + … +P(x=30)

      Karena P(x=1) + P(x=2) + … +P(x=30) = 1,

      maka yang dihitung adalah 1 – P(x=0).

   $P(x=0) = {}_{30}C_{0} (0{,}10)^{0} (0{,}90)^{30} = 0{,}0423$, maka 1–0,0423 = 0,9577

e) Terdapat paling banyak 2 buah kategori A, berarti X = 0, 1, 2

$P(x \le 2) = P(x=0) + P(x=1) + P(x=2)$

$= 0{,}9577 + 0{,}1409 + 0{,}2270$

$= 0{,}4102$

f) $\mu = N\pi$

$= 30 \,.\, 0{,}1$

$= 3$

2.2.2 Distribusi Poisson

Variabel acak diskrit X dikatakan mempunyai *distribusi poisson* jika fungsi peluangnya berbentuk : $p(x) = P(X=x) = \dfrac{e^{-\lambda}\lambda^{x}}{x!}$

dengan x = 0, 1, 2, … dan e = 2,7183

Harga-harga $e^{-\lambda}$ dapat dilihat pada lampiran.

Parameter distribusi Poisson

$\mu = \lambda$

$\sigma = \sqrt{\lambda}$

Distribusi ini sering digunakan untuk menentukan peluang sebuah kejadian yang dalam area kesempatan tertentu diharapkan sangat jarang terjadi.

Contoh 2.3 :

Misal rata-rata ada 1,4 orang buta huruf setiap 100 orang. Diambil sebuah sampel berukuran 200 orang .

Jika x = banyak buta huruf per 200 orang, maka λ = 2,8.

Peluang tidak terdapat yang tidak buta huruf adalah :

$$p(0) = \frac{e^{-2,8}(2{,}8)^{0}}{0!} = e^{-2,8} = 0{,}0608$$

Distribusi Poisson ini dapat dianggap sebagai pendekatan kepada distribusi Binom

2.3 Distribusi Probabilitas Kontinu

Secara teoritis untuk variabel acak kontinu, kurva distribusi probabilitas populasi diwakili oleh poligon frekuensi relatif yang dimuluskan. Kurva ini dapat dinyatakan oleh suatu fungsi kontinu, misalnya f(x). Fungsi f(x) ini disebut sebagai fungsi kepadatan probabilitas (probability density function)

Secara imajinatif dapat dibayangkan bahwa fungsi kepadatan probabilitas menggambarkan besarnya probabilitas perunit interval nilai variabel acaknya

Secara matematik ini dinotasikan sebagai :

$$P\ (\ a \le X \le b) = p\ (a \le x \le b) = \int_a^b f(x)dx$$

Dengan demikian secara umum probabilitas sebuah variabel acak kontinu X mengambil nilai pada suatu interval antara x dan x + dx dapat dinyatakan dalam suatu notasi matematika

$$P\ (\ x \le X \le x{+}dx) = \int f(x)dx$$

Dari uraian diatas dapat dipahami bahwa agar sebuah fungsi f(x) dapat menjadi sebuah

2.3.1 Distribusi Normal

Distribusi normal adalah sebuah distribusi yang paling luas penggunaannya.

Karakterisik Distribusi Peluang Normal

1. Bentuk kurva normal seperti bel dan simetris.
2. Parameter σ, menunjukkan lebar dari kurva normal (semakin besar nilainya, semakin lebar).
3. Titik tertinggi dari kurva nomal terletak pada nilai rata-rata = median = modus.
4. Luas total area di bawah kurva normal adalah 1. (luas bagian di sebelah kiri μ = sebelah kanan μ).
5. Peluang suatu variabel acak normal sama dengan luas di bawah kurva normal.

Persamaan distribusi normal tergantung pada 2 parameter, yaitu μ dan σ. Persamaanya sebagai berikut :

$$f(x) = \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Dimana : μ = rata-rata (*mean*)

σ = simpangan baku (*standard deviation*)

π = 3.14159

e = 2.71828

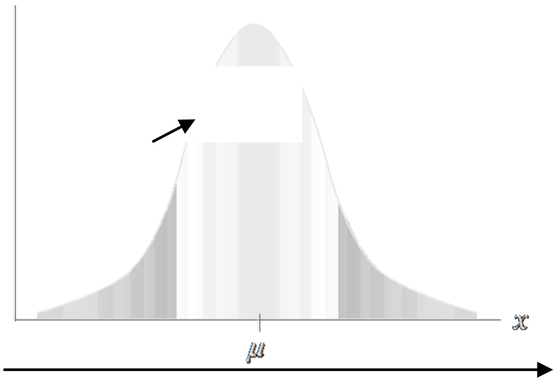
Distribusi normal dengan rata-rata $\mu = 0$ dan σ = 1 adalah distribusi normal standar . Fungsi densitasnya berbentuk :

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad \text{dimana} \quad -\infty < z < \infty$$

dengan menggunakan transformasi, rumus dapat diubah menjadi :

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

Jika digambarkan sebagai berikut :



Persentase nilai pada interval yang sering digunakan

1.**68,26%** nilai dari suatu variabel acak normal berada pada interval $\mu \pm \sigma$

2.**95,46%** nilai dari suatu variabel acak normal berada pada interval $\mu \pm 2\sigma$

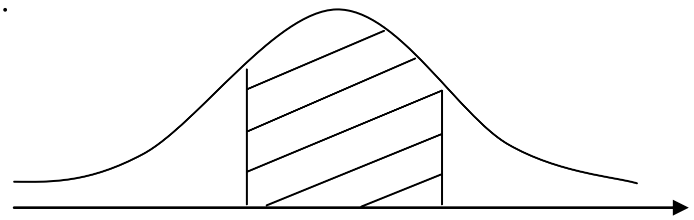
3.**99,74%** nilai dari suatu variabel acak normal berada pada interval $\mu \pm 3\sigma$

Untuk mencari peluang sebuah interval pada distribusi normal, maka fungsi distribusi itu harus diintegralkan dengan batas-batas peluang

$$p(x_1 < x < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= F(x_2) - F(x_1)$$

Atau luas daerah yang dibawah kurva dengan batasan dari x_1 sampai x_2 seperti berikut :



Contoh 2.4 :

Sebuah populasi dengan rata-rata $\mu = 12$ dan $\sigma = 2$

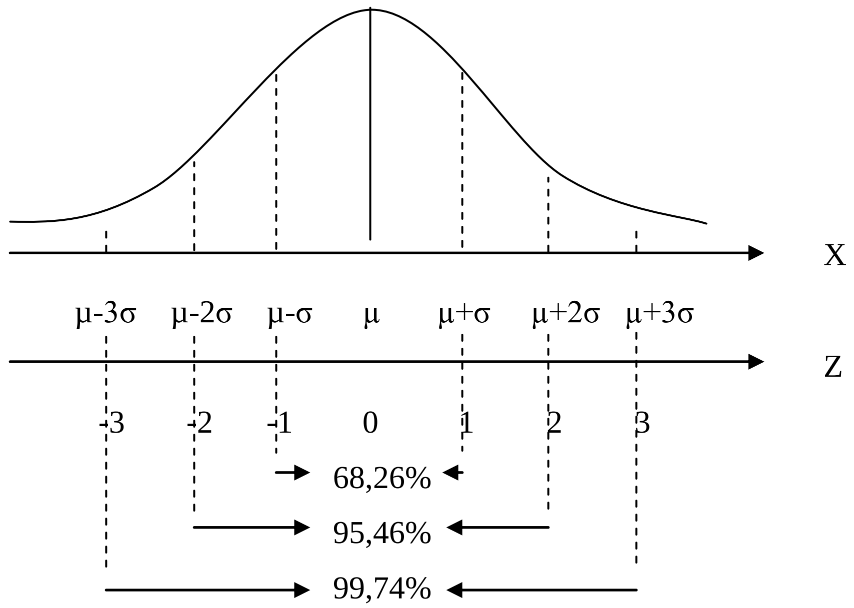
Peluang x akan berada antara 11 – 14 adalah :

$$p(11 < x < 14) = \int_{11}^{14} \frac{1}{2\sqrt{2\Pi}} e^{-\frac{1}{2}\left(\frac{x-12}{2}\right)^2}$$

Untuk mempermudah perhitungan itu, maka variabel x di transformasi menjadi angka baku z, dimana

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

atau jika di gambarkan :



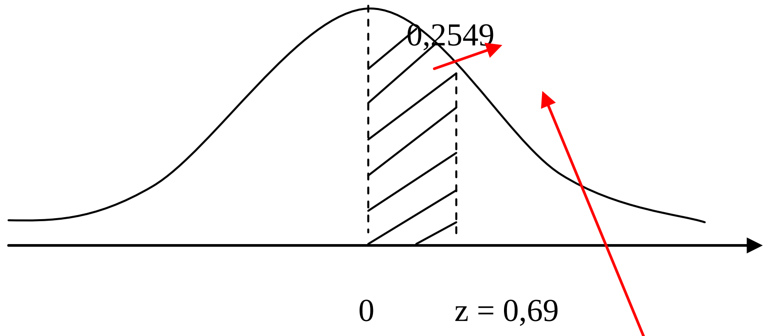
Karena nilai-nilai antara x_1 dan x_2 ditransformasikan ke z_1 dan z_2, maka luas daerah antara x_1 dan x_2 sama dengan luas daerah z_1 dan z_2 yang dapat dilihat dalam tabel Normal Baku.

Keterangan :

Angka yang tertera di tabel adalah luas daerah di bawah kurva dengan batasan dari 0 sampai titik z tertentu

Contoh 2.5 :

P (0 < z < 0,69)



z	**,00**	**,01**	**,02**	**,03**	**,04**	**,05**	**,06**	**,07**	**,08**	**,09**
,0	,0000	,0040	,0080	,0120	,0160	,0199	,0239	,0279	,0319	,0359
,1	,0398	,0438	,0478	,0517	,0557	,0596	,0636	,0675	,0714	,0753
,2	,0793	,0832	,0871	,0910	,0948	,0987	,1026	,1064	,1103	,1141
,3	,1179	,1217	,1255	,1293	,1331	,1368	,1406	,1443	,1480	,1517
,4	,1554	,1591	,1628	,1664	,1700	,1736	,1772	,1808	,1844	,1879
,5	,1915	,1950	,1985	,2019	,2054	,2088	,2123	,2157	,2190	,2224
,6	,2257	,2291	,2324	,2357	,2389	,2422	,2454	,2486	,2518	,2549
,7	,2580	,2612	,2642	,2673	,2704	,2734	,2764	,2794	,2823	,2852
,8	,2881	,2910	,2939	,2967	,2995	,3023	,3051	,3078	,3106	,3133
,9	,3159	,3186	,3212	,3238	,3264	,3289	,3315	,3340	,3365	,3389

Jadi peluang z berada diantara 0 sampai 0,69 adalah 0,2549

Untuk mencari nilai-nilai yang lain dapat dilakukan cara sebagai berikut :

1. P (z > 0,69)

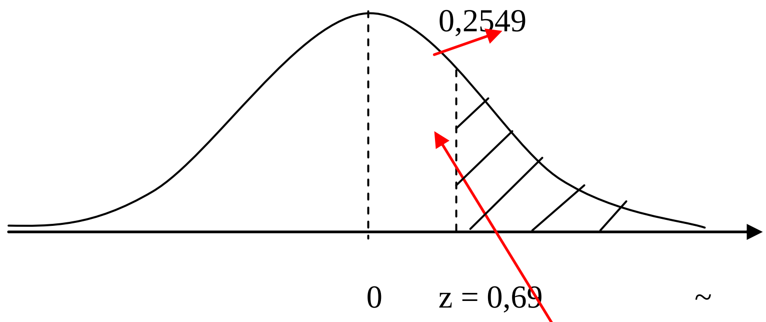
   Luas keseluruhan daerah di bawah kurva = 1

   Karena simetri, kiri dan kanan sama, maka luas dari 0 ke ~ adalah 0,5

   Sehingga luas daerah yang di arsir = P (z > 0,69)

   = 0,5 – 0,2549

   = 0,2451



z	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
,0	,0000	,0040	,0080	,0120	,0160	,0199	,0239	,0279	,0319	,0359
,1	,0398	,0438	,0478	,0517	,0557	,0596	,0636	,0675	,0714	,0753
,2	,0793	,0832	,0871	,0910	,0948	,0987	,1026	,1064	,1103	,1141
,3	,1179	,1217	,1255	,1293	,1331	,1368	,1406	,1443	,1480	,1517
,4	,1554	,1591	,1628	,1664	,1700	,1736	,1772	,1808	,1844	,1879
,5	,1915	,1950	,1985	,2019	,2054	,2088	,2123	,2157	,2190	,2224
,6	,2257	,2291	,2324	,2357	,2389	,2422	,2454	,2486	,2518	,2549
,7	,2580	,2612	,2642	,2673	,2704	,2734	,2764	,2794	,2823	,2852
,8	,2881	,2910	,2939	,2967	,2995	,3023	,3051	,3078	,3106	,3133
,9	,3159	,3186	,3212	,3238	,3264	,3289	,3315	,3340	,3365	,3389

Jadi peluang nilai z lebih dari 0,69 = 0,2451

2. P (-0,69 < z < 0,69)

   Sama saja dengan P (-0,69 < z < 0) + P (0 < z < 0,69)

   Luas daerah P (0 < z < 0,69) = 0,2549

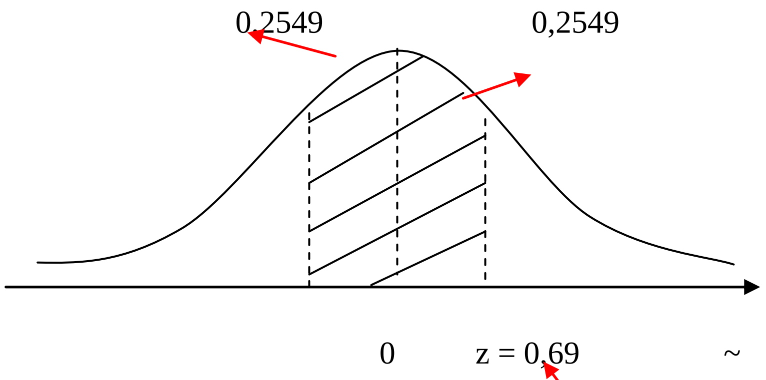
   Karena simetri, kiri dan kanan sama, maka luas

   P (-0,69 < z < 0) = P (0 < z < 0,69)

   = 0,2549

   Sehingga luas daerah yang di arsir = P (-0,69 < z < 0,69)

   = 0,2549 + 0,2549 = 0,5098



z	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
,0	,0000	,0040	,0080	,0120	,0160	,0199	,0239	,0279	,0319	,0359
,1	,0398	,0438	,0478	,0517	,0557	,0596	,0636	,0675	,0714	,0753
,2	,0793	,0832	,0871	,0910	,0948	,0987	,1026	,1064	,1103	,1141
,3	,1179	,1217	,1255	,1293	,1331	,1368	,1406	,1443	,1480	,1517
,4	,1554	,1591	,1628	,1664	,1700	,1736	,1772	,1808	,1844	,1879
,5	,1915	,1950	,1985	,2019	,2054	,2088	,2123	,2157	,2190	,2224
,6	,2257	,2291	,2324	,2357	,2389	,2422	,2454	,2486	,2518	,2549
,7	,2580	,2612	,2642	,2673	,2704	,2734	,2764	,2794	,2823	,2852
,8	,2881	,2910	,2939	,2967	,2995	,3023	,3051	,3078	,3106	,3133
,9	,3159	,3186	,3212	,3238	,3264	,3289	,3315	,3340	,3365	,3389

Jadi peluang nilai z antara -0,69 sampai 0,69 adalah ,05098

3. P (0,52 < z < 0,69)

   Luas daerah P (0 < z < 0,52) = 0,1985

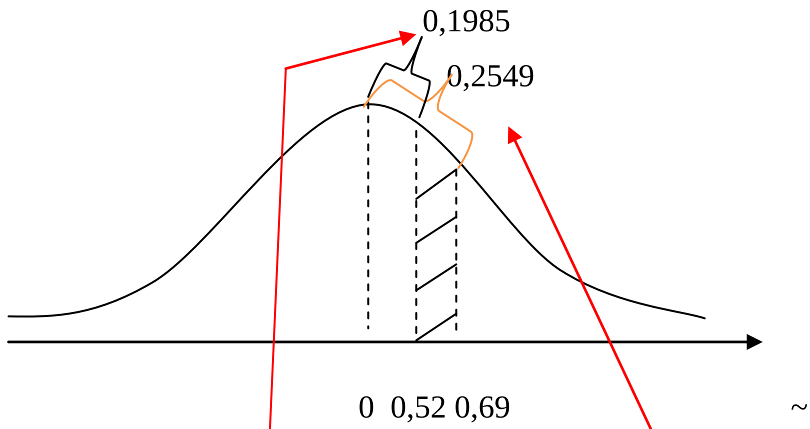
   Luas daerah P (0 < z < 0,69) = 0,2549

   Sehingga luas daerah yang di arsir

   = P (0 < z < 0,69) - P (0 < z < 0,52)

   = 0,2549 – 0,1985

   = 0,0564



z	,00	,01	,02	,03	,04	,05	,06	,07	,08	,09
,0	,0000	,0040	,0080	,0120	,0160	,0199	,0239	,0279	,0319	,0359
,1	,0398	,0438	,0478	,0517	,0557	,0596	,0636	,0675	,0714	,0753
,2	,0793	,0832	,0871	,0910	,0948	,0987	,1026	,1064	,1103	,1141
,3	,1179	,1217	,1255	,1293	,1331	,1368	,1406	,1443	,1480	,1517
,4	,1554	,1591	,1628	,1664	,1700	,1736	,1772	,1808	,1844	,1879
,5	,1915	,1950	,1985	,2019	,2054	,2088	,2123	,2157	,2190	,2224
,6	,2257	,2291	,2324	,2357	,2389	,2422	,2454	,2486	,2518	,2549
,7	,2580	,2612	,2642	,2673	,2704	,2734	,2764	,2794	,2823	,2852
,8	,2881	,2910	,2939	,2967	,2995	,3023	,3051	,3078	,3106	,3133
,9	,3159	,3186	,3212	,3238	,3264	,3289	,3315	,3340	,3365	,3389

Jadi peluang nilai z antara 0,52 sampai 0,69 adalah ,00564

2.4. Latihan Soal

1. Berdasarkan suatu survey diketahui 2 dari 5 laki-laki dewasa punya peluang menderita Osteoporosis. Jika disuatu kantor ada 5 orang laki-laki, hitunglah probabilitas bahwa 5 orang tersebut :
   a. Tidak ada satupun yang menderita Osteoporosis
   b. Paling sedikit 3 orang menderita Osteoporosis
   c. Hanya 2 orang menderita Osteoporosis
2. Menurut pendapat seorang ahli mengatakan bahwa 4 dari 7 wanita berpotensi mengalami anemia. Dari 10 orang wanita, tentukan probabilitasnya
   a. Hanya satu orang wanita yang meng alami anemia
   b. Lebih dari 7 orang wanita mengalami anemia
   c. Paling sedikit 2 orang wanita mengalami anemia

3. Jika 15% barang yang diproduksi suatu mesin pabrik diketahui rusak, berapa probabilitasnya dari 4 barang yang diproduksi :
   a. Semua rusak
   b. Paling banyak 2 rusak
   c. Paling sedikit 3 rusak
4. Probabilitas bahwa seorang balita akan menderita reaksi buruk akibat imunisasi adalah 0,001. Hitunglah bahwa dari 2000 balita yang diimunisasi,
   a. Tidak ada satupun balita yang menderita reaksi buruk akibat imunisasi.
   b. Hanya 2 balita yang menderita reaksi buruk akibat imunisasi
   c. Lebih dari 3 balita yang menderita reaksi buruk akibat imunisasi
5. Diketahui bahwa rata-rata 1 dari 1500 mobil yang lewat jalan tol Krapyak mengalami kerusakan ban. Apabila pada hari tertentu lewat 4500 mobil di jalan tol Krapyak, berapa probabilitas bahwa :
   a. Hanya satu mobil yang mengalami kerusakan ban
   b. Kurang dari 2 mobil mengalami kerusakan ban
   c. Paling sedikit 3 mobil mengalami kerusakan ban
6. Seorang broker real estate mengatakan bahwa 2 dari 40 rumah yang ditawarkan akan terjual dalam setiap minggunya. Jika rumah yang tersedia 80 rumah, tentukan probabilitas bahwa dalam waktu satu minggu akan terjual :
   a. Hanya satu rumah
   b. Lebih dari 5 rumah
7. Suatu studi tentang pertumbuhan dari tanaman cactus jenis tertentu dianggap berdistribusi normal dengan rata-rata tumbuh 44,8 mm dengan deviasi standar 4,7 mm selama jangka waktu 12 bulan. Tentukan probabilitas untuk jenis cactus tersebut rata-rata pertumbuhannya:
   a. Antara 43,8 dan 45,7 cm
   b. Kurang dari 42,8
8. Pada Penyelengaraan Pemilihan Miss Celebrity 2008 di kota Semarang yang diselenggarakan oleh SCTV diketahui bahwa rata-rata tinggi badan calon Miss Celebrity yang mengikuti audisi berdistribusi normal dengan rata-rata 170 cm dan simpangan baku 4,8 cm. Tentukan :
   a. Berapa probabilitas calon Miss Celebrity yang mempunyai tinggi lebih dari 176,4 cm

b. Berapa probabilitas calon Miss Celebrity yang mempunyai tinggi antara 167,3 dan 172,6
c. Jika 15% calon Miss Celebrity dengan tinggi badan tertinggi, berhak mengikuti final Pemilihan Miss Celebrity, tentukan tinggi minimal supaya calon peserta berhak mengikuti final Pemilihan Miss Celebrity 2008.

9. Pada Penerimaan Siswa Baru (PSB) di SD Negeri 01 Semarang, diketahui bahwa umur rata-rata anak yang mendaftar adalah 6,2 tahun dengan simpangan baku 2. Jika data diatas berdistribusi normal. Tentukan :
   a. Berapa probabilitas anak yang berumur kurang dari 6,5 tahun
   b. Berapa probabilitas anak yang berumur antara 6,4 tahun dan 6,8 tahun.
   c. Jika 10 % anak dengan umur terendah/termuda tidak diterima . Berapa batas nilai minimal agar seorang anak bisa diterima di SD Negeri 01 Semarang.
10. Suatu variable random mempunyai distribusi normal dengan mean $\mu = 80$ dan simpangan baku $\sigma = 4{,}8$. Berapa probabilitasnya bahwa variable random akan mempunyai nilai :
    a. Kurang dari 87,2
    b. Lebih dari 76,4
    c. Antara 81,2 dan 86,0
    d. Antara 71,6 dan 88,4
11. Panjang ikan sardine yang diterima suatu pabrik pengalengan ikan mempunyai panjang rata-rata 4,54 inci dan simpangan baku 0,25 inci. Apabila distribusi panjang ikan sardine tersebut mendekati distribusi normal, berapa persentase dari ikan-ikan tersebut yang panjangnya adalah :
    a. Lebih dari 5 inci
    b. Kurang dari 4 inci
    c. 4,4 sampai 4,6 inci
12. Suatu mesin pengisi minuman ringan diatur sedemikian rupa sehingga rata-rata mengisi setiap botol 200 milimeter. Jika volume minuman tersebut berdistribusi normal dengan simpangan baku 15 milimeter, ditanyakan :
    a. Berapaa bagian yang berisi lebih dari 224 milimeter
    b. Berapa probabilitas seluruh botol akan berisi 191 sampai 209 milimeter
    c. Berapa banyak botol minuman yang berisi melebihi 230 milimeter bila produksi diketahui 1000 botol.

d. Di bawah nilai berapa untuk diperoleh 25 % isi terendah.

13. Pengemudi taksi berdasarkan pengalamannya mengetahui bahwa jumlah penumpang yang ia antarkan untuk sore hari rata-rata 23,7 orang dengan deviasi standar 4,2. Jika dianggap jumlah penumpang berdistribusi normal, hitunglah probabilitasnya bahwa waktu sore hari pengemudi taksi tersebut mengantarkan :
    a. 20 penumpang
    b. Paling sedikit 18 penumpang
    c. Paling banyak 25 penumpang
    d. 15 sampai 21 penumpang

BAB 3

ANALISIS REGRESI

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa dapat menjelaskan arti dan kegunaan analisis regresi secara umum

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa dapat menjelaskan konsep dasar analisis regresi, menjelaskan peubah bebas dan tak bebas, menjelaskan regresi populasi dan regresi sampel, serta mampu menghitung persamaan regresi linier sederhana dan regresi linier ganda.

3.1 Regresi Linier

Suatu model persamaan matematik yang memberikan kemungkinan untuk meramalkan suatu nilai-nilai perubah tak bebas dari nilaI-nilai satu atau lebih perubah bebas dinamakan persamaan regresi.

Dalam kehidupan ini seringkali berhadapan dengan berbagai gejala yang meliputi berbagai variabel. Sebagai contoh adalah :

1. berat badan dalam taraf tertentu tergantung pada tinggi badannya,
2. produktivitas kerja pada taraf tertentu tergantung pada efisiensi dan efektivitas kerjanya,
3. produksi padi tergantung pada kesuburan tanah, teknologi, banyak curah hujan,
4. investasi bergantung pada suku bunga bank,
5. hubungan musim kemarau dengan kebutuhan payung dan es, dan sebagainya.

Berdasarkan contoh di atas, maka tampaklah mana variabel bebas (yang mempengaruhi) dan variabel terikat atau tergantung (yang dipengaruhi). Variabel yang mempengaruhi ini dalam analisis regresi disebut sebagai variabel bebas atau **variabel prediktor**, dengan lambang X; sedangkan variabel yang dipengaruhi disebut variabel takbebas atau variabel respon atau **variabel kriterium** dengan lambang Y.
Jika kita mempunyai data yang terdiri atas dua atau lebih variabel, adalah sewajarnya untuk mempelajari cara bagaimana variabel-variabel itu berhubungan. Hubungan yang diperoleh biasanya dinyatakan dalam persamaan matematik yang menyatakan **hubungan fungsional** antara variabel-variabel.

Hubungan fungsional antara satu variabel bebas dengan satu variabel tak bebas disebut **analisis regresi tunggal**, sedangkan hubungan fungsional yang lebih dari satu variabel disebut **analisis regresi ganda**.

Bentuk umum persamaan regresi linier adalah :

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

Di mana :

$$\begin{aligned} \hat{Y} &= \text{variabel tak bebas} \\ X &= \text{variabel bebas} \\ a &= \text{konstanta} \\ b &= \text{koefisien arah regresi linier} \end{aligned}$$

Bentuk persamaan regresi tersebut sering dibaca sebagai regresi X atas Y, artinya regresi X sebagai variabel bebasnya dengan Y sebagai variabel terikatnya. Sebaliknya ada pula persamaan regresi yang dibaca sebagai regresi Y atas X.

Koefisien arah regresi linier dinyatakan dengan huruf b yang juga menyatakan perubahan rata-rata variabel Y untuk setiap variabel X sebesar satu bagian. Maksudnya ialah bila harga b positif, maka variabel Y akan mengalami kenaikan atau pertambahan. Sebaliknya bila b negatif, maka variabel Y akan mengalami penurunan.

Contoh 3.1 :

Persamaan regresi antara pengunjung toko (X) dengan pembeli (Y) ialah :

$$Y = 9 + 0{,}50\ X.$$

Maknanya ialah : karena b positif, maka hubungan fungsionalnya juga menjadi positif. Selanjutnya kita dapat mengatakan bahwa jika setiap pengunjung (X) bertambah dengan 30 orang, maka rata-rata pembeli (Y) akan bertambah menjadi Y = 9 + 0,50 . 30 = 24 orang. Dan akhirnya kita dapat menyimpulkan bahwa semakin banyak pengunjung, semakin banyak pula pembelinya.

3.2 Metode Kuadrat Terkecil untuk Regresi Linier :

Metoda diagram pencar dapat dipakai menduga bentuk regresi apakah linier atau tidak. Jika tidak betul-betul yakin, lebih baik ditentukan dengan cara lain, misalnya dengan cara kuadrat terkecil (least squares). Adapun langkah-langkah untuk menyelesaikan regresi linier adalah :

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

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

$$a = \frac{\sum y}{n} + b\ \frac{\sum x}{n}$$

Dari persamaan regresi :

Selanjutnya masukkan nilai a dan b ke dalam persamaan regresi : $Y = a + bX$

Contoh 3.2 :

Setiap biaya promosi (x) yang dikeluarkan oleh suatu perusahaan akan selalu berpengaruh terhadap keuntungan (y) pada tiap tahun. Dari data suatu perusahaan tertentu diperoleh data dalam jutaan rupiah sebagai berikut :

TAHUN	BIAYA PROMOSI	KEUNTUNGAN
1	1,5	3,6
2	1,0	2,8
3	2,0	4,3
4	2,8	5,4
5	0,4	1,9
6	1,3	2,9

Tentukanlah :

i. Persamaan garis regresinya

ii. Berapa keuntungan yang diperkirakan jika biaya promosi 10 Juta

Jawab :

Dari data tersebut kita dapat menentukan variabel bebasnya (x) adalah biaya promosi sedangan variabel tak bebasnya (y) adalah keuntungan.

Buat tabel untuk menghitung nilai b dan a

x	**y**	**x.y**	**x^2**
1,5	3,6	5,4	2,25
1,0	2,8	2,8	1,00
2,0	4,3	8,6	4,00
2,8	5,4	15,12	7,84
0,4	1,9	0,76	0,16
1,3	2,9	3,77	1,69
$\sum x = 9$	**$\sum y = 20{,}9$**	**$\sum xy = 36{,}45$**	**$\sum x^2 = 16{,}94$**

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

$$= \frac{(6)(36{,}45) - (9)(20{,}9)}{(6)(16{,}94) - (9)^2} = \frac{218{,}70 - 188{,}1}{101{,}64 - 81}$$

$$= \frac{30{,}60}{20{,}64} = 1{,}48$$

$$a = \frac{\sum y}{n} + b\frac{\sum x}{n}$$

$$= \frac{20{,}9}{6} + (1{,}48)\frac{9}{6} = \frac{7{,}56}{6} = 5{,}7$$

i. Jadi persamaan garis regresinya adalah ***y = 5,7 + 1,48 x***

ii. Jika x = 10 maka y = 5,7 + 1,48 (10) = 20,5

jadi keuntungan yang diperkirakan adalah 20,5 jt

3.3 Regresi Linier Ganda

Bentuk persamaan regresi linier ganda dengan 2 variabel bebas adalah :

$$Y = a + b_1 X_1 + b_2 X_2$$

Sedangkan bentuk persamaan regresi dengan 3 variabel independen adalah:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 \quad \ldots\text{.dst}$$

Untuk memperoleh koefisien regresi a, b_1 dan b_2 dapat digunakan metode *ordinary least square* (OLS) yang pada prinsipnya adalah meminimumkan jumlah kuadrat deviasi di sekitar garis regresi. Nilai koefisien regresi a, b_1 dan b_2 dapat dipecahkan secara simultan dari tiga persamaan berikut.

$$\Sigma Y = na + b_1 \Sigma X_1 + b_2 \Sigma X_2$$

$$\Sigma X_1 Y = a \Sigma X_1 + b_1 \Sigma X_1^2 + b_2 \Sigma X_1X_2$$

$$\Sigma X_2 Y = a \Sigma X_2 + b_1 \Sigma X_1X_2 + b_2 \Sigma X_2^2$$

Contoh 3.3 :

Suatu penelitian yang dilakukan di Hero Supermarket bertujuan untuk mengetahui hubungan dan pengaruh variabel harga dan pendapatan terhadap permintaan minyak goreng. Berikut adalah hasil penelitiannya.

No. Sampel	Permintaan Minyak (liter/bulan)	Harga Minyak (Rp. ribu/ liter)	Pendapatan (Rp. Juta/bulan)
1	3	8	10
2	4	7	10
3	5	7	8
4	6	7	5
5	6	6	4
6	7	6	3
7	8	6	2
8	9	6	2
9	10	5	1
10	10	5	1

Berdasarkan data tersebut, hitunglah koefisien regresinya !

Jawab :

Untuk mendapatkan koefisien regresi, sesuai dengan persamaan (a), (b) dan (c), perlu dihitung lebih dulu nilai-nilai sebagai berikut.

Y	X_1	X_2	Y X_1	Y X_2	X_1^2	X_2^2	X_1 X_2
3	8	10	24	30	64	100	60
4	7	10	28	40	49	100	70
5	7	8	35	40	49	64	56
6	7	5	42	30	49	25	35
6	6	4	36	24	36	16	24
7	6	3	42	21	36	9	18
8	6	2	48	16	36	4	12
9	6	2	54	18	36	4	12
10	5	1	50	10	25	1	5
10	5	1	50	10	25	1	5
$\Sigma Y = 68$	**$\Sigma X_1 = 63$**	**$\Sigma X_2 = 46$**	**$\Sigma Y X_1 = 409$**	**$\Sigma Y X_2 = 239$**	**$\Sigma X_1^2 = 405$**	**$\Sigma X_2^2 = 324$**	**$\Sigma X_1 X_2 = 317$**

Dengan menggabungkan persamaan (a), (b) dan (c) dengan hasil perhitungan di atas dapat diperoleh persamaan berikut.

$$68 = 10a + 63\, b_1 + 46\, b_2$$

$$409 = 63a + 405\, b_1 + 317\, b_2$$

$$239 = 46a + 317\, b_1 + 324\, b_2$$

Untuk mendapatkan nilai koefisien a, b_1 dan b_2 dapat dilakukan dengan cara substitusi yang memerlukan perhitungan ketelitian yang baik. Cara lebih sederhana adalah sebagai berikut.

$$A = n\, \Sigma X_1\, Y - \Sigma X_1\, \Sigma Y \qquad = 10 \,.\, 409 - 63 \,.\, 68 = -194$$

$$B = n\, \Sigma (X_2)^2 - (\Sigma X_2)^2 \qquad = 10 \,.\, 324 - (46)^2 = 1124$$

$$C = n\, \Sigma X_1\, X_2 - \Sigma X_1\, \Sigma X_2 \qquad = 10 \,.\, 317 - 63 \,.\, 46 = 272$$

$$D = n\, \Sigma X_2\, Y - \Sigma X_2\, \Sigma Y \qquad = 10 \,.\, 239 - 46 \,.\, 68 = -736$$

$$E = n\, \Sigma (X_1)^2 - (\Sigma X_1)^2 \qquad = 10 \,.\, 405 - (63)^2 = 81$$

$$F = EB - C^2 \qquad = 81 \,.\, 1124 - (272)^2 = 17060$$

$$b_1 = \frac{AB - CD}{F} = \frac{(-194).1124 - (272.(-736))}{17060} = -1{,}015$$

$$b_2 = \frac{DE - AC}{F} = \frac{(-736).81 - (-194).272}{17060} = -0{,}41$$

$$a = \frac{\sum Y - b_1 \sum X_1 - b_2 \sum X_2}{n} = \frac{68 - (-1{,}015 \times 69) - (-0{,}41 \times 46)}{10} = 15{,}086$$

Dengan memasukkan nilai a, b_1 dan b_2, maka persamaan regresinya dapat dinyatakan sebagai berikut.

$$Y = 15{,}086 - 1{,}015\, X_1 - 0{,}41\, X_2$$

3.4 Latihan Soal :

1. Diketahui data sebagai berikut :

X	1	2	3	4	5	6	7	8
Y	5	6	5	8	9	7	4	3

   a. Tentukan persamaan regresinya !

   b. Tentukan nilai bagi Y jika diketahui X = 4.5 !

2. Data dibawah ini menunjukkan besarnya biaya advertensi (% daribiaya total) dan laba usaha bersih (% dari total penjualan) dari sampel random 6 tokotekstil :

Biaya (x)	**Laba (y)**
1,5	3,6
1,0	2,8
2,8	5,4
0,4	1,9
1,3	2,9
2,0	4,3

   a) Tentukan garis regresinya dengan metode least square

   b) Untuk biaya advertensi 1,2 (% daribiaya total) perkirakan besarnya laba bersih (% dari total penjualan)

3. Data dibawah ini menunjukkan umur dan harga mobil sedan bekas merk tertentu

Umur (tahun)	**Harga (ribuan rupiah)**
1	1.795
4	985
10	295
2	1.295
5	795
6	995
8	845
1	1.625

   a) Hitunglah garis regresinya yang memungkinkan kita untuk meramalkan harga berdasarkan umur mobil.

   b) Ramalkan harga mobil sedan bekas yang sudah berumur 3 tahun.

c) Untuk mobil yang sudah berumur 20 tahun bagaimana pendapat anda ?

4. Data dibawah ini menunjukkan umur anak (X1), berat waktu lahir (X2), dan panjang badan anak tersebut (Y) dari sampel random 9 anak.

Umur (hari)	**Berat lahir (kg)**	**Panjang badan (cm)**
78	2,75	57,5
69	2,15	52,8
77	4,41	61,3
88	5,52	67,0
67	3,21	53,5
80	4,32	62,7
74	2,31	56,2
94	4,30	68,5
102	3,71	69,2

a. Hitunglah persamaan regresinya dengan metode least square dari Y atas X1 dan X2. Kemudian hitunglah dugaan nilai Y berdasarkan nilai X1 dan X2 nya

b. Perkirakan panjang badan rata-rata bagi anak yang berumur 75 hari dan berat badan waktu lahir 3,15 kg.

BAB 4

ANALISIS KORELASI

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa dapat menjelaskan arti dan kegunaan analisis korelasi secara umum

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa mampu menjelaskan dengan tepat konsep dasar analisis korelasi serta mampu menghitung koefisien korelasi dan koefisien determinasi dengan tepat

4.1 Pendahuluan

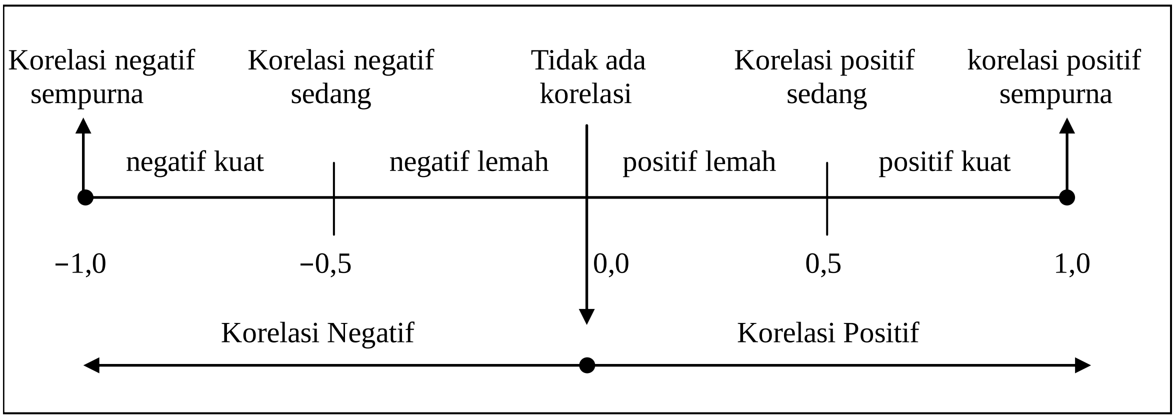
Korelasi adalah suatu alat analisis yang dipergunakan untuk mencari hubungan antara variabel independen/bebas dengan variabel dipenden/takbebas. Apabila beberapa variabel independen/bebas dihubungkan dengan satu variabel dependen/tak bebas disebut korelasi berganda. Dan apabila satu variabel independen/bebas berhubungan dengan satu variabel dependent/takbebas disebut korelasi parsial .

Hubungan antara dua variabel dapat karena hanya kebetulan saja dapat pula memang merupakan hubungan yang sebab akibat. Dua varibel berkorelasi apabila perubahan yang lain secara teratur, dengan arah yang sama atau arah yang berlawanan. Dalamanalisa korelasi disamping mengukur kesesuaian garis regresi terhadap data sampel atau disebut ***Koefisien Determinasi atau Koefisien Penentu***, juga mengukur keeratan hubungan antara variabel atau disebut ***Koefisien Korelasi***.

Dengan kata lain, analisa regresi menjawab bagaimana pola hubungan variabel-variabel dan analisa korelasi menjawab bagaimana keeratan hubungan yang diterangkan dalam persamaan regresi.

Kedua analisa ini biasanya dipakai bersama-sama. Koefisien korelasi dilambangkan dengan r dan koefisien determinasi dilambangkan dengan r^2.

Harga r bergerak antara –1 dan +1 dengan tanda negatif menyatakan adanya korelasi tak langsung atau korelasi negatif dan tanda positif menyatakan adanya korelasi langsung atau korelasi positif. r=0 menyatakan tidak ada hubungan linier antara variabel X dan Y.



4.2 Koefisien Korelasi Rank

Koefisen korelasi ini mengukur kedekatan hubungan antara dua variabel ordinal. Koefisien korelasi ini dinamakan koefisien korelasi pangkat atau koefisien korelasi Spearman, yang disimbolkan dengan r.

Pasangan data hasil pengamatan (X_i , Y_i) kita susun menurut urutan besar nilainya dalam tiap variabel. Kemudian kita bentuk selisih atau beda peringkat X_i dan peringkat Y_i yang data aslinya berpasangan. Beda ini disimbolkan dengan b_i, maka koefisien korelasi peringkat r dihitung dengan rumus:

$$r = 1 - \frac{6 \sum b_i^2}{n\,(n^2 - 1)}$$

Contoh 4.1 :

Data berikut adalah penilaian 2 orang juri terhadap 8 orang peserta perlombaan.

Peserta	Juri I	Juri II
A	70	80
B	85	75
C	65	55
D	50	60
E	90	85
F	80	70
G	75	90
H	60	65

Tentukan Koefisien Korelasi rank !

Jawab :

Peserta	**Juri I**	**Juri II**	**Beda (b_i)**	b_i^2
A	5	3	2	4
B	2	4	-2	4
C	6	8	2	4
D	8	7	1	1
E	1	2	-1	1
F	3	5	-2	4
G	4	1	3	9
H	7	6	1	1
Jumlah	-	-	-	28

Koefisien Korelasi Rank :

$$r = 1 - \frac{6 \sum b_i^2}{n(n^2 - 1)} = 1 - \frac{6(28)}{8(64-1)} = 0,667$$

4.3 Koefisien Korelasi Pearson (Product Moment)

Untuk sekumpulan data (X_i, Y_i) berukuran n, koefisien korelasi dapat dihitung dengan rumus:

$$r = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{\left\{n(\sum X_i^2) - (\sum X_i)^2\right\}\left\{n(\sum Y_i^2) - (\sum Y_i)^2\right\}}}$$

Contoh 4.2 :

Diketahui data jumlah SKS dan IPK mahasiswa sebagai berikut :

Jumlah SKS (X)	IPK (Y)
10	3,00
10	2,50
15	2,00
10	1,50
5	1,00

Tentukan nilai koefisien korelasi dengan metode product moment dan jelaskan artinya!

Jawab :

Buat tabel penolong untuk menghitung r

No	X_i	Y_i	$X_i\,Y_i$	X_i^2	Y_i^2
1	10	3,00	30	100	9,00
2	10	2,50	25	100	6,25
3	15	2,00	30	225	4,00
4	10	1,50	15	100	2,25
5	5	1,00	5	25	1,00
n=5	$\sum X_i = 50$	$\sum Y_i = 10$	$\sum X_iY_i = 105$	$\sum X_i^2 = 550$	$\sum Y_i^2 = 22{,}5$

$$r = \frac{5\,(105) - (50)\,(10)}{\sqrt{\left\{5\,(550) - (50)^2\right\}\left\{5\,(22{,}5) - (10)^2\right\}}} = \frac{25}{\sqrt{(250)\,(12{,}5)}} = 0{,}447$$

Dari hasil ini ternyata didapat korelasi positif antara jumlah sks (X) dan IPK yang didapat (Y).

4.4 Koefisien Determinasi

Koefisien determinasi merupakan ukuran untuk mengetahui kesesuaian atau ketepatan antara nilai dugaan dengan data sampel. Koefisien determinasi didefinisikan sebagai berikut.

Koefisien determinasi *adalah bagian dari keragaman total variabel tak bebas Y (variabel yang dipengaruhi atau dependent) yang dapat diterangkan atau diperhitungkan oleh keragaman variabel bebas X (variabel yang mempengaruhi, independent)*

Jadi koefisien determinasi adalah kemampuan variabel X mempengaruhi variabel Y. Semakin besar koefisien determinasi menunjukkan semakin baik kemampuan X mempengaruhi Y.

$$\text{Koefisien Determinasi} = r^2 = \frac{[\, n \sum X_i Y_i - (\sum X_i)(\sum Y_i)\,]^2}{\sqrt{\left\{n(\sum X_i^2) - (\sum X_i)^2\right\}\left\{n(\sum Y_i^2) - (\sum Y_i)^2\right\}}}$$

4.5 Koefisien Korelasi Ganda

Untuk 2 variabel bebas (X_1 dan X_2) maka r dihitung dengan rumus:

$$r_{y\,x_1\,x_2} = \sqrt{\frac{r^2_{y\,x_1} + r^2_{y\,x_2} - 2\, r_{y\,x_1}\, r_{y\,x_2}\, r_{x_1\,x_2}}{1 - r^2_{x_1\,x_2}}}$$

dimana :

$r_{y\,x_1\,x_2}$ = Koefisien korelasi ganda antara variable X_1 dan X_2 secara bersama-sama dengan variable Y

$r_{y\,x_1}$ = Koefisien korelasi X_1 dengan Y

$r_{y\,x_2}$ = Koefisien korelasi X_2 dengan Y

$r_{x_1\,x_2}$ = Koefisien korelasi X_1 dengan X_2

Contoh 4.3 :

Misalkan melakukan pengamatan terhadap 10 keluarga mengenai:

X_1 = pendapatan dalam ribuan rupiah

X_2 = jumlah keluarga dalam satuan jiwa

Y = pengeluaran untuk membeli barang A dalam ratusan rupiah

X_1	10	2	4	6	8	7	4	6	7	6
X_2	7	3	2	4	6	5	3	3	4	3
Y	23	7	15	17	23	22	10	14	20	19

Akan dibuktikan ada hubungan linier positif dan signifikan antara variabel X_1 dan X_2 secara bersama-sama dengan variabel Y.

Jawab :

No	**X_1**	**X_2**	**Y**	**X_1Y**	**X_2Y**	**X_1X_2**	**X_1^2**	**X_2^2**	**Y^2**
1	10	7	23	230	161	70	100	49	529
2	2	3	7	14	21	6	4	9	49
3	4	2	15	60	30	8	16	4	225
4	6	4	17	102	68	24	36	16	289
5	8	6	23	184	138	48	64	36	529
6	7	5	22	152	110	35	49	25	484
7	4	3	10	40	30	12	16	9	100
8	6	3	14	84	42	18	36	9	196
9	7	4	20	140	80	28	49	16	400
10	6	3	19	114	57	18	36	9	361
Jumlah	**60**	**40**	**170**	**1121**	**737**	**267**	**406**	**182**	**3162**

Dari tabel diperoleh:

$n = 10$, $\sum X_1 = 60$, $\sum X_2 = 40$, $\sum Y = 170$, $\sum X_1Y = 1122$, $\sum X_2Y = 737$, $\sum X_1 X_2 = 267$, $\sum X_1^2 = 406$, $\sum X_2^2 = 182$, $\sum Y^2 = 3162$

$$r_{y\,x_1} = \frac{n \sum X_1 Y - (\sum X_1)(\sum Y)}{\sqrt{\left\{n(\sum X_1^2) - (\sum X_1)^2\right\}\left\{n(\sum Y^2) - (\sum Y_1)^2\right\}}} =$$

$$r_{y\,x_1} = \frac{10\,(1122) - (60)\,(170)}{\sqrt{\left\{10\,(406) - (60)^2\right\}\left\{10\,(3162) - (170)^2\right\}}}$$

$$r_{y\,x_1} = \frac{1020}{\sqrt{460 \text{ x } 2720}} = \frac{1020}{1118{,}57}$$

$$r_{y\,x_1} = 0{,}912$$

$$r_{y\,x_2} = \frac{n \sum X_2 Y - (\sum X_2)(\sum Y)}{\sqrt{\left\{n(\sum X_2^2) - (\sum X_2)^2\right\}\left\{n(\sum Y^2) - (\sum Y)^2\right\}}} =$$

$$r_{y\,x_2} = \frac{10\,(737) - (40)\,(170)}{\sqrt{\left\{10\,(182) - (40)^2\right\}\left\{10\,(3162) - (170)^2\right\}}}$$

$$r_{y\,x_2} = \frac{570}{\sqrt{220 \text{ x } 2720}} = \frac{570}{773{,}56}$$

$$r_{y\,x_2} = 0{,}74$$

$$r_{x_1 x_2} = \frac{n \sum X_1 X_2 - (\sum X_1)(\sum X_2)}{\sqrt{\{n(\sum X_1^2) - (\sum X_1)^2\}\{n(\sum X_2^2) - (\sum X_2)^2\}}} =$$

$$r_{x_1 x_2} = \frac{10(267) - (60)(40)}{\sqrt{\{10(406) - (60)^2\}\{10(182) - (40)^2\}}}$$

$$r_{x_1 x_2} = \frac{270}{\sqrt{460 \text{ x } 220}} = \frac{270}{318{,}12}$$

$$r_{x_1 x_2} = 0{,}85$$

$$r_{y\, x_1 x_2} = \sqrt{\frac{r^2_{y\, x_1} + r^2_{y\, x_2} - 2\, r_{y\, x_1}\, r_{y\, x_2}\, r_{x_1 x_2}}{1 - r^2_{x_1 x_2}}}$$

$$= \sqrt{\frac{(0{,}912)^2 + (0{,}74)^2 - 2(0{,}912)(0{,}74)(0{,}85)}{1 - (0{,}85)^2}}$$

$$= \sqrt{0{,}8354}$$

$$= 0{,}914$$

Kesimpulan: Terdapat hubungan yang signifikan antara X_2 bersama-sama dengan X_2 dengan Y.

Atau : Terdapat hubungan yang signifikan antara pendapatan dan jumlah keluarga dengan pengeluaran untuk membeli barang A.

4.6 Latihan Soal :

1. Data dibawah ini menunjukkan jumlah pemakaian pupuk (X) dan hasil panen padi yang diperoleh (Y):

Pupuk (kg)	**Hasil Panen (kw)**
20	8
40	9
50	11
70	11
100	12
110	14
120	15
150	16

Hitung koefisien korelasi dengan metode product moment dan Jelaskan artinya .

2. Dua orang ibu rumah tangga diminta untuk mengemukakan tingkat preferensinya terhadap sabun mandi berbagai merk. Hasilnya adalah sebagai berikut :

Merk Sabun Mandi	Ny. Witono	Ny. Hartoko
A	3	5
B	5	6
C	8	4
D	12	9
E	10	8
F	7	12
G	9	11
H	1	3
I	4	1
J	6	2
K	2	10
L	11	7

Hitunglah nilai koefisien rank

3. Tabel dibawah ini menunjukkan berat badan, tinggi badan, dan umur dari sampel random 12 anak laki-laki. Berat badan diukur dalam pound, tinggi badan diukur dalam inci, dan umur diukur dalam tahun.

Berat Badan (X1)	Tinggi Badan (X2)	Umur (Y)
64	57	8
71	59	10
53	49	6
67	62	11
55	51	8
58	50	7
77	55	10
57	48	9
56	52	10
51	42	6
76	61	12
68	57	9

Hitung koefisien korelasi antara variabel X_1 dan X_2 secara bersama-sama dengan variabel Y.

BAB 5

DISTRIBUSI SAMPLING

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa dapat memahami hubungan nilai sampel dan populasi dan menentukan distribusi sampling yang tepat untuk digunakan

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa dapat membentuk distribusi sampling dan membuktikan hubungan nilai sampel dengan populasi, menghitung distribusi rata-rata suatu sampel, menghitung distribusi proporsi dari suatu sampel, menghitung distribusi sampling selisih rata-rata serta menghitung distribusi sampling selisih proporsi

5.1. Pendahuluan

Untuk mempelajari populasi kita memerlukan sampel yang diambil dari populasi yang bersangkutan. Meskipun kita dapat mengambil lebih dari sebuah sampel berukuran n dari populasi berukuran N, pada prakteknya hanya sebuah sampel yang biasa diambil dan digunakan untuk hal itu. Sampel yang diambil ialah sampel acak dan dari sampel tersebut nilai-nilai statistiknya dihitung untuk digunakan seperlunya. Untuk itu diperlukan sebuah teori yang dikenal dengan nama distribusi sampling. Distribusi sampling biasanya diberi nama tergantung pada nama statistk yang digunakan. Misalnya distribusi sampling rata-rata, distribusi sampling proporsi, distribusi simpangan baku dan sebagainya.

5.2. Distribusi Sampling Rata-rata

Jika sebuah populasi mempunyai rata-rata μ dan simpangan baku σ yang besarnya terhingga, maka untuk ukuran sampel acak n cukup besar, distribusi rata-rata sampel mendekati distribusi normal dengan rata-rata $\mu_{\bar{x}} = \mu$ dan simpangan baku $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Jika ukuran populasi kecil dibandingkan dengan ukuran sampel, yaitu (n/N) > 5%

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}} \quad ,$$

dan jika ukuran populasi besar dibandingkan dengan ukuran sampel, yaitu (n/N) $\leq$ 5%

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Distribusi normal yang didapat dari distribusi rata-rata perlu distandarkan agar daftar distribusi normal baku dapat digunakan. Ini perlu untuk perhitungan-perhitungan. Untuk ini digunakan transformasi :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

Contoh 5.1 :

Tinggi badan mahasiswa rata-rata mencapai 165 cm dan simpangan baku 8,4 cm. Untuk itu diambil sampel acak terdiri atas 45 mahasiswa. Tentukanlah berapa peluang tinggi rata-rata ke-45 mahasiswa tersebut :

a. Antara 160 cm dan 168 cm
b. Paling sedikit 166 cm

Jawab :

Jika ukuran populasi tidak dikatakan besarnya, selalu dianggap cukup besar untuk berlakunya teori. Ukuran sampel n = 45 tergolong sampel besar sehingga dalil limit pusat berlaku. Jadi rata-rata $\bar{x}$ untuk tinggi mahasiswa akan mendekati distribusi normal dengan :

$$\text{Rata-rata} \qquad \mu_{\bar{x}} = 165 \text{ cm}$$

$$\text{Simpangan Baku} \quad \sigma_{\bar{x}} = \frac{8,4}{\sqrt{45}} = 1,252 \text{ cm}$$

a. Dengan $\bar{x}$ = 160 cm dan $\bar{x}$ = 168 cm didapat :

$$z_1 = \frac{160 - 165}{1,252} = -3,99 \text{ dan } z_2 = \frac{168 - 165}{1,252} = 2,40$$

Penggunaan daftar distribusi normal baku memberikan luas kurva

= 0,5 + 0,4918 = 0,9918

Peluang rata-rata tinggi ke-45 mahasiswa antara 160 cm dan 168 cm adalah 0,9918

b. Rata-rata tinggi paling sedikit 166 cm memberikan angka z paling sedikit

$$z = \frac{166 - 165}{1,252} = 0,80$$

Dari daftar normal baku, luas kurva 0,5 – 0,2881 = 0,2119. Peluang yang dicari = 0,2119

Apabila dari populasi diketahui variannya dan perbedaan antara rata-rata dari sampel ke sampel diharapkan tidak lebih dari sebuah harga d yang ditentukan, maka berlaku hubungan :

$$\sigma_{\bar{x}} \leq d$$

Dari rumus ini ukuran sampel yang paling kecil sehubungan dengan distribusi rata-rata dapat ditentukan

Untuk contoh diatas, misalkan harga-harga $\bar{x}$ dari sampel yang satu dengan sampel lainnya diharapkan tidak mau lebih dari 1 cm. Jika populasi cukup besar, maka :

$$\frac{\sigma}{\sqrt{n}} \leq d \text{ yang menghasilkan } \frac{8,4}{\sqrt{n}} \leq 1 = \frac{70,56}{n} \leq 1 \text{ atau } n >= 70,56$$

Paling sedikit perlu diambil sampel terdiri atas 71 mahasiswa

5.3. Distribusi Sampling Proporsi

Misalkan populasi diketahui berukuran N yang didalamnya terdapat peristiwa A sebanyak Y diantara N. Maka didapat parameter proporsi peristiwa A sebesar µ = (Y/N)

Jika dari suatu populasi diambil sampel acak berukuran n dan dimisalkan didalamnya ada peristiwa A sebanyak x. Sampel ini memberikan statistik proporsi peristiwa

A = p= x/n. Jika semua sampel yang mungkin diambil dari populasi itu maka didapat sekumpulan harga-harga statistik proporsi. Dari kumpulan ini kita dapat menghitung rata-ratanya,diberi symbol μ_p dan simpangan bakunya diberi simbol σ_p

Jika ukuran populasi kecil dibandingkan dengan ukuran sampel, yaitu (n/N) > 5% , maka :

$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi\,(1-\pi)}{n}}\sqrt{\frac{N-n}{N-1}}$$

dan jika ukuran populasi besar dibandingkan dengan ukuran sampel, yaitu (n/N) $\leq$ 5%, maka :

$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi\,(1-\pi)}{n}}$$

σ_p dinamakan kekeliruan baku proporsi atau galat baku proporsi. Untuk ukuran sampel n cukup besar distribusi proporsi mendekati distribusi normal. Untuk perhitungan, daftar distribusi normal baku dapat digunakan dan untuk itu diperlukan transformasi

$$z = \frac{p - \pi}{\sigma_p}$$

Jika perbedaan antara proporsi sampel yang satu dengan yang lainnya diharapkan tidak lebih dari sebuah harga d yang ditentukan, maka berlaku

Contoh 5.2 :

Ada petunjuk kuat bahwa 10 % anggota masyarakat tergolong dalam golongan A. Sebuah sampel acak terdiri atas 100 orang telah diambil.

a. Tentukan peluangnya bahwa dari 100 orang itu akan ada paling sedikit 15 orang dari golongan A.
b. Berapa orang harus diselidiki agar persentase golongan A dari sampel yang satu dengan yang lainnya diharapkan berbeda paling besar dengan 2 % ?

Jawab :

Populasi yang dihadapi berukuran cukup besar dengan π = 0,10 dan 1 - π = 0,90

a. Untuk ukuran sampel 100, diantaranya paling sedikit 15 tergolong kategori A, maka paling sedikit p = 0,15.

Kekeliruan bakunya adalah :

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{0{,}10 \text{ x } 0{,}90}{100}} = 0{,}03$$

Bilangan z paling sedikit $= \dfrac{0{,}15 - 0{,}10}{0{,}03} = 1{,}67$

Dari daftar normal baku luasnya = 0,5 – 0,4525 = 0,0475

Peluang dalam sampel itu akan ada paling sedikit 15 kategori A adalah 0,0475.

b. π = 0,10 dan 1 - π = 0,90 sedangkan d = 0,02 maka :

$$\sqrt{\frac{0{,}1 \text{ x } 0{,}9}{n}} \leq 0{,}02 \text{ yang menghasilkan } n \geq 225$$

Paling sedikit sampel harus berukuran 225

5.4. Distibusi Sampling Selisih Rata-Rata

Misalkan kita punya dua populasi masing-masing berukuran N_1 dan N_2 . Populasi kesatu mempunyai rata-rata μ_1 dan simpangan baku σ_1 , sedangkan populasi kedua mempunyai rata-rata μ_2 dan simpangan baku σ_2. Dari setiap populasi secara independen kita ambil sampel-sampel acak berukuran n_1 dari populasi kesatu dan berukuran n_2 dari populasi kedua. Untuk membedakan, populasi kesatu dimisalkan mempunyai variabel X dan populasi kedua mempunyai variabel Y . Dari sampel-sampel ini, kemudian dihitung rata-ratanya. Dan didapat rata-rata sampel : $\bar{x}_1, \bar{x}_2, , ..., 0, \bar{x}_k$, dan $\bar{y}_1, \bar{y}_2, , ..., 0, \bar{y}_k$ dengan k = banyak sampel yang dapat diambil dari populasi kesatu dan r = banyak sampel yang dapat diambil dari populasi kedua.

Bentuklah sekarang semua selisih antara rata-rata dari sampel-sampel dalam kumpulan kesatu dan rata-rata dari sampel-sampel dalam kumpulan kedua . Didapat kumpulan selisih rata-rata yang bentuk umumnya : $\bar{x}_i - \bar{y}_j$ dengan i = 1, 2, 3, ….,k dan j = 1,2,3, …, r

Kumpulan selisih rata-rata sampel akan membentuk distribusi selisih rata-rata. Dari kumpulan ini kita dapat menghitung rata-ratanya, diberi simbol $\mu_{\bar{x}-\bar{y}}$, dan menghitung simpangan bakunya, diberi simbol $\sigma_{\bar{x}-\bar{y}}$. Ternyata bahwa, untuk N_1 dan N_2 cukup besar dan sampel-sampel acak diambil secara independen satu sama lain, didapat hubungan :

$$\mu_{\bar{x}-\bar{y}} = \mu_1 - \mu_2$$

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Untuk ukuran-ukuran sampel cukup besar, maka selisih rata-rata $\bar{x} - \bar{y}$ akan mendekati distribusi normal. Untuk membuat distribusi normal ini menjadi distribusi normal baku digunakan transformasi

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}-\bar{y}}}$$

Contoh 5.3 :

Rata-rata tinggi mahasiswa laki-laki 163 cm dan simpangan bakunya 5,2 cm, sedangkan untuk mahasiswa perempuan, parameter tersebut berturut-turut 152 cm dan 4,9 cm. Dari kedua kelompok mahasiswa itu masing-masing diambil sampel acak, secara independen, berukuran sama, ialah 140 orang. Berapa peluang rata-rata tinggi mahasiswa laki-laki paling sedikit 10 cm lebihnya dari rata-rata tinggi mahasiswa perempuan?

Jawab :

Misalkan $\bar{x}$ dan $\bar{y}$ masing-masing menyatakan rata-rata tinggi dari sampel untuk mahasiswa laki-laki dan perempuan. Yang ditanyakan adalah peluang $\bar{x} - \bar{y}$ paling sedikit 10 cm. Dari yang diketahui didapat μ_1 = 163 cm, σ_1 = 5,2 cm, μ_2 = 152 cm, σ_2 = 4,9 cm, dan $n_1 = n_2 = 140$, maka :

Rata - rata $\mu_{\bar{x}-\bar{y}} = (163 - 152)\ \text{cm} = 11\ \text{cm}$

Simpangan Baku $\sigma_{\bar{x}-\bar{y}} = \sqrt{\dfrac{(5,2)^2}{140} + \dfrac{(4,9)^2}{140}}\ \text{cm} = 0,6038\ \text{cm}$

$$z = \frac{10 - 11}{0,6038} = -1,66$$

Luas daerah normal baku yang diperlukan adalah 0,5 + 0,4515 = 0,9515. Jadi peluang yang dicari = 0,9515

5.5 Distribusi Sampling Selisih Proporsi

Misalkan ada dua populasi masing-masing berdistribusi Binom, kedua-duanya berukuran cukup besar. Didalam kedua populasi itu ada peristiwa A dengan proporsi π_1 untuk populasi kesatu dan π_2 untuk populasi kedua. Dari kedua populasi itu, secara independen diambil sampel-sampel acak berukuran n_1 dari populasi kesatu dan berukuran n_2 dari populasi kedua. Untuk peristiwa A, didapat kumpulan proporsi

$p_1 = \dfrac{x_i}{n_1}, i = 1, 2, ..., k$ dan $p_2 = \dfrac{y_i}{n_2}, j = 1, 2, ..., r$

Dengan x_i = adanya peristiwa A dalam sampel yang diambil dari populasi kesatu, y_j = adanya peristiwa A dalam sampel yang diambil dari populasi kedua, k dan r masing-masing banyak sampel yang mungkin diambil dari populasi kesatu dan kedua.

Selisih proporsi $(p_1 - p_2)$ dapat dibentuk sehingga terdapat kumpulan selisih proporsi. Dari kumpulan ini dapat dihitung rata-ratanya dan diberi simbol μ_{sp} dan simpangan bakunya diberi simbol σ_{sp}, dengan sp = $(p_1 - p_2)$ = selisih antara proporsi sampel kesatu dan proporsi sampel kedua. Ternyata untuk ini berlaku :

$$\mu_{sp} = \pi_1 - \pi_2$$

$$\sigma_{sp} = \sqrt{\frac{\pi_1 (1 - \pi_1)}{n_1} + \frac{\pi_2 (1 - \pi_2)}{n_2}}$$

Untuk ukuran sampel cukup besar distribusi selisih proporsi akan mendekati distribusi normal. Agar supaya distribusi normal ini menjadi distribusi normal baku maka diperlukan transformasi :

$$z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sigma_{sp}}$$

Contoh 5.4 :

Ada petunjuk kuat bahwa calon A akan mendapat suara 60 % dalam pemilihan. Dua buah sampel acak secara independen telah diambil masing-masing terdiri atas 300 orang. Tentukan peluangnya akan terjadi perbedaan persentase tidak lebih dari 10 % akan memilih A.

Jawab :

Kedua sampel diambil dari sebuah populasi, jadi kita anggap dua populasi yang sama, sehingga $\pi_1 = \pi_2 = 0{,}6$. Jika x = banyak orang yang memilh A dalam sampel kesatu, dan y = banyak orang yang memilih B dalam sampel kedua, maka yang dicari adalah peluang $p_1 - p_2 < 10$ % atau $p_2 - p_1 < 10$ %

Setelah digabungkan menjadi -10 % < $p_1 - p_2$ < 10 %

$\mu_{sp} = 0{,}6 - 0{,}6 = 0$

$$\sigma_{sp} = \sqrt{\frac{0{,}6 \text{ x } 0{,}4}{300} + \frac{0{,}6 \text{ x } 0{,}4}{300}} = 0{,}04$$

Bilangan z yang perlu adalah :

$$z_1 = \frac{-0{,}1 - 0}{0{,}04} = -2{,}50 \quad \text{dan} \quad z_2 = \frac{0{,}1 - 0}{0{,}04} = 2{,}50$$

Luas daerah normal baku yang diperlukan = 2 (0,4938) = 0,9876

5.6 Latihan Soal

1. Suatu populasi terdiri dari 5 angka yaitu : 2, 3, 6, 8, dan 11. Dari populasi ini kemudian ditarik semua sampel yang beranggotakan 2 dengan pengembalian yang mungkin dapat diambil dari populasi tersebut. Hitunglah :
   a. Mean populasi
   b. Deviasi standar populasi
   c. Mean dari distribusi sampling harga mean
   d. Deviasi standar dari distribusi sampling harga deviasi standar.
2. Diketahui 500 bola yang mempunyai berat rata-rata 5,02 ons dan deviasi standar 0,3 ons. Hitunglah probabilitasnya bahwa dari sampel random 100 bola mempunyai berat keseluruhan antara :

a. Antara 4,96 dan 5,00 ons
b. Lebih dari 5,10 ons

3. Tinggi badan mahasiswa suatu Perguruan Tinggi berdistribusi normal dengan mean 68 cm dan deviasi standar 3 cm. Apabila terdapat 25 mahasiswa suatu PT, tentukan peluang rata-rata tinggi antara :
   c. Terletak antara 66,8 dan 68,3 cm
   d. Kurang dari 66,4
4. Bola lampu hasil produksi Pabrik A mempunyai umur rata-rata 1400 jam dengan deviasi standar 200 jam, sedangkan bola lampu hasil produksi Pabrik B mempunyai umur rata-rata 1200 jam dengan deviasi standar 100 jam. Jika sampel random sebanyak 125 bola lampu diambil dari masing-masing merk diuji, berapa probabilitasnya bahwa merk A mempunyai umur rata-rata paling sedikit :
   a. 160 jam lebih lama daripada merk B.
   b. 250 jam lebih lama daripada merk B.
5. Misalkan bahwa rata-rata mahasiswa Indonesia 162 cm dengan simpangan baku 6,5 cm. Jika ada 49 mahasiswa, tentukan peluang rata-rata tinggi mahasiswa :
   a. Paling rendah 155 cm
   b. Paling tinggi 175 cm
   c. Antara 158 cm dan 172 cm
   d. Kurang dari 160 cm
6. Dari pengalaman memperlihatkan bahwa 10% anggota masyarakat menderita penyakit A. Penelitian dilakukan terhadap 900 orang, tentukan :
   a. Rata-rata dan simpangan baku untuk proporsi yang menderita penyakit A
   b. Peluang sampel itu akan berisikan anggota yang menderita penyakit A :
      1. Antara 80 dan 95 orang
      2. Lebih dari 98 orang
      3. Paling banyak 75 orang
7. Merk lampu A rata-rata menyala 1400 jam dan merk lampu B rata-rata menyala 1300 jam. Simpangan bakunya masing-masing 160 jam dan 125 jam. Dari tiap populasi diambil sebuah sampel acak berukuran 85 dari lampu A dan 100 dari lampu B. Tentukan peluang rata-rata menyala lampu dalam sampel dari A paling sedikit 50 jam lebihnya dari rata-rata menyala lampu dalam sampel dari B.

8. Pengalaman mencatat bahwa 65% dari penduduk ternyata menyenangi pemimpin A. Dua buah sampel acak telah diambil masing-masing berukuran 250. Tentukan bagaimana peluangnya bahwa kedua sampel itu akan memperlihatkan perbedaan presentase lebih dari 12% yang menyenangi pemimpin A.
9. Diketahui bahwa produk yang dihasilkan mesin tertentu 2%-nya rusak. Berapa probabilitasnya bahwa dari pengiriman sebanyak 400 produk itu.
   a. 3% atau lebih ternyata rusak
   b. 2% atau kurang ternyata rusak

BAB 6

PENAKSIRAN PARAMETER

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa dapat memahami hubungan nilai sampel dan populasi dan menentukan distribusi sampling yang tepat untuk digunakan

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa dapat menjelaskan penaksiran titik dan interval parameter populasi, mengetahui jenis penaksiran parameter populasi, menggunakan penaksiran nilai rata-rata, selisih rata-rata, proporsi dan selisih proporsi yang sesuai dengan kasus serta menghitung jumlah sampel yang dibutuhkan

6.1 Pendahuluan

Dalam membuat taksiran (pendugaan) sangat diperlukan konsep probabilitas karena sangat berguna dalam pembuatan keputusan pada kondisi ketidakpastian. Setiap orang selalu pernah membuat suatu dugaan, contoh hari ini cuaca mendung, maka dugaan kita bahwa hari ini akan hujan. Seorang Manajer juga harus melakukan dugaan-dugaan. Seringkali mereka dituntut untuk membuat dugaan yang rasional dalam kondisi yang penuh ketidakpastian tanpa informasi yang lengkap. Agar dugaan yang dilakukan dapat menghasilkan suatu dugaan yang baik, maka mereka harus menguasai konsep pendugaan secara statistik, contoh: manajemen memutuskan untuk memproduksi barang pada tingkat tertentu berdasarkan kemungkinan permintaan yang akan terjadi terhadap barang tersebut. Pertimbangan yang dilakukan dapat berdasarkan pengalaman yang lalu (data histories), kondisi alam (musim hujan, musim kemarau), pesaing, dan lain sebagainya. Dalam analisis statistik, penarikan kesimpulan merupakan bagian yang sangat penting. Kesimpulan yang diambil mengenai sekelompok sampel akan digeneralisasikan terhadap populasinya. Generalisasi kesimpulan tersebut mengandung risiko bahwa akan terdapat kekeliruan atau ketidaktepatan.

Kriteria taksiran (pendugaan) yang baik, yaitu:

1. Tidak bias (*Unbiasedness*),

   Artinya statistik sampel yang digunakan sebagai penduga harus sama atau mendekati parameter populasi penduga

2. Efisiensi (*Efficiency)*,

   Artinya statistik sampel memiliki deviasi standar yang kecil

3. Konsistensi (*Consistency*),

   Artinya jika ukuran sampel meningkat maka statistik sampel akan semakin mendekati parameter populasinya.

4. Kecukupan (*Sufficiency)*,

   Artinya suatu taksiran dikatakan memiliki kecukupan jika taksiran tersebut dapat memberikan informasi yang cukup mengenai sifat populasinya.

Ada dua jenis taksiran (pendugaan) yang dilakukan terhadap populasi, yaitu:

1. **Penaksiran Titik (*Point Estimation*)**

   Penaksiran titik mengandung pengertian bahwa suatu parameter (misal μ) akan ditaksir hanya dengan menggunakan satu bilangan saja (misalnya dengan $\bar{x}$). Penaksiran titik sering mengalami kekeliruan, sehingga probabilitas suatu penaksiran titik tersebut tepat adalah sangat kecil atau mendekati nol. Sehingga penaksiran titik jarang digunakan.

   Taksiran titik untuk rata-rata populasi (μ) dan proporsi populasi (π) menggunakan rata-rata sample ($\bar{x}$) dan proporsi sample (p) yang dapat dihitung dengan menggunakan rumus:

   $$\bar{x} = \frac{\sum x}{n} \quad \text{dan} \quad p = \frac{x}{n}$$

2. **Penaksiran Interval (*Interval Estimation*)**

   Penaksiran interval merupakan interval nilai (*range*) yang nilai parameter populasi berada di dalamnya.Tujuan membuat penaksiran interval adalah mengurangi kesalahan penaksiran.

Penaksiran interval memiliki batas-batas tertentu sehingga penaksiran akan berada di antaranya. Batas-batas tersebut adalah batas bawah taksiran (*lower limit estimate*) yang merupakan nilai taksiran parameter populasi terendah dan batas atas taksiran (*upper limit estimate*) merupakan nilai taksiran parameter populasi tertinggi.

Batas-batas dalam penaksiran dengan interval harus ditunjang dengan adanya derajat keyakinan/kepastian yang biasanya dinyatakan dengan prosentase. Derajat keyakinan tersebut disebut dengan ***Confidence Coefficient***, besarnya derajat keyakinan sama dengan **1 -** α (α = tingkat kesalahan duga), misalnya: derajat keyakinan 90% maka α= 10%; derajat keyakinan 95% maka α= 5%.

Sedangkan batas-batasnya dinamakan ***Confidence Interval***.

Penaksiran interval dibedakan menjadi 2 yaitu:

1. Penaksiran rata-rata untuk data yang bersifat kontinu
2. Penaksiran proporsi untuk data yang bersifat diskrit

Penaksiran dilakukan terhadap angka-angka statistik atau Penaksiran dilakukan terhadap angka-angka statistik atau angka-angka yang diperoleh dari sampel. Sampel yang digunakan untuk perhitungan dibedakan antara sampel kecil (n< 30) dan sampel besar (n>=30), pembedaan sampel tersebut digunakan untuk pemilihan tabel distribusi yang akan digunakan dalam perhitungan.

Apabila sampel kecil maka digunakan ***Tabel Distribusi Student "t"*** dengan degree of freedom (df) atau derajat kebebasan = n-1. Apabila sampel besar maka digunakan ***Tabel Distribusi Normal Standart.***

6.2 Menaksir Rata-Rata μ

Misalkan kita mempunyai sebuah populasi berukuran N dengan rata-rata μ dan simpangan baku σ. Dari populasi ini parameter rata-rata μ akan ditaksir. Untuk keperluan ini ,ambil sebuah sampel acak berukuran n, lalu hitung statistik yang perlu ialah $\bar{x}$ dan s. Titik taksiran untuk rata-rata μ ialah $\bar{x}$. Dengan kata lain nilai μ besarnya ditaksir oleh harga $\bar{x}$ yang didapat dari sampel.

Untuk memperoleh taksiran yang lebih tinggi derajat kepercayaannya, digunakan interval taksiran atau selang taksiran disertai nilai koefisien kepercayaan yang dikehendaki.

a. Simpangan baku σ diketahui dan populasinya berdistribusi normal

$$\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

b. Simpangan baku σ diketahui dan populasinya berdistribusi normal,jika (n/N)>5%

$$\bar{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Dengan α = koefisien kepercayaan dan $z_{\frac{\alpha}{2}}$ = bilangan z didapat dari tabel normal baku untuk peluang ½ α

c. Simpangan baku σ tidak diketahui dan populasinya berdistribusi normal

$$\bar{x} - t_{\frac{\alpha}{2};\, n-1} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2};\, n-1} \cdot \frac{s}{\sqrt{n}}$$

d. Simpangan baku σ tidak diketahui dan populasinya berdistribusi normal, jika (n/N) > 5%

$$\bar{x} - t_{\frac{\alpha}{2};\, n-1} \cdot \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} < \mu < \bar{x} + t_{\frac{\alpha}{2};\, n-1} \cdot \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Dengan α = koefisien kepercayaan dan $t_{\frac{\alpha}{2}}$ = nilai t didapat dari daftar distribusi student /distribusi t dengan derajat kebebasan dk = n -1

Contoh 6.1 :

Dari sampel 100 orang pedagang premium eceran di Semarang diperkirakan nilai rata-rata isi premium botol 1 liter yang dijual yaitu sebeser 0.86 liter, dengan standar deviasi 0,3. Dengan tingkat kepercayaan 95%, lakukan penaksiran interval kepercayaan (confidence level) nilai rata rata isi premiun botol 1 liter tersebut !

Jawab :

Dengan tingkat kepercayaan 95%, $z_{\frac{\alpha}{2}} = z_{0,475} = 1,96$ dan $\frac{\sigma}{\sqrt{n}} = \frac{0,3}{\sqrt{100}} = 0,03$, maka nilai rata-rata isi botol premium yang sesungguhnya adalah :

$$\bar{x} - z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}}$$

$$0,86 - 1,96 . 0,03 < \mu < 0,86 + 1,96 . 0,03$$

$$0,86 - 0,058 < \mu < 0,86 + 0,058$$

$$0,802 < \mu < 0,918$$

Contoh 6.2 :

Dari suatu peternakan ayam, setiap kandang yang berisi 20 ekor ayam diketahui bahwa rata-rata ayam bertelur adalah 20 telur setiap bulan setiap ekornya, dengan simpangan baku 2 ekor. Hitung tingkat kepercayaan 95%, untuk rata-rata bertelur populasi ayam yang sesungguhnya !

Jawab :

Dengan tingkat kepercayaan 95%, $t_{\frac{\alpha}{2}; n-1} = 2,093$ dan $\frac{s}{\sqrt{n}} = \frac{2}{\sqrt{20}} = 0,4472$, maka rata-rata bertelur populasi ayam yang sesungguhnya adalah :

$$\bar{x} - t_{\frac{\alpha}{2};n-1} . \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2};n-1} . \frac{s}{\sqrt{n}}$$

$$20 - 2,093 . 0,4772 < \mu < 20 + 2,093 . 0,4772$$

$$20 - 0,94 < \mu < 20 + 0,94$$

$$19,06 < \mu < 20,94$$

6.3 Menaksir Proporsi

Misalkan kita mempuyai sebuah populasi berukuran N dimana terdapat proporsi π untuk peristiwa A yang ada di dalam populasi itu. Sebuah sampel acak berukuran n diambil dari populasi ini. Misalkan terdapat x peristiwa A, sehingga proporsi sampel untuk peristiwa A = (x/n). Jadi titik taksiran untuk π adalah x/n .

Untuk memperoleh taksiran yang lebih tinggi derajat kepercayaannya, digunakan interval taksiran atau selang taksiran disertai nilai koefisien kepercayaan yang dikehendaki.

a. Jika (n/N) $\leq$ 5%

$$p - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{pq}{n}} \; \langle \; \pi \; \langle \; p + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{pq}{n}}$$

b. Jika (n/N) > 5%

$$p - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}} \; \langle \; \pi \; \langle \; p + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{pq}{n}} \sqrt{\frac{N-n}{N-1}}$$

Dengan α = koefisien kepercayaan dan $z_{\frac{\alpha}{2}}$ = bilangan z didapat dari tabel normal baku untuk peluang $\frac{1}{2}\alpha$

Contoh 6.3:

Kita ingin menaksir ada berapa persen anggota masyarakat yang berumur di atas 15 tahun yang termasuk golongan kaya raya. Untuk ini sebuah sampel acak berukuran n = 1200 diambil yang menghasilkan 504 golongan kaya raya.

Jawab:

Persentase golongan kaya raya dalam sampel = 504/1200 x 100 % = 42 %

Titik taksiran adalah 42 %.

dengan p = 0,42 q = 0,58 dan $z_{0,475}$ = 1,96, maka:

$$0{,}42 - 1{,}96 \sqrt{\frac{0{,}42.0{,}58}{1200}} < \pi < 0{,}42 + 1{,}96 \sqrt{\frac{0{,}42.0{,}58}{1200}}$$

atau: $0{,}39 < \pi < 0{,}45$

Kita yakin sebesar 95% bahwa persentase anggota masyarakat yang kaya raya akan ada dalam interval 39 % dan 45 %.

Contoh 6.4:

Untuk meningkatkan pelayanan kepada konsumen, PT PSK Jaya di Tangerang melakukan survei kepuasan pelanggan. Dari 3000 pelanggan pada bulan Agustus ternyata 2100 orang menyatakan puas dan sisanya kurang puas. Buatlah interval keyakinan tentang kepuasan konsumen dengan menggunakan tingkat keyakinan 95%.

Jawab :

Persentase pelanggan puas = $2100/3000$ x 100 % = 70 %

Titik taksiran adalah 70 %.

dengan p = 0,70 q = 0,30 dan $z_{0,475}$ = 1,96, maka:

$$0{,}70 - 1{,}96 \sqrt{\frac{0{,}70.0{,}30}{2100}} < \pi < 0{,}70 + 1{,}96 \sqrt{\frac{0{,}70.0{,}30}{2100}}$$

atau: $0{,}684 < \pi < 0{,}716$

Kita yakin sebesar 95 % bahwa persentase kepuasan konsumen akan ada dalam interval 68,4 % dan 71,6 %.

Contoh 6.5:.

Seorang pejabat bank akan memperkirakan berapa persen para nasabah yang tidak puas dengan pelayanan yang diberikan oleh para pegawainya. Untuk maksud tersebut, dilakukan penelitian terhadap 250 orang nasabah yang dipilih secara acak. Ternyata ada 60 orang yang tidak puas. Dengan tingkat keyakinan 95%, buatlah pendugaan interval persentase para nasabah yang tidak puas.

Jawab :

Persentase nasabah tidak puas = $60/250$ x 100 % = 24 %

Titik taksiran adalah 24 %.

dengan p = 0,24 q = 0,76 dan $z_{0,475}$ = 1,96, maka:

$$0{,}24 - 1{,}96 \sqrt{\frac{0{,}24.0{,}76}{2100}} < \pi < 0{,}24 + 1{,}96 \sqrt{\frac{0{,}24.0{,}76}{2100}}$$

atau: $0{,}24 - 0{,}053 < \pi < 0{,}24 + 0{,}053$

$0{,}187 < \pi < 0{,}293$

Kita yakin sebesar 95 % bahwa persentase kepuasan nasabah bank akan ada dalam interval 18,7 % dan 29,3 %.

6.4 Menaksir Selisih Rata-rata

Misalkan kita mempuyai dua buah populasi, kedua-duanya berdistribusi normal. Rata-rata dan simpangan bakunya masing-masing μ_1 dan σ_1 untuk populasi kesatu, μ_2 dan

σ_2 untuk populasi kedua. Dari masing-masing populasi secara independen diambil sebuah sampel acak dengan ukuran n_1 dan n_2. Rata-rata dan simpangan baku dari sampel-sampel itu berturut-turut $\bar{x}_1$, s_1, dan $\bar{x}_2$, s_2.

Akan ditaksir selisih rata-rata ($\mu_1 - \mu_2$).

$$(\bar{x}_1 - \bar{x}_2) - z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \; \langle \; \mu_1 - \mu_2 \; \langle \; (\bar{x}_1 - \bar{x}_2) + z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Dengan α = koefisien kepercayaan dan $z_{\frac{\alpha}{2}}$ = bilangan z didapat dari tabel normal baku untuk peluang ½ α

a. Jika σ tetapi tidak diketahui besarnya. Maka besarnya s dinyatakan dengan rumus :

$$s^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}$$

Dan

$$(\bar{x}_1 - \bar{x}_2) - t_{\frac{\alpha}{2}} . s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \; \langle \; \mu_1 - \mu_2 \; \langle \; (\bar{x}_1 - \bar{x}_2) + t_{\frac{\alpha}{2}} . s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Dengan α = koefisien kepercayaan dan $t_{\frac{\alpha}{2}}$ = nilai t didapat dari daftar distribusi student /distribusi t dengan derajat kebebasan dk = $n_1 + n_2 - 2$

6.5 Menaksir Selisih Proporsi

Misalkan kita mempunyai dua buah populasi, dengan parameter untuk peristiwa yang sama masing-masing π_1 dan π_2. Dari populasi ini secara independen masing-masing diambil sebuah sampel acak berukuran n_1 dari populasi kesatu dan n_2 dari populasi kedua. Proporsi untuk peristiwa yang diperhatikan dari sampel-sampel itu adalah $p_1 = x_1/n_1$ dan $p_2 = x_2/n_2$ dengan x_1 dan x_2 berturut-turut menyatakan banyaknya peristiwa yang diperhatikan yang didapat didalam sampel kesatu dan kedua.

Akan ditentukan interval taksiran untuk $\pi_1 - \pi_2$ sebagai berikut :

$$(p_1 - p_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} \; \langle \; \pi_1 - \pi_2 \; \langle \; (p_1 - p_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Dengan α = koefisien kepercayaan dan $z_{\frac{\alpha}{2}}$ = bilangan z didapat dari tabel normal baku untuk peluang ½ α

Contoh 6.6 :

Sampel acak yang satu terdiri dari 500 pemudi dan satu lagi 700 pemuda yang mengunjungi sebuah pameran telah diambil. Ternyata bahwa 325 pemudi dan 400 pemuda menyenangi pameran itu. Tentukanlah interval kepercayaan 95 % untuk perbedaan persentase pemuda dan pemudi yang mengunjungi pameran dan menyenanginya.

Jawab :

Persentase pemudi yang menyenangi pameran = $p_1 = 325/500$ x 100 % = 65 %

Persentase pemuda yang menyenangi pameran = $p_2 = 400/700$ x 100 % = 57 %

q_1 = 35 % dan q_2 = 43 %; n_1 = 500 dan n_2 = 700

$$\sqrt{\frac{p_1\,q_1}{n_1} + \frac{p_2\,q_2}{n_2}} = \sqrt{\frac{0,65.0,35}{500} + \frac{0,57.0,43}{700}} = 0,0284$$

Sehingga diperoleh:

$$(p_1 - p_2) - z_{\frac{\alpha}{2}} \sqrt{\frac{p_1\,q_1}{n_1} + \frac{p_2\,q_2}{n_2}} \;<\; \pi_1 - \pi_2 \;<\; (p_1 - p_2) + z_{\frac{\alpha}{2}} \sqrt{\frac{p_1\,q_1}{n_1} + \frac{p_2\,q_2}{n_2}}$$

$(0,65 – 0,57) – (1,96)\,(0,0284) < \pi_1 - \pi_2 < (0,65 – 0,57) + (1,96)\,(0,0284)$

Atau : $0,024 < \pi_1 - \pi_2 < 0,136$

Jadi 95 % yakin bahwa perbedaan persentase pemudi dan pemuda yang mengunjungi pameran dan menyenanginya akan ada dalam interval yang dibatasi oleh 2,4 % dan 13,6 %

Contoh 6.7 :

PT. Reksadana Duit menawarkan portopolio baru untuk investasi. Untuk produk baru ini, perusahaan perlu mengetahui kemampuan investor dalam menghadapi resiko. Untuk keperluan tersebut diambil sampel masing-masing 120 investor tua dan muda. Hasil survei menunjukkan bahwa sebanyak 80 orang kaum tua dan 60 orang kaum muda setuju untuk menerima resiko lebih besar. Buatlah interval keyakinan untuk melihat selisih proporsi dan kemampuan menghadapi resiko tersebut dengan tingkat keyakinan 90%.

Jawab :

$n_1 = 120$ $\quad\quad$ $n_2 = 120$

Persentase kaum tua = $p_1 = 80/120 \times 100\ \% = 67\ \%$

Persentase kaum muda = $p_2 = 60/120 \times 100\ \% = 50\ \%$

$q_1 = 33\ \%$ dan $q_2 = 50\ \%$ $\quad\quad$ $Z_{0,45} = 1,65$

$$\sqrt{\frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2}} = \sqrt{\frac{0,67.0,33}{120} + \frac{0,50.0,50}{120}} = 0,063$$

Sehingga diperoleh:

$$(0,67 – 0,50) – (1,65)\ (0,063) < \pi_1 - \pi_2 < (0,67 – 0,50) + (1,65)\ (0,063)$$

$$0,17 – 0,104 < \pi_1 - \pi_2 < 0,17 + 0,104$$

atau: $$0,066 < \pi_1 - \pi_2 < 0,274$$

Jadi 90 % yakin bahwa perbedaan persentase kaum tua dan kaum muda yang berani mengambil resiko ada dalam interval yang dibatasi oleh 6,6 % dan 27,4 %

Contoh 6.8 :

BKKBN melakukan penelitian di dua daerah (D_1 dan D_2) untuk mengetahui apakah ada perbedaan antara persentase penduduk yang setuju KB di daerah tersebut. Kemudian akan dibuat pendugaan interval mengenai besarnya selisih persentase tersebut. Di daerah D_1 dan D_2 masing-masing dilakukan wawancara terhadap 120 orang, antara lain menanyakan apakah mereka setuju KB atau tidak.

Dari D_1 ada 90 orang dan D_2 ada 78 orang yang setuju KB. Buatlah pendugaan interval dari perbedaan persentase tentang pendapat penduduk yang setuju dengan KB, di kedua daerah tersebut dengan tingkat keyakinan sebesar 90%.

Jawab :

$n_1 = 120$ $n_2 = 120$

Persentase daerah $D_1 = p_1 = 90/120 \times 100\ \% = 75\ \%$

Persentase daerah $D_2 = p_2 = 78/120 \times 100\ \% = 65\ \%$

$q_1 = 25\ \%$ dan $q_2 = 35\ \%$ $Z_{0,45} = 1,65$

$$\sqrt{\frac{p_1\, q_1}{n_1} + \frac{p_2\, q_2}{n_2}} = \sqrt{\frac{0,75.0,35}{120} + \frac{0,65.0,35}{120}} = 0,059$$

Sehingga diperoleh:

$$(0,75 - 0,65) - (1,65)\,(0,059) < \pi_1 - \pi_2 < (0,75 - 0,65) + (1,65)\,(0,059)$$

$$0,10 - 0,097 < \pi_1 - \pi_2 < 0,10 + 0,097$$

atau: $0,003 < \pi_1 - \pi_2 < 0,197$

Jadi 90 % yakin bahwa perbedaan persentase daerah D_1 dan daerah D_2 tentang persentase penduduk yang setuju KB ada dalam interval yang dibatasi oleh 0,3 % dan 19,7 %

6.6 Menetapkan Jumlah Sampel

Untuk melakukan penaksiran yang lebih teliti maka kesalahan penaksiran dari parameter populasi sebenarnya akan semakin kecil. Hal tersebut hanya dimungkinkan apabila jumlah sampel semakin besar. Dengan demikian maka jumlah sampel yang harus diambil dipengaruhi oleh tiga faktor, yaitu :

1. Interval kepercayaan yang diinginkan
2. Batas maksimum kesalahan yang dibolehkan (E)
3. Variansi populasi

A Jumlah Sampel Untuk Menduga Nilai Rata-Rata apabila Deviasi Standar Populasi Diketahui :

Apabila E adalah batas maksimum kesalahan perkiraan nilai rata-rata, maka Interval kepercayaan $\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \langle\ \mu\ \langle\ \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ dapat dinyatakan dengan rumusan sebagai berikut $\bar{x} - E \ \langle\ \mu\ \langle\ \bar{x} + E$ yang mana :

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\sqrt{n} = \left(\frac{z_{\alpha/2}.\sigma}{E}\right)$$

$$n = \left(\frac{z_{\alpha/2}.\sigma}{E}\right)^2$$

Contoh 6.9 :

Diketahui : $\sigma = 10\,;\, E = 2\,;\, \alpha = 0{,}95$

Ditanyakan jumlah sampel yang dibutuhkan :

Jawab :

$$n = \left(\frac{1{,}96\,(10)}{2}\right)^2 = (9{,}8)^2 = 96{,}04 = 97$$

B Jumlah Sampel Untuk Menduga Proporsi

Apabila E adalah batas maksimum kesalahan yang diperkenankan bagi pendugaan Interval kepercayaan proporsi, maka berikut $p - E \langle \mu \langle p + E$ yang mana :

$$E = z_{\alpha/2} \cdot \sqrt{\frac{pq}{n}}$$

$$\frac{E}{z_{\alpha/2}} = \sqrt{\frac{pq}{n}}$$

$$\left(\frac{E}{z_{\alpha/2}}\right)^2 = \frac{pq}{n}$$

$$n = pq\left(\frac{z_{\alpha/2}}{E}\right)^2$$

Contoh 6.10 :

Diketahui : $p = 10\,;\, E = 2\,;\, \alpha = 0{,}99$

Ditanyakan jumlah sampel yang dibutuhkan :

Jawab :

$$n = 0{,}45.0{,}55\left(\frac{2{,}58}{0{,}10}\right)^2 = 0{,}45.0{,}55\,.665{,}64 = 164{,}75 = 165$$

6.7 Latihan Soal

1. Suatu studi tentang pertumbuhan dari tanaman cactus jenis tertentu menunjukkan bahwa dari 50 tanaman yang dianggap sebagai sampel rata-rata tumbuh 44,8 mm

dengan deviasi standar 4,7 mm selama jangka waktu 12 bulan. Dengan interval konfidensi 95 %, tentukan rata-rata pertumbuhan tahunan yang sesungguhnya dari jenis cactus tersebut.

2. Sampel random sebanyak 40 drum bahan kimia ditarik dari 200 drum bahan kimia, mempunyai berat rata-rata 240,8 pound dengan deviasi standar 12,2 pound. Jika diduga bahwa berat rata-rata dari 200 drum bahan kimia tersebut adalah 240,8, tentukan dengan interval kepercayaan 95 % untuk berat rata-rata drum bahan kimia tersebut !
3. Untuk mengetahui waktu rata-rata yang diperlukan untuk merakit suatu alat mekanis tertentu, telah dilakukan perhitungan berdasarkan sampel 6 perakitan dengan waktu masing-masing 13, 14, 12, 16, 12, dan 11 menit. Buatlah interval konfidensi 95 % untuk waktu rata-rata yang sesungguhnya untuk merakit alat mekanis tersebut.
4. Sebuah sampel berupa 10 pengukuran diameter balok kayu, menunjukkan rata-rata diameter 43,8 cm dengan deviasi standar 0,6 cm. Hitunglah interval konfidensi 99 % untuk rata-rata diameter yang sesungguhnya.
5. Sampel random sebanyak 150 buah bola lampu merk A menunjukkan daya hidup rata-rata 1400 jam dengan deviasi standar 120 jam. Sampel random lain sebanyak 200 buah bola lampu merk B mempunyai daya hidup rata-rata 1200 jam dengan dengan deviasi standar 80 jam. Hitunglah interval konfidensi 95 % untuk perbedaan rata-rata daya hidup dari populasi bola lampu kedua merk itu.
6. Dua sampel masing-masing berupa 100 tanaman bibit yang tumbuh di dua tempat yang berbeda. Dari sampel pertama tinggi rata-ratanya adalah 9,8 inci dengan deviasi standar 1 inci. Dari sampel kedua mempunyai tinggi rata-rata 10,5 inci dengan deviasi standar 3 inci. Buatlah interval konfidensi 90 % untuk perbedaan tinggi dari kedua populasi.
7. Diambil sampel 12 murid yang mengikuti pelajaran matematika dengan metode modern, kemudian diambil sampel lain 10 murid yang mengikuti pelajaran matematika dengan metode konvensional. Pada akhir semester ujian dengan soal yang sama diberikan pada masing-masing kelompok. Sampel kelompok pertama mencapai nilai rata-rata 85 dengan deviasi standar 4, sedang sampel kelompok kedua mencapai nilai rata-rata 81 dengan dengan deviasi standar 5. Hitunglah interval konfidensi 90 % untuk perbedaan antara mean populasi.

8. Sebuah sampel random terdiri dari 250 lulusan SMU di kota A, 165 orang diantaranya mengatakan bahwa mereka mengharapkan dapat melanjutkan studinya ke Perguruan Tinggi Negeri. Hitunglah interval konfidensi 99% untuk proporsi yang sesungguhnya.
9. Dari sampel random sebanyak 600 wanita yang berumur 21 tahun keatas di kota B telah diwawancarai, 378 orang diantaranya mengatakan bahwa mereka lebih memilih bekerja full time daripada parttime. Hitunglah interval konfidensi 95% untuk proporsi yang sesungguhnya.
10. Sampel random sebanyak 100 butir telur telah diambil dari 1000 butir telur yang dikirim dari daerah A ke daerah B. Dari sampel tersebut diketahui 18 diantaranya pecah atau rusak. Hitunglah interval konfidensi 95% untuk proporsi telur yang pecah atau rusak dari 1000 telur tersebut.
11. Dari sampel random sebanyak 400 ibu rumah tangga di kota A, 240 diantaranya lebih menyukai sabun cuci merk Rinso daripada merk lainnya. Sampel random lain di kota B sebanyak 200 ibu rumah tangga diketahui 80 diantaranya lebih menyukai sabun cuci merk Rinso daripada merk lainnya. Estimasikan perbedaan proporsi ibu rumah tangga yang lebih menyukai sabun cuci merk Rinso dari kedua kota itu. Gunakan interval konfidensi 95 %.
12. Dari sampel random sebanyak 400 pemirsa dewasa dan 600 pemirsa remaja yang mengikuti program siaran TV tertentu, diketahui 100 pemirsa dewasa dan 300 pemirsa remaja menunjukkan bahwa mereka menyenangi jenis siaran TV tersebut. Estimasikan perbedaan proporsi pemirsa yang menyenangi program siaran TV tersebut antara semua pemirsa dewasa dan pemirsa remaja. Gunakan interval konfidensi 95 %.

BAB 7

PENGUJIAN HIPOTESA

Standar Kompetensi :
Setelah mengikuti kuliah ini, mahasiswa mampu menjelaskan dengan tepat konsep dasar hipotesis dan mampu menguji hipotesis dengan benar .

Kompetensi Dasar :
Setelah mengikuti kuliah ini, mahasiswa dapat menjelaskan pengertian dan macam-macam hipotesis dan pengujian hipotesis, menguji hipotesis uji rata-rata, menguji hipotesis uji proporsi, menguji hipotesis uji kesamaan dan perbedaan rata-rata serta menguji hipotesis uji kesamaan dan perbedaan proporsi

7.1 Pendahuluan

Hipotesis berasal dari bahasa Yunani. *Hupo* berarti Lemah atau kurang atau di bawah dan *Thesis* berarti teori, proposisi atau pernyataan yang disajikan sebagai bukti. Sehingga dapat diartikan sebagai *Pernyataan yang masih lemah kebenarannya dan perlu dibuktikan atau dugaan yang sifatnya masih sementara.* Hipotesis ini perlu untuk diuji untuk kemudian diterima atau ditolak

Pengujian Hipotesis adalah suatu prosedur yang dilakukan dengan tujuan memutuskan apakah *menerima* atau *menolak* hipotesis mengenai parameter populasi .

Penolakan suatu hipotesis bukan berarti menyimpulkan bahwa hipotesis salah dimana bukti yang tidak konsisten dengan hipotesis.

Penerimaan hipotesis sebagai akibat tidak cukupnya bukti untuk menolak dan tidak berimplikasi bahwa hipotesis itu pasti benar

Prosedur Pengujian Hipotesis

A. Menentukan formulasi hipotesis

Formula Hipotesis dibedakan 2 jenis :

1. Hipotesis nol : suatu pernyataan yang akan diuji, hipotesis tersebut tidak memiliki perbedaan/ perbedaannya nol dengan hipotesis sebenarnya.
2. Hipotesis alternatif : segala hipotesis yang berbeda dengan hipotesis nol. Pemilihan hipotesis ini tergantung dari sifat masalah yang dihadapi

   Ho : $\theta = \theta_0$ dengan beberapa kemungkinan H_1

   H_1 : $\theta < \theta o$; $\theta > \theta o$; ataukah $\theta \neq \theta o$

   satu sisi satu sisi dua sisi

B. Menentukan taraf nyata (significant level)

Taraf nyata merupakan besarnya batas toleransi dalam menerima kesalahan hasil hipotesis terhadap nilai parameter populasinya. Besarnya taraf nyata bergantung pada keberanian pembuat keputusan yang dalam hal ini berapa besarnya kesalahan yang akan ditolerir dan besarnya kesalahan tersebut disebut sebagai daerah kritis pengujian/ daerah penolakan. Taraf nyata dalam bentuk % umumnya sebesar 1%, 5% dan 10% ditulis 0,01; 0,05; 0,1.

C. Menentukan kriteria pengujian

Bentuk pembuatan keputusan dalam menerima/ menolak hipotesis nol dengan cara membandingkan nilai α tabel distribusinya dengan nilai statistiknya sesuai dengan bentuk pengujiannya.

Penerimaan Ho : nilai uji statistiknya berada di luar nilai kritis

Penolakan Ho : nilai uji statistiknya berada dalam nilai kritis

D. Menentukan nilai uji statistic

Uji statistik merupakan rumus-rumus yang berhubungan dengan distribusi tertentu dalam pengujian hipotesis. Uji statistik yang biasa digunakan antara lain : distribusi Z, t, F dan sebagainya

E. Membuat kesimpulan

Penetapan keputusan dalam penerimaan/ penolakan hipotesis nol sesuai dengan kriteria pengujiannya. Pembuatan kesimpulan dilakukan setelah membandingkan nilai uji statistik dengan α tabel / nilai kritis

7.2 Menguji Rata-rata

Umpamakanlah kita mempunyai sebuah populasi berdistribusi normal dengan rata-rata μ dan simpangan baku σ. Akan diuji mengenai parameter rata-rata μ

Untuk pasangan hipotesa :

1. Pengujian dua pihak

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

2. Pengujian satu pihak (pihak kanan)

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \succ \mu_0$$

3. Pengujian satu pihak (pihak kiri)

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \prec \mu_0$$

a. Jika σ diketahui, dengan sebuah harga yang diketahui,

Uji statistik : $z_{hitung} = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$

H_0 kita terima jika $-z_{(0,5-\alpha/2)} \prec z_{hitung} \prec z_{(0,5-\alpha/2)}$ untuk pasangan hipotesa yang pertama, $z_{hitung} \leq z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang kedua dan $z_{hitung} \geq -z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang ketiga.

b. Jika σ diketahui, dengan sebuah harga yang diketahui,

Uji statistik : $t_{hitung} = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}}$

H_0 kita terima jika $-t_{\alpha/2;\,n-1} \prec t_{hitung} \prec t_{\alpha/2;\,n-1}$ untuk pasangan hipotesa yang pertama, $t_{hitung} \leq t_{\alpha;\,n-1}$ untuk pasangan hipotesa yang kedua dan $t_{hitung} \geq -t_{\alpha;\,n-1}$ untuk pasangan hipotesa yang ketiga.

Contoh 7.1 :

Populasi ikan bandeng di pemancingan "Mania" mempunyai panjang rata-rata 80 cm dengan simpangan baku 7 cm. Setelah 2 tahun beroperasi, konsumen meragukan panjang ikan tersebut. Guna meyakinkan keabsahan hipotesis itu, seorang peneliti mengambil sampel acak 100 ekor ikan bandeng dan diperoleh hasil perhitungan panjang rata-rata ikan adalah 83 cm dan standar deviasinya tetap. Apakah ada alasan untuk meragukan bahwa rata-rata panjang ikan bandeng yang ada di pemancingan "Mania" sama dengan 80 cm pada taraf signifikan 5% ?

Jawab :

Diketahui : $\mu_0 = 80$ cm ; $\sigma = 7$ cm ; n = 100 ; $\bar{x} = 83$ cm ; $\alpha = 5\%$

a. Formula Hipotesis

Ho : $\mu = 80$

H_1 : $\mu \neq 80$

b. Taraf nyata dan nilai *z* tabel

$\alpha = 5\%$; $z_{(0,5-\alpha/2)} = z_{0,475} = 1,96$ (Uji dua arah)

c. Kriteria pengujiannya

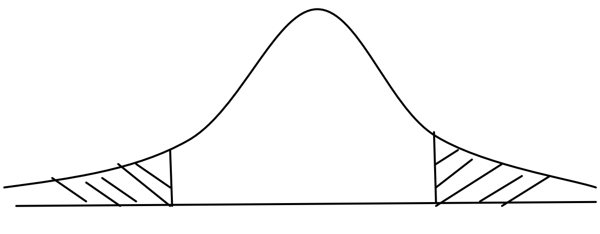
Ho diterima jika : -1,96 < Z < 1,96

d. Uji Statistik

$$z_{hitung} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

$$z_{hitung} = \frac{83 - 80}{7 / \sqrt{100}}$$

$$z_{hitung} = 4,29$$



Karena $Z_{hitung} > 1,96 \rightarrow$ Ho ditolak

e. Kesimpulan

Pada taraf nyata 5% terdapat perbedaan signifikan $\bar{x} = 83$ cm dengan $\mu_0 = 80$ cm tidak terjadi karena faktor kebetulan.

Contoh 7.2 :

Seorang peneliti ingin mengetahui apakah alat penangkap ikan rata-rata masih tetap mampu menangkap 30 ekor ikan atau lebih kecil dari itu. Data-data sebelumnya diketahui bahwa simpangan bakunya 25 ekor. Sampel yang diambil 100 alat diteliti dan diperoleh rata-rata tangkap 27 ekor. Apakah nilai tersebut masih dapat diterima sehingga alat itu mampu menangkap 30 ekor?. Ujilah dengan taraf nyata 5%.

Jawab :

Diketahui :

$\mu_0 = 30$; $\sigma = 25$; n = 100 ; $\bar{x} = 27$; $\alpha = 5\%$

a. Formula Hipotesis

   Ho : $\mu = 30$

   H_1 : $\mu < 30$

b. Taraf nyata dan nilai *Z* tabel

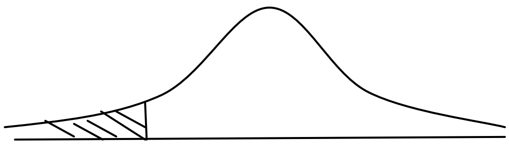
   $\alpha = 5\%$ maka $z_{(0,5-\alpha)} = z_{0,45} = 1,64$

c. Kriteria pengujiannya

   Ho diterima jika $z \geq -z_{(0,5-\alpha)}$: $z_{hitung} \geq -1,64$

d. Uji Statistik

$$z_{hitung} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{27 - 30}{25 / \sqrt{100}} = -1,2$$



-1,64

   Karena $z_{hitung} = -1,2 \geq -1,64 \rightarrow$ Ho diterima

e. Kesimpulan ; Rata-rata menangkap ikan sebesar 30 ekor.

Contoh 7.3 :

Hasil tangkapan 15 ekor ikan (kg) gurami memiliki berat rata-rata 1,208 kg dengan simpangan baku 0,02 kg. Jika taraf nyata 1%, dapatkah diyakini bahwa populasi ikan gurami rata-rata memiliki berat 1,2 kg?

Jawab :

n = 15 ; α = 0,01 ; μ_0 = 1,2 ; s = 0,02 ; $\bar{x}$ = 1,208

a. Formula Hipotesis

Ho : $\mu_0 = 1,2$

H_1 : $\mu_0 \neq 1,2$

b. Taraf nyata dan nilai *t* tabel

α = 0,01 → $t_{\frac{0,01}{2};15-1}$ = 2,977

c. Kriteria pengujiannya

Ho diterima jika : -2,977 ≤ t_{hitung} ≤ 2,977

d. Uji Statistik

$$t_{hitung} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$= \frac{1,208 - 1,2}{0,02/\sqrt{15}} = -1,47$$

e. Kesimpulan

Ho diterima, artinya populasi ikan gurami memiliki berat rata-rata 1,2 kg.

7.3 Menguji Proporsi

Untuk pasangan hipotesa :

1. Pengujian dua pihak

$$H_0 : p = p_0$$
$$H_1 : p \neq p_0$$

2. Pengujian satu pihak (pihak kanan)

$$H_0 : p = p_0$$
$$H_1 : p \succ p_0$$

3. Pengujian satu pihak

$$H_0 : p = p_0$$
$$H_1 : p \prec p_0$$

Untuk ini digunakan pendekatan oleh distribusi normal dengan statistik :

$$z_{hitung} = \frac{p - p_0}{\sqrt{p_0(1-p_0)/n}}$$

H_0 kita terima jika $-z_{(0,5-\alpha/2)} \prec z_{hitung} \prec z_{(0,5-\alpha/2)}$ untuk pasangan hipotesa yang pertama, $z_{hitung} \leq z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang kedua dan $z_{hitung} \geq -z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang ketiga.

Contoh 7.4 :

Hasil penelitian yang sudah dilakukan di suatu daerah tertinggal, dinyatakan bahwa 40% anak-anak balita menderita kekurangan gizi. Pernyataan tersebut akan diuji dengan taraf nyata 5%, untuk itu diambil sampel sebanyak 250 anak-anak balita di daerah tersebut dan diperoleh 39% diantaranya menderita kekurangan gizi. Apakah pernyataan tersebut benar?

Jawab :

n = 250 ; α = 0,05 ; p_0= 0,40 ; p = 0,39

a. Formula Hipotesis

Ho : p = 0,40

H_1 : $p \neq 0{,}40$

b. Taraf nyata dan nilai *Z* tabel

$\alpha = 0{,}05 \rightarrow z_{(0,5-\alpha/2)} = z_{(0,5-0,025)} = 1{,}96$

c. Kriteria pengujiannya

Ho diterima jika : $-1{,}96 < z_{hitung} < 1{,}96$

d. Uji Statistik

$$z_{hitung} = \frac{p - p_0}{\sqrt{p_0(1-p_0)/n}}$$

$$z_{hitung} = \frac{0{,}39 - 0{,}40}{\sqrt{0{,}40(1-0{,}40)/250}} = -0{,}33$$

e. Kesimpulan

Karena $z_{hitung} = -0{,}33 > -1{,}96$ dan $z_{hitung} = -0{,}33 < 1{,}96$ Ho diterima, artinya memang benar bahwa 40% balita menderita kekurangan gizi.

7.4 Menguji Kesamaan Dua Rata-rata

Untuk pasangan hipotesa :

a. Pengujian dua pihak

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

b. Pengujian satu pihak (pihak kanan)

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \succ \mu_2$$

c. Pengujian satu pihak

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \prec \mu_2$$

a. Jika $\sigma = \sigma_1 = \sigma_2$ dan besarnya diketahui, digunakan statistik :

$$z_{hitung} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

H_0 kita terima jika $-z_{(0,5-\alpha/2)} \prec z_{hitung} \prec z_{(0,5-\alpha/2)}$ untuk pasangan hipotesa yang pertama, $z_{hitung} \leq z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang kedua dan $z_{hitung} \geq -z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang ketiga.

b. Jika $\sigma = \sigma_1 = \sigma_2$ dan besarnya tidak diketahui, digunakan statistik :

$$t_{hitung} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\left(\frac{(n_1 - 1)\, s_1{}^2 + (n_2 - 1)\, s_2{}^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

H_0 kita terima jika $-t_{\alpha/2;\, n_1+n_2-2} \prec t_{hitung} \prec t_{\alpha/2;\, n_1+n_2-2}$ untuk pasangan hipotesa yang pertama, $t_{hitung} \leq t_{\alpha;\, n_1+n_{2-2}}$ untuk pasangan hipotesa yang kedua dan $t_{hitung} \geq -t_{\alpha;\, n_1+n_{2-2}}$ untuk pasangan hipotesa yang ketiga.

Contoh 7.5 :

Perusahaan penangkapan ikan menggunakan dua alat yang berbeda pada dua lokasi daerah penangkapan. Daerah penangkapan I terdiri dari 12 alat A sedangkan daerah penangkapan II terdiri dari 10 alat B. Waktu perendaman alat A adalah 2 jam dengan simpangan baku 0.4 jam sedangkan alat B adalah 4 jam dengan simpangan baku 0.5 jam. Yakinkah anda bahwa alat A lebih cepat perendamannya dengan taraf signifikan 1 %? (Asumsikan dua populasi berdistribusi normal dengan variansi yang sama.)

Jawab :

Diketahui :

Sampel A ; n_1 = 12 ; $\overline{x_1}$ = 2 ; s_1= 0.4

Sampel B ; n_2 = 10 ; $\overline{x_2}$ = 4 ; s_2= 0.5

a. Formula Hipotesis

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \succ \mu_2$$

b. Taraf nyata dan nilai t tabel

α = 0,01 : dk = n_1+ n_2 - 2 = 20 maka : t(0,01,20)= 2,528

c. Kriteria pengujiannya

Ho diterima jika $t_{hitung} \leq t_{\alpha;\, n_1+n_{2-2}}$: $t_{hitung} \leq 2{,}528$

d. Uji Statistik

$$t_{hitung} = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\left(\frac{(n_1-1)\,s_1{}^2 + (n_2-1)\,s_2{}^2}{n_1+n_2-2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = 11{,}45$$

karena t_{hitung} > 2,528, maka Ho ditolak

e. Kesimpulan

Waktu perendaman 2 alat memiliki perbedaan yang signifikan. Dimana alat A lebih cepat dibanding dengan alat B.

7.5 Menguji Kesamaan Dua Proporsi

Untuk pasangan hipotesa :

1. Pengujian dua pihak

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2$$

2. Pengujian satu pihak (pihak kanan)

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \succ p_2$$

3. Pengujian satu pihak

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \prec p_2$$

Untuk ini digunakan pendekatan oleh distribusi normal dengan statistik :

$$Z_{\text{hitung}} = \frac{p_1 - p_2}{\sqrt{p(1-p)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

Dimana : $p_1 = \dfrac{X_1}{n_1}$; $p_2 = \dfrac{X_2}{n_2}$ dan $p = \dfrac{X_1 + X_2}{n_1 + n_2}$

H_0 kita terima jika $-z_{(0,5-\alpha/2)} \prec z_{\text{hitung}} \prec z_{(0,5-\alpha/2)}$ untuk pasangan hipotesa yang pertama, $z_{\text{hitung}} \leq z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang kedua dan $z_{\text{hitung}} \geq -z_{(0,5-\alpha)}$ untuk pasangan hipotesa yang ketiga.

Contoh 7.6 :

Seorang ahli farmakologi mengadakan percobaan dua macam obat anti hipertensi. Obat pertama diberikan pada 100 ekor tikus dan ternyata 60 ekor menunjukkan perubahan tekanan darah. Obat kedua diberikan pada 150 ekor tikus dan ternyata 85 ekor berubah tekanan darahnya. Apakah ada perbedaan antara obat pertama dan obat kedua? Ujilah dengan derajat kebebasan 5%.

Jawab :

$$p_1 = \frac{60}{100} = 0{,}60\,;\, p_2 = \frac{85}{150} = 0{,}56\,;\, p = \frac{60+85}{100+150} = 0{,}58;\,(1\text{-}p) = 0{,}42$$

a. Formula Hipotesis

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2$$

b. Taraf nyata dan nilai *t* tabel

α = 0,05 → $z_{(0,5-\alpha/2)} = z_{(0,5-0,025)}$ = 1,96

c. Kriteria pengujiannya

Ho diterima jika : -1,96 < z_{hitung} < 1,96

d. Uji Statistik

$$Z_{hitung} = \frac{p_1 - p_2}{\sqrt{p.(1\text{-}p)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

$$Z_{hitung} = \frac{0{,}60 - 0{,}56}{\sqrt{0{,}52.0{,}48\left[\frac{1}{100} + \frac{1}{150}\right]}} = 0{,}66$$

e. Kesimpulan

Ho diterima, artinya tidak ada perbedaan antara obat pertama dan obat kedua

Contoh 7.7 :

Dua orang perawat A dan B masing-masing telah bekerja selama 7 dan 10 tahun. Kepala Puskesmas beranggapan persentase melakukan kesalahan perawat A lebih besar daripada B, untuk menguji hipotesis tersebut diambil sampel sebanyak 50 pasien yang dirawat oleh perawat A dan 60 pasien oleh perawat B, dari sampel tersebut perawat A membuat 15% kesalahan perawatan dan perawat B 12%. Apakah klaim kepala puskesmas tersebut betul? Ujilah dengan derajat kebebasan 5%.

Jawab :

$$p_A = \frac{0{,}15}{50} = 0{,}003\,;\, p_B = \frac{0{,}12}{60} = 0{,}002\,;\, p = \frac{0{,}15+0{,}12}{50+60} = 0{,}002;\, q = 0{,}998$$

a. Formula Hipotesis

$$H_0 : p_A = p_B$$
$$H_1 : p_A \succ p_B$$

b. Taraf nyata dan nilai *t* tabel

$\alpha = 0,05 \rightarrow z_{(0,5-\alpha)} = z_{(0,5-0,05)} = 1,64$

c. Kriteria pengujiannya

Ho diterima jika $z_{\text{hitung}} \leq 1,64$

d. Uji Statistik

$$Z_{hitung} = \frac{p_1 - p_2}{\sqrt{p.q\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

$$Z_{\text{hitung}} = \frac{0,003 - 0,002}{\sqrt{0,002.0,998\left[\frac{1}{50} + \frac{1}{60}\right]}} = 0,125$$

e. Kesimpulan

Ho diterima, artinya klaim kepala puskesmas tidak benar.

7.6 Latihan Soal :

1. Sampel random 12 murid dari pendidikan sekretaris, dalam tes mengetik rata-rata kecepatannya dapat mencapai 73,8 kata per menit dengan deviasi standar 7,9 kata. Dengan taraf nyata 0,01 ujilah pendapat bahwa murid-murid dari pendidikan sekretaris tersebut rata-rata dapat mengetik kurang dari 75 kata per menit.
2. Suatu pabrik cat tembok mengatakan bahwa setiap kaleng cat tembok hasil produksinya dapat menyapu tembok rata-rata seluas 10 m^2. Untuk menyakinkan apakah pendapat tersebut benar, sejumlah kaleng cat tembok tersebut dicoba hasilnya adalah (m^2) : 9,8 ; 10,2 ; 10 ; 9,4 ; 9,6 ; 10,5. Bagaimana kesimpulannya bila dipergunakan taraf signifikansi 0,05 ?
3. Sampel random sebanyak 30 kaleng susu Dancow dengan label “berat bersih 400 gram” mempunyai berat bersih rata-rata 398 gram dengan deviasi standar 6 gram. Dengan taraf nyata 0,01 ujilah pendapat bahwa berat bersih susu Dancow tersebut rata-rata kurang dari 400 gram.
4. Suatu proses produksi hanya akan menguntungkan apabila dapat menaikkan produksi rata-ratanya menjadi lebih besar dari 15 unit setiap jamnya. Untuk dapat mengambil keputusan apakah akan menggunakan mesin baru atau tidak diadakan percobaan dengan 9 mesin baru dan ternyata menghasilkan rata-rata 16,5 unit untuk setiap jam dengan deviasi standar 2,8 unit. Bagaimana keputusan yang harus diambil bila dipergunakan taraf signifikansi 0,05?

5. Seorang ahli gizi berpendapat bahwa 75% dari murid-murid SD didaerah A menderita kekurangan gizi. Suatu sampel telah dilaksanakan dengan meneliti 300 murid SD dan ternyata diketahui 206 anak menderita kekurangan gizi. Dengan taraf signifikansi 0,01 ujilah apakah murid SD yang kekurangan gizi kurang dari 75%.?
6. Dua belas pohon jeruk varietas tertentu yang dipilih secara random mempunyai tinggi rata-rata 13,6 kaki dengan deviasi standar 1,2 kaki, dan 15 pohon jeruk dengan varietas lainnya yang dipilih secara random mempunyai tinggi rata-rata 12,7 kaki dengan deviasi standar 1,5 kaki. Apakah perbedaan mean kedua sampel tersebut signifikan. Pergunakan taraf signifikansi 0,01.
7. Diduga sekurang-kurangnya 60% penduduk di suatu daerah mendukung perkara perdata oleh suatu kota tetangga yang berdekatan. Kesimpulan apakah yang Anda tarik bila hanya 110 diantara 200 orang yang diambil secara random mendukung perkara tersebut? Gunakan taraf nyata 4%
8. Sebuah pabrik rokok memproduksi dua merek rokok yang berbeda. Ternyata 56 orang diantara 200 perokok menyukai merek A dan 29 diantara 150 perokok menyukai merek B. Dapatkah kita menyimpulkan pada taraf nyata 0,06 bahwa merek A terjual lebih banyak daripada merek B?
9. Manajer pemasaran sebuah produk aditif bahan bakar mengatakan bahwa jumlah rata-rata produk aditif yang terjual adalah 1500 botol. Seorang karyawan di pabrik ingin menguji pernyataan manajer pemsaran dgn mengambil sampel selama 36 hari. Dia mendapati bahwa jumlah penjualan rata-ratanya adalah 1450 botol. Dari catatan yang ada, deviasi standart penjualan 120 botol. Dengan menggunakan $\alpha = 0,01$, apakah kesimpulan yang dapat ditarik oleh karyawan tersebut ?
10. Dari dua populasi normal yang bebas ditarik dua sampel random berukuran $n_1 = 35$ dan $n_2 = 50$ yang menghasilkan rata-rata 85 dan 78 dengan simpangan baku 5,4 dan 3,6. Ujilah hipotesis pada taraf nyata 5% bahwa $\mu_1 = \mu_2$ dengan alternatifnya $\mu_1 \neq \mu_2$

DAFTAR PUSTAKA

Anto D

ajan. 1972.*Pengantar Metode Statistik Jilid I.*LP3ES Jakarta

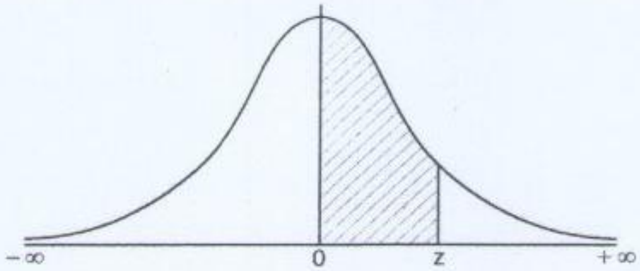
Anto Dajan. 1974.*Pengantar Metode Statistik Jilid2*.LP3ES Jakarta

Djarwanto PS, Kumpulan Soal dan Penyelesaiaannya , BPEE, Yogyakarta

Sudjana, Metoda Statistika, Penerbit Tarsito, Bandung

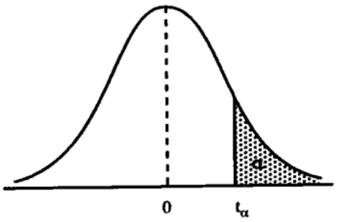
Lind, Douglas, William marchal and Samuel Wathern., Statistical Techniques in Business and Economics, 14th Edition., Mc Graw Hill 2010

Tabel Distribusi Normal Standart



z	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

Tabel Distribusi Student



v	α = 0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	22.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf.	1.282	1.645	1.960	2.326	2.576

*Tabel diambil dari Tabel IV R. A. Fisher, *Statistical Methods for Research Workers,* Oliver & Boyd Ltd., Edinburgh, dengan izin pengarang dan penerbit.