DISTRIBUSI TEORITIS

Distribusi teoritis merupakan alat bagi kita untuk menentukan apa yang dapat kita harapkan, apabila asumsi-asumsi yang kita buat benar. Distribusi teoritis memungkinkan para pembuat keputusan untuk memperoleh dasar logika yang kuat di dalam keputusan, dan sangat berguna sebagai dasar pembuatan ramalan berdasarkan informasi yang terbatas atau pertimbangan-pertimbangan teoritis, dan berguna pula untuk menghitung probabilitas terjadinya suatu kejadian.

Setiap kejadian yang dapat dinyatakan sebagai perubahan nilai suatu variabel, umumnya mengikuti suatu distribusi teoritis tertentu dan apabila sudah ketahuan jenis distribusinya, kita dengan mudah dapat mengetahui besarnya nilai probabilitas terjadinya kejadian tersebut.

Beberapa distribusi teoritis yang akan dibahas dalam bab ini, antara lain Distribusi binomial, Distribusi Poisson, Distribusi Hipergeometrik, Distribusi Multinomial, Distribusi Normal, Distribusi Kai-Kuadrat(Chi-Square), Distribusi F, dan Distribusi t.

DISTRIBUSI BINOMIAL.

Distribusi binomial atau distribusi Bernoulli adalah suatu distribusi teoritis yang menggunakan variabel acak diskrit yang terdiri dari dua kejadian yang berkomplemen, seperti sukses-gagal, baik-cacat.

Pada umumnya suatu eksperimen dapat dikatakan eksperimen Binomial apabila memenuhi syarat sebagai berikut.

1. Setiap percobaan menghasilkan dua kejadian:
   (a) Lulus (sukses) – Tidak Lulus (gagal)
   (b) Senang (sukses) – Tidak Senang (gagal)
   (c) Puas (sukses) – Tidak Puas (gagal)
   (d) Setuju (sukses) – Tidak Setuju (gagal)
   (e) Barang Bagus (sukses) – Barang Rusau (gagal)
2. Setiap eksperimen mempunyai dua hasil yang dikatagorikan menjadi “sukses” dan “gagal”.
3. Probabilitas suatu kejadian untuk suskes atau gagal adalah tetap untuk setiap kejadian. P(p), peluang sukses, P(q) peluang gagal, dan P(p) + P(q)= 1.
4. Probabilitas sukses sama pada setiap eksperimen.
5. Eksperimen tersebut harus bebas satu sama lain, artinya hasil eksperimen yang satu tidak mempengeruhi hasil eksperimen lainnya.

<u>Contoh:</u>

Suatu eksperimen Binomial, yang terdiri dari pengambilan satu bola secara acak dari kotak yang berisi 30 bola merah(= 30M) dan 70 bola putih(= 70P). Y adalah variabel acak dengan nilai sebagai berikut.

$$Y = \begin{cases} 1, \text{kalau bola merah yang terambil} \\ 0, \text{kalau bola putih yang terambil} \end{cases}$$

P(M) = p = probabilitas untuk mendapat bola merah (sukses)

$= \frac{30}{100} = 0{,}30$

P(P) = q = probabilitas untuk mendapat bola putih (gagal)

$= \frac{70}{100} = 0{,}70$

E(Y) = 1(p) + 0(q)
= 1(0,3) + 0(0,7)
= 0,3

DISTRIBUSI POISSON.

Distribusi poisson adalah pengembangan dari distribusi binomial yang mampu mengkakulasikan distribusi probabilitas dengan kemungkinan sukses (p) sangat kecil dan jumlah eksperimen (n) sangat besar (misal 100 atau lebih)

Karena distribusi poisson biasanya melibatkan jumlah n yang besar, dengan p kecil, distribusi ini biasanya digunakan untuk menghitung nilai probabilitas suatu kejadian dalam suatu selang waktu dan daerah tertentu. Contoh : Banyaknya bakteri dalam air yang bersih, Banyaknya presiden yang meninggal karena kecelakaan lalulintas, Banyaknya dering telepon dalam satu jam di suatu kantor.

Distribusi Poisson memiliki ciri-ciri sebagai berikut.

Banyaknya hasil percobaan yang terjadi dalam suatu interval waktu atau suatu daerah tertentu tidak tergantung pada banyaknya hasil percobaan yang terjadi pada interval waktu atau daerah lain yang terpisah.

Rumus untuk menyelesaikan distribusi Poisson adalah sebagai berikut.

$$P_r(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Dimana λ = rata-rata distribusi (Lambda)
= 0, 1, 2, 3, …(menuju tak hingga)
e = konstanta 2,71828

<u>Contoh:</u>

Seorang yang akan menjual mobil mewahnya memasang iklan pada suatu surat kabar yang dapat mencapai 100.000 pembaca. Dengan anggapan nilai probabilitas bahwa seorang yang membaca iklan tersebut berminat akan membeli mobilnya sebesar p = 1/50.000. Jika dari 100.000 pembaca ada dua orang yang berminat membeli mobil tersebut (p = 0,00002) dan X = banyaknya pembaca yang berminat pada mobil tersebut, berapakah P(X = 0), P(X = 1), P(X = 2), P(X = 3), P(X = 4), ,,,,?

Persoalan ini sebetulnya dapat dipecahkan dengan menggunakan fungsi Binomial, karena persoalannya hanya mencari probabilitas x “sukses” dari n = 100.000 eksperimen dimana probabilitas sukses p = 1/50.000. Akan tetapi karena n terlalu besar dan p terlalu kecil, fungsi Poisson dapat digunakan sebagai suatu pendekatan yang lebih sederhana.

Apabila λ = rata-rata distribusi = E(X) = np = $\frac{100.000}{50.000}$ = 2, (secara rata-rata dapat diharapkan 2 (dua) orang pembaca yang menanyakan keadaan mobil), maka setelah dilakukan perhitungan, kita akan memperoleh sebagai berikut.

$P_r(X = 0) = \dfrac{2^0 e^{-2}}{0!} = 0{,}1353$

$P_r(X = 1) = \dfrac{2^1 e^{-2}}{1!} = 0{,}2707$

$P_r(X = 2) = \dfrac{2^2 e^{-2}}{2!} = 0{,}2707$

$P_r(X = 3) = \dfrac{2^3 e^{-2}}{3!} = 0{,}1804$

$P_r(X = 4) = \dfrac{2^4 e^{-2}}{4!} = 0{,}0902$

$P_r(X = 5) = \dfrac{2^5 e^{-2}}{5!} = 0{,}0361$

$P_r(X = 6) = \dfrac{2^6 e^{-2}}{6!} = 0{,}0120$

$P_r(X = 7) = \dfrac{2^7 e^{-2}}{7!} = 0{,}0034$

$P_r(X = 8) = \dfrac{2^8 e^{-2}}{8!} = 0{,}0009$

$P_r(X = 9) = \dfrac{2^9 e^{-2}}{9!} = 0{,}0002$

Perhitungan ini dapat juga dilihat pada tabel Poisson, dimana x = 0, 1, 2, …, 9. Misalnya kita ingin melihat distribusi probabilitas bahwa 5 orang pembaca berminat pada mobil tersebut (p(5) dengan λ atau rata-rata distribusi = 2, perhatikan potongan tabel Poisson berikut.

x	1,0	2,0	3,0	4,0	5,0	6,0
0	0,3679	0,1353	0,0498	0,0183	0,0067	0,0025
1	0,3679	0,2707	0,1494	0,0733	0,0337	0,0149
2	0,1839	0,2707	0,2240	0,1465	0,0842	0,0446
3	0,0613	0,1804	0,2240	0,1954	0,1404	0,0892
4	0,0153	0,0902	0,1680	0,1954	0,1755	0,1339
5	0,0031	0,0361	0,1008	0,1563	0,1755	0,1606
6	0,0005	0,0120	0,0504	0,1042	0,1462	0,1606

Perhatikan kolom 2, dengan λ = 2,0, telusuri ke bawah sampai ke baris x = 5. Disana kita akan menemukan angka 0,0361. Artinya probabilitas 5 orang berminat dari 100.000 pembaca adalah 0,0361, probabilitas 6 orang berminat adalah 0,0120, dan seterusnya.

DISTRIBUSI HIPERGEOMETRIK.

Distribusi hipergeometrik sangat erat kaitannya dengan distribusi binomial. Perbedaan antara distribusi hipergeometrik dengan binomial adalah bahwa pada distribusi hipergeometrik, percobaan tidak bersifat independent. Artinya antara percobaan yang satu dengan yang lainnya saling berkait. Selain itu probabilitas "SUKSES" berubah (tidak sama) dari percobaan yang satu ke percobaan lainnya.

Untuk mencari probabilitas x sukses dalam ukuran sample n, kita harus memperoleh x sukses dari r sukses dalam populasi, dan n – x gagal dari N – r gagal. Sehingga fungsi probabilitas hipergeometrik dapat dituliskan sebagai berikut.

$$p(x) = \frac{{}^{r}C_{x}\;{}^{N-r}C_{n-x}}{{}^{N}C_{n}}, 0 \le x \le r$$

dimana :

p(x) = probabilitas x sukses (atau jumlah sukses sebanyak x) dalam n percobaan.
n = jumlah percobaan
N = Jumlah elemen dalam populasi
r = jumlah elemen dalam populasi berlabel "SUKSES"
x = Jumlah elemen berlabel "SUKSES" diantara n elemen percobaan.

Terdapat dua persyaratan yang harus dipenuhi oleh sebuah distribusi Hipergeometrik:

1. Percobaan diambil dari suatu populasi yang terbatas, dan percobaan dilakukan tanpa pengembalian.
2. Ukuran sampel n harus lebih besar dari 5% dari populasi N.

Dari rumus diatas, perhatikan bahwa

$${}^{r}C_{x} = \frac{r!}{x!(r-x)!}$$

$${}^{N-r}C_{n-x} = \frac{(N-r)!}{(n-x)(N-r-n+x)!}$$

$${}^{N}C_{n} = \frac{N!}{n!(N-n)!}$$

<u>Contoh :</u>

Sebuah anggota komite terdiri dari 5 orang, dimana 3 adalah wanita dan 2 laki-laki. Misalkan 2 orang dari 5 orang anggota komite tersebut dipilih untuk mewakili delegasi dalam sebuah konvensi/pertemuan,

(i) Berapa probabilitas bahwa dari pemilihan secara acak didapat 2 orang wanita?
(ii) Berapa probabilitas dari 2 orang yang terpilih adalah 1 laki-laki dan 1 wanita?

Penyelesaian :
Kita dapat menggunakan distribusi hipergeometrik dalam kasus ini, dengan n = 2, N = 5, r = 3 dan x = 2, x = jumlah wanita terpilih.

(i) p(2) = $$\frac{{}^{3}C_{2}\ {}^{2}C_{0}}{{}^{5}C_{2}} = \frac{\left(\frac{3!}{2!1!}\right)\left(\frac{2!}{2!0!}\right)}{\left(\frac{5!}{2!3!}\right)} = \frac{3}{10} = 0{,}3$$

Jadi probabilitas 2 orang wanita terpilih adalah 0,3

(ii) p(1) = $$\frac{{}^{3}C_{1}\ {}^{2}C_{1}}{{}^{5}C_{2}} = \frac{\left(\frac{3!}{1!2!}\right)\left(\frac{2!}{1!1!}\right)}{\left(\frac{5!}{2!3!}\right)} = \frac{3.2}{10} = \frac{6}{10} = 0{,}6$$

Jadi probabilitas terpilih 1 orang wanita dan 1 laki-laki = 0,6

DISTRIBUSI MULTINOMIAL.

Kalau pada distribusi binomial hasil sebuah percobaan hanya dikatagorikan 2 macam yaitu "sukses" dan "gagal", maka dalam distribusi multinomial, sebuah percobaan akan menghasilkan beberapa kejadian (lebih dari 2) yang saling meniadakan/saling lepas. Misalkan ada sebanyak ***k*** kejadian dalam sebuah percobaan, katakan kejadian-kejadian B_1, B_2, …, B_k. Jika percobaan diulang sebanyak n kali dan peluang terjadinya setiap kejadian B konstan/tetap dari setiap percobaan dengan $P(B_i) = P_i$ untuk i = 1, 2, 3, …, k, dan X_1, X_2, X_3, …X_k menyatakan jumlah terjadinya kejadian B_i (i = 1, 2, …, k dalam n percobaan)

Fungsi Distribusi Multinomial

$$p(x_1, x_2, x_3, ...x_k) = \left[\frac{n!}{x_1!\, x_2!\, x_3! ....x_k!}\right] p_1^{x_1}\, p_2^{x_2}\, p_3^{x_3} .....p_k^{x_k}$$

untuk nilai-nilai

X_1 = 0, 1, 2, ……; X_k = 0, 1, 2, … dan $\sum_{i=1}^{k} x_i = n$ Dimana :

X_1, X_2, ……, X_k menyatakan jumlah dari kejadian B_1, B_2, …….B_k
n menyatakan jumlah percobaan.
P_1, p_2, ..,p_k adalah probabilitas terjadinya kejadian B_1, B_2, ….B_k

Contoh :

Proses pembuatan pensil dalam sebuah pabrik melibatkan banyak buruh dan proses tersebut terjadi berulang-ulang. Pada suatu pemeriksaan terakhir yang dilakukan telah memperlihatkan bahwa 85% produksinya adalah "baik", 10%

ternyata “tidak baik tetapi masih bisa diperbaiki” dan 5% produksinya “rusak dan harus dibuang”. Jika sebuah sample acak dengan 20 unit dipilih, berapa peluang jumlah unit “baik” sebanyak 18, unit “tidak baik tetapi bisa diperbaiki” sebanyak 2 dan unit “rusak” tidak ada?

Penyelesaian :
Misalkan,
X_1 = banyaknya unit “baik”
X_2 = banyaknya unit yang “tidak baik tetapi bisa diperbaiki”
X_3 = banyaknya unit yang “rusak dan harus dibuang”

$X_1 = 18$, $X_2 = 2$, dan $X_3 = 0$ (syarat $x_1 + x_2 + x_3 = n = 20$)
dan $p_1 = 0{,}85$, $p_2 = 0{,}1$ dan $p_3 = 0{,}05$ maka :

$$p(x1,\ x2,\ x3) = \left[\frac{n!}{x1!\,x2!\,x3!}\right](p1)^{x1}(p2)^{x2}\,(p3)^{x3}$$

$$p(18,\ 2,\ 0) = \left[\frac{20!}{18!\,2!\,0!}\right](0{,}85)^{18}\,(0{,}1)^{2}\,(0{,}05)^{0}$$

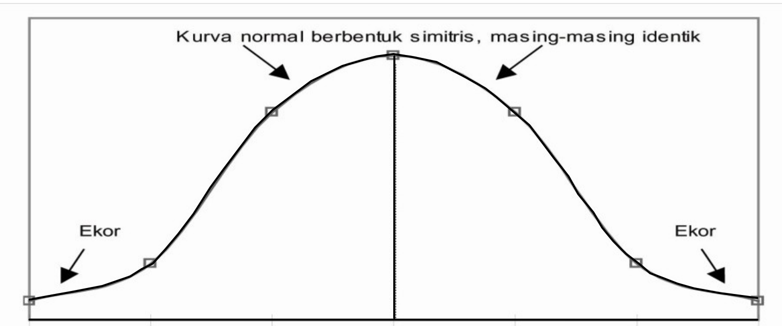
$= 190\,(0{,}85)^{18}\,(0{,}01)\,(1)$
$= 0{,}102$

Jadi peluangnya sebesar 0,102.

DISTRIBUSI NORMAL.

Di antara sekian banyak distribusi, barangkali distribusi normal merupakan distribusi yang secara luas banyak digunakan dalam berbagai penerapan. Distribusi normal merupakan distribusi kontinu yang mensyaratkan variabel yang diukur harus kontinu misalnya tinggi badan, berat badan, dan sebagainya.

Karakteristik Distribusi Kurva Normal :



1. Kurva berbentuk genta (μ= Md = Mo)
2. Kurva berbentuk simetris
3. Kurva normal berbentuk asimptotis
4. Kurva mencapai puncak pada saat X= μ
5. Luas daerah di bawah kurva adalah 1; ½ di sisi kanan nilai tengah dan ½ di sisi kiri.

Persamaan matematika bagi distribusi probabilitas acak normal tergantung pada dua parameter, yaitu μ dan σ atau nilai tengah dan simpangan bakunya. Fungsi kepadatan probabilitas normal dapat dituliskan sebagai berikut.

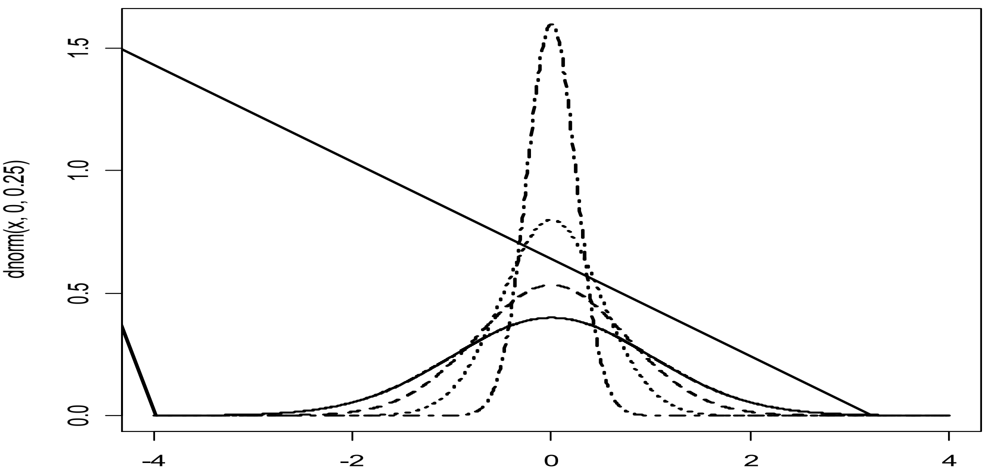
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$, untuk $-\infty < X < \infty$

di mana :

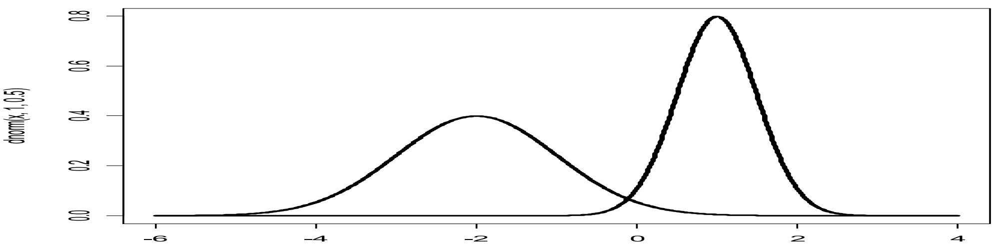
π = 3,14159

e = 2,71828

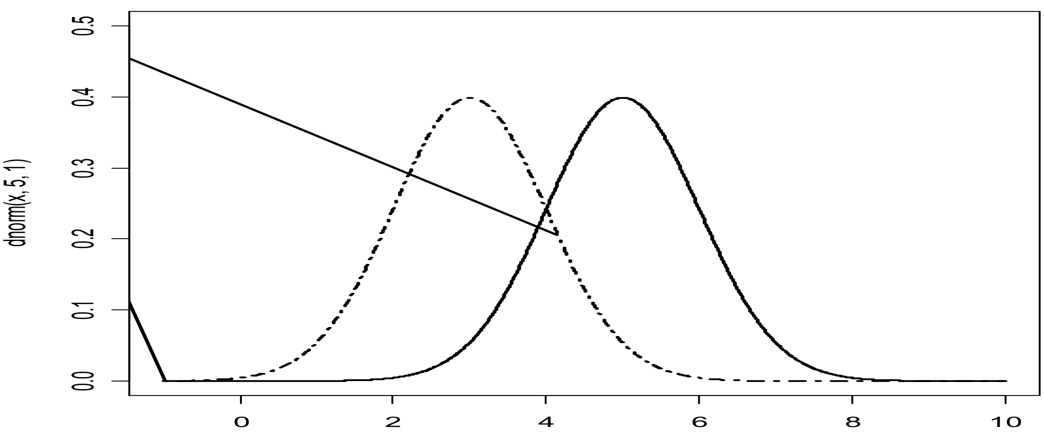
Bila nilai-nilai μ dan σ diketahui, maka kita dapat menggambarkan kurva normal itu dengan pasti. Bagaimanapun bentuk dan ketinggian dari kurva normal sangat tergantung pada dua variabel ini.



Distribusi kurva normal dengan μ sama dan σ berbeda



Distribusi kurva normal dengan μ dan σ berbeda



Distribusi kurva normal dengan μ berbeda dan σ sama

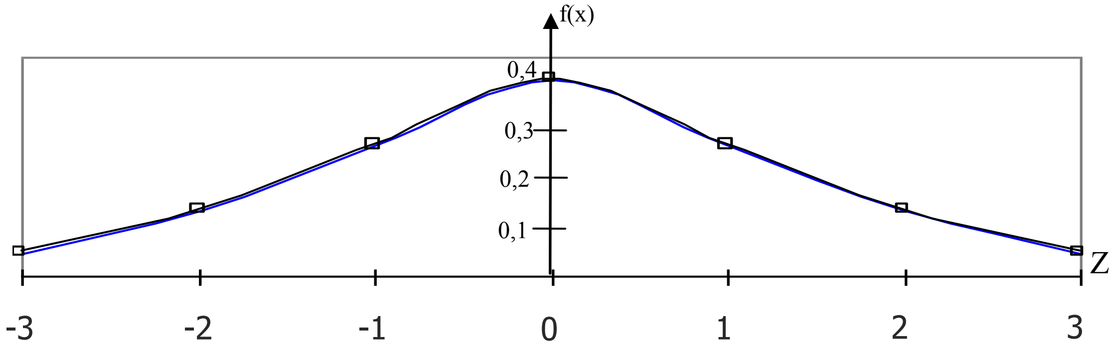
Distribusi Normal Standar.

Keluarga distribusi normal memiliki jumlah yang banyak sekali, akibat pengaruh rata-rata dan simpangan baku. Akan tetapi, untuk mencari probabilitas suatu interval dari variabel random kontinu dapat dipermudah dengan menggunakan batuan distribusi normal standar.

Distribusi normal standar adalah distribusi normal yang memiliki rata-rata (μ) = 0 dan simpangan baku (σ) = 1. Bentuk fungsinya adalah.

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2}$$

Dalam bentuk diagram atau kurva (disebut kurva normal standar), distribusi normal standar digambarkan :



Kurva distribusi normal standar

Dari bentuk kurva distribusi normal standar tersebut, dapat diketahui sifat - sifat distribusi terebut, yaitu :

1. kurva simetris terhadap sumbu Y.
2. mempunyai titik tertinggi (0, $\frac{1}{\sqrt{2\pi}}$), dengan $\frac{1}{\sqrt{2\pi}}$ = 0,4
3. cekung kebawah untuk interval X = -1 sampai X = +1 dan cekung ke atas untuk nilai X di luar interval tersebut.
4. meluas atau melebar tanpa batas ke kiri dan ke kanan serta mendekati sumbu X secara cepat begitu bergerak dari X = 0 ke kiri maupun ke kanan.
5. luas seluruh daerah di bawah kurva dan di atas sumbu X sebesar 1 unit.

Untuk mengubah distribusi normal umum menjadi distribusi normal standar, gunakan nilai Z. Bentuk rumusnya adalah :

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

Dimana :

Z = variabel normal standar

X = nilai variabel random

μ = rata-rata variabel random

σ = simpangan baku variabel random

Nilai Z adalah angka atau indeks yang menyatakan penyimpangan suatu nilai variabel random (X) dari rata-rata (μ) dihitung dalam satuan simpangan baku (σ).

Contoh Soal:

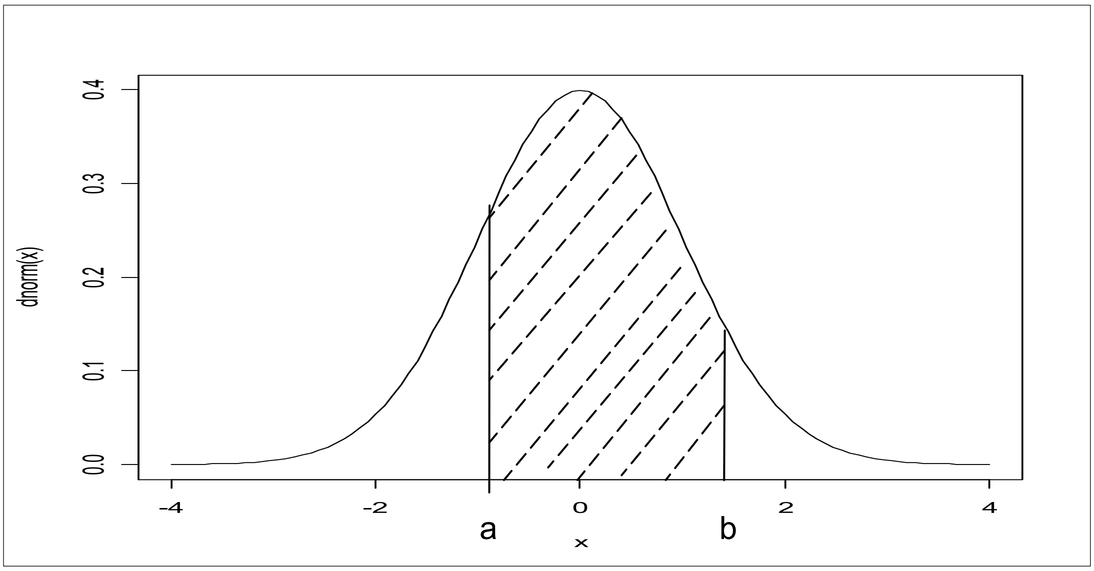
Harga saham di BEJ mempunyai nilai tengah (X)=490,7 dan standar deviasinya 144,7. Berapa nilai Z untuk harga saham 600?

Jawab:
Diketahui: Nilai μ = 490,7 dan σ = 144,7

Maka nilai Z =(X - μ) / σ
Z = (600 – 490,7)/144,7
Z = 0,76

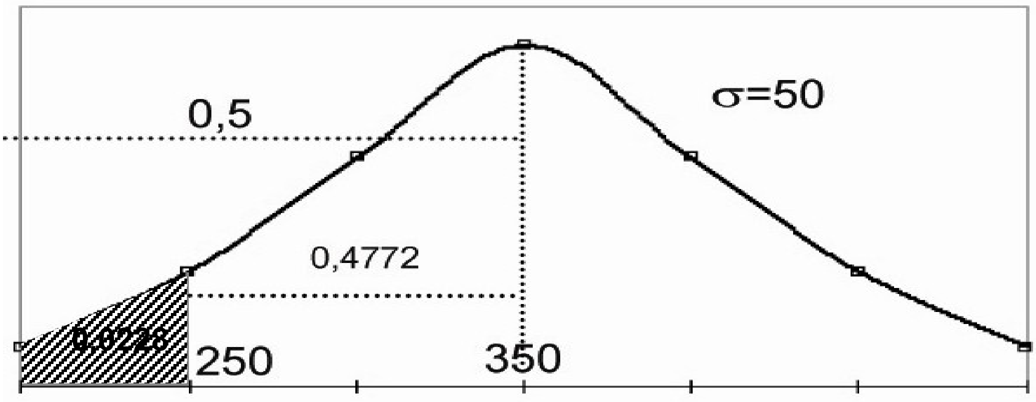
Luas Di Bawah Kurva Normal.
Luas daerah kurva normal antara x = a dan x = b dinyatakan sbb:

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



Luas daerah P(a<x<b)= luas daerah di arsir

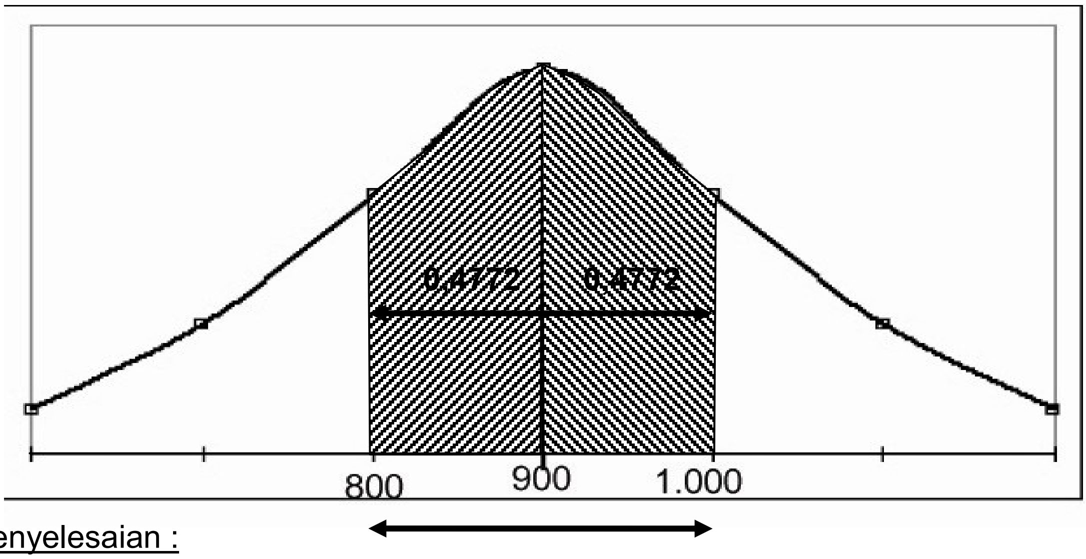
Contoh :
PT GS mengklaim berat buah mangga "B" adalah 350 gram dengan standar deviasi 50 gram. Bila berat mangga mengikuti distribusi normal, berapa probabilitas bahwa berat buah mangga mencapai kurang dari 250 gram, sehingga akan diprotes oleh konsumen.



Penyelesaian :

- Transformasi ke nilai z
  AP(x< 250); P(x=250) = (250-350)/50=-2,00 Jadi P(x<250)=P(z<-2,00)
- Lihat pada tabel luas di bawah kurva normal
  P(z<-2,00)=0,4772
- Luas sebelah kiri nilai tengah adalah 0,5. Oleh sebab itu, nilai daerah yang diarsir menjadi 0,5 – 0,4772=0,0228. Jadi probabilitas di bawah 250 gram adalah 0,0228 (2,28%). Dengan kata lain probabilitas konsumen protes karena berat buah mangga kurang dari 250 gram adalah 2,28%.

Contoh:
PT Work Electric, memproduksi Bohlam Lampu yang dapat hidup 900 jam dengan standar deviasi 50 jam. PT Work Electric ingin mengetahui berapa persen produksi pada kisaran antara 800-1.000 jam, sebagai bahan promosi bohlam lampu. Hitung berapa probabilitasnya!



Penyelesaian :
P(800<X<1.000)?

- Hitung nilai Z
  Z_1 = (800-900)/50 = -2,00;
  Z_2 = (1.000-900)/50 = 2,00
- Jadi: P(800<X<1.000) =P(-2,00<Z<2,00);
  P(-2,00<Z) = 0,4772 dan P(Z>2,00) = 0,4772
  Sehingga luas daerah yang diarsir adalah = 0,4772+0,4772= 0,9544.
  Jadi P(800<X<1.000) = P(-2,00 < Z<2,00) = 0,9544.
  Jadi 95,44% produksi berada pada kisaran 800-1.000 jam. Jadi jika PT Work Electric mengklaim bahwa lampu bohlamnya menyala 800-1.000 jam, mempunyai probabilitas benar 95,44%, sedang sisanya 4,56% harus dipersiapkan untuk garansi.

DISTRIBUSI KAI-KUADRAT (χ^2 = Chi Square).
Distribusi kai-kuadrat sangat berguna sebagai kriteria untuk pengujian hipotesis mengenai varians dan juga untuk uji ketepatan penerapan suatu fungsi apabila digunakan untuk data hasil observasi atau data empiris. Dengan demikian, kita dapat menentukan apakah distribusi pendugaan berdasarkan sampel hampir sama atau mendekati distribusi teoritis, sehingga kita dapat menyimpulkan bahwa

populasi dari mana sampel itu kita pilih mempunyai distribusi yang kita maksud (misalnya, suatu populasi mempunyai distribusi binomial, Poisson, atau Normal).

Misalnya sebuah dadu yang mempunyai 6 mata (mata 1, 2, 3, 4, 5, 6) dilemparkan ke atas sebanyak 300 kali. Dalam jangka panjang, kita harapkan untuk melihat masing-masing mata tersebut muncul dengan frekuensi yang sama, yaitu masing-masing muncul 50 kali. Dalam prakteknya, frekuensi mata dadu yang muncul sekitar 50, walaupun dadu itu termasuk "fair dice". Dengan menggunakan kai-kuadrat, kita dapat menentukan apakah suatu dadu dapat diatakan "fair" setelah membandingkan frekuensi dari masing-masing mata dadu tersebut.

Apabila Z_i = N(0,1) = variabel normal dengan rata-rata 0 dan variens sama dengan 1, atau E(Z) = 0, σ_Z^2= 1, maka jumlah $Z_1^2 + Z_2^2 + \ldots + Z_k^2$ sama dengan χ_k^2 dengan derajat kebebasan sebesar k.

$$\chi_k^2 = \sum Z_i^2$$

Kalau suatu himpunan yang terdiri n variabel acak X = {X_i}, dimana $X_i \sim n(\mu, \sigma^2)$ untuk semua i (i = 1, 2, …, n), maka kita dapat memperoleh variabel Z seperti yang dimaksud di atas, dengan rumus.

$$Z_i = \frac{X_i - \mu}{\sigma} \sim \text{N90,1)}, \qquad i = 1, 2, \ldots, n.$$

$$\chi_n^2 = \sum \left(\frac{x_i - \mu}{\sigma} \right)^2$$

χ_n^2 =kai-kuadrat dengan derajat kebebasan sebesar n.

Apakah yang dimaksud dengan derajat kebebasan.
Misalnya kita diminta untuk menentukan 5 nilai X, yaitu X_1, X_2, X_3, X_4, X_5, dimana syaratnya sudah ditentukan bahwa rata-ratanya. $\overline{X}$= 5. Jadi, jumlah kelima nilai X tersebut adalah 25.

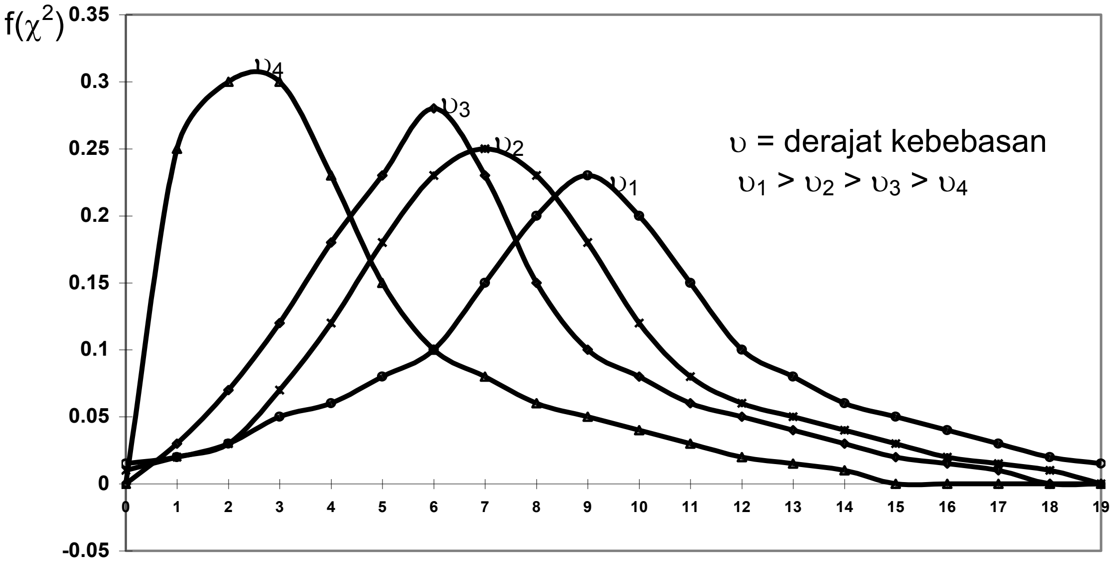
Kalau nilai x_1, x_2, x_3, dan x_4 kita tentukan misalnya x_1 = 4, x_2 = 5, x_3 = 6, dan x_4 = 7, maka nilai x_5 tidak bebas lagi untuk menentukannya. Nilai x_5 harus membuat jumlah kelima nilai x tersebut menjadi 25.

Dengan demikian $X_5 = 25 - (X_1 + X_2 + X_3 + X_4) = 25 - (4 + 5 + 6 + 7) = 3$. Jadi X_5 = 3. kita mempunyai 4 kebebasan (satu kali tidak) di antara 5 pilihan, dengan kata lain hanya mempunyai derajat kebebasan sebanyak 4 yaitu (5 – 1). Kalau harus memilih dari n elemen derajat kebebasannya = n – 1. Di dalam soal kita hanya memperkirakan satu penduga (=X), sehingga derajat kebebasan ada (n – 1). Apabila harus memperkirakan k penduga, maka derajat kebebasan ada (n – k).

Kalau k = 2, yaitu a dan b(a penduga dari A dan b penduga dari B), dalam persoalan regresi Y = A + BX + ε, maka derajat kebebasannya (n – 2).

Bentuk kurva kai-kuadrat sangat dipengaruhi oleh besar kecilnya nilai derjat kebebasan. Makin kecil nilai derajat kebebasan, bentuk kurvanya makin menceng kekanan dan makin besar nilai derajat kebebasan ($n \to \infty$), bentuk kurvanya makin mendekati bentuk funsi normal. Kai-kuadrat merupakan fungsi kontinu dan nilainya tidak pernah negatif. Nilai rata-ratanya makin meningkat kalau nilai derajat kebebasan juga makin meningkat.

Kurva Kai-Kuadrat dengan derajat kebebasan $x = \left(\chi_x^2\right)$



$$E\left(\chi_\upsilon^2\right) = \mu = \upsilon$$

Rata-rata kai-kuadrat denga derajat kebebasan sebesar υ adalah sama dengan υ.

$$Var\left(\chi_\upsilon^2\right) = \sigma^2 = 2\upsilon$$

Untuk keperluan perhitungan nilai χ^2, suatu tabel kai-kuadrat telah dibuat menurut berbagai nilai derajat kebebasan. Dalam tabel, derajat kebebasan sering diberi simbol υ, r, atau n dan sering disngkat d.o.f atau d.f. Dalam membaca tabel kai-kuadrat, agar diperhatikan simbol (notasi) dibagian atas yang digunakan dalam tabel tersebut. Tabel kai-kuadrat memuat nilai χ^2, dan bukan nilai probabilitas seperti halnya tabel distribusi normal.

Untuk υ > 100, distribusi kai-kuadrat mendekati distribusi normal, dimana variabel Z sebagai variabel normal baku dapat diperoleh dengan cara berikut.

$$Z = \sqrt{2\chi^2} - \sqrt{2\upsilon - 1}$$

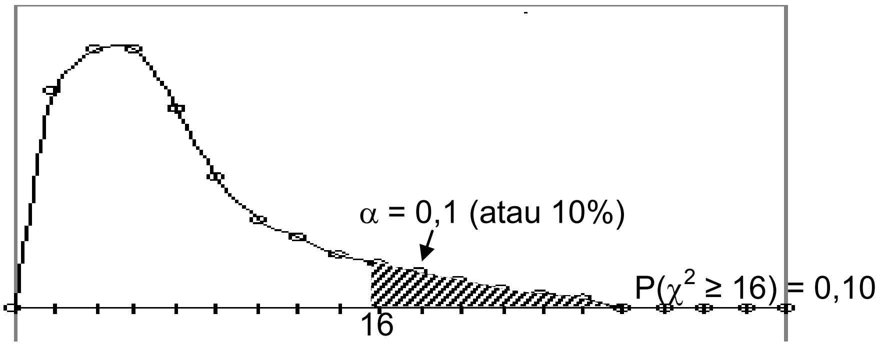
Nilai χ^2 dengan nilai derajat kebebasan yang berbeda dan tingkat probabilitas yang berlainan dapat dilihat dalam tabel χ^2.

Cara Membaca Tabel χ^2.

Misalkan α = probabilitas bahwa kai-kuadrat mengambil nilai sama atau lebih besar dari nilai yang terdapat pada tabel kai-kuadrat dengan derajat kebebasan sebesar υ. Nilai kai-kuadrat dari tabel diberi simbol $\chi_{\alpha,\upsilon}^2$.

$$P\left(\chi^2 \geq \chi_{\alpha,\upsilon}^2\right) = \alpha$$

Kalau υ = 10, dan α = 10% maka luas daerah dari kurva kai-kuadrat berdasarkan tabel kai-kuadrat terletak di sebelah kanan dari $\chi_{(0,10),(n)}^2$ = 16,00



Sebagian dari tabel Kai-Kuadrat adalah sebagai berikut.

Tabel Distribusi Kai-Kuadrat

υ \ α	90% (0,90)	50% (0,50)	10% (0,10)	5% (0,05)
1	0,01	0,45	2,71	3,84
5	1,61	4,35	9,24	11,10
10	4,87	9,34	16,00	18,30
20	12,40	19,30	28,40	31,40
30	20,60	29,30	40,30	43,80
40	29,10	39,30	51,80	55,80

Tabek Kai-Kuadrat memuat nilai hasil perhitungan kumulatif.
Misalnya dari tabel diatas, untuk derajat kebebasan sebesar υ = 10, luas kurvanya sebesar 90% terletak disebelah kanan titik di mana χ^2 = 4,87.

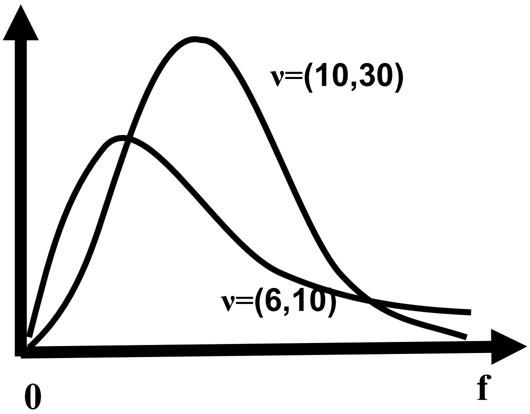
DISTRIBUSI F.
Salah satu perbandingan yg dilakukan dalam statistik adalah perbandingan variabilitas atau variansi dari dua buah sampel. Statistik yg dipergunakan dalam membandingkan variansi 2 buah sampel dinamakan distribusi F.

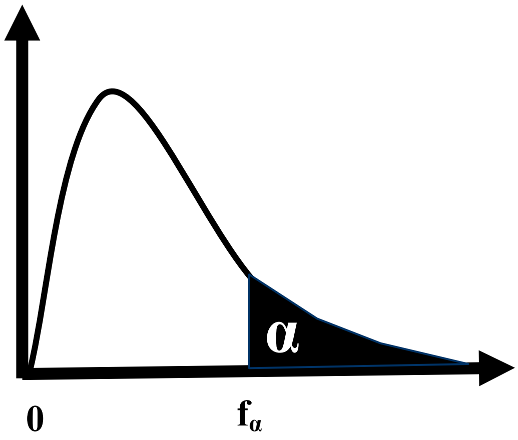
Jika S_1^2 dan S_2^2 adalah variansi dari 2 buah sampel random yg tak saling bergantung (independen) dengan ukuran n_1 dan n_2 yg diambil dari populasi **normal** dengan variansi σ_1^2 dan σ_2^2, maka statistik F:

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

Mengikuti distribusi F dengan derajat kebebasan $v_1=n_1-1$ dan $v_2=n_2-1$. Distribusi F bersifat asimetrik . bentuknya bergantung pada derajat kebebasannya

Perbandingan variansi : Distribusi F





Jika fα (v_1,v_2) menyatakan nilai kritis f dengan luas ekor kanan α untuk derajat kekebasan v_1,v_2, maka: (perhatikan urutan v_1 dan v_2)

$$f_{1-\alpha}(\nu_2,\nu_1) = \frac{1}{f_\alpha(\nu_1,\nu_2)}$$

Karena ada dua derajat kebebasan yg menentukan bentuk Distribusi F maka, tabel distribusi lebih terbatas, hanya ditabelkan nilai kritis F untuk beberapa nilai luas ekor kanan yg populer dipakai (misalnya α=5%)

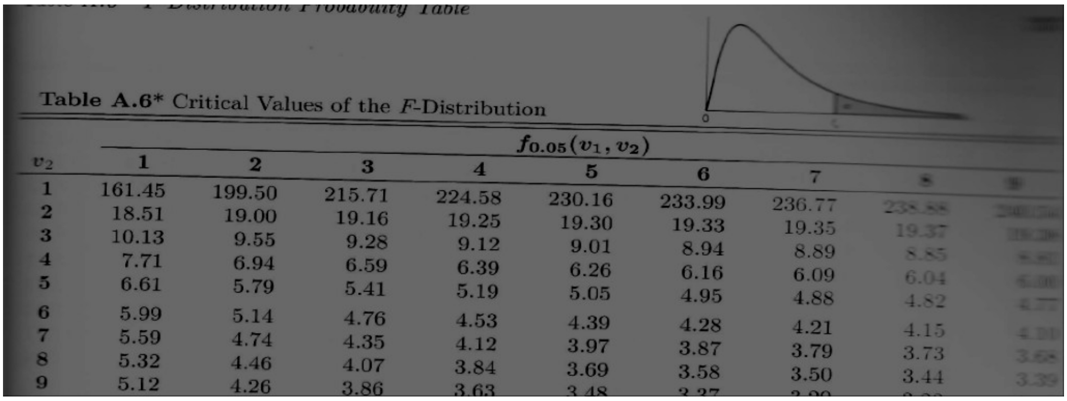


Table A.6* Critical Values of the F-Distribution

v_2	$f_{0.05}(v_1, v_2)$ 1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	[illegible]
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	[illegible]
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	[illegible]
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	[illegible]
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	[illegible]
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	[illegible]
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	[illegible]
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.68
9	5.12	4.26	3.86	3.63	3.48	[illegible]	[illegible]	[illegible]	3.39

Nanti distribusi F akan dipakai untuk memeriksa kesamaan rata-rata dari beberapa grup sampel yg diambil secara independen. Ada dua faktor yg akan menentukan apakah perbedaan rata-rata sampel memang nyata atau tidak yaitu:

1. Variasi di dalam sampel (within)
2. Variasi antar sampel (between)

DISTRIBUSI t.

Distribusi t selain digunakan untuk menguji suatu hipotesis juga untuk membuat pendugaan interval. Biasanya, distribusi t digunakan untuk menguji hipotesis mengenai nilai parameter, paling banyak dari 2 populasi (lebih dari 2, harus digunakan F), dan dari sampel yang kecil misalnya n < 100, bahkan seringkali n ≤ 30. Untuk n yang cukup besar (n ≥ 100, atau mungkin cukup n > 30) dapat digunakan distribusi normal, maksudnya tabel normal dapat digunakan sebagai pengganti tabel t.

Kalau Z = N(0,1) = variabel normal dengan rata-rata 0 dan simpangan baku 1, dam χ^2_υ = kai-kuadrat dengan derajat kebebasan υ, maka variabel t dapat diperoleh dengan cara berikut.

$$t = \frac{Z}{\sqrt{\frac{\chi^2_\upsilon}{\upsilon}}}$$

Artinya, fungsi mempunyai distribusi t dengan derajat kebebasan sebesar υ.
Variabel t dapat mengambil nilai negatif maupun positif, oleh karena pada dasarnya variabel t ini berasal dari variabel normal, padahal kita ketahui variabel normal selain mengambil nilai positif juga negatif. Variabel t juga mempunyai kurva yang simetris terhadap t = 0.

Tabel t, seperti tabel distribusi normal, dapat digunakan untuk mencari nilai variabel t apabila nilai probabilitas α sudah diketahui, atau sebaliknya. Untuk menggunakan tabel t harus ditentukan terlebih dahulu besarnya nilai α dan υ. Oleh karena kurva t simetris, maka kita boleh hanya mencari nilai t sebelah kanan titik 0.

Jikalau sampel kecil (n<30) maka S^2 akan berfluktuasi cukup besar dari sampel ke sampel sehingga perlu statistik yg lebih baik. Jika sampel kecil akan tetapi berasal dari distribusi normal, maka statistik t berikut ini:

$$T = \frac{\overline{X} - \mu}{S / \sqrt{n}}$$

Tabel distribusi student diberikan untuk nilai kritis t yg terkait dengan luas ekor kanan dari distribusi t, untuk berbagai nilai derajat kebebasan yg berbeda.

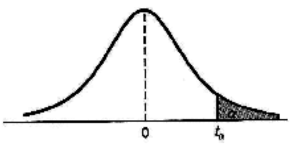


Table A.4 Critical Values of the *t*-Distribution

	α						
v	**0.40**	**0.30**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262

Contoh :
Seorang peneliti menyatakan rata-rata hasil panen setelah diberi pupuk adalah 500 gram per mm pupuk yg diberikan. Dia kemudian mengambil sampel 25 batch panen, dan memutuskan dia akan puas dengan klaimnya jikalau ternyata nilai t dari sampel terletak antara $-t_{0.05}$ s/d $t_{0.05}$. Peneliti tsb mengasumsikan bahwa bobot hasil panen mengikuti distribusi normal. Ternyata sampelnya memiliki rata-rata 518 gram dengan standard deviasi sampel 40. Apakah dia akan puas dengan klaimnya?

<u>Penyelesaian :</u>
Ini adalah persoalan distribusi student t. Ukuran sampel n=25, sehingga derajat kebebasan v=n-1=25-1=24. Dari tabel diketahui bahwa untuk v=24, maka $t_{0.05}$ = 1.711, sedangkan hasil sampelnya memberikan

$$T = \frac{\overline{X} - \mu}{S / \sqrt{n}} = \frac{518 - 500}{40 / \sqrt{25}} = 2.25$$