DIKTAT KULIAH

PROBABILITAS DAN STATISTIKA

TEP4413

Oleh

Nur Hayati, S.ST, MT

PROGRAM STUDI TEKNIK ELEKTRO FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH YOGYAKARTA

2017

PRAKATA

Bismillaahirrahmaanirrahiim.

Alhamdulillah dengan mengucapkan puji syukur ke hadirat Allah SWT atas segala rahmat dan karunia Nya serta ucapan Sholawat dan salam bagi Nabi Besar Muhammad SAW, karena berkat rahmat dan ridho-Nya penulis dapat menyelesaikan diktat kuliah berjudul "Pobabilitas dan Statistika". Penulis menyadari dengan sepenuh hati, bahwa tanpa bantuan, dukungan, dan bimbingan serta doa yang tulus dari berbagai pihak sangatlah sulit untuk menyelesaikan buku ini. Untuk itu, dalam kesempatan yang berharga dan terbatas ini ijinkan penulis menyampaikan ucapan terima kasih dengan tulus kepada:

1. Allah SWT atas segala rahmat dan karunia Nya serta ucapan Sholawat dan salam bagi Nabi Besar Muhammad SAW, karena berkat rahmat dan ridho-Nya saya dapat menyelesaikan penulisan disertasi ini.
2. Orangtua khususnya ibu atas doa dan dukungan yang luar biasa, yang selalu sabar menemani penulis.
3. Rekan-rekan staf pengajar dan staf administrasi DTE FT UMY atas semua diskusi, dukungan dan kerjasamanya.
4. Semua pihak yang telah membantu dan mendoakan penulis selama ini, yang dengan rasa hormat tidak dapat ditulis satu per satu di ruang terbatas ini.

Semoga Allah membalas kebaikan bapak, ibu, saudara dan semua pihak dengan pahala yang berlipat ganda, keberkahan, kebahagiaan dan kemuliaan dunia dan akhirat. Penulis menyadari bahwa buku teks ini masih jauh dari sempurna. Untuk itu segala kritik dan saran yang bersifat membangun akan penulis terima dengan lapang dada. Akhirnya, semoga buku teks ini dapat bermanfaat dalam proses belajar-mengajar di Jurusan Teknik Elektro Fakultas Teknik Universitas Muhammadiyah Yogyakarta. Aamiin Yaa Robbal 'Alaamiin.

Yogyakarta, Juni 2017

Penulis

DAFTAR ISI

BAB 1. KONSEP DASAR PROBABILITAS

A. Eksperimen Acak

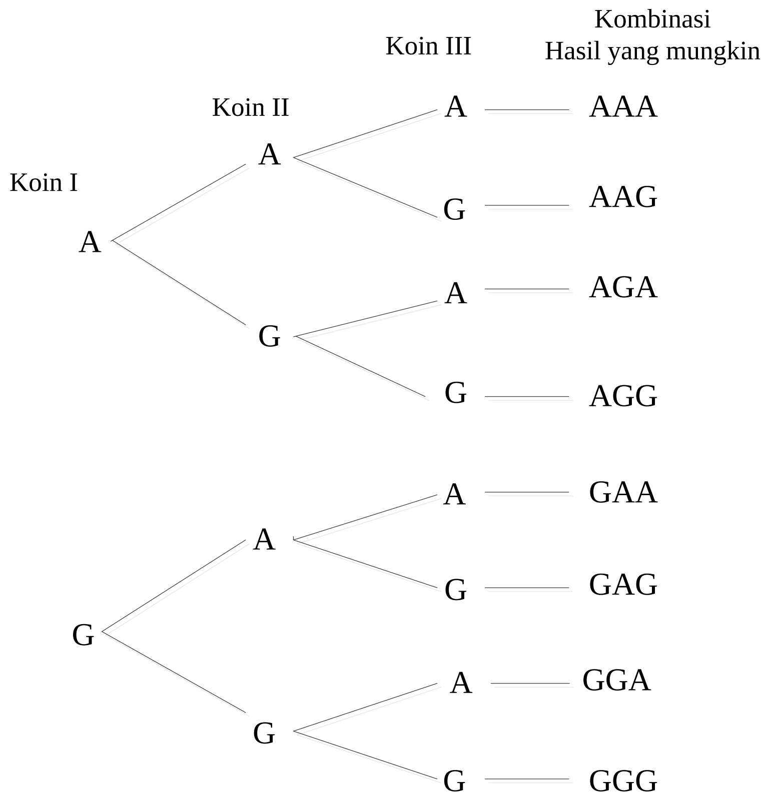
Eksperimen merupakan bagian yang penting dalam bidang sains dan engineering. Eksperimen bermanfaat karena peneliti dapat mengasumsi-kan bahwa jika melakukan percobaan/ eksperimen tertentu dalam kondisi yang kurang lebih identik maka hasil yang diperoleh juga hampir sama.Dalam kondisi diatas peneliti bisa mengendalikan nilai dari variabel-variabel yang mempengaruhi hasil dari eksperimen. Namun, dalam beberapa eksperimen ada kondisi dimana peneliti tidak dapat memastikan ataupun mengendalikan nilai dari variabel-variabel tertentu, sehingga hasil dari eksperimen pertama berbeda dengan hasil setelahnya meskipun mayoritas kondisinya sama.Eksperimen inilah yang disebut sebagai eksperimen acak

Contoh eksperimen acak adalah sebagai berikut. Jika sebuah dadu dilempar ke atas, maka kemungkinan mata dadu yang muncul adalah 1, 2, 3, 4, 5 atau 6. Percobaan pelemparan sebuah koin. Setelah koin dilempar dan jatuh, maka kemungkinan yang muncul adalah "Gambar" atau "Angka". Jika duah koin dilempar keatas maka kemungiinannya adalah : AA, AG, GG, GA. Suatu pabrik memproduksi sejenis produk kesehatan. Kemungkinan produk yang dihasilkan adalah produk yang "cacat" dan "tidak cacat".

B. Ruang Sampel dan Titik Sampel

Ruang Sampel yang dilambangkan dengan S (set) merupakan himpunan dari semua hasil yang mungkin terjadi dari eksperimen acak. Setiap hasil ekperimen atau kemungkinan-kemungkinan yang akan muncul dalam ruang sampel disebut sebagai titik sampel. Sehingga titik sampel merupakan unsur atau anggota dari ruang sampel. Contoh ruang sampel dan titik sampel adalah sebagai berikut. Jika sebuah dadu dilempar ke atas, maka tentukan ruang sampel kemungkinan mata dadu yang muncul adalah 1, 2, 3, 4, 5 atau 6. Sehingga ruang sampelnya dinyatakan dengan S = {1, 2, 3, 4, 5,6}. Suatu pabrik memproduksi sejenis produk kesehatan. Kemungkinan produk yang dihasilkan adalah produk yang "cacat" dan "tidak cacat". Sehingga ruang sampel dari sebuah produk yang dihasilkan oleh pabrik dapat dinyatakan dengan S = {Cacat, Tidak Cacat}. Sebuah koin dilempar ke atas. Setelah jatuh, maka kemungkinan sisi yang muncul paling atas adalah "Gambar" atau "Angka". Sehingga ruang sampel percobaan tersebut dapat dinyatakan dengan S = {Angka, Gambar}.

Selanjutnya setelah mengetahui ruang sampel dan titik sampelnya maka hal yang bisa dilakukan adalah melakukan penghitungan titik sampel. Sebagai contoh terdapat tiga buah koin (uang logam) dilemparkan sekali, maka berapa banyaknya titik sampel dalam ruang sampel ? Pada percobaab tersebut koin I dapat menghasilkan 2 hasil yang mungkin, yaitu Angka (A) atau gambar (G), koin II dapat menghasilkan 2 hasil yang mungkin, A atau G demikian pula koin III dapat menghasilkan 2 hasil yang mungkin, A atau G. Sehingga jumlah titik sampel yang dihasilkan adalah (2).(2).(2) atau sama dengan 8 sebagaimana diilustrasikan oleh gambar berikut.



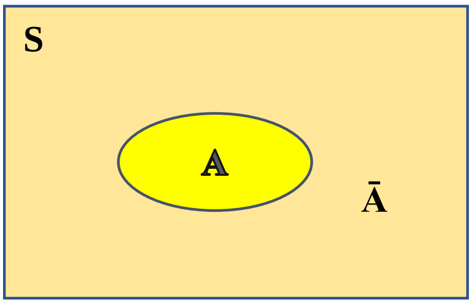
Gambar 1.1 Diagram pohon perhitungan titik sampel

C. Permutasi dan Kombinasi

Permutasi merupakan susunan dari suatu himpunan obyek yang dapat dibentuk yang memperhatikan urutan, sedangkan kombinasi merupakan susunan dari suatu himpunan obyek yang dapat dibentuk tanpa memperhatikan urutan. Banyaknya permutasi n obyek berlainan adalah n! (n faktorial) atau n x (n-1) x (n-2) x ...... x 4 x 3 x 2 x 1. Banyaknya permutasi n obyek berlainan bila diambil r sekaligus adalah $P_r^n = \frac{n!}{(n-r)!}$ sedangkan banyaknya permutasi n benda berlainan yang disusun melingkar adalah (n-1)!. Sementara itu jumlah kombinasi ari n obyek yag berlainan jika diambil sebanyak r adalah $C_r^n = \frac{n!}{r!\,(n-r)!}$

D. Kejadian

Kejadian merupakan salah satu sub himpunan (subset dari ruang sampel atau biasa disebut sebagai himpunan bagian dari ruang sampel. Kejadian dilambangkan dengan himpunan A dimana anggota-anggota dari A adalah titik sampel. Pada kasus pelemparan satu buah dadu dengan himpunan semesta S={1,2,3,4,5,6} dan A={2} dapat diartikan bahwa A adalah kejadian muncul mata dadu 2 sehingga A $\subseteq$ S dimana A merupakan himpunan bagian dari S dan 2 $\in$ A dibaca dengan 2 elemen A dimana 2 disebut sebagai titik sampel. Sebagaimana kejadian maka S merupakan kejadian pasti karena salah satu dari elemen S pasti muncul sedangkan himpunan kosong (Ø atau { }) disebut sebagai kejadian mustahil karena dari Ø tidak mungkin muncul. Dengan menggnakan operasi-operasi himpunan terhadap kejadian-kejadian dalam S, maka akan diperoleh kejadian-kejadian lain dalam S. Selanjutnya akan dibahas sekilas tentang diagram venn. Diagram venn merupakan gambaran dari hubungan antara kejadian dan ruang sampel. Perhatikan gambar diagram venn berikut.



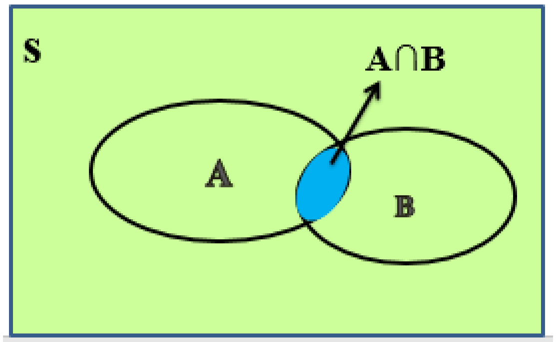
S = ruang sampel

A = himpunan bagian dari S

Ā = komplemen dari A

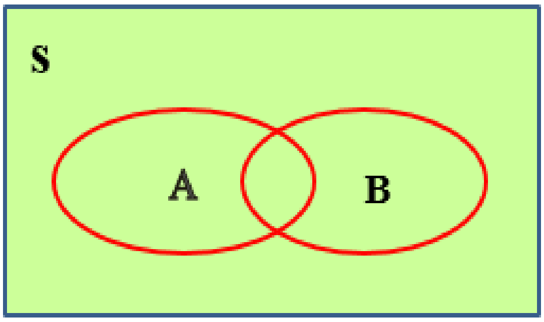
Gambar 1.2 Diagram Venn

Komplemen suatu kejadian A terhadap ialah himpunan semua unsur S yang tidak termasuk A. Komplemen dinyatakan dengan lambang Ā atau A'. Misal suatu ruang sampel S={buku, pensil, jurnal, majalah, koran}, jika A = {buku,pensil, jurnal} maka komplemen dar A atau ditulis dengan Ā atau A'={ majalah, koran}. Dengan menggunakan operasi-operasi himpunan terhadap kejadian- kejadian dalam S, maka akan diperoleh kejadian-kejadian lain dalam S. Jenis lain kejadian tersebut diantaranya adalah irisan dan gabungan. Irisan adalah interseksi antara dua kejadian. Misalnya jika A dan B adalah dua buah kejadian maka irisan keduanya sering ditulis A ∩ B : kejadian yang unsurnya termasuk A dan B atau dilambangkan dengan A ∩ B = { x : x ∈ A dan x ∈ B} atau dinyatakan bahwa A irisan B = x sedemikian rupa sehingga x elemen A dan x elemen B.



Gambar 1.3 Diagram venn Irisan

Operasi himpunan terhadap kejadian- kejadian dalam S dalam bentuk lainya adalah gabungan. Gabungan antara dua kejadian A dan B adalah kejadian-kejadian yang mengandung semua unsur yang termasuk A atau B atau keduanya. Gabungan kedua kejadian tersebut sering ditulis $A \cup B$ dimana secara lengkap dapat dituliskan sebagai $A \cup B = \{ x : x \in A,\ x \in B$ atau $x \in AB\}$ yang dinyatakan dengan A gabungan B = x sedemikian rupa sehingga x elemen A, x elemen B atau x elemen AB.



Gambar 1.3 Diagram venn Gabungan

E. Konsep dan Perumusahan Probabilitas

Dalam eksperimen acak selalu ada ketidakpastian mengenai apakah suatu kejadian khusus akan atau tidak akan terjadi. Untuk memudahkan dalam mengukur peluang atau probabilitas atas kejadian khusus tersebut, dimana dengan ukuran ini seorang peneliti dapat mengharapkan munculnya kejadian, maka probabilitas dinyatakan dalam angka pecahan antara 0 sampai 1 atau dalam persentase. Probabilitas 0 atau 0% dimana semakin mendekati 0 atau nilai semakin kecil maka kemungkian munculnya peluang semakin sedikit atau bahkan tidak akan terjadi. Probabilitas 1 atau 100% artinya semakin besar kemungkinan terjadinya suatu event atau bahkan pasti terjadi.

Contoh penulisan probabilitas dalam desimal atau persentase adalah sebagai berikut. Pada hari Jumat adalah penutupan bursa saham, maka kebanyakan investor berusaha meraih keuntungan melalui penjualan saham atau yang biasanya diistilahkan profit taking, sehingga probabilitas menjual mencapai 0,7 sedangkan membeli 0,3. Melihat kondisi kesiapan mahasiswa yang mengikuti ujian Kalkulus II, maka mahasiswa yang mempunyai probabilitas untuk lulus 70% dan tidak lulus 30%.

Terdapat dua pendekatan dalam perumusan probabilitasatau suatu kejadian yaitu pendekatan klasik dan pendekatan relatif. Pada pendekatan klasik diasumsikan bahwa semua peristiwa mempunyai kesempatan yang sama untuk terjadi (*equally likely*). Rumusan untuk probabilitas dapat dinyatakan sebagai berikut

$$\text{Probabilitas} = \frac{\text{Jumlah kemungkinan hasil (peristiwa)}}{\text{Jumlah total kemungkinan hasil}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots..\ldots\ldots\ldots\ldots\ldots\ldots(1.1)$$

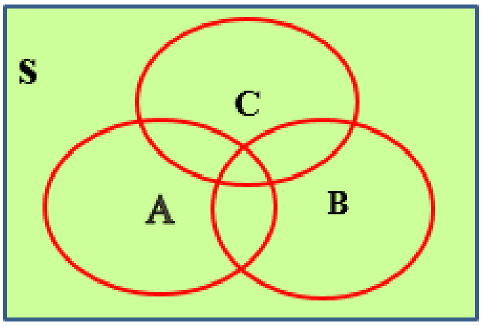
Bila kejadian A terjadi sebanayk m cara dari seluruh n cara maka probabiltas kejadian E ditulis dengan P(A) adalah = m/n. Pada suatu percobaan hanya 1 peristiwa yang terjadi, dan peristiwa lain tidak mungkin terjadi pada waktu yang bersamaan maka dikenal sebagai peristiwa saling lepas. “Peristiwa saling lepas (mutually exclusive) adalah terjadinya suatu peristiwa sehingga peristiwa yang lain tidak terjadi pada waktu yang bersamaan”. Pada suatu percobaan atau kegiatan semua hasil mempunyai probabilitas yang sama, dan hanya satu peristiwa yang terjadi maka peristiwa ini dikenal dengan lengkap terbatas kolektif (*collection exhaustive*). Pada kegiatan mahasiswa belajar semua hasil ada yang sangat memuaskan, memuaskan dan terpuji. Jumlah hasil ada 3 dan hanya 1 peristiwa yang terjadi, maka probabilitas setiap peristiwa adalah 1/3(pasti).

Pada pendekatan frekuensi relatif jika suatu eksperimen diulang sebanyak *n* kali, dimana *n* sangat besar sampai mendekati nilai tak hingga, dan jika suatu kejadian terjadi sebanyak *m* kali, maka probabilitas dari kejadian tersebut sama dengan nilai limit dari *m/n* atau $\lim_{n\to\infty}\left(\frac{m}{n}\right)$. Probabiltas tersebut disebut probabilitas empiris. Meskipun mudah dan berguna dalam praktik, namun secara mamtematis perumusan konsep probabilitas dengan konsep frekuensi relative mempunyai kelemahan karena suatu nilai limit yang benar-benar mungkin sebenarnya tidak ada. Sedangkan pendekatan klasik memiliki kelemahan karena menuntut syarat hasil mempunyai kesempatan/kemungkinan yang sama untuk muncul. Oleh karena itu konsep probabilitas modern dikembangkan dengan pendekatan aksiomatik.

Pada pendekatan aksiomatik, aksioma-aksioma probabilitas jika diketahui bahwa suatu dan kejadian A dengan ruang sampel S dan peluang kejadian A pada S adalah P(A) = $\frac{n(A)}{n(S)} = \frac{m}{n}$, maka dapat diselidiki sifat dari P(A) adalah :

- Aksioma 1 ➔ $0 < P(A) < 1$, A merupakan himpunan bagian dari S yaitu A ∈ S, maka banyaknya A selalu lebih sedikit dari banyaknya anggota S, yaitu $n(A) \leq n(S)$, sehingga $0 < \frac{n(A)}{n(S)} < 1$ atau $0 < P(A) < 1$. Disebut probabilitas kemungkinan, artinya kejadian atau peristiwa tersebut dapat atau tidak dapat terjadi.
- Aksioma 2 ➔ P (A) = 0, Dalam hal A adalah himpunan kosong, A= Ø, artinya A tidak terjadi pada S maka n(A)=0 sehingga $P(A) = \frac{n(A)}{n(S)} = \frac{0}{n} = 0$. Disebut juga sebagai probabilitas kemustahilan, artinya kejadian atau peristiwa tersebut tidak akan terjadi.
- Aksioma 3 ➔ P(A) = 1. Dalaam hal A=S, maksimum banyaknya anggota A sama dengan banyaknya anggota S, maka n(A) = n(S) = n sehingga $P(A) = \frac{n(A)}{n(S)} = \frac{n}{n} = 1$. Disebut juga sebaga probabilitas kepastian, artinya kejadian atau peristiwa tersebut pasti terjadi.

Perumusan probabilitas pada kejadian majemuk sekurang-kurangnya terdapat dua kejadian atau lebih. Peluang gabungan tiga kejadian merupakan contoh probabilitas kejadian majemuk. Gambar diagram venn serta rumusan untuk peluang gabungan tiga kejadian adalah sebagai berikut.



Gambar 1.4 Diagram venn gabungan tiga kejadian

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$ …………(1.2)

Keterangan:

P(A) : Peluang kejadian A

P(B) : Peluang kejadian B

P(C) : Peluang kejadian C

P (A∩B) : Peluang kejadian A irisan B

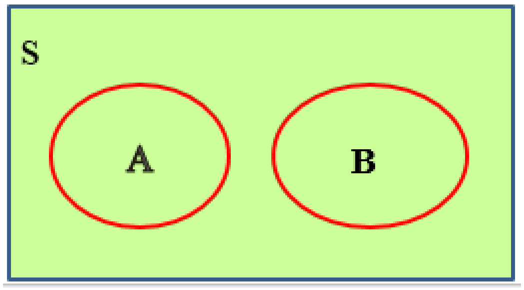
P (A∩C) : Peluang kejadian A irisan C

P (B∩C) : Peluang kejadian B irisan C

P (A∩B∩C) : Peluang kejadian A irisan B irisan C

Contoh soal peluang mahasiswa lulus kalkulus adalah 2/3 dan peluang mahasiswa lulus Bahasa inggris adalah 4/9. Bila peluang lulus sekurang-kurangnya satu matakuliah diatas adalah 4/5, berapa peluang ia lulus kedua mata kuliah.

Jenis probabilitas kejadian lainnya probabilitas kejadian saling lepas. Dua buah kejadian dinyatakan saling lepas bila A dan B kejadian sembarang pada ruang sampel S dan berlaku $A \cap B = \emptyset$. Kejadian saling lepas disebut juga kejadian saling bertentangan atau saling terpisah (*mutually exclusive*). Dua kejadian A dan B saling lepas artinya kejadian A dan B tidak mungkin terjadi secara bersamaan. Diagram venn serta rumus umum kejadia saling lepas adalah sebagai berikut.



Gambar 1.5 Diagram venn kejadian saling lepas

$P(A \cup B) = P(A) + P(B)$ ....(1.3)

karena $P(A \cap B) = 0 \rightarrow A \cap B = \emptyset$ maka rumusan tersebut dapat dinyatakan dalam bentuk lain

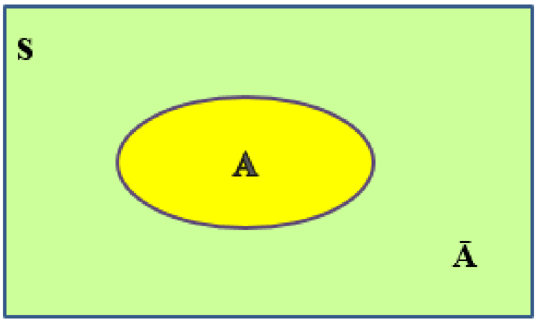
$P(A \cup B \cup) = P(A) + P(B) + P(C)$ ....(1.4)

Contoh soal

1. Bila A dan B dua kejadian saling lepas, dengan P(A) = 0,3 dan P(B)= 0,25 tentukanlah P(AUB) !
2. Pada pelemparan dua buah dadu tentukanlah probabilitas munculnya muka dadu dengan jumlah 7 atau 11. Hint : buat tabel kemungkinan munculnya dua muka dadu.

Probabilitas selanjutnya adalah probabilitas dua kejadian saling komplementer. Ingat kembali teori himpunan, jika $A \in S$, maka komplemen A bisa ditulis dengann symbol A’. Kejadian A’ adalah kumpulan titik sampel yang merupakan titik sampel S tetapi bukan merupakan titik sampel A. Jika $A' = \{ x : x \in S \text{ dan } x \notin S \}$ maka $A \cap A' = \emptyset$ dan $A \cup A' = S$.

Diagram venn dan rumus umum probabilitas kejadian saling komplementer adalah sebagai berikut.



Gambar 1.6 Diagram venn kejadian saling komplementer

P(A') = 1 - P(A) ……………………………………………..……………………….…(1.5)

Contoh soal pada pelemparan dua dadu jika A adalah kejadian munculnya muka dadu sama, hitunglah probabilitas munculnya muka dua dadu yang tidak sama!

Pembahasan probabilitas berikutnya yaitu probabilitas dua kejadian saling bebas (independen). Dalam kehidupan sehari-hari ada istilah "tidak saling mempengaruhi, independen atau tidak ada sangkut pautnya. Dalam hal tersebut maka terjadinya kejadian yang satu tidak dipengaruhi oleh terjadinya kejadian yang lain. Dua kejadian A dan B dalam ruang sampel S dikatakan independen jika kejadian A tidak mempengaruhi kejadian B dan demikian puula sebaliknya. Rumus umum probabilitas dua kejadian saling bebas atau independen adalah

P(A ∩ B) = P(A).P(B) ……………………………………………..……………………… (1.6)

Sedangkan jika 3 kejadian maka

P(A ∩ B ∩C) = P(A).P(B).P(C) ……………………………………………..…………… (1.7)

Contoh soal peluang bahwa sebuah robot berkaki akan "hidup" selama 25 tahun adalah $\frac{3}{5}$ dan peluang bahwa robot beroda akan "hidup" selama 25 tahun adalah $\frac{2}{3}$ Tentukan peluang bahwa keduanya akan hidup selama 25 tahun.

Selanjutnya sebuah probabiltas dikatan probabilitas bersyarat terjadi bila A dan B adalah dua kejadian sedemikian rupa sehingga P(A)> 0. Dimana kejadian B terjadi dengan syarat kejadian A terjadi terlebih dahulu atau dikatakan bahwa kejadian B bersyarat A yang ditulis B|A. Probabilitas terjadinya kejadian B dengan syarat A telah terjadi disebut sebagai kejadian bersyarat yang ditulis ddengan P(B|A).

$P(B|A) = \frac{P(B\cap A)}{P(A)}$ **;** P(A)>0 maka P(B∩A) = P(B|A) . P(A) ……..……………………… (1.8)

Catatan : probabilitas bersyarat untuk kejadian saling bebas adalah P(B|A) =P(B) dan P(A|B)=P(A) karena pada kejadian saling bebas P(A∩B) = P(A).P(B) maka

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{P(A).P(B)}{P(A)} = P(B) \quad \ldots\ldots (1.9)$$

Pada probabilitas bersyarat tiga kejadian, jika dua kejadian P(A∩B) = P(A). P(B|A) maka untuk tiga kejadian rumusannya adalah sebagai berikut

$$P(A\cap B\cap C) = P(A).\ P(B|A) . P(C \mid A\cap B) \quad \ldots\ldots (1.10)$$

A,B dan C semua nya terjadi sehingga probabilitas kejadian A dikalikan dengan probabilitas kejadian B dengan syarat A telah terjadi, dikalikan dengan probabilitas terjadinya C dengan syarat A dan B keduanya telah terjadi.

Contoh soal

1. Peluang suatu penerbangan regular berangkat tepat pada waktunya P(D) = 0.83. Peluang penerbangan itu mendarat tepat pada waktunya P(A) = 0,92 dan peluang penerbangan tsb berangkat dan mendarat tepat pada waktunya adalah P(A∩D) = 0,78. Hitunglah peluang pesawat pada penerbangan tsb yang memiliki kondisi mendarat tepat waktu bila diketahui bahwa pesawat tersebut berangkat tepat waktu serta kondisi berangkat tepat waktu bila diketahui bahwa pesawat tersebut mendarat tepat waktu.

Jawab

- Probabilitas berangkat tepat waktu P(D) = 0.83
- Probabilitas mendarat tepat waktu P(A) = 0,92
- Probabilitas berangkat dan mendarat tepat waktu P(A ∩ D) = 0,78
- Maka mendarat tepat waktu bila berangkat tepat waktu yaitu $P(A|D) = \frac{P(A\cap D)}{P(D)} = \frac{0,78}{0,83} = 0,94$
- Berangkat tepat waktu bila mendarat tepat waktu adalah $P(D|A) = \frac{P(D\cap A)}{P(A)} = \frac{0,78}{0,92} = 0,85$

2. Misalkan kita mempunyai sebuah kotak berisi 20 resistor, dan 5 diantaranya rusak. Bila 2 resistor diambil secara acak (satu-satu) tanpa pengembalian, berapa peluang resistror tersebut yang terambil keduanya rusak?

Jawab

- A = kejadian terambilnya resistor rusak pada pengambilan pertama dan B = kejadian terambilnya resistor rusak pada pengambilan kedua bila terpilih resistor rusak pada pengambilan pertama

- Peluang terambilnya setiap resistor sama yaitu masing-masing 1/20 maka P(A) = 5/20 =1/4 dan P(B|A)=4/19
- Peluang terambilnya resistror rusak pada pengambilan pertama dan kedua P(A∩B) yaitu $P(B|A)=\frac{P(A\cap B)}{P(A)}$ ➔ P (A∩B) = P(B|A). P(A) = 4/19 . 1/4 = 1/19

3. Sebuah dadu ideal dilemparkan sebanyak 2 kali. Tentukan probabilitas munculnya 4,5,6 pada pelemparan pertama dan 1,2,3 atau 4 pada pelemparan kedua.

Jawab

- Diketahui jika A adalah kejadian munculnya 4,5 atau 6 pada pelemparan pertama dan jika B adalah kejadian munculnya 1,2,3 atau 4 pada pelemparan kedua. Fakta yang ada bahwa pelemperan dadu yang kedua tidak dipengaruhi (independen) terhada pelemaran ertama sehingga kedua kejadian tersebut merupakan **bersyarat** (pelemparan kedua teradi dengan syarat pelemparan pertama sudah dilakukan) dan kejadian **saling bebas**
- Sehingga P(B|A)=P(B) maka P(A∩B) = P(A). P(B|A) = P(A).P(B) = 3/6 * 4/6 = 12/36 =1/3

F. Teorema Bayes

Bila B1, B2, B3,….., Bn adalah kejadian saling lepas (saling meniadakan) dalam ruang sampel S dan A kejadian lain sembarang dalam S, maka probabilitas kejadian bersyarat B**i** | A dapat dirumuskan

$$P(Bi \mid A) = \frac{P(Bi\cap A)}{P(A)} = \frac{P(A|Bi) \,.\, P(Bi)}{\sum_{i=1}^{n} P(A|Bi) \,.\, P(Bi)} \quad \ldots\ldots\ldots (1.11)$$

Bila A1, A2, A3,….., An adalah kejadian saling lepas (saling meniadakan) dalam ruang sampel S dan B kejadian lain sembarang dalam S, maka probabilitas kejadian bersyarat A**i** | B dapat dirumuskan

$$P(Ai \mid B) = \frac{P(Ai\cap B)}{P(B)} = \frac{P(B|Ai) \,.\, P(Ai)}{\sum_{i=1}^{n} P(B|Ai) \,.\, P(Ai)} \quad \ldots\ldots\ldots (1.12)$$

Contoh soal misalkan ada 3 kotak masing-masig berisi 2 bola. Kotak 1 berisi 2 bola merah kotak 2 berisi 1 bola merah dan 1 bola putih dan kotak 3 berisi 2 bola putih. Dengan mata tertutup anda diminta mengambil satu kotak secara acak dan kemudian mengambil 1 bola secara acak dari kotak yang terambil itu. Bila bola yang terambil ternyata berwarna merah. Berapakah peluangnya bola tersebut terambil dari masing-masing kotak (kotak 1, kotak 2 dan kotak 3) ?

BAB 2. VARIABEL ACAK dan DISTRIBUSI PROBABILITAS

A. Peubah Acak dan Ruang Sampel

Peubah acak atau biasa disebut sebagai variabel acak merupakan suatu fungsi yang mengaitkan suatu bilangan real pada setiap unsur dalam ruang sampel. Misalnya untuk setiap titik dari ruang sampel dipasang sebuah bilangan sehingga terdefinisikan sebuah fungsi pada ruang sampel tersebut. Fungsi ini disebut sebagai peubah/variabel acak (variabel stokastik) atau lebih tepatnya fungsi acak (fungsi stokastik). Peubah acak biasanya dinyatakan dalam huruf capital seperti X atau Y. Nilai setiap peubah acak dinyatakan dalam huruf kecil padanannya misal x atau y.

Contoh sederhana dari variabel/peubah acak yaitu misalkan terdapat dua bola diambil satu demi satu tanpa pengembalian dari suatu kantong berisi empat bola merah (M) dan empat bola hitam (H). Bila Y menyatakan jumlah bola merah yang diambil maka berapa nilai y yang mungkin dari peubah acak Y? Perhatikan tabel berikut.

Tabel 2.1 Ruang sampel beserta nilai peubah acak y

Ruang_sampel	**y**
MM	2
MH	1
HM	1
HH	0

Contoh lain dari penentuan peubah acak yaitu jika terdapat suatu kotak yang berisi 4 uang logam ratusan (R) dan 2 logam lima puluhan (L). Tiga uang diambil secara acak. Maka diperoleh ruang sampel yang mungkin adalah sebagai berikut S = {RRR, RRL, RLR, RLL, LRR, LRL, LLR}. Apabila dari percobaan pengambilan 3 uang logam tersebut, ditetapkan peubah acak X yang menyatakan jumlah uang logam ratusan yg muncul, maka nilai x yang diperoleh adalah sebagaimana tabel 2.2 di bawah ini.

Tabel 2.2 Ruang sampel dengan nilai peuah acak x

Ruang_sampel	x
RRR	3
RRL	2
RLR	2
LRR	2
RLL	1
LRL	1
LLR	1

Peubah acak terbagai menjadi beberapa jenis diantaranya adalah sebagaimana berikut. Peubah acak diskrit merupakan jenis peubah acak yang fungsi distribusinya diskrit atau peubah acak yang memiliki himpunan kemungkinan hasilnya terhitung. Sehingga ruang sampel diskrit diartikan sebagai ruang sampel yang mengandung titik yang berhingga banyaknya atau sederetan anggota yang jumlahnya sebanyak bilangan bulat. Sementara itu peubah acak kontinyu didefinisikan sebagai peubah acak yang fungsi distribusinya kontinyu atau peubah acak yg dapat memperoleh semua nilai pada skala kontinyu. Dengan demikian maka ruang sampel kontinyu adalah ruang sampel mengandung titik sampel yang tak berhingga banyaknya dan jumlahnya sebanyak titik pada sepotong garis. Sehingga 2 contoh sebelumnya yaitu pengambilan bola merah dan pengambilan tiga uang logam merupakan peubah diskrit.

B. Distribusi Probabilitas

Pada contoh sebelumnya dimana terdapat ruang sampel S={RRR, RRL, RLR, RLL, LRR, LRL, LLR} yang merupakan kumpulan semua hasil yang mungkin terjadi pada pengambilan 3 uang logam, dapat ditentukan probabilitas dari nilai-nilai variabel acak X, sebab titik sampel – titik sampel S mempunyai nilai probabilitas. Pada ruang sampel tersebut jika variabel acak X yang menyatakan jumlah uang logam ratusan (R) yg muncul, maka nilai dari X adalah X=1, X=2, X=3. Nilai X=1 berkaitan dengan titik sampel (RLL) atau (LRL) atau (LLR) dengan probabilitas P(X=1) = { P(RLL) + P(LRL) + P(LLR) } = 1/7+1/7+1/7 =3/7. Sementara itu untuk nilai X=2 berkaitan dengan titik sampel (RRL) atau (RLR) atau (LRR) dengn probabilitas P(X=2) = { P(RRL) + P(RLR) + P(LRR) } = 1/7+1/7+1/7 = 3/7. Selanjutnya nilai X=3 berkaitan dengan titik sampel (RRR) dengan probabilitas P(X=3)=P (RRR) = 1/7. Distribusi probabilitas X atau Pasangan nilai-nilai variabel acak X dengan probabilitas dari nilai-nilai X, yaitu P(X=x) dapat dinyatakan dalam tabel 2.3 berikut.

Tabel 2.3 Tabel distribusi probabiltas P(X=x)

X=x	1	2	3
P(X=x)	3/7	3/7	1/7

Pada distribusi probabilitas diskrit misalkan X adalah variabel acak diskrit, dan misalkan kemungkinan nilai-nilainya adalah x1,x2,x3,... disusun dalam suatu urutan tertentu. Misalkan juga nilai-nilai tsb memiliki probabilitas yang dinyatakan oleh P(X=xk) = *f(xk)* dimana *k*=1,2,3, dan seterusnya. Bentuk persamaan **P(X=x) = *f(x)*** disebut sebagai **fungsi probabilitas** atau dikenal juga sebagai **distribusi probabilitas.** Secara umum bahwa rumus ***f(x)*** adalah **fungsi probabilitas jika *f (x)* ≥ *0*** dan $\sum_x \boldsymbol{f(x) = 1}$.

Pada fungsi distribusi untuk variabel acak jika X adalah variabel acak dan P(X=x) adalah distribusi probabilitas dari X, maka fungsi **f(x) = P(X=x)** disebut sebagai **fungsi probabiltas X** atau fungsi frekuensi X atau fungsi padat peluang X. Sifat-sifat fungsi f(x) untuk variabel acak X diskrit adalah P(X=x) = f(x), f(x) ≥ 0 serta $\sum_x f(x) = 1$. Sedangkan sifat-sifat fungsi f(x) untuk variabel acak X kontinyu adalah P(a < X < b) = $\int_a^b f(x)dx$, f(x) ≥0 serta $\int_{-\infty}^{\infty} f(x)dx = 1$. Jika variabel acak X mempunyai distribusi probabilitas f(x) maka fungsi distribusi kumulatif atau disingkat fungsi distribusi dari X yaitu F(x) dirumuskan dengan F(x) = P(X ≤ x) dimana $\sum_{X \le x} f(x)$ bila X diskrit dan $\int_{-\infty}^{\infty} f(x)dx$ bila X kontinyu. Sifat-sifat fungsi distribusi kumulatif F(x) adalah 0 ≤ F(x) ≤ 1, jika x1 < x2 maka F(x1)<F(x2) dikatakan fungsi F(x) monoton tidak turun (*nondecreasing*), F(x) diskontinyu dari kiri tetapi kontinyu dari kanan [artinya $\lim_{h\to 0+} F(x+h) = F(x)$ untuk semua x] serta sifat keempat $\lim_{x\to -\infty} F(x) = F(-\infty) = 0$ dan $\lim_{x\to +\infty} F(x) = f(+\infty) = 1$.

Contoh pada pelemparan tiga uang logam, bila X menyatakan benyaknya muncul gambar (g) tentukan

a. Nilai-nilai variabel acak X
b. Distribusi probabilitas X ➔ P(X=x)
c. Gambarlah distribusi probabilitas X
d. Nilai-nilai dari fungsi distribusi probabilitas / fungsi probabilitas ➔ f(x)
e. Nilai-nilai dari fungsi distribusi kumulatif / fungsi distribusi ➔ F(x)

Jawaban atas contoh soal tersebut di atas yaitu Ruang sampel S = {(ggg), (gga), (gag), (agg), (aag), (aga), (gaa), (aaa)}

a. Karena X menyatakan banyaknya muncul gambar pada S maka nilai-nilai dari X adalah X=0, X=1, X=2, X=3

b. Probabilitas

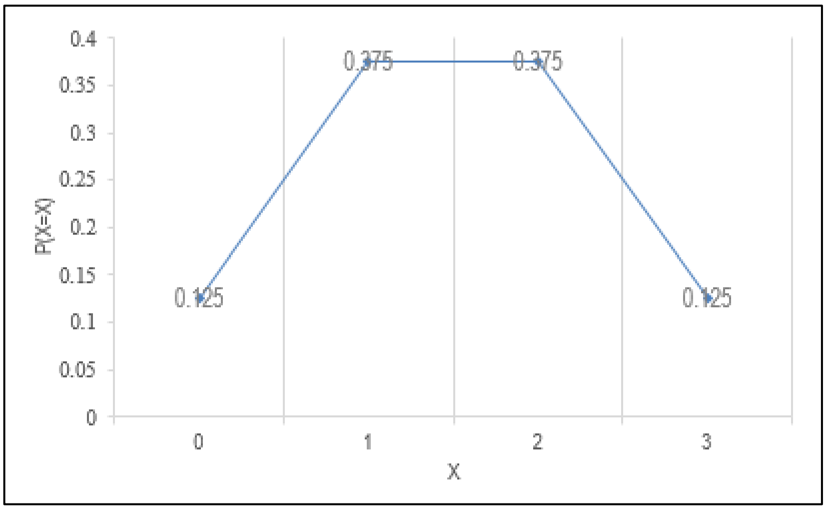
P(X=0) = P(aaa) = 1/8

P(X=1) = P(aag) + P(aga) + P(gaa) = 1/8+1/8+1/8 =3/8

P(X=2) = P(gga) + P(gag) + P(agg) = 1/8+1/8+1/8 =3/8

P(X=3) = P(ggg) = 1/8

c. Gambar distribusi probabilitas X



Gambar 2.1 Gambar distribusi probabilitas X

d. f(x) = P(X=x), maka nilai-nilai f(x) adalah

f(0) = P(X=0) = 1/8

f(1) = P(X=1) = 3/8

f(2) = P(X=2) = 3/8

f(3) = P(X=3) = 1/8

e. Nilai-nilai dari distribusi kumulatif F(x) adalah

F(0) = P(X ≤ 0) = P(X=0) = f(0) =1/8

F(1) = P(X ≤ 1) = P(X=0) + P(X=1) = 1/8 + 3/8 = 4/8

F(2) = P(X ≤ 2) = P(X=0) + P(X=1) + P(X=2) = 1/8 + 3/8 + 3/8 = 7/8

F(3) = P(X ≤ 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 1/8 + 3/8 + 3/8 + 1/8 = 8/8

BAB 3. HARAPAN MATEMATIS

A. Nilai Harapan Matematis

Bila variabel acak X yang mempunyai nilai x1, x2,x3,……, xn serta memiliki fungsi probabilitas f(x) = P(X=x), maka harapan atau ekspektasi matematis dari X yang ditulis E(X) didefinikan sebagai berikut

E(X) = x**1** P(X = x**1**) + x**2** P(X = x**2**) + x**3** P(X = x**3**) + ..…. + x**n** P(X = x**n**) ……………..(3.1)

Untuk nilai X diskrit ataupun kontinyu maka rumusan untuk nilai ekspektasi matematis adalah

$$\mu = E(X) = \begin{cases} \sum_{x} x\, f(x) & ;\ \text{jika X diskret} \\ \int_{-\infty}^{\infty} x\, f(x)\, dx & ;\ \text{jika X kontinu} \end{cases} \qquad \ldots\ldots (3.2)$$

E(X) adalah Ekspektasi atau harapan matematik atau nilai harapan dari peubah acak X dan juga banyak yang menyebutnya rata-rata peubah acak X atau rata-rata distribusi probabilitas X atau sering juga disebut sebagai mean dari X dilambangkan dengan **µx.** Mean atau Ekspektasi dari X memberikan satu nilai tunggal yang bertindak sebagai wakil atau rata-rata dari nilai-nilai X, sehingga sering diebut sebagai "ukuran tendensi sentral". Ukuran-ukuran yang lain adalah Median dan Modus. Modus dari suatu variabel acak diskrit merupaka nilai yang paling sering muncul atau dengan kata lain memiliki probabilitas terbesar untuk terjadi. Median merupakan suatu nilai x dimana $P(X < x) \leq \frac{1}{2}$ dan $P(X>x) \leq \frac{1}{2}$ atau biasa disebut titik tengah dari interval. Selanjutnya harapan matematis memiliki sifat-sifat yaitu :

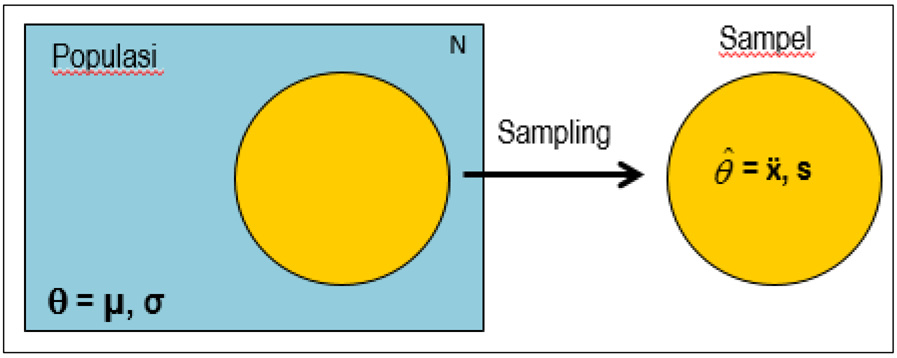
1. *E(c)* = c
2. *E(bX) = bE(X)*
3. *E(a + bX) = a + bE(X)* dimana *a,b,c* adalah suatu konstanta

Sedangkan teorema ekspektasi matematis diantaranya adalah

- Teorema 1, Jika c adalah suatu konstanta sembarang maka E(cX) = cE(X)
- Teorema 2, Jika X dan Y adalah variabel acak sembarang maka E(X + Y) = E(X) + E(Y)
- Teorema 3, Jika X dan Y adalah variabel acak yg independen maka E(XY) = E(X) E(Y)

B. Kegunaan Nilai Harapan Matematis

Salah satu cabang ilmu statistika yaitu statistika inferensia, merupakan Statistika yang berkenaan dengan cara penarikan kesimpulan berdasar- kan data yang diperoleh dari sampel untuk menggambarkan karakteristik atau ciri suatu populasi. Penarikan kesimpulan biasanya dilakukan pada penelitian atau studi dengan memakai metode survei yang memakai data dari sampel. Namun demikian, hasil perhitungan yang diperoleh diperluas untuk menggambarkan atau menyimpukan karakteristik populasi. Sehingga antara sampel dan populasi ada keterkaitan yang sangat erat, sebagaimana ilustrasi berikut.



Gambar 3.1 Hubungan antara Populasi dan Sampel

Populasi menggambarkan sesuatu yang sifatnya ideal atau teoritis sedangkan sampel menggambarkan sesuatu yang sifatnya nyata atau empiris. Populasi dan sampel masing-masing mempunyai karakteristik atau ciri yang dapat diukur. Karakteristik yang diukur dari populasi disebut parameter misalnya mean, standar deviasi, proporsi dan lain sebagainya.

Kegunaan utama harapan matematis adalah untuk menentukan Mean (μ) dan variansi (σ^2) serta standar deviasi (σ) dari parameter populasi yang dirumuskan sebagai berikut.

1. Mean Populasi μ = E(X) .... (3.3)
2. Variansi Populasi

$$\sigma^2 = \mathrm{E}\,\{(\mathrm{X}-\mu)^2\} = \begin{cases} \sum(x-\mu^2)\, f(x); \text{jika X diskrit} \\ \int_{-\infty}^{\infty}(x-\mu)^2\, f(x); \text{jika X Kontinyu} \end{cases} \quad \text{.............................. (3.4)}$$

Jika μ adalah konstanta maka variansi disederhakana menjadi $\sigma^2 = \mathrm{E}(X^2) - \mu^2$

3. Standar Deviasi $\sigma = \sqrt{E\{(X-\mu)^2\}}$ .... (3.5)

C. Rumusan Parameter Sampel : Nilai Harapan Matematis

Rumusan parameter sampel yang akan dibahas meliputi nilai harapan matematis, variansi dan standar deviasi. Rumusan variansi perubah acak X diskret atau dilambangkan dengan sigma kuadrat yaitu

$$\sigma^2 = E(X^2) - \mu^2 \quad \text{(3.6)}$$

Bukti

$$\sigma^2 = \sum_x (x-\mu)^2 f(x) = \sum_x (x^2 - 2x\mu + \mu^2) f(x)$$

$$= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x) \quad \text{(3.7)}$$

Karena

$$\mu = \sum_x x f(x) \quad \text{dan} \quad \sum_x f(x) = 1 \quad \text{(3.8)}$$

Maka diperoleh

$$\sigma^2 = \sum_x x^2 f(x) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \quad \text{(3.9)}$$

Contoh soal 1

Pada pelemparan tiga uang logam. Tentukan harapan matematis munculnya “muka” pada tiap pelemparan, jika X menyatakan banyaknya muncul muka.

Jawab

Diketahui :

S = {(m,m,m), (m,m,b), (m,b,m), (b,m,m), (b,b,m), (b,m,b), (m,b,b), (b,b,b) }

X = {0,1,2,3} ;

Maka P(X=x)

P(x = 0) = 1/8 ;

P(x = 1) = 3/8 ;

P(x = 2) = 3/8 ;

P(x = 3) = 1/8 ;

Sehingga :

$E(X) = \sum_{x=0}^{3} x\, f(x) \quad = \sum_{x=0}^{3} x\, P(X=x)$

E(X) = (0) P(x=0) + (1) P(x=1) + (2) P(x=2) + (3) P(x=3)

E(X) = (0)(1/8) + (1)(3/8) + (2)(3/8) + (3)(1/8)

E(X) = (0+3+6+3)/8 = 12/8 = 1,5

Contoh soal 2

Pada pelemparan dua dadu. Tentukan harapan matematis munculnya jumlah muka dua dadu, jika X menyatakan jumlah muka dua dadu.

Jawab

X=x	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$E(X) = \sum_{x=2}^{12} x\, f(x) \quad = \sum_{x=2}^{12} x\, P(X=x)$

E(X) = (2) P(x=2) + (3) P(x=3) + (4) P(x=4) +…..+ (12) P(x=12)

E(X) = 252/36 = 7

Contoh soal 3

Misalkan suatu permainan dimainkan dengan menggunakan sebuah dadu yang diasumsikan ideal. Dalam permainan ini seorang pemain akan menang $20 jika muncul angka 2, $40 jika muncul angka 4, kalah $30 jika muncul angka 6 ; sementara pemain tersebut tidak akan menang atau kalah jika angka lain muncul. Tentukan ekspektasi jumlah uang yang akan dimenangkan!

Jawab

- Misalkan jika X adalah variabel acak yang menyatakan jumlah uang yang dimenangkan pada suatu lemparan
- Jumlah yang mungkin dimenangkan jika angka yang muncul berturut-turut 1,2,3,4,5,6 adalah x1,x2,x3,x4,x5 dan x6 ➔ munculnya muka dadu 1 adalah x1, muka dadu 2 adalah x2 dan seterusnya. Sementara probabilitasnya adalah f(x1) , f(x2), …… f(x6). Maka fungsi probabilitas untuk X adalah

Xi	**0**	**+20**	**0**	**+40**	**0**	**-30**
f(xi)	1/6	1/6	1/6	1/6	1/6	1/6

- $E(X) = x\mathbf{1}.\ f(x\mathbf{1}) + x\mathbf{2}.\ f(x\mathbf{2}) + x\mathbf{3}.\ f(x\mathbf{3}) + x\mathbf{4}.\ f(x\mathbf{4}) + x\mathbf{5}.\ f(x\mathbf{5}) + x\mathbf{6}.\ f(x\mathbf{6})$
- $E(X) = 0.\ (1/6) + 20.\ (1/6) + 0.\ (1/6) + 40.(1/6) + 0.\ (1/6) + (-30)\ (1/6) = 5$
- Maka pemain tersebut dapat diekspektasikan akan memenangkan $5 Oleh karena itu, dalam suatu permainan yang adil, pemain ini diekpektasikan harus membayar $5 agar dapat bermain dalam permainan tersebut.

Contoh Soal 4

Jika seseorang membeli sebuah lotere, maka ia dapat memenangkan hadiah pertama sebesar Rp. 50.000.000,- atau hadiah kedua Rp. 20.000.000,- masing-masing dengan probabilitas 0,001 dan 0,003. Berapa seharusnya harga yang fair untuk lotere tersebut?

Jawab

Misalkan Variabel X menyatakan nilai kemenangan orang tersebut, maka nilai X1= 50.000.000 dengan Probabilitas P(X=x1) =0,001 dan X2=20.000.000 dengan Probabilitas P(X=x2) = 0,003

Sehingga nilai Harapan Matematis X adalah

$$\begin{aligned} E(X) &= X\mathbf{1}\ P(X=x\mathbf{1}\ ) + X\mathbf{2}\ P(X=x\mathbf{2}\ ) \\ &= (50.000.000)\ (0{,}001) + (20.000.000)\ (0{,}003) \\ &= 110.000 \end{aligned}$$

Jadi harga yang fair dari lotere tersebut adalah Rp. 110.000,-

Contoh Soal 5

Pada pengiriman 6 pesawat TV berisi 2 rusak. Sebuah hotel membeli 3 pesawat TV secara acak dari kiriman tersebut. Bila X menyatakan banyaknya TV yang rusak yang dibeli hotel. Tentukan Nilai Harapan X dan Simpangan baku X! (Catatan : Sebelum mulai menghitung, cari dulu distribusi probabilitas X)

Jawab

Misalkan

TV baik = 4 ➔ { b1, b2,b3,b4} sedangkan TV rusak ; = 2 ➔ {r1, r2}

X menyatakan jumlah TV rusak yang dibeli hotel

Dengan memanfaatkan rumus kombinasi maka diperoleh data TV yang dibeli dengan kriteriia

- 3 baik dan 0 rusak maka 4C3 . 2C0 = $\binom{4}{3}.\binom{2}{0}$ = 4.1
- 2 baik dan 1 rusak maka 4C2 . 2C1 = $\binom{4}{2}.\binom{2}{1}$ = 6.2
- 1 baik dan 2 rusak maka 4C1 . 2C2 = $\binom{4}{1}.\binom{2}{2}$ = 4.1
- Seluruh Kombinasi dari opsi yang ada adalah 6C3 = $\binom{6}{3}$ = 20

Bila X menyatakan jumlah TV **rusak** yang dibeli hotel, maka nilai-nilai X adalah = x1, x2, x3 bernilai 0,1,dan 2 masing-masing dengan probabilitas

P(X = 0) = $\frac{4.1}{20}$ = 1/5

P(X = 1) = $\frac{6.2}{20}$ = 3/5

P(X = 2) = $\frac{4.1}{20}$ =1/5

Sehingga Nilai harapan $E(X)$ $= 0(\frac{1}{5}) + 1(\frac{3}{5}) + 2\left(\frac{1}{5}\right) = 1$

Maka $E(X^2) = 0^2(\frac{1}{5}) + 1^2(\frac{3}{5}) + 2^2\left(\frac{1}{5}\right) = \frac{7}{5}$

Variasi $\sigma^2 = E(X^2) - E(X)$ $= \frac{7}{5} - 1 = \frac{2}{5} = 0{,}4$

Simpangan baku σ $= \sqrt{\sigma^2} = \sqrt{0{,}4} = 0{,}63$

Contoh soal 6

Fungsi kepadatan dari suatu variabel acak X ditentukan oleh

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & selain\ di\ atas \end{cases}$$

Maka nilai Ekspektasi dari X adalah

Jawab

Ingat kembali rumus integral $\int a\, x^2\, \text{dx} = \frac{a}{n+1} x^{n+1} + C$

$$\int (ax+b)^n \equiv \frac{1}{a(n+1)}(ax+b)^{n+1} + C$$

Sehingga diperoleh :

$$E(X) \int_{-\infty}^{\infty} x.f(x)dx = \int_0^2 x\left(\frac{1}{2}x\right)dx = \int_0^2 \frac{x^2}{2}dx = \frac{x^3}{6}\Big|_0^2 = \frac{4}{3}$$

Soal tambahan

Dalam suatu bisnis tertentu, seorang dapat memperoleh keuntungan sebesar Rp. 3.000.000,- dengan probabilitas 0,6 atau menderita kerugian sebesar Rp. 1.000.000,- dengan probabilitas 0,4. Tentukan nilai harapannya

BAB 4. DISTRIBUSI PROBABILITAS DISKRIT

A. Distribusi Binomial

Perhatikan kembali setiap hasil percobaan statistik pada pembahasan sebelumnya, dari semua percobaan hasil-hasil yang ada dapat dibedakan menjadi 2 jenis, sebagaimana contoh berikut ini. Pada pelemparan sebuah uang logam kita dapat melemparakan sebanyak 10 kali atau 100 kali dan seterusnya sesuai kepentingan. Kemudian hasil –hasil yang muncul dibedakan menjadi 2 yaitu kejadian munculnya muka dan bukan muka. Pada Pelemparan sebuah dadu sebanyak 100 kali akan diperoleh hasil kemungkinan munculnya 6 dan bukan enam. Secara singkat hasil-hasil yang muncul pada percobaan statistic dapat dibedakan menjadi dua yaitu kejadian sukses dan kejadian gagal. Selain itu kejadian sukses dan gagal dari satu percobaan ke percobaan lainnya tersebut merupakan saling bebas. Suatu percobaan statistik disebut percobaan Binomial jika mempunyai ciri-ciri :

1. Percobaan diulang sebanyak n kali
2. Hasil percobaan dibedakan menjadi 2 yaitu kejadian sukses (S) dan kejadian gagal (G)
3. Probabilitas terjadinya kejadian sukses (S) dan gagal (G) yaitu : P(S)= p dan P(G)=1-p = q adalah tetap pada tiap kali percobaan diulang
4. Semua hasil yang muncul bebas satu sama lain

B. Perumusan Distribusi Binomial

Jika Peluang sukses (S) dalam suatu experimen adalah p→ prob(S)= p dan Peluang gagal (G) adalah q = 1 –p → prob(G) = q. Maka pada 1 kali eksperimen peluang sukses p dan peluang gagal q. Untuk 2 kali ekserimen peluang sukses kemudian sukses (S,S) : pp, peluang sukses kemudian gagal (S,G) : pq, peluang gagal kemudian sukses (G,S) : qp, peluang gagal kemudian gagal (G,G): qq. Sedangkan untuk 3 kali eksperimen peluang sukses dan gagal dapat dilihat pada tabel berikut.

Tabel 3.1 Sukses-Gagal dalam 3 kali eksperimen

jumlah sukses	cara sukses	jumlah cara sukses	probabilitas	
3	SSS	1	1 *ppp*	1 p^3q^0
2	SSG, SGS, GSS	3	3 *ppq*	3 p^2q^1
1	SGG, GSG, GGS	3	3 *pqq*	3 p^1q^2
0	GGG	1	1 *qqq*	1 p^0q^3

Pada Sukses-Gagal dalam 3 kali atau 5 kali eksperimen dimana untuk 3 kali eksperimen diperoleh peluang sukses pada experimen ke-3 qqp dan peluang sukses di salah satu experiment pqq+ qpq+ qqp. Untuk 5 kali eksperimen diperoleh peluang sukses 2 kali *ppqqq*+ *pqpqq*+ ... + *qqqpp*, atau disederhanakan dalam bentuk rumus $\binom{5}{2}p^2q^3 = 10p^2q^3$. Maka peluang mendapatkan *x* kali suksesdari *n* kali experimen adalah

$\mathrm{f(x)} = \mathrm{P(X = x)} = f_x\ \mathrm{(x, n, p)} = \binom{n}{x}p^x(1-p)^{n-x}$ ……………………………..…… (4.1)

x= 0,1,2,….,n ; x adalah koefsien binomial.

Distribusi binomial mempunyai nilai rata-rata, variansi, simpangan baku sebagai berikut :

Rata-rata (Ekspektasi Matematik) $\boldsymbol{\mu = n.p}$ …………………………………………….. (4.2)

Variansi $\boldsymbol{\sigma^2 = n.p.q}$ ……………….…………………………………………….…. (4.3)

Simpangan baku $\boldsymbol{\sigma = \sqrt{npq}}$ ……………….………………………………….…. (4.4)

Ada kalanya perhitungan probabilitas ditribusi binomial lebih mudah dilakukan dengan memakai distribusi kumulatif. Bila ada n percobaan terdapat paling tidak sebanyak r sukses, maka distribusi binomial kumulatif P(X ≥ r) dapat dirumuskan seagai berikut :

P(X ≥ r) = fx (r,n,p) + fx (r+1,n,p) + ….. + fx (n,n,p) ……………….…………….. (4.5)

P(X ≥ r) = $\sum_{x=r} fx\ (r, n, p)$ ……………….…………………………………….……. (4.6)

Contoh soal 1. Diketahui suatu percobaan statistic yang diulang sebanyak n=4 dengan P(S) =2/3 dan P(G) =1/3 tetap pada setiap percobaan. Misalkan X banyaknya sukses. Tentukan P(X=0), P(X=1), P(X=2), P(X=3) dan P(X=4).

Jawab 1

$P(X=x) = \binom{n}{x}p^x q^{n-x} = \binom{4}{x}(2/3)^x(1/3)^{4-x}$; x= 0,1,2,3,4

Maka diperoleh

- $P(X=0) = \binom{4}{0}(2/3)^0(1/3)^4 = 1/81$
- $P(X=1) = \binom{4}{1}(2/3)^1(1/3)^3 = 8/81$
- $P(X=2) = \binom{4}{2}(2/3)^2(1/3)^2 = 24/81$

- $P(X=3) = \binom{4}{3}(2/3)^3(1/3)^1 = 32/81$
- $P(X=4) = \binom{4}{4}(2/3)^4(1/3)^0 = 16/81$
- Perhatikan bahwa $\sum_x P(x) = 1$

Contoh Soal 2. Setiap tahun dalam 5 tahun dilakukan pemilihan acak untuk menetapkan alokasi dana kepada 1 dari 4 kegiatan (A,B,C,D). Setiap kali dilakukan pemilihan, masing-masing kegiatan memiliki peluang yang sama untuk terpilih (mendapatkan dana). Berapa persen peluang kegiatan A mendapatkan dana 3x? Berapa persen peluang kegiatan A mendapatkan dana 5x, 4x, 3x, 2x, 1x, 0x?

Jawab 2

Setiap kali pemilihan diperoleh probabilitas (As) = probabilitas kegiatan A terpilih P(As)= ¼= 0.25 = p dan Probabilitas (Ag) = probabilitas kegiatan A tak terpilih P(Ag) = 1 –p= 0.75 = q

Dalam 5 kali pemilihan, maka peluang terpilih (sukses) 3 kali adalah

$$f_X(x;n,p) = f_X(3;5,0.25) = \binom{5}{3} 0.25^3\ 0.75^2 = 0.088$$

- Dalam 5 kali pemilihan (n = 5)

koefisien binomial

jumlah sukses	jumlah cara	peluang terjadi
0	1	0.237
1	5	0.396
2	10	0.264
3	10	0.088
4	5	0.015
5	1	0.001
	∑ =	**1.000**

Contoh Soal 3. Pengalaman menunjukkan bahwa pada setiap pencetakan kertas koran, dari 1500 lembar yang dicetak telah terjadi kerusakan sebanyak 150 lembar. Bila dicetak sebanyak 10 lembar, tentukanlah probabilitas dari variabel acak X, bila X menyatakan banyaknya kertas yang rusak pada pencetakan.

Jawab 3.

Diketahui n = 10 ; X = 0,1,2,3,……,10 (banyaknya kertas yang rusak pada penstensilan)

P(S) = P (Kertas rusak) = 150/1500 = 1/10 = p

P(G) = P (Kertas tidak rusak) = 1-1/10 = 9/10 =q

Sehingga untuk n=10 ; p = 0,1 dan q = 0,9

P (X=x) = $\binom{10}{x}(0,1)^{x}(0,9)^{10-x}$; x= 0,1,2,3,…., 10

Maka untuk perhitungan tersebut gunakan tabel distribusi binomial

Contoh soal 4. Bila sekeping uang logam dilemparkan sebanyak 6 kali, Hitunglah probabilitas memperoleh 5 muka dan probabilitas paling sedikit 5 muka.

Jawab 4

Diketahui n= 6 ; p = 0,5 dan q = 0,5

P (X=x) = $\binom{6}{x}(0,5)^{x}(0,5)^{6-x}$; x= 0,1,2,3,3,4,5,6

P (5muka) = $\binom{6}{5}(0,5)^{5}(0,5)^{1}$; x=0,09375

P(X ≥ 5) = 0,109 (lihat tabel distribusi binomial kumularif untuk n = 6, p =0,5 dan r =5)

Contoh soal 5. Bila variabel acak X mempunyai distribusi binomial dengan n = 15 ; p= 0,3 Tentukan :

a. P(X = 10)

b. P(X ≤ 7)

c. P(X ≥ 7)

d. Nilai rata-rata

e. Variansi

f. Simpangan baku

C. Distribusi Poisson

Pada perumusan tentang distribusi binomial yaitu $\mathrm{P(X = x)} = \binom{n}{x}p^x(1-p)^{n-x}$ dengan x=0,1,2,….n. Bila *n kecil* dan *p besar,* maka perhitungan probabilitas nilai variabel acak X tidak mengalami masalah, karena nilai P(X=x) dapat dihitung secara langsung atau menggunakan table untuk bilangan n, nilai p dan nilai x tertentu. Namun, Jika *n besar* dan *p kecil sekali* maka probabilitas nilai X tidak bisa atau sulit dihitung baik scr langsung ataupun dgn menggunakan table, sebab table hanya ada nilai maks n=25 dan minimum p = 0,01. Sehingga untuk melakukan perhitungan probabilitas distribusi binomial dengan kasus n besar dan p sangat kecil dilakukan dengan memakai pendekatan distribusi Poisson. Jika n besar dan p kecil sedemikian rupa sehingga np → μ atau μ = np tetap, maka distribusi binomial dapat didekati dengan memakai distrribusi poisson yg dirumuskan sebagai berikut.

$$f(x) = P(X = x) = \frac{e^{-\mu}\mu^x}{X!} \quad \text{………………………………………………………...} \quad (4.7)$$

Dimana X = 0,1,2,….n ; μ disebut parameter ; e = 2,71828.

Rata-rata $\mu = np$ …………………..………………………………….……………... (4.8)

Variansi $\sigma^2 = np$ ……………………………………………….……………... (4.9)

Simpangan baku $\sigma = \sqrt{np}$ ………………………………………………………. (4.10)

Contoh soal 6. Bila variabel acak X mempunyai distribusi binomial dengan n = 100; p = 0,005 Tentukan P(X = 15)

Jawab 6

$f(x) = P(X = x) = \binom{100}{x}(0{,}005)^x(0{,}995)^{100-x}$ dimana x = 0,1,2,3,……………,100

Maka $f(15) = P(X = 15) = \binom{100}{15}(0{,}005)^{15}(0{,}995)^{85}$

Jika menggunakan rumusan binomial sebagaimana diatas maka akan sulit dihitung sehingga perlu pendekatan poisson sebagai berikut.

μ = np

= 100. 0,005

μ = 0,5

$P(X=x) = \frac{e^{-\mu}\mu^{x}}{X!} = \frac{e^{-0,5}(0,5)^{15}}{X!}$; x = 0,1,2,3,…..,100

$P(X=15) = \frac{e^{-0,5}(0,5)^{15}}{15!} = 0,0000$

Hasilnya lihat dengan menggunakan table sehingga diperoleh nilai yang paling sesuai

Contoh Soal 7. Bila 5 uang logam dilempar sebanyak 128 kali, hitunglah probabilitas munculnya 5 angka sebanyak 0,1,2,3,4, dan 5 dari seluruh pelemparan.

Jawab 7.

n = 128 ;

Probabiltas munculnya satu angka adalah ½ dan probabiltas munculnya lima angka adalah ½. ½. ½. ½ .½ = 1/32 =p

Karena munculnya angka saling bebas (independen) atau tidak saling mempengaruhi.

Maka q = 1 – p = 1 – 1/32 = 31/32 dan dengan menggunakan distribusi binomial diperoleh

$P(X = x) = \binom{128}{x}\left(\frac{1}{32}\right)^{x}\left(\frac{31}{32}\right)^{128-x}$; x = 0,1,2,….128

Dengan distribusi Poisson μ = np = 128 . $\left(\frac{1}{32}\right)$ = 4

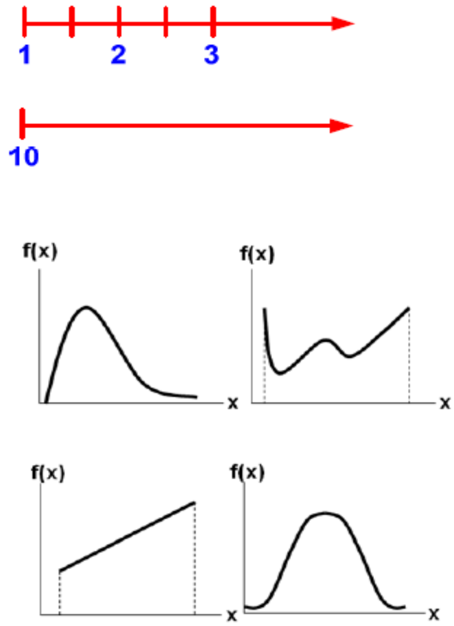
$P(X=x) = \frac{e^{-4}(4)^{x}}{X!}$

X	$\boldsymbol{P(X = x) = \binom{128}{x}\left(\frac{1}{32}\right)^{x}\left(\frac{31}{32}\right)^{128-x}}$	$P(X=x) = \frac{\boldsymbol{e^{-4}(4)^{x}}}{\boldsymbol{X!}}$
0	0,0172	0,0183
1	0,0711	0,0732
2	0,1457	0,1464
3	0,1974	0,1952
4	0,1990	0,1952
5	0,1592	0,1562

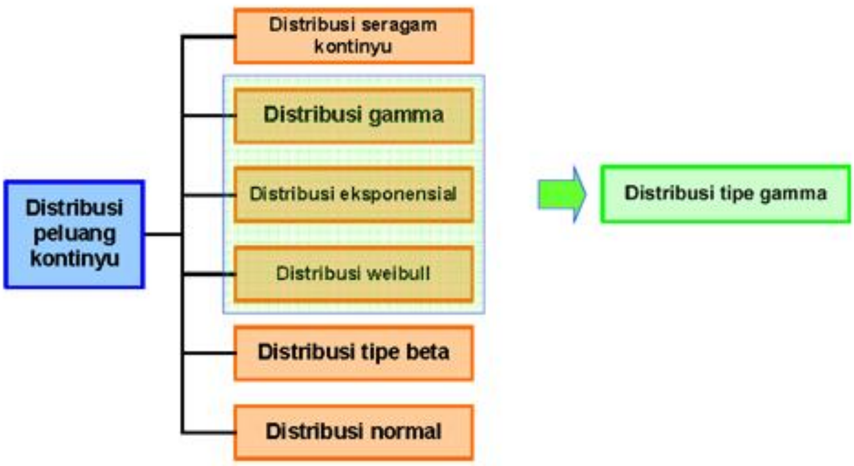
BAB 5. DISTRIBUSI PROBABILITAS KONTINYU

A. Distribusi probabilitas kontinyu

Berbeda dengan variabel random diskrit, sebuah variabel random kontinyu adalah variabel yg dapat mencakup nilai pecahan maupun mencakup range/rentang nilai tertentu. Karena terdapat bilangan pecahan yang jumlahnya tidak terbatas, kita tidak dapat menuliskan semua nilai yg mungkin bersama dengan probabilitasnya masing – masing dalam bentuk tabel. Namun dapat dituliskan dalam bentuk fungsi kepadatan probabilitas (*Probability Density Function : pdf*). Plot untuk fungsi seperti ini disebut kurva probabilitas dan nilai probabilitasnya dinyatakan sebagai luas suatu kurva yang bernilai positif.



Gambar 5.1 Contoh kurva probabilitas diskrit dan kontinyu



Gambar 5.2 Jenis distribusi peluang kontinyu

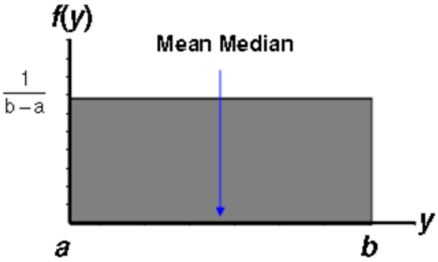
B. Distribusi Seragam Kontinyu

Distribusi seragam kontinyu memiliki ciri yaitu variabel random seragam Y = salah satu nilai dalam interval $a \leq y \leq b$ serta setiap Y memiliki nilai peluang seragam dalam selang $a \leq y \leq b$.

Diberikan oleh :

$$f(y) = \begin{cases} \frac{1}{b-a} & \text{, jika y bernilai } a \leq y \leq b \\ 0 & \text{, jika y bernilai lainnya} \end{cases}$$

$$\mu = \frac{a+b}{2} \quad \text{dan} \quad \sigma^2 = \frac{(b-a)^2}{12}$$



Contoh soal 1. Sebuah mesin roll menghasilkan lembaran baja dengan ketebalan berkisar antara $150 \leq y \leq 200$. Tentukan fungsi distribusi peluang, rata – rata, dan variansi dari ketebalan baja jika dianggap menganut distribusi seragam.

Jawab 1.

$$f(y) = \frac{1}{b-a} = \frac{1}{200-150} = \frac{1}{50}$$

$$\mu = \frac{a+b}{2} = \frac{150+200}{2} = 175$$

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(200-150)^2}{12} = \frac{2500}{12}$$

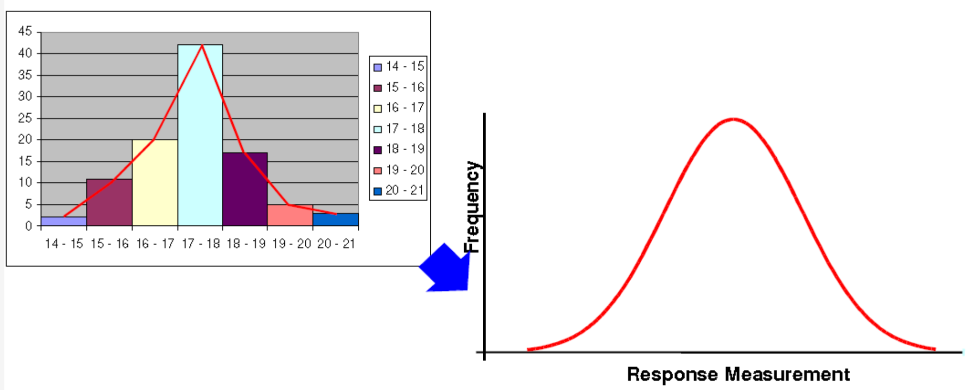
C. Distribusi Normal

Distribusi normal disebut juga “Gaussian Distribution” (sesuai dengan nama penemunya Carl Gauss). Diantara sekian banyak distribusi, distribusi normal merupakan distribusi yang secara luas banyak digunakan dalam berbagai penerapan. Distribusi normal merupakan distribusi kontinyu yang mensyaratkan variabel yang diukur harus kontinyu misalnya tinggi badan, berat badan, skor IQ, jumlah curah hujan, isi botol coca cola, hasil ujian, dan lain lain. Contoh data berdistribusi normal yaitu Jika terdapat 100 orang sampel yang diambil secara acak, setiap orang diminta untuk mengerjakan suatui tugas tertentu. Hasil pengamatan terhadap waktu yang mereka gunakan untuk menyelesaikan tugas tersebut disajikan dalam tabel berikut.

Tabel 5.1 Contoh data berdistribusi normal

Waktu (detik)	Frekuensi	Frekuensi
14-15	2	0,02
15-16	11	0,11
16-17	20	0,2
17-18	42	0,42
18-19	17	0,17
19-20	5	0,05
20-21	3	0,03
Total	100	1

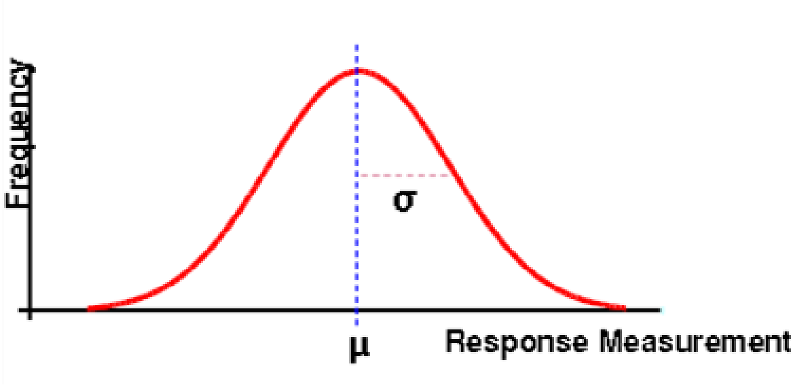
Misalkan percobaan tersebut diulang kembali, kali ini jumlah sampel yang digunakan adalah 5000 orang. Lalu histogram frekuensi relatifnya dibuat dengan lebar kelas yg dibuat kecil (sehingga jumlah kelas menjadi banyak). Maka histogram tersebut akan terdiri atas kotak persegi panjang yg ramping dalam jumlah yang banyak. Dengan semakin banyaknya sampel yg diambil dan lebar interval kelas yang kecil, maka histogram frekuensi relatif yg dihasilkan akan semakin mendekati bentuk kurva normal.



Gambar 5.3. Contoh histogram dan kurva normal distribusi normal

Distribusi Normal f(x) didefinisikan pada interval terbuka $-\infty < x < +\infty$. Distribusi normal dengan parameter $\mu \, dan \, \sigma^2$ biasanay ditulis N(μ, σ^2). Dengan memperhatikan persamaan umum & grafik distribusi normal f(x), tampak bahwa bentuk kurva normal ditentukan oleh dua parameter yaitu rata-rata (μ) dan simpangan baku (σ). Rumus untuk persamaan umum distribusi normal adalah

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{...} \quad (5.1)$$

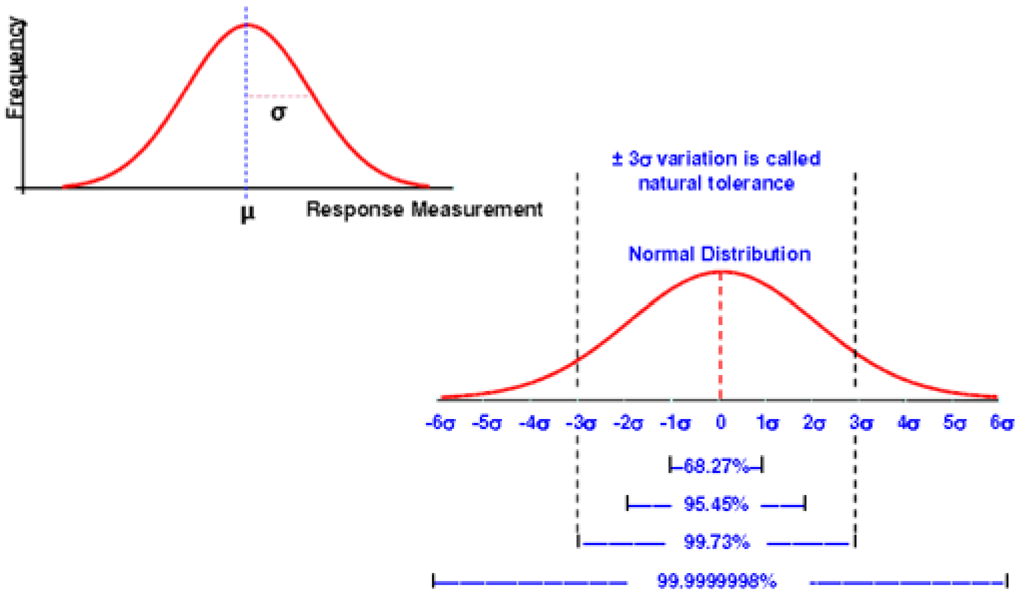


Gambar 5.4. Contoh kurva distribusi Normal

Dari gambar diatas, bila nilai σ mengecil, maka bentuk kurva akan lebih rapat & runcing dan sebagian besar nilai x akan berkumpul atau mendekati nilai rata-rata μ. Sebaliknya makin besar nilai σ, maka bentuk kurva akan lebih renggang dan tumpul, dimana sebagian besar nilai x menjauhi nilai rata-rata μ. Perhatikan gambar 5.5 berikut, σ paling runcing (warna merah dengan σ = 0,5) < σ kuning (σ=0,75) < σ ungu (1,5) dengan nilai rata-rata μ sama yaitu 0.

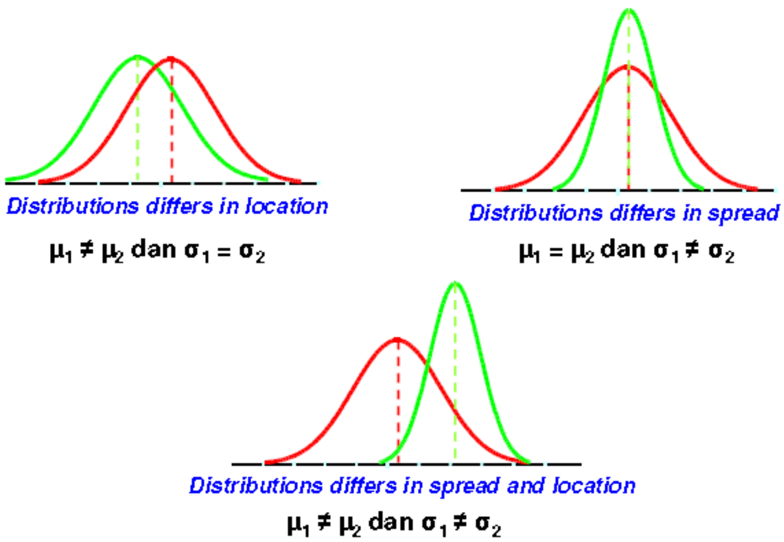
Ciri-ciri Distribusi Normal

1. Kurva berbentuk garis lengkung yang halus dan menyerupai genta/ lonceng ;
2. Kedua ekor/ ujungnya semakin mendekati sumbu absisnya tetapi tidak pernah memotong ;
3. Distribusi normal memiliki dua parameter, yaitu μ dan σ yang masing – masing menentukan lokasi dan bentuk distribusi ;
4. Titik tertinggi kurva normal berada pada rata – rata ;
5. Distribusi normal adalah distribusi yang simetris ;
6. Simpangan baku (standar deviasi = σ), menentukan lebarnya kurva. Makin kecil σ, maka bentuk kurva makin runcing ;
7. Total luas daerah dibawah kurva normal adalah 1 ;
8. Jika jarak dari masing – masing nilai X diukur dengan σ, maka kira – kira 68% berjarak 1σ, 95% berjarak 2σ, dan 99% berjarak 3σ.



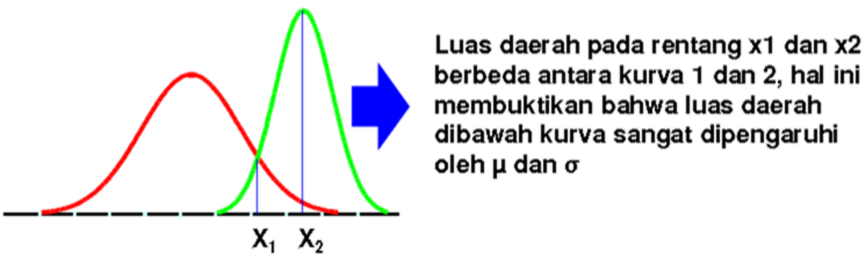
Gambar 5.5 Pendekatan ciri-ciri kurva distribusi normal

Persamaan matematika bagi distribusi probabilitas acak normal tergantung pada dua parameter, yaitu μ dan σ. Bila kedua nilai tersebut diketahui, maka kita dapat menggambarkan kurva normal tersebut dengan pasti.



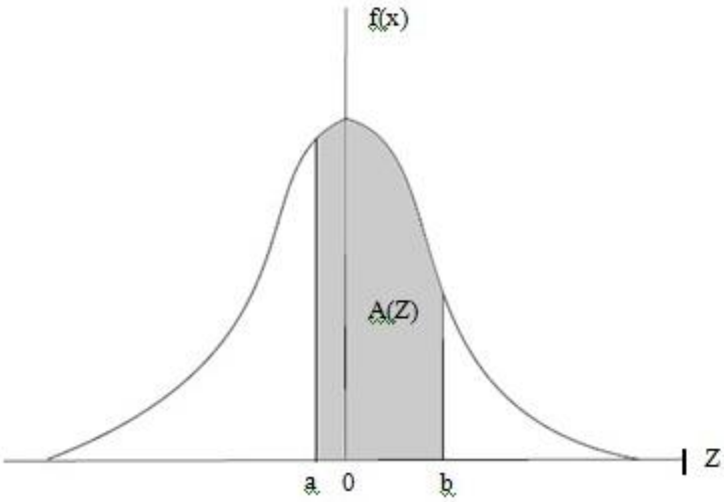
Gambar 5.6 Beberapa contoh perbandingan antara dua jenis kurva normal

Karena persamaan kurva normal tergantung pada nilai – nilai μ dan σ, maka kita akan memiliki bermacam – macam bentuk kurva.



Gambar 5.6 Contoh kurva normal dengan nilai x berbeda

Probabilitas distribusi normal f(x) pada interval a<x<b, ditentukan dengan memakai luas daerah di abwah kurva f(x) sebagaimana ditunjukkan oleh gambar berikut.



Gambar 5.7 Grafik Probabilitas Distribusi Normal f(x) pada interval a<x<b

Probabilitas P(a < x < b) ditunjukkan oleh luas daerah yang diarsir, yang dibatasi oleh kurva f(x), sumbu X, garis tegak X=a dan X=b. Dari Grafik sebelumnya Probabilitas P(a<x<b) dihitung dengan memakai integral dari fungsi f(x) yang diabatasi oleh x=a dan x=b, yaitu

$$P(a < x < b) = \int_a^b f(x)dx = \int_a^b \left(\frac{1}{\sigma\sqrt{2\pi}}\right) e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \text{ .... (5.2)}$$

Akan tetapi secara matematis bentuk integral dari fungsi f(x) tersebut sulit dipecahkan secara langsung dengan teknik integral. Oleh karena itu , penyelesaiannya dilakukan dengan memakai transfor- masi nilai-nilai X menjadi nilai-nilai baku Z yaitu :

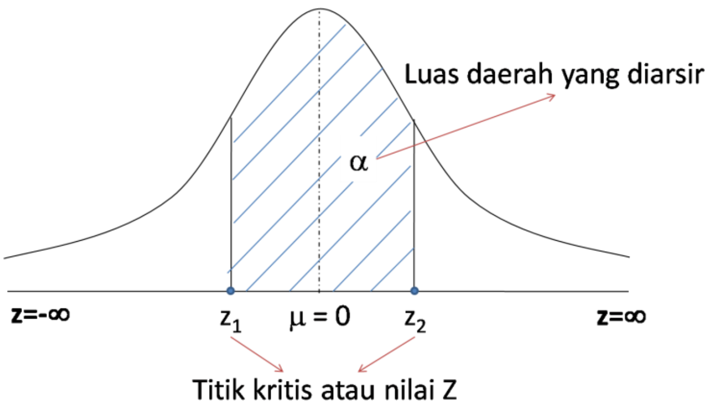
$$Z = \frac{x-\mu}{\sigma} \text{ ....(5.3)}$$

Dengan transformasi tersebut diperoleh distribusi normal Z yang dengan rata-rata μ = 0 dan simpangan baku σ = 1 atau ditulis N(0,1). Distribusi normal Z sseperti ini disebut sebagai distribusi normal standard. Dengan demkian fungsi distribusi f(x) berubah menjadi fungsi distribusi f(Z) yaitu :

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} \text{ ; dimana } -\infty < Z < +\infty \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5.4)$$

Nilai Z *(standard units)* = angka yang menunjukkan penyimpangan suatu variabel acak X dari *mean* (μ) di-hitung dalam satuan standar deviasi (σ). Untuk mengetahui berbagai luas dibawah lengkungan kurva normal standar sudah tersedia tabel luas kurva normal standar.

Berdasarkan fungsi distribusi Z, Probabilitas nilai-nilai Z terletak antara z1<Z<z2 ditunjukkan oleh luas daerah yang diarsir.



Gambar 5.8 Probabilitas fungsi distribusi Z

Selanjutnya Probabilitas P(z1<Z<z2) dihitung dengan rumus sebagai berikut

$$Luas\ daerah\ di\ bawah\ kurva = P(z_1 < Z < z_2) = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2} dz \quad \ldots\ldots\ldots\ldots\ldots\ldots (5.5)$$

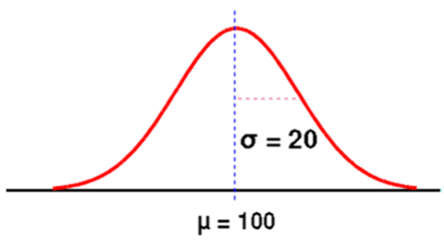
Berdasarkan integral dari fungsi distribusi normal standard Z tsb, probabilitas P(z1<Z<z2) dihitung dengan memakai "Tabel Distribusi Normal Standar".

Contoh Soal 1. Misalkan dimiliki kurva normal dengan μ = 100 dan σ = 20. Hitunglah :

a. Luas kurva normal antara 100 – 125 atau P(100 ≤ x ≤ 125)

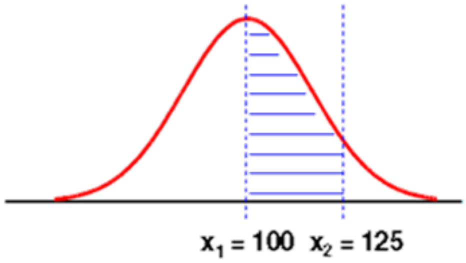
b. Luas kurva normal antara 80 – 100 atau P(80 ≤ x ≤ 100)

c. Luas kurva normal antara 75 – 120 atau P(75 ≤ x ≤ 120)



Jawab 1.

a.

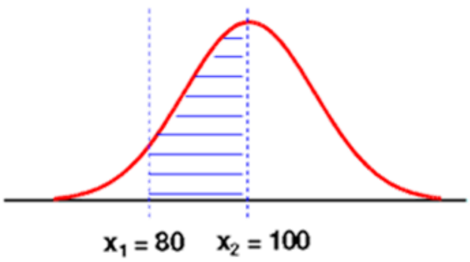


$$z_1 = \frac{100 - 100}{20} = \frac{0}{20} = 0 \longrightarrow \text{Menurut tabel luasnya} = 0{,}5000$$

$$z_2 = \frac{125 - 100}{20} = \frac{25}{20} = 1{,}25 \longrightarrow \text{Menurut tabel luasnya} = 0{,}8944$$

MAKA, Luas kurva normal antara 100 – 125 = 0,8944 – 0,5000 = 0,3944

b.

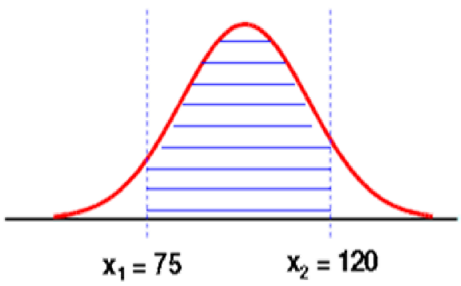


$$z_1 = \frac{80 - 100}{20} = \frac{-20}{20} = -1 \longrightarrow \text{Menurut tabel luasnya} = 0{,}1587$$

$$z_2 = \frac{100 - 100}{20} = \frac{0}{20} = 0 \longrightarrow \text{Menurut tabel luasnya} = 0{,}5000$$

MAKA, Luas kurva normal antara 80 – 100 = 0,5000 – 0,1587 = 0,3413

c.



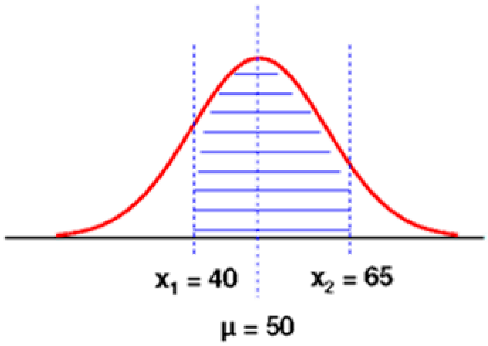
$$z_1 = \frac{75 - 100}{20} = \frac{-25}{20} = -1{,}25 \longrightarrow \text{Menurut tabel luasnya} = 0{,}1056$$

$$z_2 = \frac{120 - 100}{20} = \frac{20}{20} = 1 \longrightarrow \text{Menurut tabel luasnya} = 0{,}8413$$

MAKA, Luas kurva normal antara 75 – 120 = 0,8413 – 0,1056 = 0,7357

Contoh Soal 2. Misalnya seorang sarjana teknik mesin menyelidiki hasil panenan padi untuk merancang sebuah mesin perontok padi. Dari 300 orang petani di suatu daerah diketahui hasil panenan rata – rata sebesar 50 kwintal dengan deviasi standar sebesar 10 kwintal. Peneliti tersebut telah menge- cek distribusi hasil panenan dan dinyatakan memiliki distribusi normal. Tentukan probabilitas hasil panenan berkisar antara 40 sampai 65 kwintal.

Jawab 2.



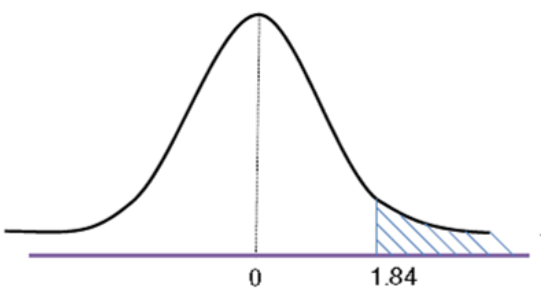
$$z_1 = \frac{40-50}{10} = \frac{-10}{10} = -1 \longrightarrow$$ **Menurut tabel luasnya = 0,1587**

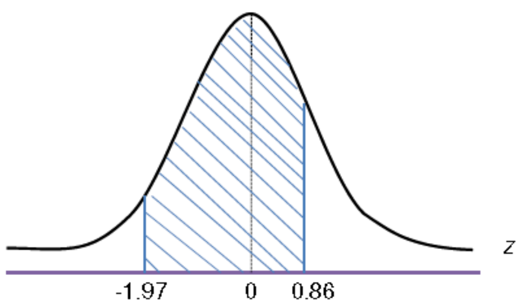
$$z_2 = \frac{65-50}{10} = \frac{15}{10} = 1{,}5 \longrightarrow$$ **Menurut tabel luasnya = 0,9332**

MAKA, Luas kurva normal antara 40 – 65 = 0,9332 – 0,1587 = 0,7745

Contoh Soal 3. Diberikan distribusi normal baku, hitunglah daerah di bawah kurva yang dibatasi:

a. sebelah kanan z = 1.84
b. antara z = -1.97 dan z = 0.86





Jawab 3

a. Luas sebelah kanan = 1 – luas sebelah kiri z = 1.84 (lihat gambar). Dari tabel luas sebelah kiri = 0.9671, jadi Luas sebelah kanan = 1 – 0.9671 = 0.0329

b. Luas daerah antar batas tersebut adalah luas di sebelah kiri z = 0.86 dikurangi dengan luas di sebelah kiri z = -1.97.

Dari tabel diperoleh 0.8051 - 0.0244 = 0.7807

Contoh Soal 4. Sebuah mesin pembuat resistor dapat memproduksi resistor dengan ukuran rata-rata 40 ohm dengan standard deviasi 2 ohm. Misalkan ukuran tersebut mempunyai distribusi normal, tentukan peluang resistor mempunyai ukuran lebih dari 43 ohm.

Jawab 4.

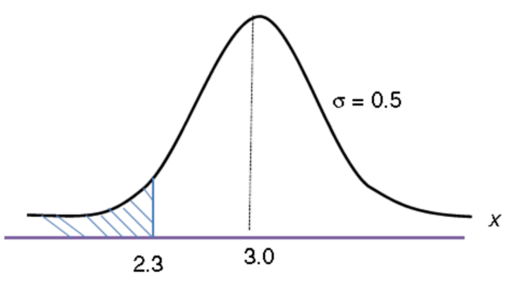
$$Z = \frac{(43-40)}{2} = 1.5$$

sehingga dapat dihitung:

$$P(X > 43) = P(Z > 1.5) = 1 - P(Z < 1.5)$$
$$= 1 - 0.9332 = 0.0668$$

Contoh Soal 5. Suatu jenis batere mobil rata-rata berumur 3 tahun dengan simpangan baku 0.5 tahun. Bila dianggap umur bater berdistribusi normal, carilah peluang suatu batere berumur kurang dari 2,3 tahun.

Jawab 5.



Untuk menghitung $P(X < 2.3)$, hitunglah luas di bawah kurva normal sebelah kiri titik 2.3. Ini sama saja menghitung luas daerah sebelah kiri z padanannya: $z = (2.3 - 3.0)/0.5 = -1.4$ dan dari tabel normal baku diperoleh:

$$P(X < 2.3) = P(Z < -1.4) = 0.0808$$

Contoh soal 6. Terdapat 200 orang mahasiswa yang mengikuti ujian Kalkulus di sebuah Prodi, diperoleh bahwa nilai rata-rata adalah 60 dan simpangan baku (standard devisasi) adalah 10. Bila distribusi nilai menyebar secara normal, berapa:

a. persen yang mendapat A, jika nilai A $\geq$ 80;

b. persen yang mendapat nilai C, jika nilai C terletak pada interval $56 \leq C \leq 68$;

c. persen yang mendapat nilai E jika nilai E < 45

Misalkan X adalah peubah acak yang menyatakan nilai ujian Kalkulus.

Jawab 6.

$$Z = \frac{X - 60}{10}$$

(a) $z = (80 - 60)/10 = 2$

$P(X \geq 80) = P(Z \geq 2) = 1 - 0.9772 = 0.0288 = 2.28\%$

(b) $z_1 = (56 - 60)/10 = -0.4$ dan $z_2 = (68 - 60)/10 = 0.8$

$P(56 \leq X \leq 68) = P(-0.4 \leq Z \leq 0.8) = P(Z < 0.8) - P(Z < -0.4)$

$= 0.4435 = 44.35\%$

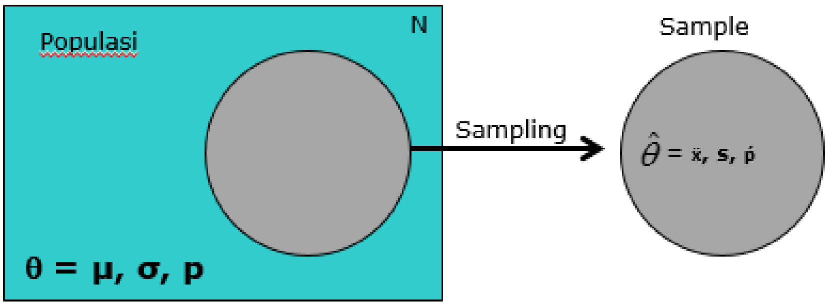
(c $z = (45 - 60)/10 = -1.5$

$P(X \leq 45) = P(Z \leq -1.5) = 0.0688 = 6.68\%$

BAB 6. DISTRIBUSI PENCUPLIKAN / DISTRIBUSI SAMPEL

A. Pengertian dan Konsep Dasar

Istilah-istiah yang sering digunakan dalam distribusi meliputi informasi berikut. Populasi merupakan kumpulan/himpunan dari keseluruhan elemen yang menjadi obyek penelitian/penyelidikan, sedangkan sampel adalah sebagian obyek dari populasi. Elemen merupakan sesuatu yang menjadi obyek penyelidikan. Sementara itu pengertian dari sensus adalah cara mengumpulkan data dari populasi, dan sampling adalah cara mengumpulkan data dari sampel. Kelemahan dari metode sensus dalam pengumpulan data yaitu dibutuhkan biaya mahal, waktu lama, tenaga yang besar. Kelemahan sensus diatasi dengan teknik sampel (sampling). Sedangkan keuntungan teknik sampel adalah biaya yang rendah serta waktu yang pendek tanpa mengurangi keakuratan. Karateristik populasi merupakan sifat-sifat atau ciri-ciri yang diamati dalam suatu populasi. Selanjutnya pengertian dari parameter adalah karakteristik yang dihitung dari populasi, sedangkan statistik adalah karakteristik yang dihitung dari sampel.



Gambar 6.1 Grafik hubungan antara populasi dan sampel

Terdapat gap antara populasi dan sampel yang disebut sebagai kesalahan (penyimpangan). Sebab kesalahan sampel: kesalahan pemilihan sampel, kesalahan hitung, dan lain lain. Sampel yang representatif memiliki ciri: ukuran tertentu yang memakai syarat, kesalahan terkecil, dan dipilih dengan prosedur yang benar berdasarkan teknik sampel tertentu.

Tabel 6.1 Perbedaan karakteristik populasi dan sampel

Karakteristik populasi	**Karakteristik Sampel**
Ukuran N	Ukuran n
Parameter	Statistik
Mean μ	Mean X
Simpangan baku σ	Simpangan baku S
Populasi berhingga dan tak berhingga	Sampel besar dan kecil

Sampel merupakan potret dari populasi, dimana dengan adanya sampel diharapkan bisa memperoleh gambaran yang sesungguhnya mengenai karakteristik populasi. Hasil perhitungan dari data sampel akan dipakai untuk menyimpulkan parameter populasi. Sebagai contoh jika jumlah mahasiswa di UMY N= 20.000 (populasi mahasiswa) dan hanya n=100 yang pilih dari 100 sampel mahasiswa yg terpilih diharapkan mampu mewakili 20.000 populasi dalam penelitian.

Parameter populasi ditulis dengan huruf latin θ, di mana θ bisa berupa rata-rata populasi μ, simpangan baku populasi σ, proporsi populasi p. Sedangkan statistik dari sampel ditulis $\hat{\theta}$ (θ topi), bisa berupa rata-rata sampel **X**, simpangan baku sampel S, proporsi sampel $\hat{p}$.

B. Teknik Pengambilan Sampel

Teknik pengambilan sampel menurut proses memilihnya terbagi menjadi dua bagian yaitu pengambilan sampel dengan pengembalian, dimana ukuran populasi adalah tetap. Sesuai untuk ukuran populasi terbatas serta selanjutnya adalah pengambilan sampel tanpa pengembalian, dimana ukuran populasi akan berkurang, sesuai untuk populasi tak terbatas. Sementara itu, teknik pengambilan sampel menurut probabilitasnya terbagi menjadi empat yaitu teknik pengambilan sampel acak sederhana, sistematik, stratifikasi dan kluster.

Teknik pengambilan sampel acak sederhana diartikan sebagai pengambilan sampel sebanyak n sedemikian rupa sehingga setiap unit dalam populasi mempunyai kesempatan yang sama untuk dipilih menjadi sampel dan setiap ukuran sampel n juga mempunyai kesempatan yang sama untuk terambil. Metode yang bisa digunakan pada pengambilan ini adalah dengan cara undian (digoncang seperti arisan) atau tabel bilangan acak. Sedangkan teknik pengambilan sampel acak sistematik dilakukan dengan mengambil setiap unsur ke-k dalam populasi untuk dijadikan sampel, dengan titik awal ditentukan secara acak diantara k unsur yang pertama. Tenik ini sering digunakan karena dapat menarik kesimpulan yang tepat mengenai parameter populasi sebab sampelnya menyebar secara merata di seluruh populasi. Sementara itu, teknik pengambilan sampel acak statifikasi dilakukan dengan membagi populasi menjadi beberapa kelompok / sub populasi (strata), sehingga setiap kelompok akan memiliki nilai variabel yang tidak terlalu bervariasi (relatif homogen). Selanjutnya dari setiap strata dipilih sampel melalui proses pengambilan secara sederhana. Sedangkan pada teknik pengambilan sampel acak stratifikasi dilakukan dengan membagi populasi menjadi beberapa strata (tingkatan) kemudian sampel diambil secara acak dari setiap tingkatan. Teknik ini dilakukan bila populasinya heterogen. Cara pengambilan sampel untuk setiap tingkatan tidak sama, harus sebanding dengan jumlah anggota setiap tingkatan (proporsional). Rumusan umum teknik pengambilan sampel acak stratifikasi adalah $n_i = \frac{N_i}{N}n.$ Terakhir, teknik pengambilan sampel acak kluster dilakukan dengan mengambil beberapa kluster (kelompok) secara acak kemudian semua atau sebagian dari anggota masing-masing kelompok diambil secara acak sebagai sampel.

C. Distribusi Sampel

Statistik yang dihitung dari sampel digunakan untuk menyimpulkan parameter populasi. Dalam hal ini berarti dilakukan generalisasi data. Generasilasi akan lebih baik dan meyakinkan jika sampel diambil secara berulang-ulang dan acak sehingga diperoleh banyak sampel acak yang berlainan dari populasi yang sama. Data dari sampel dihitung karakteristiknya dan karena sampel bersifat acak, maka statistik dari sampel juga bersifat acak sehingga secara keseluruhan statistik yang diperoleh akan membentuk suatu distribusi yaitu distribusi statistik. Distribusi probabilitas dari suatu statistik disebut distribusi sampel.

Pada distribusi sampel, statistik sampel yang diperoleh bersifat acak (variabel acak) yang mengikuti suatu distribusi tertentu. Bila statistik yang dihitung dari sampel adalah rata-rata maka akan diperoleh distribusi sampel rata-rata. Bila statistik yang dihitung dari sampel adalah proporsi maka akan diperoleh distribusi sampel proporsi.

D. Distribusi Sampel Rata-rata

Pada distribusi sampel rata-rata, bila populasi **terbatas** berukuran N dengan rata-rata μx dan simpangan baku σx diambil sampel berukuran n secara berulang **tanpa pengembalian**, maka diperoleh rumusan sebagai berikut.

Rata-rata untuk distribusi sampel rata rata $\mu_{\bar{x}} = \mu_x$ …………………………….…… (6.1)

Simpangan baku pada sampel rata rata $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ ………………………………. (6.2)

dimana $\sqrt{\frac{N-n}{N-1}}$ disebut sebagai faktor koreksi.

Pada distribusi sampl rata-rata, bila populasi **tak terbatas** dengan rata-rata μx dan simpangan baku σx diambil sampel berukuran n **tanpa pengembalian**, maka diperoleh rumusan sebagai berikut.

Rata-rata untuk distribusi sampel rata rata $\mu_{\bar{x}} = \mu_x$ ……………………………….…. (6.3)

Simpangan baku pada sampel rata rata $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$ ………………..…………………….... (6.4)

Karena N besar sekali (populasi tak terbatas) maka perbandingan $\frac{N-n}{N-1} = 1$ dan $\sqrt{\frac{N-n}{N-1}} = 1$

Pada distribusi sampl rata-rata, bila sampel diambil cukup besar, n>=30, maka **distribusi sampelnya akan mendekati distribusi normal** sehingga variabel random Z dapat dihitung dengan rumus sebagai berikut.

$$Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu_x}{\sigma_{\bar{x}}} \quad \text{………………………………………..…………………….. (6.5)}$$

Contoh soal 1. Suatu sampel acak sebesar n = 10 diambil tanpa pengembalian dari populasi **tak terbatas** yang mempunyai rata-rata $\mu x = 5{,}5$ dan simpangan baku $\sigma x = 2{,}92$. Tentukan rata-rata dan simpangan baku dari sampel rata-rata tersebut.

Jawab 1.

Cara menjawab pertanyaan ini gunakan rumus 6.3 dan 6.4 sebagaimana berikut.

Rata-rata untuk distribusi sampel rata rata $\mu_{\bar{x}} = \mu_x$

Simpangan baku pada sampel rata rata $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{2{,}92}{\sqrt{10}} = 0{,}92$

Contoh soal 2. Kecepatan maksimum 2000 mobil mempunyai rata-rata 135,5 km/jam dengan simpangan baku 5,2 km/jam. Jika sampel sebesar 150 mobil dipilih secara acak tanpa pengembalian, hitung probabilitas kecepatan maksimum rata-rata dari 150 mobil tersebut yang lebih besar dari 136,1 km/jam!

Jawab 2.

- Populasi terbatas N= 2000
- Sampel besar (n>30), n= 150
- Kecepatan rata-rata $\mu x = 135{,}5$ Km/Jam
- Simpangan baku $\sigma x = 5{,}2$ Km/jam
- Distribusi sampel kecepatan maksimum rata-rata merupakan distribusi normal sehingga statistik Z mempunyai distribusi normal standar $Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu_x}{\sigma_{\bar{x}}}$

  Lihat tabel distribusi normal 1,46 ➔ 0,9279

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}} = \frac{5{,}2}{\sqrt{150}} . \sqrt{\frac{2000 - 150}{2000 - 1}} = 0{,}41$$

$$Z = \frac{\bar{X} - \mu_X}{\sigma_{\bar{X}}} = \frac{136{,}1 - 135{,}5}{0{,}41} = 1{,}46$$

Jadi probabilitas kecepatan maksimum rata-rata mobil yang lebih besar dari 136,1 km/jam adalah P(X>136,1) = P(Z>1,46) = 1- 0,9279 = 0,0721.

E. Distribusi Sampel Proporsi

Pada distribusi sampel proporsi, bila populasi berukuran N mengandung jenis p sebanyak X, maka proporsi p adalah X/N. Jika dari populasi tersebut diambil sampel berukuran n yang juga mengandung proporsi x/n dan sampel diambil berulang maka statistik yang bersifat acak. Sehingga mempunyai distribusi sampel proporsi dengan rata-rata dan simpangan baku sebagai berikut.

Rumus umum proporsi $\hat{p} = \frac{x}{n}$

Rata-rata pada distribusi sampel proporsi $\mu_{\hat{p}} = \mu_p = \frac{X}{N}$ ………………..…………..... (6.6)

Simpangan baku sampel proporsi bila populasi terbatas $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ ……... (6.7)

Simpangan baku sampel proporsi bila populasi tak terbatas $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ ……......... (6.8)

Variabel random Z pada distribusi sampel proporsi $Z = \frac{\hat{p}-p}{\sigma_{\hat{p}}}$ ………………................ (6.9)

Contoh soal 3. Diketahui sebanyak 10% dari ibu-ibu rumah tangga di Bandung memakai detergen A untuk mencuci pakaiannya. Jika dari populasi tersebut diambil sampel berukuran 100. Tentukan rata-rata dan simpangan baku dari populasi ibu-ibu rumah tangga yang memakai detergen A! Bila dari sampel tersebut ternyata terdapat paling sedikit 15 ibu rumah tangga yang memakai detergen A, tentukan probabilitasnya!

Jawab 3.

- p = 10% = 0,1
- Populasi ibu-ibu rumah tangga di Bandung ➔ N populasi tak terbatas
- Sampel besar (n>30) ➔ n= 100
- Rata-rata μp(topi) dan Simpangan baku σp(topi) [$\hat{p}$ => $p(topi)$] ibu-ibu rumah tangga yang memakai detergen A ➔ Rata-rata μp(topi) dan Simpangan baku σp(topi) distribusi sampel proporsi
- Bila x Paling sedikit 15 (x≥15) dalam sampel n = 100 maka p(topi) = x/n = 15/100 = 0,15

Rata-rata = 0,1

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0,1.0,9}{100}} = 0,03$$

Proporsi yang memakai detergen A adalah 15/100 = 0,15

$$Z = \frac{\hat{p}-p}{\sigma_{\hat{p}}} = \frac{0,15-0,1}{0,03} = 1,67$$

- Dari tabel distribusi normal untuk nilai Z=1,67 ➔ 0.9525
- Sehingga P(Z>1,67) = 1- 0.9525 = 0,0475

F. Distribusi Sampel Beda Dua Rata-rata

Pada distribusi sampel beda dua rata-rata, jika terdapat 2 populasi dimana populasi 1 sebanyak N1, mempunyai rata-rata μ1 dan simpangan baku σ1 kemudian Populasi 2 sebanyak N2 mempunyai rata-rata μ2 serta simpangan baku σ2. Dari populasi 1 diambil sampel acak sebanyak n1 dengan rata-rata X1 dan dari populasi 2 sampel acak sebanyak n2 dengan rata-rata X2 dimana kedua sampel tersebut dianggap saling bebas. Dari sampel X1 dan X2 dapat dibuat sampel baru yang juga bersifat acak, yaitu sampel beda dua rata-rata. Rumus umum untuk nilai rata-rata dan simpangan baku dari distribusi sampel beda dua rata-rata adalah sebagai berikut.

Rata-rata pada distribusi sampel beda dua rata-rata $\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2$ …………..... (6.10)

Simpangan baku pada distribui sampel beda dua rata-rata bila populasi terbatas adalah

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} \cdot \sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{(N_1+N_2)-1}} \quad \text{……..…………………………..….. (6.11)}$$

Simpangan baku bila populasi tak terbatas $\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$ ………………....... (6.12)

Variabel random Z pada distribui sampel beda dua rata-rata yaitu

$$Z = \frac{(\bar{X}1-\bar{X}2)-(\mu_1-\mu_2)}{\sigma_{\bar{x}_1-\bar{x}_2}} \quad \text{…………….. (6.13)}$$

Contoh soal 4. Di suatu universitas diketahui rata-rata tinggi badan mahasiswa laki-laki adalah 164 cm dengan simpangan baku 5,3 cm. Sedangkan mahasiswa perempuan tinggi badannya rata-rata 153 cm dengan simpangan baku 5,1 cm. Dari dua populasi tersebut diambil sampel acak yang saling bebas masing-masing 150 orang. Berapa probabilitas rata-rata tinggi mahasiswa laki-laki paling sedikit 12 cm lebihnya daripada rata-rata tinggi mahasiswa perempuan?

Jawab 4.

Diketahui populasi 1: $\mu_1 = 164$ cm, $\sigma_1 = 5{,}3$ cm dan sampel 1 : $n_1 = 150$ orang

populasi 2: $\mu_2 = 153$ cm, $\sigma_2 = 5{,}1$ cm dan sampel 2 : $n_2 = 150$ orang

Misal $\overline{X}_1$ = rata-rata tinggi badan mahasiswa laki-laki

$\overline{X}_2$ = rata-rata tinggi badan mahasiswa perempuan

Rata-rata: $\mu_{\bar{X}_1-\bar{X}_2} = \mu_1 - \mu_2 = 164 - 153 = 11$ cm

Simpangan baku: $\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} = \sqrt{\frac{5{,}3^2}{150} + \frac{5{,}1^2}{150}} = 0{,}6$

$$Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - 11}{0{,}6}$$

Karena rata- rata tinggi badan mahasiswa laki-laki **paling sedikit 12 cm lebihnya** dari pada rata-rata tinggi mahasiswa perempuan, maka $(\bar{X}1 - \bar{X}2) \geq 12$ sehingga $Z = \frac{12-11}{0,6} = 1,67$
Dari tabel untuk nilai Z=1,67 ➔ 0.9525 sehingga P(Z≥1,67) = 1- 0.9525 = 0,0475.

G. Distribusi Sampel Beda Dua Proporsi

Pada distribusi sampel beda dua proporsi, terdapat dua populasi dimana populasi 1 berukuran N1 terdapat jenis X1 dengan proporsi X1/N1 dan populasi 2 berukuran N2 terdapat jenis X2 dengan proporsi X2/N2. Bila populasi 1 diambil sampel acak berukuran n1 maka sampel ini akan mengandung jenis x1 dengan proporsi x1/n1. Demikian juga dengan populasi 2 diambil sampel acak berukuran n2 maka sampel ini akan mengandung jenis x2 dengan proporsi x2/n2. Sampel 1 dan 2 dapat membentuk sampel acak baru yaitu sampel beda dua proporsi. Rumus umum distribusi sampel beda dua proporsi yaitu sebagai berikut.

Rata-rata pada distribusi sampel beda dua proporsi $\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2$ …………..... (6.14)

Simpangan baku pada distribui sampel beda dua proporsi bila populasi terbatas adalah

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1\,(1-p_1)}{n_1} + \frac{p_2\,(1-p_2)}{n_2}} \cdot \sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{(N_1+N_2)-1}} \quad \text{……..…………..……….} \quad (6.15)$$

Simpangan baku bila populasi tak terbatas $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1\,(1-p_1)}{n_1} + \frac{p_2\,(1-p_2)}{n_2}}$ ……. (6.16)

Variabel random Z pada distribui sampel beda dua rata-rata yaitu

$$Z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sigma_{\hat{p}_1-\hat{p}_2}} \quad \text{………………..} \quad (6.17)$$

Contoh soal 5. Terdapat 5% barang di gudang timur cacat, sedangkan barang yang cacat di gudang barat sebanyak 10%. Bila diambil sampel acak sebanyak 200 barang dari gudang timur dan 300 barang dari gudang barat. Tentukan probabilitas persentase barang yang cacat dalam gudang barat 2% lebih banyak dibanding gudang timur!

Jawab 5.

$$\text{Gudang barat}: n_1 = 300, p_1 = 0,1$$
$$\text{Gudang timur}: n_2 = 200, p_2 = 0,05$$
$$\hat{p}_1 = \text{proporsi barang yang cacat di gudang barat dalam sampel}$$
$$\hat{p}_2 = \text{proporsi barang yang cacat di gudang timur dalam sampel}$$
$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \sqrt{\frac{0,1(0,9)}{300} + \frac{0,05(0,95)}{200}} = 0,023$$
$$Z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sigma_{\hat{p}_1-\hat{p}_2}} = \frac{(\hat{p}_1-\hat{p}_2)-(0,1-0,05)}{0,023}$$

Karena barang cacat di gudang barat 2% lebih banyak daripada di gudang timur, maka $(\hat{p}_1 - \hat{p}_2) > 0{,}02$ sehingg a diperoleh $Z = \frac{0{,}02-0{,}05}{0{,}023} = -1{,}3$ dari tabel distribusi normal untuk nilai Z= -1,3 ➔ 0.0968 sehingga $(\hat{p}_1 - \hat{p}_2) > 0{,}02$= P(Z>-1,3) =1- 0.0968 = 0,9032 = 90,32%.

Soal Latihan.

1. Pada suatu pengiriman barang yang terdiri dari 2000 tube elektronika telah diketahui terdapat 600 unit tube yang tidak memenuhi standar mutu. Jika sampel acak sebanyak 500 unit dipilih dari populasi tersebut tanpa pengembalian, berapakah probabilitas sampel populasi yang tidak memenuhi standar mutu :
   a. akan kurang dari 150/500
   b. antara 144/500 sampai dengan 145/500
   c. lebih besar dari 164/500

2. Besi baja yang diproduksi perusahaan A mempunyai rata-rata daya regang sebesar 4500 lbs dan variansi sebesar 40000 lbs, sedangkan yang diproduksi perusahaan B mempunyai ratarata daya regang sebesar 4000 lbs dan variansi sebesar 90000 lbs. Misalkan sampel random sebanyak 50 diambil dari perusahaan A dan sampel random sebanyak 100 diambil dari perusahaan B, berapakah probabilitas rata-rata daya regang beda dua rata-rata dari dua sampel itu yang lebih besar dari 600 lbs?

BAB 7. TEORI PENDUGAAN (TEORI ESTIMASI)

A. Pengantar Teori Estimasi

Pada proses pengambilan sampel, tujuan utama pengambilan sampel dari suatu populasi adalah untuk memperoleh informasi mengenai parameter populasi atau singkatnya untuk mengetahui parameter populasi itu sendiri. Parameter populasi seperti rata-rata dan simpangan baku, sedangkan contoh parameter dalam praktiknya yaitu rata-rata nilai ujian bahasa inggris mahasiswa UMY. Median nilai ujian kalkulus mahasiswa UMS.

Pada kondisi tertentu sering kali parameter populasi tidak diketahui, meskipun distribusi populasi diketahui. Sebagaimana misalnya suatu populasi mempunyai distribusi normal tetapi parameter rata-rata dan simpangan baku tidak diketahui atau misalkan ada suatu populasi yang memiliki distribusi binomial, tetapi parameter proporsi p tidak diketahui. Oleh Karena parameter populasi tidak diketahui, maka ada dua cara untuk mengetahui parameter populasi yang dipelajari dalam statistika inferensia yaitu cara pendugaan (penaksiran/estimasi) serta pengujian hipotesis. Dua cara ini didasarkan pada besaran yang dihitung dari sampel. Sebagai tambahan informasi bahwasanya terdapat dua jenis statistik yaitu statistika deskriptif dan inferensia. Statistika Deskriptif adalah statistika yang berkenaan dengan metode atau cara mendeskripskan, menggambarkan, menjabarkan atau mengurangi data. Sedangkan, statistika Inferensia adalah statistika yang berkenaan dengan cara penarikan kesimpulan berdasarkan data yang diperoleh dari sampel untuk menggambarkan karakteristik atau ciri dari suatu populasi. Pada statistika inferensia, semua metode yang berhubungan dengan analisis sebagian data atau juga sering disebut dengan sampel untuk kemudian sampai pada estimasi atau penarikan kesimpulan mengenai keseluruhan data. Dalam statistika inferensia, statistik $\hat{\theta}$ inilah yang dipakai untuk menduga parameter θ dari populasi. Parameter $\hat{\theta}$ adalah penduga, sedangkan θ adalah sesuatu yang diduga. Statistik $\hat{\theta} = \bar{X}$ dipakai untuk menduga parameter $\theta = \mu$, statistik $\hat{\theta} = S$ dipakai untuk menduga parameter $\theta = \sigma$ dan statistik $\hat{\theta} = \hat{p}$ dipakai untuk menduga parameter $\theta = p$.

B. Penduga yang Baik

Pada saat tujuan statistik adalah untuk memperoleh gambaran yang baik mengenai populasi, maka statistik $\hat{\theta}$ yang dipakai untuk menduga parameter θ haruslah merupakan suatu penduga yang baik. Ciri-ciri penduga yang baik terdiri atas tiga hal yaitu bahwa $\hat{\theta}$ merupakan penduga tidak bias, penduga yang efisien serta merupakan penduga yang konsisten. Penduga tak bias artinya penduga yang dengan tepat mengenai sasaran atau apabila nilai penduga sama dengan nilai yang diduganya. Penduga Efisien artinya bila ada lebih dari satu penduga, maka penduga yang efisien adalah penduga yang mempunyai variansi paling kecil. Sedangkan penduga yang konsisten artinya jika ukuran sampel yang diambil semakin bertambah maka nilai penduga akan mendekati parameternya (bila sampel semakin besar, maka nilai *tetha* topi akan semakin mendekati nilai *tetha).*

C. Pendugaan Titik

Teori pendugaan terbagi menjadi dua jenis yaitu pendugaan titik (estimasi titik) dan pendugaan interval (estimasi interval). Pada pendugaan titik, bila nilai parameter θ dari populasi hanya diduga dengan memakai satu nilai statistik $\hat{\theta}$ dari sampel yang diambil dari populasi tersebut. Pada pendugaan interval (estimasi interval), bila nilai parameter θ dari populasi diduga dengan memakai beberapa nilai statistik $\hat{\theta}$ yang berada dalam suatu interval, misalnya $\hat{\theta}1 < \theta < \hat{\theta}2$. Rumusan umum untuk pendugaan titik adalah sebagai berikut.

Rata-rata pendugaan titik $\overline{X} = \frac{\sum X}{n}$ .... (7.1)

Variansi pendugaan titik $S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$ .... (7.2)

Proporsi pendugaan titik $\hat{p} = \frac{X}{n}$ .... (7.3)

Contoh pemanfaatan pendugaan titik yaitu ketika seseorang ingin menduga berapa sesungguhnya rata-rata tinggi badan orang Indonesia, untuk itu diambil sampel acak sebanyak 1000 orang dan kita ukur tinggi badan masing-masing. Misal diperoleh rata-rata tingginya X=164cm. Nilai rata-rata ini digunakan untuk menduga rata-rata tinggi badan orang indonesia yang sesungguhnya. Oleh karena hanya memakai satu nilai saja X=164 sebagai penduga maka X=164 cm disebut sebagai penduga titik. Dalam praktiknya terdapat kelemahan pada penggunaan pendugaan titik seperti sulitnya menentukan nilai derajat kepercayaan serta jika sampel yang berbeda maka nilai statistik yang dihasilkan juga beda karena hanya satu maka kita akan ragu penduga mana yang baik.

D. Pendugaan Interval

Pada pendugaan atau estimasi interval Bila nilai parameter θ dari populasi diduga dengan memakai beberapa nilai statistik $\hat{\theta}$ yang berbeda dalam suatu interval, misalnya $\hat{\theta}1<\theta<\hat{\theta}2$, maka statistik $\hat{\theta}$ disebut penduga interval. Sebagaimana contoh pada kejadian sebelumnya dimana tinggi badan orang Indonesia diduga pada interval $160 <\theta<166$ atau

155<θ<169, tinggi orang indonesia diduga pada interval tersebut. Dalam pendugaan interval semakin lebar interval, semakin besar kepercayaan. Dalam praktek interval yang harus dipakai adalah interval yang sempit tetapi mempunyai derajat kepercayaan yang dapat diterima. Derajat kepercayaan penduga $\hat{\theta}$ meliputi nilai koefisien kepercayaan yang ditulis dengan α dimana 0<α<1 dan dinyatakan dalam bentuk probabilitas. Kondisi ini dapat digambarkan sebagai berikut. Pada contoh sebelumnya, rata-rata tinggi badan orang indonesia diduga berada pada interval 160<θ<166 dengan probabilitas 0,95 maka dituliskan P(160< θ<166)=0,95. Bila rata-rata tinggi orang indonesia diduga berada pada interval 155< θ <169 dengan probabilitas 0,99 maka bisa dituliskan P(155< θ<169)=0,99. Dalam statistika biasanya dipilih interval yang lebih pendek, tetapi dengan probabilitas yang tinggi atau kepercayaan yang tinggi. Dari kondisi diatas maka lebih baik memilih P(160<θ<166)=0,95 karena terkadang dengan adanya keterbatasan dalam ukuran sampel, pemilihan interval harus dengan mengorbankan derajat kepercayaan karena interval yang sempit dengan probabilitas yang tinggi sulit dicapai sekaligus.

Rumus umum pendugaan interval dilakukan dengan mengambil sampel acak secara berulang maka akan diperoleh distribusi statistik θ sehingga probabilitas dari interval $\hat{\theta}1<\theta<\hat{\theta}2$ akan sama dengan nilai tertentu yang diinginkan, sebagaimana persamaan berikut.

$P\left[\hat{\theta}1 < \theta < \hat{\theta}2\right] = 1 - \propto$; $\text{dimana } 0 < \propto < 1$ ……………………………………..… (7.4)

- α disebut koefisien kepercayaan
- 1-α derajat kepercayaan
- $P\left[\hat{\theta}1 < \theta < \hat{\theta}2\right]$ adalah interval kepercayaan

Sesuai dengan persamaan tersebut diatas, maka dari contoh sebelumnya bahwa P(160<θ<166)= 0,95, maka ungkapan yang tepat sekarang kita percaya 95% bahwa parameter populasi θ akan terletak antara 160 sampai dengan 166 berdasarkan sampel yang diambil dr populasi tersebut. Jadi BUKAN, diungkapkan dengan probabilitas sama dengan 0.95 bahwa parameter populasi θ terletak antara 160 sampai 166 berdasar sampel yang diambil dari populasi itu.

E. Pendugaan Parameter Populasi dengan Sampel Besar (n ≥ 30)

Pada pendugaan populasi dengan sampel beesar terbagi atas empat jenis sampel yaitu sampel rata-rata, beda sua rata-rata, proporsi dan beda dua proporsi. Pada jenis sampel rata-rata, rumusan umum interval kepercayaan (1-α) untuk menduga rata-rata μ, bila σ diketahui adalah sebagai berikut.

$$P\left[\bar{X} - Z_{\frac{\alpha}{2}} . \sigma_{\bar{x}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} . \sigma_{\bar{x}}\right] = 1-\alpha \quad \ldots\ldots (7.5)$$

$\bar{X}$: rata-rata distribusi sampel rata2

$Z_{\frac{\alpha}{2}}$: nilai dari tabel distribusi normal kumulatif

$\sigma_{\bar{x}}$: simpangan baku distribusi sampel rata2

α : koefisien kepercayaan

Simpangan baku untuk populasi terbatas $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ sedangkan populasi tak terbatas adalah dengan menjadikan faktor koreksi sama dengan satu atau $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$ dimana nilai $\sigma_{\bar{x}}$ simpangan baku dari distribusi sampel rata-rata ; Bila $\sigma_{\bar{x}}$ tidak diketahui, maka dapat digunakan penduga dari σ yaitu S.

Selanjutnya rumus untuk menentukan jumlah sampel dengan dugaan meleset, bila $\bar{X}$ merupakan penduga untuk rata-rata populasi, maka dapat dipercaya bahwa (1-α)x100% bahwa kesalahanya akan lebih dari suatu besaran tertentu e yang ditetapkan sebelumnya dengan syarat, yaitu

$$n = \left(\frac{Z_{\frac{\alpha}{2}} . \sigma}{e}\right)^2 \quad \ldots\ldots (7.6)$$

Rumus umum interval kepercayaan (1-α) pada pendugaan parameter proporsi untuk menduga proporsi P adalah :

$$P\left[\hat{p} - Z_{\frac{\alpha}{2}} . \sigma_{\hat{p}} < P < \hat{p} + Z_{\frac{\alpha}{2}} . \sigma_{\hat{p}}\right] = 1-\alpha \quad \ldots\ldots (7.7)$$

dimana $P = \frac{X}{N}$ dan $\hat{p} = p = \frac{x}{n}$ dengan nilai simpangan baku untuk populasi terbatas adalah $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} . \sqrt{\frac{N-n}{N-1}}$ serta simpangan baku tidak terbatas $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$.

Rumus umum interval kepercayaan (1-α) pada pendugaan parameter beda dua rata-rata untuk menduga beda dua rata-rata μ_1 - μ_2 adalah :

$$P\left[(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} . \sigma_{\bar{x}1-\bar{x}2} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} . \sigma_{\bar{x}1-\bar{x}2}\right] = 1-\alpha \quad \ldots\ldots (7.8)$$

Dimana nilai simpangan baku untuk populasi terbatas dihitung dengan rumus

$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} . \sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{(N_1+N_2)-1}}$ dan untuk populasi tak terbatas $\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$

Rumus umum interval kepercayaan (1-α) pada pendugaan parameter beda dua proporsi untuk menduga beda dua proporsi p_1 - p_2 adalah :

$$P\left[(\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}}.\sigma_{\hat{p}_1-\hat{p}_2} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}}.\sigma_{\hat{p}_1-\hat{p}_2}\right] = 1-\alpha \text{ ……… (7.9)}$$

Dimana simpangan baku pada distribui sampel beda dua rata-rata bila populasi terbatas adalah $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.\sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{(N_1+N_2)-1}}$ dan simpangan baku bila populasi tak terbatas dihitung dengan rumusan $\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$.

Contoh soal 1. Diketahui bahwa dari populasi para pegawai suatu perusahaan diambil sampel sebanyak 100 orang dan dicatat gaji tahunan masing-masing. Rata-rata dan simpangan baku gaji mereka adalah $\overline{X}$ = Rp 30.000.000,- dan S = Rp 6.000.000,-. Jika nilai interval kepercayaan untuk menduga sebesar 95%. Berapa sesungguhnya rata-rata gaji para pegawai di perusahaan tersebut?
Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa kita percaya 95% bahwa rata-rata gaji tahunan yang sesungguhnya dari para pegawai di perusahaan itu berkisar antara Rp 28.824.000,-sampai dengan Rp 31.176.000,-

Contoh soal 2. Pada suatu sampel acak ukuran n = 500 orang disuatu kota ditemukan bahwa 340 orang diantaranya suka menonton TV untuk acara dunia dalam berita. Hitunglah interval kepercayaan 95% untuk menduga berapa proporsi sesungguhnya penduduk di kota itu yang suka menonton TV untuk acara dunia dalam berita tersebut.

Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa kita percaya 95% bahwa proporsi penduduk dikota tersebut yang sesungguhnya suka menonton TV untuk acara dunia dalam berita antara 64,1% sampai dengan 71,9%.

Contoh soal 3. Pada ujian kalkulus diberikan dua kelompok mahasiswa, yaitu mahasiswa perempuan sebanyak 75 orang dan mahasiswa laki-laki sebanyak 50 orang. Kelompok mahasiswa perempuan memperoleeh nilai rata2 82 dengan simpangan baku 8, sedangkan kelompok mahasiswa laki-laki memperoleh rata-rata 76 dan simpangan baku 6. Bila µ1 menyatakan rata2 nilai ujian kelompok mahasiswa perempuan dan µ2 menyatakan rata2 nilai ujian kelompok laki2, buatlah interval kepercayaan 96% untuk menduga berapa sesungguhnya beda rata-rata dua kelompok mahasiswa tersebut

Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa kita percaya 96% bahwa beda sesungguhnya nilai rata-rata ujian kalkulus dua kelompok mahasiswa terletak antara 3,429 sampai dengan 8,571

Contoh soal 4. Suatu Survei diadakan terhadap pengunjung PRJ. Untuk itu diambil dua kelompok sampel. Sampel pertama adalah pengunjung perempuan sebanyak 500 orang an ketika mereka ditanya sebanyak 325 orang mengatakan puas dengan pameran di PRJ. Sedangkan sampel kedua terdiri atas pengunjung pria sebanyak 700 orang, dan 400 orang diantara-nya menyatakan puas dengan pameran di PRJ. Buatlah interval kepercayaan 95% untuk menduga berapa sesungguhnya beda dua populasi pengunjung yang puas dengan pameran di PRJ.

Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa kita percaya 95% bahwa beda proporsi sesungguhnya pengunjung yang puas dengan pameran di PRJ adalah antara 2% sd 14%.

Contoh soal 5. Suatu sampel acak sebesar 500 keluarga konsumen golongan masyarkat A dan 600 keluarga konsumen golongan masyarakat B telah dipilih untuk suatu penelitian. Dari golongan A ternyata200 menyatakan senang terhadap suatu hasil produksi tertentu, sedangkan dari B, 150 keluarga menyatakan senang terhadap barang hasil produksi tersebut. Tentukan 95% selang kepercayaan untuk selisih proporsi sesungguhnya kedua golongan konsumen tersebut!

n_a = 500 $\quad n_b$ = 600

P_a = 200/500 $\quad P_b$ = 150/600

= 0.4 $\quad$ = 0.125

95% selang kepercayaan → $100\% - \alpha = 95\%$

$\alpha = 5\%$

½ $\alpha = 2,5\%$

$Z_{0,025}$ = - 1,96

(1- α) 100% selang kepercayaan untuk $P_a - P_b$ adalah

$$(P_a - P_b) \quad \pm \quad Z_{\frac{1}{2}a}\sqrt{\frac{P_a(1-P_a)}{n_a}} + \sqrt{\frac{P_b(1-P_b)}{n_b}}$$

$$(0.4 - 0.25) \quad \pm \quad 1.96\sqrt{\frac{0.4(1-0.4)}{500}} + \sqrt{\frac{0.25(1-0.25)}{600}}$$

$$(P_a - P_b) \quad \pm \quad Z_{\frac{1}{2}a}\sqrt{\frac{P_a(1-P_a)}{n_a}} + \sqrt{\frac{P_b(1-P_b)}{n_b}}$$

$$(0.4 - 0.25) \quad \pm \quad 1.96\sqrt{\frac{0.4(1-0.4)}{500}} + \sqrt{\frac{0.25(1-0.25)}{600}}$$

$$0.15 \quad \pm \quad 1.96\sqrt{0.00048 + 0.00031}$$

$$0.15 \quad \pm \quad 0.0551$$

Jadi 95% selang kepercayaan untuk selisih proporsi sesungguhnya kedua golongan konsumen tersebut adalah <0,0949 $< p_a - p_b <$ 0,2051.

F. Pendugaan Parameter Populasi dengan Sampel Kecil (n < 30)

Suatu sampel dikatakan sampel kecil jika jumlah sampelnya kurang dari tiga puluh (n <30). Pada kondisi ini maka tabel yang digunakan untuk melihat nilai distribusinya adalah jenis tabel distribusi student atau biasa disebut sebagai tabel distribusi t. Variabel random yang digunakan juga bukan nilai Z melainkan variabel random t. Oleh karena itu rumusan umum pada pendugaan dengan sampel kecil hanya berbeda pada nilai variabel random beserta tabel distribusi yang digunakan. Pembahasan pertamauntuk samppe kecil dimulai dari pendugaan parameter distribusi sampel rata-rata. Interval kepercayaan (1-α) untuk menduga rata-rata μ dengan sampel kecil (n<30) yang diambil dari suatu populasi dimana variansi σ^2 tidak diketahui adalah sebagai berikut.

$$\mathrm{P}\left[\overline{\mathrm{X}} - \mathrm{t}_{(\frac{\propto}{2},v)} . \sigma_{\overline{\mathrm{x}}} < \mu < \overline{\mathrm{X}} + \mathrm{t}_{(\frac{\propto}{2},v)} . \sigma_{\overline{\mathrm{x}}}\right] = 1 - \propto \quad \text{.................................... (7.10)}$$

$\mathrm{v} = \mathrm{n} - 1$ dimana v aalah derajat kebebasan distribusi t .... (7.11)

$\mathrm{t}_{(\frac{\propto}{2},v)}$ nilai t diperoleh dari table distribusi student (tabel distribusi t)

Nilai simpangan baku σ^2 untuk populasi terbatas $\sigma_{\overline{\mathrm{X}}} = \frac{\sigma_{\mathrm{x}}}{\sqrt{\mathrm{n}}}\sqrt{\frac{\mathrm{N}-\mathrm{n}}{\mathrm{N}-1}}$ atau $\sigma_{\overline{\mathrm{X}}} = \frac{\mathrm{S}}{\sqrt{\mathrm{n}}}\sqrt{\frac{\mathrm{N}-\mathrm{n}}{\mathrm{N}-1}}$

sedangkan untuk populasi tak terbatas adalah $\sigma_{\overline{\mathrm{X}}} = \frac{\sigma_{\mathrm{x}}}{\sqrt{\mathrm{n}}} = \frac{\mathrm{S}}{\sqrt{\mathrm{n}}}$

Contoh soal 6. Suatu sampel acak sebanyak 15 mahasiswa diambil dari populasi mahasiswa di suatu universitas. Ke-15 mahasiswa diberikan tes bahasa inggris dan diperoleh nilai rata-rata mereka adalah 75 dengan simpangan baku 8. Buatlah interval kepercayaan 95% untuk menduga kemampuan bahasa inggris semua mahasiswa di universitas tersebut.

Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa dengan interval kepercayaan 95% untuk pendugaan kemampuan rata-rata bahasa inggris semua mahasiswa di universitas itu P(70,6<μ<79,4)=0,95. Artinya kita percaya 95% bahwa kemampuan rata-rata bahasa inggris semua mahasiswa di universitas itu terletak di antara 70,6 sampai dengan 79,4.

Pada pendugaan parameter beda dua rata-rata (μ1 - μ2), misalkan diketahui dua populasi masing-masing mempunyai rata-rata μ_1 dan μ_2 , dan distribusinya mendekati normal. Misalkan variansi dua populasi itu sama yaitu $\sigma_1^2 = \sigma_2^2 = \sigma^2$ tetapi tidak diketahui berapa besarnya, maka Interval kepercayaan (1-α) untuk menduga dua rata-rata (μ1 - μ2), dengan sampel kecil (n<30) yang diambil dari suatu populasi dimana variansi sama namun tidak diketahui nilainya adalah sebagai berikut.

$$P\left[(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\propto}{2},v)}.\sigma_{\overline{x}1-\overline{x}2} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + t_{(\frac{\propto}{2},v)}.\sigma_{\overline{x}1-\overline{x}2}\right] = 1-\propto \ldots (7.12)$$

Dimana nilai simpangan baku $\sigma_{\overline{x}_1-\overline{x}_2}$ dan simapngan baku gabungan S_p dihitung dengan rumus

Simpangan baku $\sigma_{\overline{x}_1-\overline{x}_2} = S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ ………………………………………….... (7.13)

Simpangan baku gabungan $S_p = \sqrt{\frac{(n_1-1)S_1{}^2+(n_2-1)S_2{}^2}{n_1+n_1-2}}$ …………………………….. (7.14)

Derajat kebebasan $v = n_1 + n_2 - 2$ ……………………………………….……... (7.15)

Bila variansi dua populasi itu tidak sama besarnya yaitu $\sigma_1^2 \neq \sigma_2^2$ dan kedua variansi tidak diketahui nilainya, maka interval kepercayaan (1-α) untuk beda dua rata-rata (μ1 - μ2) dari dua populsai tersebut secara rumus umum akan sama dengan persamaan 7.12 yaitu sebagai berikut $P\left[(\overline{X}_1 - \overline{X}_2) - t_{(\frac{\propto}{2},v)}.\sigma_{\overline{x}1-\overline{x}2} < \mu_1 - \mu_2 < (\overline{X}_1 - \overline{X}_2) + t_{(\frac{\propto}{2},v)}.\sigma_{\overline{x}1-\overline{x}2}\right] = 1-\propto$

Akan tetapi, nilai dari simpangan baku dan derajat kebebasan untuk kondisi variansi dua populasi yang tidak sama besarnya adalah sebagai berikut.

Simpangan baku $\sigma_{\overline{x}_1-\overline{x}_2} = \sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}}$ …………………………………..………….. (7.16)

Derajat kebebasan $v = \dfrac{\left(\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}\right)^2}{\left\{\frac{\left(S_1{}^2/n_1\right)^2}{n_1-1}+\frac{\left(S_2{}^2/n_2\right)^2}{n_2-1}\right\}}$ ……………………………………..…….. (7.17)

Contoh soal 7. Seorang dosen ingin membandingkan hasil belajar mahasiswa dalam mata kuliah kalkulus di suatu universitas berdasarkan dua metode mengajar yang dipakai. Untuk itu dibuat dua kelas, yaitu kelas A dan kelas B. Di kelas A terdiri atas 12 mahasiswa yang diajar dengan metode biasa. Di kelas B terdiri atas 10 mahasiswa yang diajar dengan metode baru. Setelah diadakan tes akhir semester untuk mata kuliah kalkulus, ternyata mahasiswa di kelas A memperoleh nilai rata-rata 85 dengan simpangan baku 4, sedangkan mahasiswa di kelas B memperoleh nilai rata-rata 81 dengan simpangan baku 5. Buatlah interval kepercayaan 80% untuk menduga beda kemampuan mahasiswa dalam mata kuliah kalkulus yang diajar dengan dua metode tersebut di universitas tersebut. Diasumsikan bahwa dua populasi mempunyai distribusi mendekati normal dan variansinya sama!

Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa interval kepercayaan untuk perbedaan kemampuan mahasiswa dalam mata kuliah kalkulus yang diajar dengan dua mettode di universitas itu adalah P(1,456<μ1-μ2<6,544)=0,80.

Contoh soal 8. Catatan selama 10 tahun terakhir menunjukkan bahwa rata-rata curah hujan di bogor selama bulan januari adalah 4,93 cm dengan simpangan baku 1,14 cm. Sedangkan didaerah Tangerang pada bulan yang sama rata-rata curah hujannya adalah 2,64 cm dengan simpangan baku 0,66 cm. Buatlah interval kepercayaan 90% untuk beda rata-rata curah hujan yang sebenarnya selama bulan januari di dua daerah tersebut. Diasumsikan bahwa data berasal dari dua populasi normal dengan variansi 1 dan variansi 2 yang berbeda dan kedua nilainya tidak diketahui.

Berdasarkan hasil perhitungan dengan rumus yang ada diperoleh hasil bahwa interval kepercayaan untuk perbedaan kemampuan mahasiswa dalam mata kuliah kalkulus yang diajar dengan dua mettode di universitas itu adalah $P(1,556<\mu1-\mu2<3,024)=90\%$.

Soal Latihan 1. Suatu sampel acak berukuran 10, mempunyai rata-rata 9,5 dan simpangan baku 3,24. Buatlah interval kepercayaan 90% untuk menduga rata-rata dari populasi tersebut.

Soal Latihan 2. Seorang pimpinan perusahaan ingin mengetahui perbedaan rata-rata gaji bulanan karyawan di perusahaan A dan perusahaan B. Untuk itu diambil sampel secara acak masing2 sebanyak 9 orang kemudian dilakukan wawancara satu persatu. Hasil wawancara menunjukkan bahwa gaji

BAB 9. Pengujian Hipotesis

A. Pengantar Pengujian Hipotesis

Secara pengertian hipotesis merupakan suatu asumsi atau anggapan yang bisa benar atau bisa salah mengenai sesuatu hal, dan dibuat untuk menjelaskan sesuatu hal tersebut sehingga memerlukan pengecekan lebih lanjut. Asumsi atau anggapan itu seringkali dipakai sebagai dasar dalam memutuskan atau menetapkan sesuatu dalam rangka menyusun perencanaan atau kepentingan lainnya baik dalam bidang ekonomi, bisnis, pendidikan, dan lain sebagainya. Contoh dari asumsi adalah penyusunan RAPBN dimana didalamnya terdapat informasi yang berkaitan dengan parameter-parameter dengan prediksi nilai tertentu, misalnya pertumbuhan ekonomi 4,5% per tahun, harga minyak mentah dunia 350ribu per barel, tingkat inflasi mencapai 8% per tahun, nilai tukar 13.000 / dolar atau penerimaan negara dari pajak 190 T. Sementara contoh nyata dalam dunia elektrronika yang dekat dengan kehidupan mahasiswa elektro misalkan, pengujian hipotesis terkait dengan kualitas mesin solder 1 lebih baik dari solder 2, metode baru dapat menghasilkan output yang lebih tinggi atau pemanfaatan komponen yang baru lebih aman dan memberikan hasil yang lebih akurat.

Bila hipotesis ini dikaitkan dengan parameter populasi, maka hipotesis ini disebut hipotesis statistik. Hipotesis statistik adalah suatu asumsi atau anggapan atau pernyataan yang mungkin benar atau mungkin salah mengenai parameter satu populasi atau lebih. Pengujian hipotesis merupakan langkah-langkah atau prosedur yang dilakukan dengan tujuan untuk memutuskan apakah menolak atau menerima parameter populasi. Pada pengujian hipotesis yang dilakukan adalah menguji Parameter satu populasi yaitu Θ sama dengan nilai tertentu $\Theta 0$ ($\Theta = \Theta 0$) atau tidak ($\Theta \neq \Theta 0$). Jika terdapat 2 populasi masing-masing dengan parameter $\Theta 1$ dan $\Theta 2$, maka perlu di uji apakah $\Theta 1 = \Theta 2$ atau tidak ($\Theta 1 \neq \Theta 2$). Contoh pengujian hipotesis yaitu pada percobaan pelemparan sebuah uang logam dengan populasi tak terbatas, uji hipotesis bahwa uang logam setimbang, dimana Probabilitas angka = Probabilitas gambar $P(A)=P(G)=1/2$ atau $\Theta = ½$. Percobaan tersebut diuji dengan sampel sejumlah 100 kali lemparan. Jika hasilnya muncul angka 47 kali, maka P angka = 0,47 sehingga kesimpulan uang logam setimbang, kita menerima P angka = P gambar. Sedangkan jika hasil akhir muncul angka 45 kali, maka P angka = 0,45 sehingga kesimpulan uang logam setimbang, kita menerima P angka = P gambar. Selanjutnya jika pelemparan menyebabkan munculnya angka sebanyak 25 kali, maka P angka = 0,25 dan peneliti tidak berani menyimpulkan bahwa uang logam tersebut setimbang karena tidak cukup alasan untuk menerima P angka = P gambar.

B. Rumusan Pengujian Hipotesis

Pada suatu hipotesis yang teah dibuat maka terdapat dua kemungkinan tindak lanjut yaitu pertama, menolak hipotesis atau menyimpulkan bahwa hipotesis tidak benar dan yang kedua menerima hipotesis dimana tidak cukup informasi dari sampel bahwa hipotesis harus kita tolak, artinya walaupun hipotesis itu kita terima tidak berarti hipotesis itu benar. Pembuatan rumusan pengujian hipotesis hendaknya membuat pernyataan hipotesis yang diharapkan akan ditolak disebut hipotesis nol (Ho). Penolakan hipotesis nol akan menjurus pada penerimaan hipotesis alternatif / tandingan ditulis H1.

Contoh kegiatan pengujian meliputi hipotesis bahwa jenis obat baru lebih efektif menurunkan berat badan. Rumusan hipotesis yang dihasilkan yaitu

- Hipotesis nol, Ho : obat baru=obat lama
- H Alternatif H1 : obat baru lebih baik dari obat lama

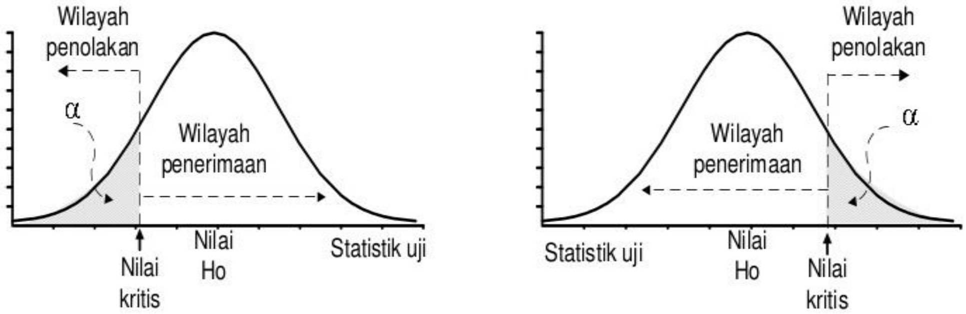
Contoh lain, misalkan dokter menyatakan lebih dari 60% pasien menderita sakit paru2 karena merokok. Maka hipotesisnya :

- Hipotesis nol Ho : p : 60% = 0,6
- H Alternatif H1 : p $\neq$ 0,6

Dalam membuat hipotesis, peneliti tidak lepas dari suatu kesalahan. Kesalahan-kesalahan tersebut dalam statistic diklasifikasikan menjadi dua jenis yang meliputi kesalahan jenis I dan jenis II. Apapun yang diperoleh dari sampel merupakan perkiraan sebagai dasar keputusan menolak atau menerima hipotesis nol. Menolak atau menerima mengandung ketidakpastian sehingga menimbulkan resiko oleh pembuat keputusan. Kesalahan jenis I muncul akibat menolak Ho, padahal Ho benar dan harusnya diterima. Probabilitas melakukan kesalahan jenis I disebut taraf nyata atau keberartian ditulis α. A=P(kesalahan jenis I)=P (menolak Ho/Ho benar). Kesalahan Jenis II muncul akibat menerima Ho, padahal Ho salah dan harusnya ditolak. Probabilitas melakukan kesalahan jenis II disebut β. β= P(kesalahan jenis II) = P(menerima Ho/Ho salah). α dan β terdapat hubungan yaitu memperkecil α maka nilai β membesar dan sebaliknya. solusi sampel diperbesar.

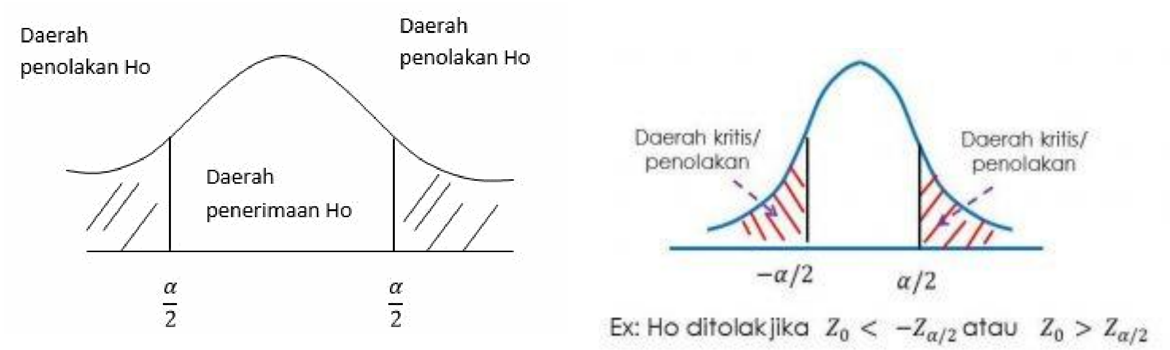
C. Uji Satu Arah dan Uji Dua Arah

Pada uji satu arah, bila hipotesis nol H_0: $\theta = \theta_0$ dilawan dengan Hipotesis alternatif H_1: $\theta > \theta_0$ atau H_1: $\theta < \theta_0$. Uji satu arah ditandai dengan adanya satu daerah penolakan hipotesis nol yang bergantung pada nilai kritis tertentu. Nilai kritis diperoleh dari tabel untuk nilai α yang telah dipilih sebelumnya. Sebagai catatan sederhana bahwa pada pengujian satu arah baik kiri ataupun kanan terjadi jika ada keterangan lebih kecil atau lebih besar dan atau kurang dari atau lebih dari. Nilai Kritis merupakan nilai yang membatasi darah penolakan dan penerimaan H0 atau Zα sebagaimana tampak pada gambar di bawah ini.



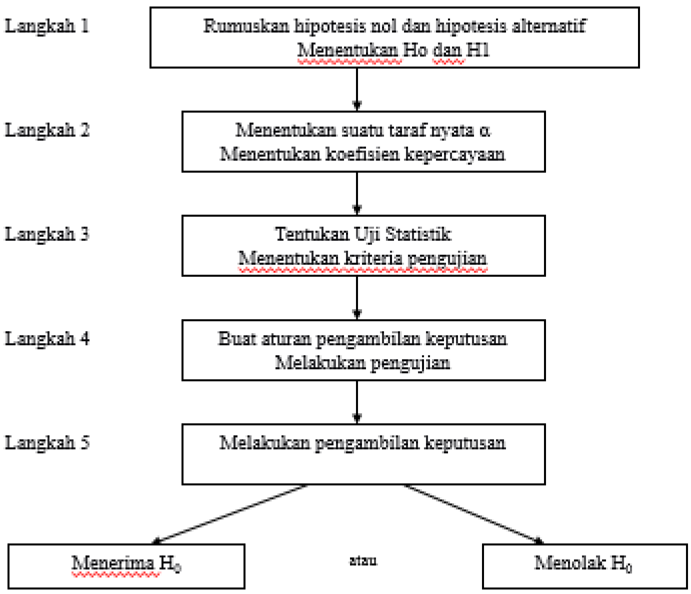
Gambar 9.1 Daerah Penerimaan dan Penolakan Ho pada uji satu arah

Pada uji dua arah bila hipotesis nol H_0: $\theta = \theta_0$ dilawan dengan hipotesis alternatif H_1 : $\theta \neq \theta_0$. Uji dua arah ditandai dengan adanya dua daerah penolakan hipotesis nol yang juga bergantung pada nilai kritis tertentu. Nilai kritis ini diperoleh dari tabel untuk nilai $\alpha/2$ yang telah dipilih sebelumnya. Catatan sederhana untuk menentukan jenis pengujian dua arah yaitu jika tidak ada keterangan kurang dari, lebih dari, lebih besar dan atau lebih kecil maka pengujian dilakukan dua arah



Gambar 9.2 Contoh beberapa jenis daerah penerimaan dan penolakan H0 pada uji dua arah

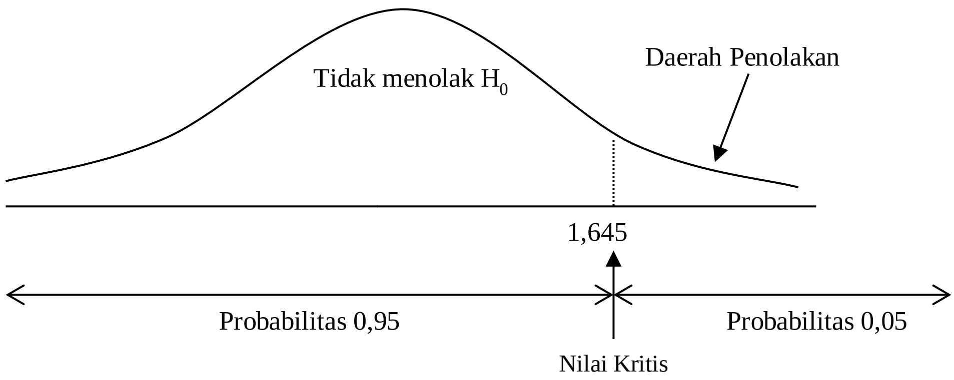
Proses dalam menguji suatu hipotesis terbagi atas 5 langkah, sebagaiman tampak pada diagram pada gambar 9.3. Langkah tersebut meliputi perumusan hipotesis, penentuan taraf nyata dan uji statistik, penentuan aturan dalam pengambilan keputusan dengan melakukan pengujian dan terakhir adalah mengambil keputusan. Detail penjelasan pada setiap langkah adalah sebagai berikut.



Gambar 9.3 Diagram proses pengujian hipotesis

Langkah pertama adalah merumuskan hipotesis yang akan diuji yang meliputi hipotesis nol dan hipotesis alternatif. Hipotesis nol disebut H0 (dibaca H nol) sedangkan hipotesis alternatif menggambarkan apa yang akan disimpulkan jika menolak hipotesis nol. Hipotesis alternatif ditulis H1 (dibaca H satu). Langkah kedua yaitu menentukan taraf nyata. Taraf nyata diberi tanda α (*alpha*), disebut juga tingkat resiko karena menggambarkan resiko yang harus dipikul bila menolak hipotesis nol padahal hipotesis nol sebetulnya benar. Tidak ada satu taraf nyata yang diterapkan untuk semua penelitian yang menyangkut penarikan sampel. Kita harus mengambil suatu keputusan untuk memakai taraf 0,05 (disebut taraf 5 persen), taraf 0,01, atau taraf yang lain antara 0 dan 1. Pada umumnya pada proyek penelitian menggunakan taraf 0,05, sedangkan untuk pengendalian mutu dipilih 0,01, dan untuk pengumpulan jajak pendapat ilmu-ilmu sosial dipakai 0,10. Selanjutnya langkah ketiga yaitu melakukan uji statistik. Proses uji statistik merupakan penentuan suatu nilai berdasarkan informasi dari sampel, dan akan digunakan untuk menentukan apakah akan menerima atau menolak hipotesis. Ada bermacam-macam uji statistik, di sini kita akan menggunakan uji statistik seperti z, *student-t.* Langkah ke empat dilakukan dengan menentukan suatu aturan pengambilan keputusan. Aturan pengambilan keputusan merupakan pernyataan mengenai kondisi di mana hipotesis nol ditolak dan kondisi di mana hipotesis nol tidak ditolak. Gambar 9.4 berikut menggambarkan daerah penolakan untuk suatu uji taraf nyata.

Distribusi Sampling bagi Statistik z



Gambar 9.4 contoh daerah penolakan untuk suatu uji taraf nyata

Perhatikan dalam gambar diatas, tampak bahwa daerah di mana hipotesis nol tidak ditolak mencakup daerah di sebelah kiri 1,645. Daerah penolakan adalah di sebelah kanan dari 1,645. Diterapkan suatu uji satu arah dengan taraf nyata yang dipilih adalah 0,05. Nilai 1,645 memisahkan daerah-daerah dimana hipotesis nol ditolak dan di mana hipotesis nol tidak ditolak. Nilai 1,645 dinamakan nilai kritis.

Langkah terakhir dalam uji statistik atau langkah kelima adalah mengambil keputusan untuk menolak atau menerima / tidak menolak hipotesis nol. Keputusan menolak hipotesis nol karena nilai uji statistik terletak di daerah penolakan. Perlu juga diperhatikan bahwa keputusan untuk menolak atau tidak adalah keputusan yang diambil oleh peneliti yang sedang melakukan penelitian. Hasil ini merupakan rekomendasi berdasarkan bukti-bukti sampel yang dapat diberikan peneliti kepada manajer puncak sebagai pembuat keputusan, tetapi keputusan akhir biasanya tetap diambil oleh manajer puncak tersebut.

D. Pengujian Parameter Berbagai Jenis Sampel

Parameter yang digunakan pada penentuan uji hipotesis ini meliputi jenis sampel baik sampel besar dengan $n \geq 30$ maupun sampel kecil, jenis populasi terbatas atau tak terbatas. rata-rata sampel, simpangan baku, derajat kebebasan, serta nilai proporsi. Rincian rumusan pengujian parameter untuk sampel rata-rata, beda dua rata serta proporsi adalah sebagai berikut. Setiap rumusan dibedakan berdasarkan jenis sampel untuk hipotesis nol atau H0 serta jenis pengujian satu atau dua arah pada H1 untuk memudahkan dalam penentuan daerah kritis.

Pengujian parameter rata-rata:

H_0	Uji Statistik	H_1	Daerah Kritis
Sampel besar $\mu = \mu_0$	$Z_h = \dfrac{\bar{X} - \mu_0}{\sigma_{\bar{x}}}$,$\sigma_{x\ bar}$ diketahui dan berbeda nilai pada populasi yang terbatas dan tak terbatas	$\mu < \mu_0$ $\mu > \mu_0$ ----------- $\mu \neq \mu_0$	Pengujian satu arah $Z < -Z_\alpha$ atau $Z > Z_\alpha$ ------------------------ Pengujian dua arah $Z < -Z_{\alpha/2}$ dan $Z > Z_{\alpha/2}$
Sampel kecil $\mu = \mu_0$	$t_h = \dfrac{\bar{X} - \mu_0}{\sigma_{\bar{x}}}$ Dimana $v = n-1$ σ_x tidak diketahui; $\sigma_x \rightarrow S_x$ baik populasi terbatas dan tak terbatas	$\mu < \mu_0$ $\mu > \mu_0$ --------- $\mu \neq \mu_0$	Pengujian satu arah $t < -t_{\alpha,v}$ Atau $t > t_{\alpha,v}$ ------------------------ Pengujian dua arah $t < -t_{\alpha/2,v}$ & $t > t_{\alpha/2,v}$

Pengujian parameter beda dua rata-rata:

H_0	Uji Statistik	H_1	Daerah Kritis
Sampel besar $\mu_1 - \mu_2 = 0$	$Z_h = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}1-\bar{x}2}}$ σ_{x1} dan σ_{x2} diketahui serta berbeda untuk populasi terbatas dan tak terbatas	sisi kiri $\mu_1 - \mu_2 < 0$ sisi kanan $\mu_1 - \mu_2 > 0$ dua sisi $\mu_1 - \mu_2 \neq 0$	Pengujian satu arah $Z < -Z_\alpha$ atau $Z > Z_\alpha$ ------------------------ Pengujian dua arah $Z < -Z_{\alpha/2}$ dan $Z > Z_{\alpha/2}$
Sampel kecil $\mu_1 - \mu_2 = 0$	$t_h = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}1-\bar{x}2}}$ dngan ketentuan jika $\sigma_1 = \sigma_2 = \sigma$ tidak diketahui maka niai σ_{x1-x2} untuk populasi terbatas dan tak terbatas dan $v = n1 - n2 - 2$ dan jika $\sigma_1 \neq \sigma_2 \neq \sigma$ maka rumusan yg berbeda untuk σ & v	sisi kiri $\mu_1 - \mu_2 < 0$ sisi kanan $\mu_1 - \mu_2 > 0$ ----------- dua sisi $\mu_1 - \mu_2 \neq 0$	Pengujian satu arah $t < -t_{\alpha,v}$ Atau $t > t_{\alpha,v}$ ------------------------ Pengujian dua arah $t < -t_{\alpha/2,v}$ & $t > t_{\alpha/2,v}$

Cara lain untuk menguji hipotesis beda dua rata-rata adalah sebagai berikut

H_0	Uji Statistik	H_1	Daerah Kritis
$\mu_D = d_0$	$t_h = \dfrac{\bar{d} - d_0}{S_d / \sqrt{n}}$ $\bar{d} = \dfrac{\Sigma d}{n} = \dfrac{\Sigma(x_1 - x_2)}{n}$	Satu arah $\mu_D < d_0$ $\mu_D > d_0$ Dua arah $\mu_D \neq d_0$	Pengujian satu arah $t < -t_{\alpha,v}$ Atau $t > t_{\alpha,v}$ ------------------------ Pengujian dua arah $t < -t_{\alpha/2,v}$ & $t > t_{\alpha/2,v}$

Dimana $\bar{d}$ = rata-raa beda atau selisih nilai dua kelompok =
$d0$ = nilai beda yg dihipotesiskan = 0
S_d = simpangan baku nilai-nilai d
$v = n - 1$

Ringkasan pengujian parameter rata-rata dan dua rata-rata dengan sampel tak terbatas

H_0	Uji Statistik	H_1	Daerah Kritis
$\mu = \mu_0$	$Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$ σ diketahui	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$Z < -Z_{\alpha}$ $Z > Z_{\alpha}$ $Z < -Z_{\alpha/2}$ dan $Z > Z_{\alpha/2}$
$\mu = \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{S / \sqrt{n}}$ $v = n - 1$ σ Tidak diketahui	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$T < -t_{\alpha,v}$ $T > t_{\alpha,v}$ $T < -t_{\alpha/2,v}$ dan $T > t_{\alpha/2,v}$

Ringkasan pengujian parameter rata-rata dan dua rata-rata dengan sampel tak terbatas

H_0	Uji Statistik	H_1	Daerah Kritis
$\mu_1 = \mu_2$	$Z = \dfrac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}}$ σ_1 dan σ_2 diketahui	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$Z < -Z_{\alpha}$ $Z > Z_{\alpha}$ $Z < -Z_{\alpha/2}$ dan $Z > Z_{\alpha/2}$
$\mu_1 = \mu_2$ Dimana $\mu_1 - \mu_2 = d_0$	$T = \dfrac{(\overline{X}_1 - \overline{X}_2) - d_0}{S_p \sqrt{1/n_1 + 1/n_2}}$ $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ $v = n1 + n2 - 2$ $\sigma_1 = \sigma_2$ dan tidak diketahui	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$T < -t_{\alpha,v}$ $T > -t_{\alpha,v}$ $T < -t_{\alpha/2,v}$ dan $T > -t_{\alpha/2,v}$

Ringkasan pengujian parameter rata-rata dan dua rata-rata dengan sampel tak terbatas

H_0	**Uji Statistik**	**H_1**	**Daerah Kritis**
$\mu_1 = \mu_2$	$t_h = \dfrac{(\bar{X}_1 - \bar{X}_2) - d_0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$ $d_0 = \mu_1 - \mu_2$ $v = \dfrac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\dfrac{\left(S_1^2/n_1\right)^2}{n_1 - 1} + \dfrac{\left(S_2^2/n_2\right)^2}{n_2 - 1}}$ $\sigma_1 \neq \sigma_2$ dan tidak diketahui	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$T < -t_{\alpha,v}$ $T > t_{\alpha,v}$ $T < -t_{\alpha/2,v}$ dan $T > t_{\alpha/2,v}$

Ringkasan pengujian parameter rata-rata dan dua rata-rata dengan sampel tak terbatas

H_0	**Uji Statistik**	**H_1**	**Daerah Kritis**
$\mu_D = d_0$	$t_h = \dfrac{\bar{d} - d_0}{S_d/\sqrt{n}}$ Dimana $\bar{d}$ = rata-raa beda atau selisih nilai dua kelompok D0 = nilai beda yang dihipotesis-kan = 0 dan v = n – 1 Pengamatan yg dipasangkan	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$T < -t_{\alpha,v}$ $T > t_{\alpha,v}$ $T < -t_{\alpha/2,v}$ dan $T > t_{\alpha/2,v}$

Pengujian Parameter Proporsi

H_0	**Uji Statistik**	**H_1**	**Daerah Kritis**
sampel besar $p = p_0$	$Z_h = \dfrac{\hat{p} - p_0}{\sigma_{\hat{p}}}$	sisi kiri $p < p_0$ sisi kanan $p > p_0$ dua arah $p \neq p_0$	Pengujian satu arah $Z < -Z_\alpha$ atau $Z > Z_\alpha$ ------------------------------- Pengujian dua arah $Z < -Z_{\alpha/2}$ dan $Z > Z_{\alpha/2}$

$$untuk-populasi-tak-terbatas$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p_0(1-p_0)}{n}}$$

$$untuk-populasi-terbatas$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p_0(1-p_0)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Pengujian Parameter beda dua proporsi

H_0	**Uji Statistik**	**H_1**	**Daerah Kritis**
Sampel besar $p_1 = p_2$	$Z_h = \dfrac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sigma_{\hat{p}1-\hat{p}2}}$ Pada umumnya p1 dan p2 tidak diketahui maka nilai simpangan baku ditaksir dgn rumus berikut :	sisi kiri $p < p_0$ sisi kanan $p > p_0$ dua arah $p \neq p_0$	Pengujian satu arah $Z < -Z_\alpha$ atau $Z > Z_\alpha$ ------------------------------- Pengujian dua arah $Z < -Z_{\alpha/2}$ dan $Z > Z_{\alpha/2}$

$$bila-populasi-tak-terbatas$$

$$\sigma_{\hat{p}1-\hat{p}2} = \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)};$$

$$bila-populasi-terbatas$$

$$\sigma_{\hat{p}1-\hat{p}2} = \sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \cdot \sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{N_1+N_2-1}}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} \qquad \hat{q} = 1 - \hat{p}$$

$$p_1 = \frac{x_1}{n_1} \qquad p_2 = \frac{x_2}{n_2}$$

Contoh soal 1. Seorang pejabat perbankan yang bertanggung jawab tetang pemberian kredit, mempunyai anggapan bahwa rata-rata modal perusahaan nasional adalah sebesar 100 juta rupiah, dengan alternatif **lebih besar** dari itu. Untuk menguji anggapan tsb, dipilih sampel secara acak sebanyak 81 unit perusahaan nasional, yang ternyata rata-rata modalnya sebesar

105 juta rupiah dengan simpangan baku diketahui sebesar 18 juta rupiah. Jika nilai α = 1% ujilah anggapan tersebut.

Jawab

Diketahui $\mu 0$ = 100 juta ; n = 81 (sampel besar) ; X = 105 juta ; S = 18 juta ; α = 1% sehingga 1- α = 99% dan Populasi tak terbatas.

a) Menentukan H0 dan H1

   H0 : U = U0 (100 juta), artinya rata2 modal perusahaan sama dengan 100 juta

   H1 : U > U0 (100 juta), artinya rata2 modal perusahaan lebih besar dari 100juta

b) Menentukan derajat keyakinan

   Nilai koefisien kepercayaan α = 1% sehingga derajat kepercayaan (1- α)= 99%

c) Menentukan Kriteria pengujian

   Nilai n= 81>30 maka menggunakan nilai Z tabel

   Dari tabel luas kurva normal jika α = 0,01 maka Z➔ -2,33

   Pengujian satu arah/sisi sebelah kanan karena nilai pada H1 : $\mu > \mu_0$

   H0 diterima jika Zh ≤ -2,33 dan H0 ditolak jika Zh > -2,33

d) Pengujian

$$Z_h = \frac{(105.000.000 - 100.000.000)}{18.000.000 / \sqrt{81}} = 2,5$$

e) Kesimpulan

   Berdasarkan pengujian, karena Zh =2,5> -2,33 maka H0 ditolak. Berarti rata-rata modal perusahaan nasional bukan 100.000.000 melainkan 105.000.000

Latihan soal 1. Sebuah populasi berupa seluruh pelat baja yang diproduksi oleh sebuah perusahaan memiliki rata-rata panjang 80 cm dengan simpangan baku 7cm. Sesudah berselang 3 tahun, teknisi perusahaan meragukan hipotesis mengenai rata2 panjang pelat baja tsb. Guna meyakinkan keabsahan hipotesis itu, diambil suatu sampel acak sebanyak 100 unit pelat baja dari populasi diatas, dan diperoleh hasil hitungan 83 cm, an standard deviasi nya tetap. Apakah ada alasan untuk meraguan bahwa rata-rata panjang pelat baja yg dihasilkan perusahaan itu sama dengan 80 cm pada taraf signifikansi α=5%?

Cara menjawab soal jenis ini adalah dengan menentukan H0 dan H1, menentukan level of significance (derajat keyakinan), menentukan kriteria pengujian, melakukan pengujian dan menarik kesimpulan.

Latihan soal 2. Jika pada contoh soal 2 sebelumnya, ditambah data bahwa teknisi perusahaan telah menemukan metode baru yang dapat memperpanjang pelat baja **paling sedikit 2cm**, sedangkan simpangan baku nya tetap. Untuk menguji hipotesis tersebut, diambil sampe secara acak sebanyak 100 unit pelat baja dari populasi, dan diperoleh rata-rata panjang pelat baja 83 cm. dengan taraf α=5%, apakah ada alasan guna menganggap bahwa hasil pelat baja dengan metode baru tsb memang lebih panjang dr hasil yg diperoleh dgn metode lama ?

Latihan soal 3. Perusahaan MAGIC yg bergerak dibidang suku cadang komputer mikro, akan memperkenalkan produk terbarunya di pasaran. Untuk itu bagian pengendalian kualitas perusahaan mengambil sampel secara acak sebanyak 170 unit suku cadang dan ditemukan ada 16 yang cacat. Dari data tsb apakah beenar produksi yg ditemukan cacat **kurang dari** 10%. Gunakan taraf signifikansi 2%

Latihan soal 4. Asosiasi real estate sedang menyiapkan brosur yg mereka rasa mungkin menarik bagi calon pembeli rumah di daerah A dan B di suatu kota. Satu hal yg menarik adalah lama waktu si pembeli tinggal di rumah yg bersangkutan. Sebuah sampel yg terdiri dari 40 rumah di daerah A memperlihatkan bahwa rata2 kepemilikan adalah 7,6 tahun dengan simpangan baku 2,3 tahun. Sedangkan sampel yg terdiri dr 55 rumah di daerah B memperlihatkan bahwa rata2 lama waktu kepemilikan adalah 8,1 tahun dengan simpangan baku 2,9 tahun.

Latihan soal 5. Pada taraf signifikansi 5% apakah kita dapat menarik kesimpulan bahwa penduduk di daerah A memiliki rumah mereka dalam waktu **lebih singkat** dari penduduk di daerah B?

BAB 10. Regresi dan Korelasi

A. Pengantar Regresi dan Korelasi

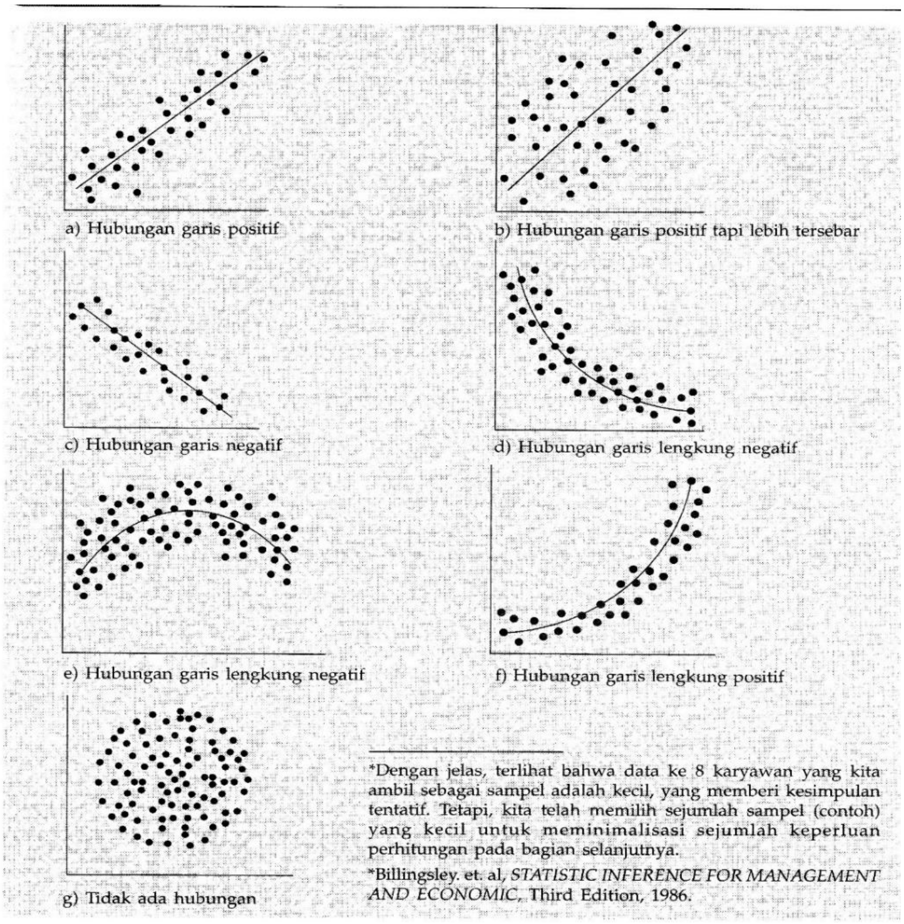
Dalam kehidupan sehari-hari sering ditemukan masalah/kejadian yang saling berkaitan satu sama lain. Manusia memerlukan analisis hubungan antara kejadian tersebut. Dalam bab ini akan membahas dua kejadian yg saling berhubungan, khususnya 2 kejadian yg dapat diukur secara matematis. Dua hal yang perlu dianalisis yaitu hubungan fungsional (persamaan matematis) dan hubungan kekuatan.

Analisis regresi merupakan suatu analissis yang digunakan untuk mempelajari dan mengukur hubungan fungsional (statistik atau persamaan matematis) yang terjadi antara dua varibel atau lebih variabel. Variabel tersebut adalah variabel X (variabel independent / variabel yang mempengaruhi / variabel yang diketahui), dan variabel Y (variabel dependent / variabel yang dipengaruhi/ variabel yang tidak diketahui). Sementara itu, analisis korelasi merupakan suatu analissis yang bertujuan untuk mengukur "seberapa kuat" atau "derajat kedekatan", suatu relasi yang terjadi antar variabel.

Pada dasarnya hubungan antar 2 variabel dapat dibedakan atas menjadi 3 yaitu hubungan searah/positif, hubungan bersifat kebalikan/negative serta tidak ada hubungan. Hubungan yang searah diartikan apabila perubahan variabel x (independent) akan mempengaruhi variabel y (dependent) yang searah atau jika variabel x bertambah, maka variabel y bertambah pula, dan sebaliknya. Contohnya adalah hubungan antara pengeluaran iklan (x) dan jumlah penjualan (y) atau hubungan antara penghasilan (X) dan pengeluaran konsumsi (Y). Selanjutnya, dua variabel dikatakan mempunyai hubungan yang bersifat kebalikan atau negatif, apabila perubahan variabel independent (x) akan mempengaruhi variabel dependent (Y) pada arah yang berlawanan. Artinya apabila variabel x bertambah, maka variabel y berkurang atau sebaliknya, jika variabel x berkurang maka variabel y bertambah. Contohnya yaitu hubungan antara usia kendaraan (X) dengan tingkat harga (Y) atau hubungan antara harga barang (x) dengan jumlah yang diminta (Y). Terakhir, dua variabel dikatakan tidak punya hubungan apabila perubahan pada variabel independent (x) **<u>tidak</u>** mempengaruhi perubahan pada variabel dependent (y). Sebagai contoh adalah hubungan antara konsumsi pangan (x) dengan tingginya gedung (y).

B. Regresi Linier Sederhana

Pada regresi linier sederhan, penggambaran garis regresi dapat dilakukan salah satunya adalah metode diagram berserak (*the scatter diagram*). Setelah ditetapkan bahwa terdapat hubungan logis di antara variabel, maka untuk mendukung analisis lebih jauh, tahap selanjutnya adalah membuat diagram pencar, yang menunjukkan titik-titik tertentu. Setiap titik memperlihatkan suatu hasil yang kita nilai sebagai varibel bebas maupun variabel tak bebas. Diagram pencar ini memiliki 2 manfaat, yaitu dapat membantu menunjukkan apakah terdapat hubungan yang bermanfaat antara dua variabel, serta membantu menetapkan tipe persamaan yang menunjukkan hubungan antara kedua variabel tersebut.Contoh gambaran umum dari diagram berserak atau diagram pencar dapat dilihat seabagaimana berikut.



Gambar 10.1 Diagram berserak (*The scatter diagram*)

Persamaan yang digunakan untuk mendapatkan garis regresi pada data diagram pencar disebut persamaan regresi. Untuk menempatkan garis regresi pada data yang diperoleh maka digunakan metode kuadrat terkecil, sehingga bentuk persamaan regresi adalah sebagai berikut:

Y’ = a + b X ………………………………………………………………………………. (10.1)

Dimana:

Y’: nilai estimasi/taksiran untuk variabel terikat (tak bebas Y)

a: titik potong garis regresi pd sumbu y (nilai estimate Y’ bila x=0)

b: gradien garis regresi (perub nilai estimasi Y’ per satuan perubahan nilai x) atau koefidien arah dari garis regresi
X: nilai variabel bebas

Kesamaan diantara garis regresi dan garis trend tidak dapat berakhir dengan persamaan garis lurus. Dalam hal ini dicari persamaan regresi yg *paling baik* untuk mewakili sebaran titik data tersebut. Suatu kriteria bahwa persamaan regresi yg paling baik adalah regresi yg mempunyai total kuadrat selisih yg paling minimum. Garis regresi) memiliki dua sifat matematis berikut Σ(Y – Y’) = 0 dan Σ(Y – Y’)^2 = nilai terkecil atau terendah. Dengan perkataan lain, garis regresi akan ditempatkan pada data dalam diagram sedemikian rupa sehingga penyimpangan (perbedaan) positif titik-titik terhadap titik-titik pencar di atas garis akan mengimbangi penyimpangan negatif titik-titik pencar yang terletak di bawah garis, sehingga hasil penyimpangan keseluruhan titik-titik terhadap garis lurus adalah nol. Untuk memperoleh total kuadrat error paling minimum, dipakailah metode kuadrat minimum. Dari persamaan regresi linear sebelumnya akan memiliki total kuadrat error minimum bila koefisien regresi a dan b dihitung dengan rumus berikut.

$$a = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n \sum X^2 - (\sum X)^2} \text{ dan } b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (10.2)$$

atau

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \quad \text{dan } a = \frac{\sum Y}{n} - b\left(\frac{\sum X}{n}\right) \text{ atau } a = \bar{Y} - b\bar{X} \quad \ldots\ldots\ldots\ldots\ldots (10.3)$$

Koefisien Regresi adalah lereng garis regresi (nilai b). Nilai b positif , menunjukkan hubungan antara variabel x dan y searah atau hubungannya positif. Nilai b negatif, menunjukkan hubungan antara variabel x dan y berlawanan arah atau hubungannya negatif. Besar kecilnya perubahan variabel x terhadap variabel y ditentukan besar kecilnya koefisien regresi. Dalam praktknya ada istilah yang disebut sebagai “kesalahan baku dari penaksiran” atau *standard error of estimation* oleh Y’=a+bX yaitu seabagai berikut.

$$S_{y'.x} = \sqrt{\frac{\sum (Y - Y')^2}{n}} \quad \text{dijabarkan menjadi } S_{y'.x} = \sqrt{\frac{\sum Y^2 - a \sum Y - b \sum XY}{n}} \quad \ldots\ldots\ldots\ldots (10.4)$$

Perbedaan mendasar antara regresi dan korelasi adalah bahwa regresi menunjukkan hubungan antara variabel satu dengan variabel lainnya. Sifat hubungan dapat dijelaskan: variabel yang satu sebagai penyebab, variabel yang lain sebagai akibat. Sedangkan korelasi tidak menunjukkan hubungan sebab akibat, akan tetapi menunjukkan hubungan antara variabel satu dengan yang lain.

C. Korelasi Linier Sederhana

Secara maematis bahwasanya koefisien korelasi (r) diartikan sebagai ukuran hubungan linier peubah X dan Y dimana nilai r berkisar antara (+1) sampai (-1). Nilai r yang (+) ditandai oleh nilai b yang (+), demikian pula nilai r yang (-) ditandai oleh nilai b yang (-). Jika nilai r mendekati +1 atau r mendekati -1 maka X dan Y memiliki korelasi linier yang tinggi. Jika nilai r = +1 atau r = -1 maka X dan Y memiliki korelasi linier sempurna. Kemudian, jika nilai r = 0 maka X dan Y tidak memiliki relasi (hubungan) linier (dalam kasus r mendekati 0, anda dapat melanjutkan analisis ke regresi eksponensial).

Koefisien determinasi sampel disimbulkan dengan R yang nilainya merupakan kuadrae dari koefisien korelasi sehingga R= r². Ukuran proporsi keragaman total nilai peubah Y yang dapat dijelaskan oleh nilai peubah X melalui hubungan linier. Penetapan dan interpretasi koefisien krelasi dan koefisien determinasi adalah esbagai berikut.

$$r = \frac{n\sum XY - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(10.5)}$$

$$R = r^2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(10.6)}$$

Simbol r adalah Koefisien Korelasi dan koefisien determinasi sampel adalah R = r²

Contoh soal regresi. Berikut adalah data Biaya Promosi dan Volume Penjualan PT BIMOIL perusahaan Minyak Goreng. Buatlah persmaan regresi linear sederhana dengan minimum kuadrat terkecil. Perkirakan Volume penjualan jika, dikeluarkan biaya promosi Rp. 10 juta.

Tahun	X Biaya Promosi (Juta Rupiah)	Y Volume Penjualan (Ratusan Juta Liter)	XY	X²	Y²
1992	2	5	10	4	25
1993	4	6	24	16	36
1994	5	8	40	25	64
1995	7	10	70	49	100
1996	8	11	88	64	121
Σ	Σx = 26	Σy = 40	Σxy = 232	Σx² =158	Σy² = 346

n = 5

Jawab. Bentuk umum persaman regresi linier sederhana: Y = a + bX

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b = \frac{(5\times 232)-(26\times 40)}{(5\times 158)-(26^2)} = \frac{1160-1040}{790-676} = \frac{120}{114} = 1.05263...$$

$$b = 1{,}053$$

$$a = \frac{\sum_{i=1}^{n} y_i}{n} - b\frac{\sum_{i=1}^{n} x_i}{n}$$

$$a = \frac{40}{5} - \left(1.05263...\times\frac{26}{5}\right) = 8-(1.05263...\times 5.2)$$

$$a = 8 - 5.4736... = 2.5263.... => 2.53$$

Sehingga Y = a + b X $\rightarrow$ Y = 2,530 + 1,053X

Y = 2,530 + 1,053 X

Jika X = 10

Y = 2,53 + 1,053 (10)

Y= 2,53 + 10,53 = 13,06 (ratusan juta liter)

Maka volume penjualan = 13.06 x 100 000 000 liter

Contoh Korelasi (Lihat soal regresi). Setelah mendapatkan persamaan Regresi Y= 2.52 + 1.053 X, hitung koefisien korelasi (r) dan koefisien determinasi (R). Gunakan data berikut.

- $\Sigma x = 26$
- $\Sigma y = 40$
- $\Sigma xy = 232$
- $\Sigma x^2 = 158$
- $\Sigma y^2 = 346$

Jawab

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]}}$$

$$r = \frac{(5\times 232) - (26\times 40)}{\sqrt{\left[(5\times 158) - (26^2)\right]\times\left[(5\times 346) - (40^2)\right]}}$$

$$r = \frac{1160 - 1040}{\sqrt{\left[790 - 676\right]\times\left[1730 - 1600\right]}} = \frac{120}{\sqrt{114\times 130}}$$

$$r = \frac{120}{\sqrt{14820}} = \frac{120}{121.73...} = 0.9857...$$

Nilai r = 0,9857 menunjukkan bahwa peubah X (biaya promosi) dan Y (volume penjualan) berkorelasi linier yang positif dan tinggi.

$R = r^2 = 0.9857...^2 = 0,97165.... = 97\,\%$

Nilai R = 97% menunjukkan bahwa 97% proporsi keragaman nilai peubah Y (volume penjualan) dapat dijelaskan oleh nilai peubah X (biaya promosi) melalui hubungan linier. Sisanya, yaitu 3 % dijelaskan oleh hal-hal lain.

DAFTAR PPUSTAKA

1. Statistika dan Probabilitas, Dr Boediono dan Dr. Ir. Wayan Koster, MM., 2014
2. Probabiltas dan Statistik Edisi Kedua oleh Schaum
3. Statistika Induktif, Drs. Danang Sunyoto, MM., 2012
4. Probabilitas dan Statistik, Ramadoni Syahputra, 2005