DISTRIBUSI CHI KUADRAT DAN APLIKASINYA

SKRIPSI

Disusun Oleh :

SRI HASTUTI WIJAYANTI

No. Mhs : 98611015

NIRM : 980051013206120015

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS ISLAM INDONESIA

JOGJAKARTA

2002

DISTRIBUSI CHI KUADRAT DAN APLIKASINYA

SKRIPSI

Diajukan untuk dipertahankan dalam Sidang Penguji

sebagai salah satu syarat untuk memperoleh Gelar

Sarjana S-1 pada Jurusan Statistika

Oleh :

SRI HASTUTI WIJAYANTI

No. Mhs : 98611015

NIRM : 980051013206120015

JURUSAN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS ISLAM INDONESIA

JOGJAKARTA

2002

Skripsi ini telah disyahkan dan disetujui untuk diuji

Pada tanggal : 15 NOVEMBER 2002

Disusun Oleh :

SRI HASTUTI WIJAYANTI

No. Mhs : 98611015

NIRM : 980051013206120015

Mengetahui :

Pembimbing I	**Pembimbing II**
(Prof . Suryo Guritno , M. Stat Ph. D)	**(R.B. Fajriya Hakim, M.Si)**

Ketua Jurusan Statistika

(R.B. Fajriya Hakim, M.Si)

Telah dipertahankan didepan team penguji tingkat sarjana

Jurusan Statistika

Fakultas Matematika dan Ilmu Pengetahuan Alam

Tanggal : 20 NOVEMBER 2001

Penguji **Tanda Tangan**

1. **Prof. Dr. Suryo Guritno M.Stat Ph. D** ....
2. **R.B Fajriya Hakim, M.Si** ....
3. **Jaka Nugraha, M.Si** ....
4. **Edi Widodo, M.Si** ....

Mengetahui

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Islam Indonesia

UNIVERSITAS ISLAM INDONESIA
YOGYAKARTA
FAKULTAS MIPA

(Jaka Nugraha, M.Si)

HALAMAN PERSEMBAHAN

Skripsi ini Kupersembahkan kepada :

- Almarhum Bapak, yang selalu memberikan do'a dan kasih sayangnya hingga di ujung usia
- Ibu, dan kakak-kakakku : Mas Joko, Mas Tanto, Mbak Tatik dan Mbak Trie, makasih atas do'a, kasih sayang dan pengertiannya
- Kedua keponakanku yang lucu-lucu Dana dan Chandra
- Andry… yang selalu hadir dan singgah dalam hari-hari, pikiran dan hatiku, makasih atas semuanya…
- N 'Neng, Lita, Teh Amoy…smoga persahabatan ini slalu tetap ada
- Almamater dan Semua sahabat-sahabatku

KATA PENGANTAR

Assalamu 'alaikum wr.wb

Puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya kepada penulis, sehingga dapat menyelesaikan Laporan Tugas Akhir.

Dalam pelaksanaan dan penyusunan tugas akhir ini, penulis sadar bahwa tanpa bantuan banyak pihak, penulis tidak akan dapat melaksanakan dan menyusun Laporan Tugas Akhir ini. Pada kesempatan ini penulis ingin menyampaikan terima kasih kepada :

1. Bapak Jaka Nugraha, M.Si selaku Dekan Fakultas MIPA Universitas Islam Indonesia
2. Bapak Prof Drs Suryo Guritno, M Stat selaku Dosen Pembimbing pertama
3. Bapak RB Fajriya Hakim, M.Si selaku Dosen Pembimbing kedua dan selaku Ketua Jurusan Statistika Fakultas MIPA Universitas Islam Indonesia
4. Ibu Rohmatul Fajriya, M.Si dan seluruh Dosen Jurusan Statistika Fakultas MIPA Universitas Islam Indonesia
5. Teman-teman Statistika 98 : Mon[illegible]

Semua pihak yang telah membantu dalam penyelesaian Tugas Akhir ini yang tidak dapat penulis sebutkan satu persatu

Semoga Allah SWT menganugerahkan balasan yang lebih baik kepada semua pihak yang telah membantu penyelesaian Tugas Akhir ini.

Akhir kata penulis mohon maaf apabila ada kekurangan dan kesalahan dalam penyusunan tugas akhir ini. Penulis berharap agar hasil dari tugas akhir ini dapat bermanfaat untuk semuanya.

Wassalamu 'alaikum wr. Wb.

Jogjakarta, 2002

Penulis

DAFTAR ISI

Halaman

DAFTAR TABEL

Halaman

DAFTAR GAMBAR

Halaman

DAFTAR LAMPIRAN

Halaman

INTISARI

Tugas akhir ini membahas secara matematis tentang distribusi Chi Kuadrat dan beberapa bentuk aplikasinya. Serta menggunakan beberapa bentuk aplikasi tersebut untuk diterapkan dalam beberapa contoh kasus dengan bantuan program SPSS.

Permasalahan yang dibahas disini adalah mengenai distribusi Chi Kuadrat Terpusat dan Distribusi Chi Kuadrat Tidak Terpusat, penurunan distribusi, fungsi pembangkit momen (fungsi kumulan, sifat penjumlahan dan grafik fungsi distribusi). Bentuk aplikasi distribusi Chi Kuadrat yang dibahas adalah uji *goodness of fit* untuk distribusi normal, uji homogenitas, dan uji independensi.

Dari hasil pembahasan dapat diketahui bahwa pada dasarnya distribusi Chi Kuadrat Terpusat maupun distribusi Chi Kuadrat Tidak Terpusat sama-sama berasal dari distribusi normal. Perbedaannya adalah untuk distribusi Chi Kuadrat Terpusat berasal dari distribusi normal (0,1) sedangkan untuk distribusi Chi Kuadart Tidak Terpusat berasal dari distribusi normal $(\mu_i, 1)$ yaitu mean selain nol. Untuk bentuk aplikasi distribusi Chi Kuadrat untuk uji independensi pada dasarnya hampir sama dengan uji homogenitas. Perbedaannya uji independensi didasarkan pada suatu sampel dari suatu populasi sedangkan uji homogenitas didasarkan pada a sampel yang diambil dari r populasi.

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Dalam masalah-masalah statistik, pada umumnya data yang diperoleh dipandang sebagai sampel random yang diambil dari populasi induk (*parent population*), yaitu populasi dari mana data tersebut diperoleh. Sampel random yang dimaksud adalah sampel yang diambil jika masing-masing individu dari populasi induk saling independen dan mempunyai peluang yang sama untuk terambil dalam sampel.

Misal diambil sampel random berukuran n dari suatu populasi. Dari sampel random tersebut dibentuk suatu statistik A. Statistik A ini merupakan variabel random karena nilai statistik ini bervariasi dari satu sampel terhadap sampel yang lain. Maka statistik ini juga mempunyai distribusi probabilitas yang disebut distribusi sampling. Seperti dikemukakan diatas bahwa data yang diperoleh biasanya berupa sampel random sehingga distribusi sampling memegang peranan penting dalam statistika. Dalam tulisan ini akan dibahas salah satu macam bentuk distribusi sampling, yaitu distribusi Chi Kuadrat, karena dalam kenyataannya distribusi Chi Kuadrat banyak digunakan sebagai alat untuk menyelesaikan berbagai permasalahan yang ada.

1.2 Rumusan Masalah

Pada pembahasan berikut penulis akan menyajikan secara matematis tentang teori distribusi Chi Kuadrat dan beberapa bentuk aplikasi dari distribusi Chi Kuadrat beserta contoh kasus yang bisa diambil dari buku, jurnal dan sumber-sumber lain yang mendukung.

1.3 Pembatasan Masalah

Permasalahan yang dibahas disini dibatasi pada pembahasan bentuk dari distribusi Chi Kuadrat, penurunan distribusi Chi Kuadrat, fungsi pembangkit momen (fungsi kumulan, sifat penjumlahan dan grafik fungsi distribusi). Bentuk aplikasi dari distribusi Chi Kuadrat yang dibahas : uji *gooness of fit* (untuk distribusi normal), uji homogenitas dan uji independensi. Sedangkan dalam contoh kasusnya, data yang digunakan terbatas pada data yang telah memenuhi asumsi-asumsi yang diperlukan dalam distribusi Chi Kuadrat.

1.4 Tujuan

Tujuan dari penulisan tugas akhir ini adalah:

- Membahas secara matematis tentang teori distribusi Chi Kuadrat dan beberapa bentuk aplikasinya.
- Menggunakan beberapa bentuk aplikasi dari distribusi Chi Kuadrat untuk diterapkan dalam beberapa contoh kasus dengan bantuan program SPSS.

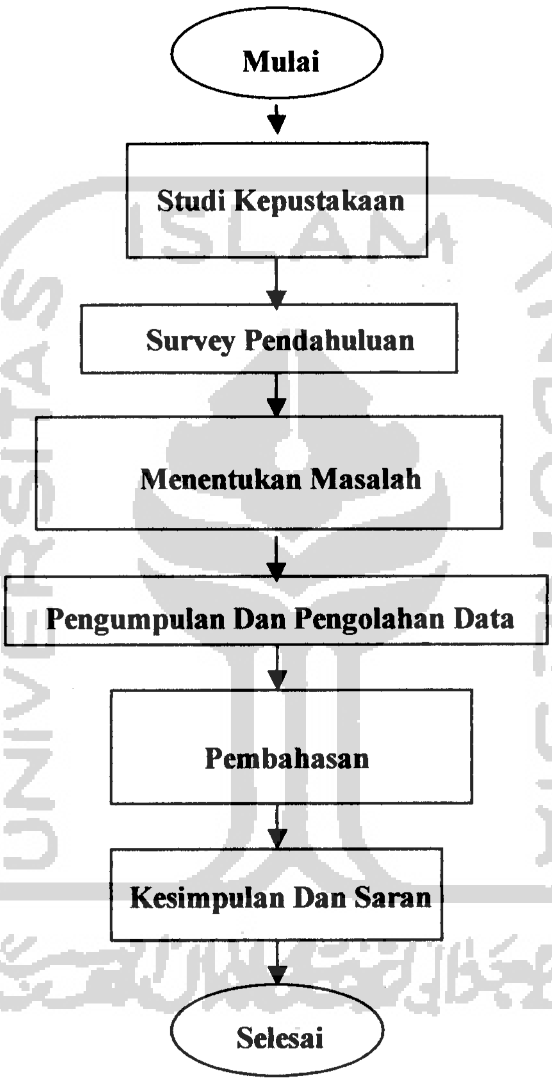
1.5 Manfaat

Dari penulisan tugas akhir ini diharapkan dapat bermanfaat untuk menambah pengetahuan dan pemahaman tentang distribusi Chi Kuadrat beserta beberapa bentuk aplikasinya.

1.6 Metodologi

1. Permasalahan yang ada dalam tugas akhir ini adalah bagaimana cara menyajikan teori-teori yang berhubungan dengan distribusi Chi Kuadrat sehingga tugas akhir ini mampu menjadi satu sumber ilmu atau tambahan pengetahuan yang mudah untuk dipahami dan dimengerti.
2. Metode yang digunakan adalah melalui studi pustaka atau literatur mengenai distribusi Chi Kuadrat dan beberapa aplikasinya. Serta diberikan contoh kasus yang dapat bersumber dari buku-buku atau jurnal yang ada.

3. Bagan alir tugas akhir adalah sebagai berikut :



Gambar 1.1 Bagan alir tugas akhir

1.7 Sistematika Penulisan

Agar penulisan ini mudah dimengerti dan memenuhi persyaratan, maka dalam penulisannya dibagi kedalam tahapan-tahapan dimana satu bab dengan bab lain merupakan suatu rangkaian yang saling melengkapi.

Sistematika tersebut adalah sebagai berikut :

BAB I. PENDAHULUAN

Dalam bab ini akan dijelaskan tentang latar belakang masalah, perumusan masalah, pembatasan masalah, tujuan, manfaat, metodologi dan sistematika penulisan tugas akhir ini.

BAB II. LANDASAN TEORI

Bab ini merupakan bagian yang berisi konsep dan prinsip dasar yang digunakan sebagai dasar pemikiran untuk membahas permasalahan yang ada, diantaranya menjelaskan tentang teori distribusi Chi Kuadrat dan sifat-sifatnya dilihat dari penurunan fungsi distribusi, fungsi pembangkit momen (fungsi kumulan, sifat penjumlahan dan grafik fungsi distribusi).

BAB III. APLIKASI DISTRIBUSI CHI KUADRAT

Pada bab tiga ini akan disajikan beberapa bentuk aplikasi dari distribusi Chi Kuadrat. Bentuk aplikasi yang dibahas tersebut antara lain Uji *Goodness Of Fit* (untuk disrtibusi normal), Uji Homogenitas dan Uji Independensi pada data tabel kategorik $a \times w$ (khususnya data tabel kategorik 2×2).

BAB IV. CONTOH KASUS APLIKASI DISTRIBUSI CHI KUADRAT

Bab ini menyajikan contoh kasus dari bentuk aplikasi dari distribusi Chi Kuadrat yang telah dibahas pada bab tiga berikut penyelesaiannya.

BAB V. KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan saran secara keseluruhan dari hasil tugas akhir yang telah disusun.

LAMPIRAN

Lampiran digunakan untuk menjelaskan data atau keterangan lain yang sifatnya terlalu terperinci untuk dimuat dibagian utama skripsi misalnya formulir, surat keterangan, data pengamatan, perhitungan-perhitungan, tabel-tabel, gambar-gambar sebagai penunjang, contoh hasil penelitian dan pengamatan.

BAB II

DISTRIBUSI CHI KUADRAT

2.1 SEJARAH DISTRIBUSI CHI KUADRAT

Sampai sekitar tahun 1960 asal mula dari distribusi Chi Kuadrat sering kali dihubungkan dengan Helmert, seorang ahli Fisika Matematik berkebangsaan Jerman yang terutama mendalami bidang Geodesi, meskipun sesungguhnya Ernst Abbe dan Irenne Jules Bienayme sudah terlebih dahulu menemukan distribusi ini.

Pada tahun 1876 Helmert membuktikan bahwa $Q = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{\sigma^2}$ berdistribusi $\chi^2_{(n-1)}$ ($n-1$ menunjukkan derajat bebasnya), dengan x_1, x_2, ...x_n adalah sampel random dari populasi Normal yang mempunyai mean μ dan variansi σ^2 yang tidak diketahui dan $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$. Untuk mencari distribusi dari Q, Helmert memperkenalkan suatu matrik transformasi ortogonal yang sampai sekarang dikenal matrik Helmert.

Sebelum Helmert menemukan distribusi ini, Bienayme seorang kebangsaan Perancis sudah lebih dahulu menemukan bahwa jumlah kuadrat dari n variabel random IID *(Independent and Identically Distributed)* Normal (0,1) merupakan variabel random yang mempunyai fungsi kepadatan probabilitas Chi Kuadrat.

Kemudian pada tahun 1866 Ernst Abbe, seorang berkebangsaan Jerman dalam usahanya untuk membuat mikroskop, pertama-tama dia menunjukkan bahwa kuantitas $\Delta = \sum_{i=1}^{n} Z_i^2$ dengan Z_i (i = 1,2 ... n), IID N(0,1) adalah variabel random berdistribusi Chi Kuadrat dengan derajat bebas n.

Pada tahun 1990, Karl Pearson menggunakan distribusi tersebut sebagai alat untuk menguji Goodness of Fit yang kemudian dikenal dengan Chi Kuadrat *Goodness of Fit Test.*

2.2 DISTRIBUSI CHI KUADRAT DAN SIFAT-SIFATNYA

Distribusi Chi Kuadrat terbagi menjadi dua bentuk yaitu : Chi Kuadrat Terpusat *(Central Chi Square Distribution)* atau yang sering disebut sebagai distribusi Chi Kuadrat (tanpa kata terpusat) dan Distribusi Chi Kuadrat Tidak Terpusat *(Non Central Chi Square Distribution)*. Akan dilihat bahwa distribusi Chi Kuadrat Terpusat merupakan bentuk khusus dari distribusi Chi Kuadrat Tidak Terpusat.

2.2.1 DISTRIBUSI CHI KUADRAT TERPUSAT

2.2.1.1 Penurunan Fungsi Distribusi

Definisi 2.2.1.1.1 : jika x adalah suatu variabel random berdistribusi Chi Kuadrat Terpusat dengan derajat bebas n, fungsi kepadatan probabilitas (fkp) dari x adalah

$$f(x) = \frac{1}{2^{n/2}\,\Gamma(n/2)}\, x^{n/2-1}\, e^{(-1/2)x} \qquad ; \qquad x > 0$$

dengan $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$

Teorema 2.2.1.1.1 :

Jika x_1 dan x_2 adalah variabel random independen yang masing-masing berdistribusi Normal dengan mean 0 dan variansi 1 (Distribusi Normal Standar N(0,1)), maka $x_1^2 + x_2^2 = r$ berdistribusi Chi Kuadrat dengan derajat bebas 2

Bukti :

$$f(x_i) = \frac{1}{\sqrt{2\pi}} e^{(-1/2)x_i^2} \qquad ; \quad x_i > 0 \text{ dan } i = 1,2$$

Karena x_1 dan x_2 saling independen, maka

$$\begin{aligned} f(x_1, x_2) &= f(x_1) f(x_2) \\ &= \frac{1}{2\pi} e^{(-1/2)(x_1^2 + x_2^2)} \end{aligned}$$

Dilakukan transformasi $x_1 = r^{1/2} \cos\theta$ ....(2.1)

$x_2 = r^{1/2} \sin\theta$ ....(2.2)

Dari (2.1) dan (2.2) diperoleh :

$r = x_1^2 + x_2^2 \qquad 0 < r < \infty$

$\operatorname{tg}\theta = \dfrac{x_2}{x_1} \qquad 0 < \theta < 2\pi$

Bukti :

Fungsi probabilitas bersama dari x_i (i = 1, 2, ..., n) adalah

$$f(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{(-1/2) x_i^2}$$

$$= \frac{1}{(\sqrt{2\pi})^n} e^{(-1/2)\sum_{i=1}^{n} x_i^2}$$

Dilakukan transformasi

$x_1 = r^{1/2} \cos\theta_1 \cos\theta_2 \ldots \cos\theta_{n-1}$

$x_2 = r^{1/2} \cos\theta_1 \cos\theta_2 \ldots \cos\theta_{n-2} \sin\theta_{n-1}$

$x_j = r^{1/2} \cos\theta_1 \cos\theta_2 \ldots \cos\theta_{n-j} \sin\theta_{n+(j-1)}$

. ...

. ...

. ...

$x_{n-1} = r^{1/2} \cos\theta_1 \sin\theta_2$

$x_n = r^{1/2} \sin\theta_1$

$$J = \frac{\delta(x_1, x_2, \ldots, x_n)}{\delta(r, \theta_1, \ldots \theta_{n-1})}$$

hasilnya sama dengan $\frac{1}{2} r^{\frac{1}{2}n-1}$ dikalikan dengan

$$\begin{vmatrix} \cos\theta_1\cos\theta_2 \ldots \cos\theta_{n-1} & \cos\theta_1\cos\theta_2 \ldots \cos\theta_{n-2} \sin\theta_{n-1} & \cdots & \sin\theta_1 \\ -\sin\theta_1 \cos\theta_2 \ldots \cos\theta_{n-1} & -\sin\theta_1 \cos\theta_2 \ldots \cos\theta_{n-2} \sin\theta_{n-1} & \cdots & \cos\theta_1 \\ -\cos\theta_1 \sin\theta_2 \ldots \cos\theta_{n-1} & -\cos\theta_1 \sin\theta_2 \ldots \cos\theta_{n-2} \sin\theta_{n-1} & \cdots & 0 \\ \ldots\ldots & \ldots\ldots & \cdots & \cdots \\ -\cos\theta_1 \cos\theta_2 \ldots \sin\theta_{n-1} & \cos\theta_1 \cos\theta_2 \ldots \cos\theta_{n-1} & \cdots & 0 \end{vmatrix}$$

$J = 1/2\, r^{(1/2)n-1} \cos^{n-1}\theta_1 \cos^{n-2}\theta_2 \ldots \cos\theta_{n-1} \sin\theta_1 \sin\theta_2 \ldots \sin\theta_{n-1}$ dikalikan dengan

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ -\mathrm{tg}\,\theta_1 & -\mathrm{tg}\,\theta_1 & -\mathrm{tg}\,\theta_1 & \cdots & \cot\theta_1 \\ -\mathrm{tg}\,\theta_2 & -\mathrm{tg}\,\theta_2 & -\mathrm{tg}\,\theta_2 & \cdots & 0 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots & \cdots \\ -\mathrm{tg}\,\theta_{n-2} & -\mathrm{tg}\,\theta_{n-2} & -\cot\theta_{n-2} & \cdots & 0 \\ -\mathrm{tg}\theta_{n-1} & -\cot\theta_{n-1} & -\mathrm{tg}\,\theta_{n-1} & \cdots & 0 \end{vmatrix}$$

$$J = \left((1/2) r^{(1/2)n-1} \cos^{n-1}\theta_1 \cos^{n-2}\theta_2 \ldots \cos\theta_{n-1} \sin\theta_1 \sin\theta_2 \ \ldots\ \sin\theta_{n-1}\right)$$
$$\left(\mathrm{tg}\,\theta_{n-1}\ \mathrm{tg}\ \theta_{n-2} \ldots \mathrm{tg}\,\theta_2 \mathrm{tg}\,\theta_1\right) + \left(\cot\theta_{n-1}\ \cot\theta_{n-2} \ldots \cot\theta_2 \cot\theta_1\right) \qquad \ldots\ldots\ldots\ldots(2.3)$$

$$\cot\theta_i + \mathrm{tg}\,\theta_i = \frac{\cos\theta_i}{\sin\theta_i} + \frac{\sin\theta_i}{\cos\theta_i} = \frac{1}{\cos\theta_i \sin\theta_i} \qquad ;\ 1 \le i \le n-1 \ldots\ldots\ldots\ldots(2.4)$$

Hasil dari (2.4) didistribusikan ke (2.3) diperoleh :

$J = 1/2 r^{(1/2)n-1}$ dikalikan dengan

$$\left[\frac{\cos^{n-1}\theta_1 \cos^{n-2}\theta_2 \ \ldots\ \cos\theta_{n-1} \sin\theta_1 \sin\theta_2 \ \ldots\ \sin\theta_{n-1}}{\cos\theta_1 \cos\theta_2 \ \ldots\ \cos\theta_{n-1} \sin\theta_1 \sin\theta_2 \ \ldots\ \sin\theta_{n-1}}\right]$$

sehingga :

$$g(r,\theta_1,\theta_2,\ldots,\theta_{n-1}) = \frac{1}{\left(\sqrt{2\pi}\right)^n} e^{-(1/2)r} \frac{1}{2} r^{(1/2)n-1} \cos^{n-2}\theta_1 \cos^{n-3}\theta_2 \ \ldots\ \cos\theta_{n-2}$$

$$0 < r < \infty\ ;\ 0 \le \theta_{n-1} \le 2\pi\ ;\ -\frac{1}{2}\pi \le \theta_i \le \frac{1}{2}\pi\ ;\ \ i = 1,2,\ldots n-2$$

Sesuai dengan sifat fungsi probabilitas, maka

$$\int \ldots \int \frac{1}{\left(\sqrt{2\pi}\right)^n} e^{-(1/2)r} \frac{1}{2} r^{(1/2)n-1} \cos^{n-2}\theta_1 \cos^{n-3}\theta_2 \ \ldots\ \cos\ \theta_{n-2}\, dr\, d\theta_1\, d\theta_2 \ldots d\theta_{n-1} = 1$$

dengan menganggap $\frac{1}{\left(\sqrt{2\pi}\right)^n} \frac{1}{2} \int \cos^{n-2}\theta_1\, d\theta_1 \ldots \int d\theta_{n-1}$

sebagai suatu konstanta k maka

$$J = \begin{vmatrix} \dfrac{\delta x_1}{\delta r} & \dfrac{\delta x_2}{\delta r} \\ \dfrac{\delta x_1}{\delta \theta} & \dfrac{\delta x_2}{\delta \theta} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{1}{2} r^{-1/2} \cos\theta & \dfrac{1}{2} r^{-1/2} \sin\theta \\ -r^{-1/2} \sin\theta & r^{-1/2} \cos\theta \end{vmatrix} = \frac{1}{2}$$

Fungsi probabilitas bersama (r,θ) dinyatakan dengan

$$g(r,\theta) = \frac{1}{2}\frac{1}{2\pi} e^{(-1/2)\,r}$$

Fungsi probabilitas marginal untuk r yaitu $g_1(r)$ diperoleh

$$g_1(r) = \frac{1}{2}\int_0^{\pi} \frac{1}{2\pi} e^{(-1/2)r} d\theta$$

$$= \frac{1}{2} e^{(-1/2)r}$$

$$= f(x) = \frac{1}{2^{2/2}\,\Gamma(2/2)} r^{2/2-1}\, e^{(-2/2)r} \quad ; \quad x > 0$$

Jadi $r = x_1{}^2 + x_2{}^2$ berdistribusi $\chi^2_{(2)}$.

Teorema dimuka dapat diperluas untuk n variabel seperti pada teorema berikut :

Teorema 2.2.1.1.2 :

Jika $x_1, x_2, \ldots x_n$, adalah variabel random independen masing-masing berdistribusi Normal Standar maka, $\sum_{i=1}^{n} x_i^2 = r$ berdistribusi Chi Kuadrat dengan derajat bebas n.

$$k\int_0^{\infty} e^{-(1/2)r}\, r^{(1/2)n-1}\, dr = 1$$

diperoleh harga $k = \dfrac{1}{2^{n/2}\,\Gamma(n/2)}$

sehingga $g(r) = \dfrac{1}{2^{n/2}\,\Gamma(n/2)} e^{-(1/2)r}\, r^{(1/2)n-1}$

Terbukti bahwa r berdistribusi $\chi^2_{(n)}$

2.2.1.2 Fungsi Pembangkit Momen

Jika x adalah suatu variabel random berdistribusi Chi Kuadrat dengan derajat bebas n, maka fungsi pembangkit momen (fpm) dari x dinyatakan dengan

$$\begin{aligned}
M_x(t) &= E\left(e^{tx}\right) \\
&= \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} e^{tx} e^{-\left(\frac{1}{2}\right)x} x^{\left(\frac{1}{2}\right)n-1}\, dx \\
&= \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\left(\frac{1}{2}\right)n-1} e^{-\frac{(1-2t)x}{2}}\, dx \\
&= (1-2t)^{-\left(\frac{1}{2}\right)n}
\end{aligned}$$

Untuk mendapatkan nilai $M_x(t) = (1-2t)^{-\left(\frac{1}{2}\right)n}$ secara rinci lihat pada lampiran 1.

- **Fungsi Kumulan**

Fungsi kumulan dari distribusi Chi Kuadrat dinyatakan dengan :

$$\begin{aligned}
K_{(x)}(t) &= \ln M_x(t) = -(1/2)n \ln(1-2t) \\
&= (1/2)n\left[2t + \frac{(2t)^2}{2} + \frac{(2t)^3}{3} + \frac{(2t)^4}{4} + \ldots\right]
\end{aligned}$$

Sehingga : k_1 = koefisien t dalam $k_x(t) = n$ (lihat lampiran 2)

$$k_2 = \text{koefisien } \frac{t^2}{2!} \text{ dalam } K_x(t) = 2n$$

$$k_3 = \text{koefisien } \frac{t^3}{3!} \text{ dalam } K_x(t) = 8n$$

$$k_4 = \text{koefisien } \frac{t^4}{4!} \text{ dalam } K_x(t) = 48n$$

Secara umum diperoleh :

$$\text{Kumulan ke - r} = K_r = \text{koefisien} \frac{t^r}{r!} \text{dalam} K_x(t)$$
$$= n2^{r-1}(r-1)!$$

Maka didapat :

$$\text{Mean } = \mu_1 = k_1 = n$$

$$\text{Variansi } = \mu_2 = k_2 = 2n$$

$$\mu_3 = k_3 = 8n$$

$$\mu_4 = k_4 + 3k_2^{\,2} = 48n + 12n^2$$

dan seterusnya

Momen ke r terhadap titik asal $(\mu' r)$

$$\mu_1' = k_1 = n$$
$$\mu_2' = k_2 + k_1^{\,2} = 2n + n^2$$
$$\mu_3' = k_3 + 3k_2 k_1 + k_1^{\,3} = 8n + 6n^2 + n^3$$

dan seterusnya

dan $z^2 = \frac{(x-\mu)^2}{\sigma^2}$ berdistribusi $\chi^2_{(1)}$. Sehingga jika $x_i(i = 1, 2, ..., k)$ adalah variabel random IID $N(\mu, \sigma^2)$, maka $\sum_{i=1}^{k} z^2 = \frac{(x_i-\mu)^2}{\sigma^2}$ berdistribusi $\chi^2_{(k)}$.

2. Pembicaraan pada nomor 1 dapat diperluas. Misal x_i adalah variabel random IID berdistribusi Normal dengan mean μ_x dan variansi $\sigma_x^{\ 2}$ $(i = 1,2,...,n)$ dan y_j adalah variabel random IID berdistribusi Normal dengan mean μ_y dan variansi $\sigma_y^{\ 2}$ $(j = 1,2,...,m)$. Masing-masing x_i dan y_j saling independen. Untuk variabel random x_i, maka

$$\sum_{i=1}^{n} z_i^{\ 2} = \sum_{i=1}^{n} \frac{(x_i - \mu_x)^2}{\sigma_x^2} \text{ berdistribusi } \chi^2_{(n)} \qquad (2.5)$$

Untuk variabel random y_j, maka

$$\sum_{j=1}^{m} w_j = \sum_{j=1}^{m} \frac{(y_j - \mu_y)}{\sigma_y^2} \text{ berdistribbusi } \chi^2_{(m)} \qquad (2.6)$$

(2.5) dan (2.6) dijumlahkan diperoleh jumlah dari $(n+m)$ variabel random berdistribusi Chi Kuadrat. Akibatnya $\sum_{i=1}^{n} z_i^{\ 2} + \sum_{j=1}^{m} w_j^{\ 2}$ berdistribusi Chi Kuadrat dengan derajat bebas $(n+m)$.

3. Misal x berdistribusi $N(\mu, \sigma^2)$ didefinisikan variansi sampel berukuran n untuk x adalah : $s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$. Jika kedua ruas dikalikan dengan n kemudian dibagi dengan σ^2 diperoleh $\frac{ns^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sigma^2}$. Misal diberikan mean dari populasi adalah μ, maka

$$\sum_{i=1}^{n}(x_i-\mu)^2=\sum_{i=1}^{n}[(x_i-\bar{x})+(\bar{x}-\mu)]^2$$
$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+(\bar{x}-\mu)^2+2\sum_{i=1}^{n}(x_i-\bar{x})^2(\bar{x}-\mu)^2$$

Karena $(\bar{x}-\mu)^2$ adalah konstanta dan $\sum_{i=1}^{n}(x_i-\bar{x})^2=0$, maka

$$2\sum_{i=1}^{n}(x_i-\bar{x})(\bar{x}-\mu)=2(\bar{x}-\mu)\sum_{i=1}^{n}(x_i-\bar{x})=0$$

sehingga $\sum_{i=1}^{n}(x_i-\mu)^2=\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{i=1}^{n}(\bar{x}-\mu)^2$
$$=\sum_{i=1}^{n}(x_i-\bar{x})^2+n(\bar{x}-\mu)^2$$

Masing-masing ruas dibagi dengan σ^2 diperoleh

$$\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}=\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{\sigma^2}+\frac{n(\bar{x}-\mu)^2}{\sigma^2}$$

Karena x berdistribusi $N(\mu,\sigma^2)$, maka $\bar{x}$ berdistribusi $N(\mu,\sigma^2/n)$

Sehingga : $\frac{n(\bar{x}-\mu)^2}{\sigma^2}=\frac{(\bar{x}-\mu)^2}{\sigma^2/n}$ berdistribusi $\chi^2_{(1)}$

Pada persamaan (2.1) sudah didapat bahwa $\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}$ berdistribusi $\chi^2_{(n)}$.

Akibatnya $\chi^2_{(n)}=\frac{ns}{\sigma^2}+\chi^2_{(1)}$, dengan sifat penjumlahan didapatkan bahwa $\frac{ns}{\sigma^2}$ berdistribusi $\chi^2_{(n-1)}$.

Bentuk dari akibat 3 mirip dengan bentuk pada akibat 1. Perbedaannya hanyalah μ yang ada pada akibat 1 diganti dengan $\bar{x}$ pada akibat 3, sehingga substitusi $\bar{x}$ untuk μ mengakibatkan berkurangnya satu derajat bebas (dari n menjadi n -1). Hal ini terjadi karena dengan penggantian μ menjadi $\bar{x}$ ada satu konstrain linier yang harus dipenuhi yaitu $\sum_{i=1}^{n}(x_i-\bar{x})=0$ dimana konstrain

$$f''(x) = f'(x)\left[\frac{\frac{1}{2}n-1}{x} - \frac{1}{2}\right] + f(x)\left[-\frac{\frac{1}{2}n-1}{x^2}\right] \quad \text{.......................................(2.9)}$$

dari (2.8) disubstitusikan ke (2.9) diperoleh :

$$f''(n-2) = -\frac{1}{4}\frac{1}{\left(\frac{1}{2}n-1\right)}f(n-2) < 0$$

sehingga harga ektrem dari fungsi distribusi Chi Kuadrat adalah harga ektrem maximum yang dicapai pada nilai $x = n-2$. (Modus distribusi Chi Kuadrat adalah $(n-2)$).

Dimuka telah diperolah bahwa $f'(x) = f(x)\left[\frac{\frac{1}{2}n-1}{x} - \frac{1}{2}\right]$. Karena $x > 0$ dan $f(x)$ adalah fungsi kepadatan probabilitas yang selalu bernilai positif, maka dari $f'(x)$ diperoleh :

a) Untuk $n \leq 2$, maka $f'(x) < 0$ untuk setiap x.

   Sehingga untuk derajat bebas 1 dan 2 , grafik distribusi Chi Kuadrat monoton turun.

b) Untuk $n > 2$ maka $f'(x) = \begin{bmatrix} > 0 \text{ jika } x > (n-2) \\ = 0 \text{ jika } x = (n-2) \\ < 0 \text{ jika } x < (n-2) \end{bmatrix}$

   yang berakhibat untuk $n > 2$:

   - Grafik f(x) monoton naik untuk $0 < x < (n-2)$.
   - Grafik f(x) monoton turun untuk $(n-2) < x < \infty$.

tambahan ada satu lagi ukuran yang disebut oleh Karl Pearson sebagai kurtosis . dengan kurtosis dapat diperoleh gambaran tentang keruncingan dan kedataran grafik. Kurtosis ini diukur dengan $\beta_2 = \frac{\mu_4}{\mu_2^2}$ atau $\gamma_2 = \beta_2 - 3$

Kurtosis dari distribusi Chi Kuadrat adalah :

$$\beta_2 = \frac{48n + 12n^2}{(2n)^2} = \frac{12}{n} + 3$$

$$\gamma_2 = \frac{12}{n}$$

distribusi Normal mempunyai $\beta_2 = 3$ atau $\gamma_2 = 0$, sehingga grafik distribusi Chi Kuadrat lebih runcing jika dibandingkan dengan distribusi Normal.

2.2.2 DISTRIBUSI CHI KUADRAT TIDAK TERPUSAT

Telah dilihat bahwa jika $x_i (i = 1.2,...n)$ adalah variabel random berdistribusi N(0,1), maka $\sum_{i=1}^{n} x_i^2$ berdistribusi Chi Kuadrat atau lengkapnya berdistribusi Chi Kuadrat Terpusat. Sekarang akan diperlihatkan distribusi dari $\sum_{i=1}^{n} x_i^2$ dengan x_i adalah variabel random independen berdistribusi $N(\mu_i, 1)$. Perbedaannya yang jelas terlihat bahwa mean dari x_i adalah μ_i bukan 0. Distribusi dari $\sum_{i=1}^{n} x_i^2$ dikenal sebagai distribusi Chi Kuadrat Tidak Terpusat. Seperti pada distribusi Chi Kuadrat Terpusat, maka distribusi Chi Kuadrat Tidak Terpusat juga melibatkan derajat bebas. Selain itu juga melibatkan parameter

$\frac{1}{2}\sum_{i=1}^{n}\mu_i^2$ yang dikenal sebagai parameter ketidakterpusatan, yang dikembangkan dengan λ. Distribusi Chi Kuadrat Tidak Terpusat dengan derajat bebas n dan parameter ketidakterpusatan λ dilambangkan dengan $\chi'^2_{(n,\lambda)}$. Jika $\mu_i = 0$ untuk semua i, sehingga $\lambda = 0$, distribusi ini menjadi Chi Kuadrat Terpusat. Fungsi kepadatan probabilitas dari $\chi'^2_{(n,\lambda)}$, adalah : $f(x) = \sum_{j=0}^{\infty}\frac{e^{-\lambda}\lambda^j}{j!}f(\chi^2_{(n+2j)})$

dengan $f(\chi^2_{(n+2j)})$ adalah fungsi kepadatan probabilitas distribusi Chi Kuadrat Terpusat dengan derajat bebas $(n+2j)$.

2.2.2.1 Penurunan Fungsi Distribusi

Disini diturunkan fungsi kepadatan probabilitas distribusi Chi Kuadrat Tidak Terpusat melalui fungsi pembangkit momen. Jika x berdistribusi $N(\mu,1)$ maka $M_{x^2}(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx^2} e^{-\frac{(x-\mu)^2}{2}}dx$. Terlebih dahulu diperlihatkan :

$$\begin{aligned}
\exp\left(tx^2 - \frac{(x-\mu)^2}{2}\right) &= \exp\left(tx^2 - \frac{x^2 - 2\mu x + \mu^2}{2}\right) \\
&= \exp-\left(\frac{1-2t}{2}\right)\left(x^2 - \frac{2\mu x}{1-2t} + \frac{\mu^2}{1-2t}\right) \\
&= \exp-\left(\frac{1-2t}{2}\right)\left(\left(x - \frac{\mu}{1-2t}\right)^2 + \frac{\mu^2}{1-2t} - \frac{\mu^2}{(1-2t)^2}\right) \\
&= \exp-\left(\frac{1-2t}{2}\right)\left(x - \frac{\mu}{1-2t}\right)^2 \exp\left(\frac{1-2t}{2}\right)\left(\frac{\mu^2(1-2t)-\mu^2}{(1-2t)^2}\right) \\
&= \exp\left(-\frac{t\mu^2}{1-2t}\right)\exp-\left[\left(\frac{1-2t}{2}\right)\left(x - \frac{\mu}{1-2t}\right)^2\right]
\end{aligned}$$

$$\text{Sehingga}: M_{x^2(t)} = \left\{\exp\left(\frac{t\mu^2}{1-2t}\right)\right\}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp-\left(\left(\frac{1-2t}{2}\right)\left(x-\frac{\mu}{1-2t}\right)^2\right)dx$$

$$= \left\{\exp\left(\frac{t\mu^2}{1-2t}\right)\right\}\frac{1}{\sqrt{1-2t}}\int_{-\infty}^{\infty}\frac{\sqrt{1-2t}}{\sqrt{2\pi}}\exp-\left(\left(\frac{1-2t}{2}\right)\left(x-\frac{\mu}{1-2t}\right)^2\right)$$

$$= \frac{1}{\sqrt{1-2t}}\exp\left(\frac{t\mu^2}{1-2t}\right) = (1-2t)^{-(\frac{1}{2})}\exp\left(\frac{t\mu^2}{1-2t}\right)$$

Jika $x_i (i = 1,2,\ldots,n)$ iid $N(\mu_i, 1)$ maka fpm dari variabel random $u = \sum_{i=1}^{n} x_i^2$ adalah :

$$M_u(t) = M_{\sum_{i=1}^{n} x_i^2}(t) = \prod_{i=1}^{n} M_{x_i^2}(t)$$

$$= (1-2t)^{-(\frac{n}{2})}\exp\left(\frac{t}{1-2t}\sum_{i=1}^{n}\mu_i^2\right)$$

$$= (1-2t)^{-(\frac{n}{2})}\exp\left(\frac{2\lambda t}{(1-2t)}\right) \text{dengan } \lambda = \frac{1}{2}\sum_{i=1}^{n}\mu_i^2$$

atau dapat juga ditulis :

$$M_u(t) = (1-2t)^{-(\frac{n}{2})}\exp\left(\lambda(-1+\frac{1}{1-2t})\right)$$

$$= (1-2t)^{-(\frac{n}{2})}e^{-\lambda}\exp\left(\frac{\lambda}{1-2t}\right)$$

$$= (1-2t)^{-(\frac{n}{2})}e^{-\lambda}\sum_{j=0}^{\infty}\left(\frac{\lambda}{1-2t}\right)^j\frac{1}{j!}$$

$$= \sum_{j=0}^{\infty}\frac{e^{-\lambda}\lambda^j}{j!}(1-2t)^{-(\frac{1}{2})(n+2j)} \qquad ; \; 2t < 1$$

Sehingga dengan menggunakan teorema ketunggalan untuk fpm, maka fkp dari u adalah :

$$f(u) = \sum_{j=0}^{\infty}\frac{e^{-\lambda}\lambda^j}{j!}f(\chi^2_{(n+2j)}), \text{ dengan}$$

$$f(\chi^2_{(n+2j)}) = \frac{1}{2^{(\frac{1}{2})(n+2j)}\Gamma^{(\frac{1}{2}(n+2j))}}e^{-\frac{1}{2}u}u^{(\frac{1}{2}(n+2j)-1)}$$

2.2.2.3 Fungsi Kumulan

Fungsi kumulan dari distribusi Chi Kuadrat Tidak Terpusat adalah :

$$K_x(t) = \ln M_x(t) = -\frac{n}{2}\ln(1-2t) + 2t\lambda\ (1-2t)^{-1}$$
$$= \frac{n}{2}(2t + \frac{(2t)^2}{2} + \ldots + \frac{(2t)^r}{r} + \ldots) + 2t\lambda\left(1 + 2t + (2t)^2 + \ldots + (2t)^r + \ldots\right)$$
$$= (n+2\lambda)t + (n+4\lambda)t^2 + \ldots + \left(\frac{2^{r-1}}{r}n + 2\lambda\, 2^{r-1}\right)t^r$$

sehingga :

$$k_1 = n + 2\lambda$$
$$k_2 = 2(n+4\lambda)$$
$$k_3 = 8(n+6\lambda)$$
$$k_4 = 48(n+8\lambda)$$

Secara umum diperoleh ke-r:

$$K_r = r!\left(\left(\frac{n}{r}\right) + 2\lambda\right)2^{r-1} = 2^{r-1}(r-1)!(n+2\lambda r)$$

Mean $= \mu_1 = k_1 = n + 2\lambda$

Varian $= \mu_2 = k_2 = 2(n+4\lambda)$

$$\mu_3 = k_3 = 8(n+6\lambda)$$
$$\mu_4 = k_4 + 3k_2{}^2 = 12n^2 + 48n + 96n\lambda + 384\lambda + 192\lambda^2$$

BAB III

APLIKASI DISTRIBUSI CHI KUADRAT

Distribusi Chi Kuadrat mempunyai kegunaan yang sangat luas. Misalkan dalam uji Goodness of Fit yaitu uji yang digunakan untuk mengetahui apakah data sampel yang diperoleh cocok dengan distribusi yang dihipotesis. Kegunaan yang lain dari distribusi Chi Kuadrat misalnya dalam Uji Homogenitas dan Uji Independensi.

Dipandang trial independen n kali, yang untuk setiap trial ada k hasil yang mungkin, yaitu $A_1, A_2, ..., A_k$.

Definisi 3.1 :

Variabel random $n_1, n_2, ..., n_k$ mempunyai distribusi Multinomial dengan n trial dan probabilitas $p_1, p_2, ..., p_k$, jika variabel-variabel itu timbul dari ekperimen yang berikut :

1. Sampel berisi n pengamatan yang independen
2. Setiap pengamatan masuk dalam salah satu kategorik atau sel k yaitu $C_1, C_2, ..., C_k$.
3. Probabilitas sebuah pengamatan masuk dalam sel ke-i adalah p_i, dan dari pengamatan satu terhadap pengamatan yang lain memiliki probabilitas yang selalu sama. Dimana $\sum_{i=1}^{k} p_i = 1$.
4. Frekuensi amatan ditunjukkan oleh $O_1, O_2, ..., O_k$ dimana $\sum_{i=1}^{k} O_i = n$

5. e_i adalah frekuensi harapan yang dihitung dengan menggunakan aturan

   $e_i = np_i \quad ; \; i = 1, 2, ..., k$ dan $\sum_{i=1}^{k} e_i = n$.

6. n_i = banyaknya C_i terjadi dalam n trial :

   $P(n_1, n_2, ..., n_k)$ = P (terjadinya C_i sebanyak n_i kali dalam n trial, i = 1,2,...,k)

   $$= \frac{n!}{n_1! n_2! ... n_k!} P_1^{n_1} P_2^{n_2} ... P_k^{n_k}$$

   dengan $n_1, n_2, ..., n_k$ bilangan bulat positif dengan $\sum_{i=1}^{k} n_i = n$

Teorema 3.1 :

Misal suatu sampel random berukuran n dimana elemen-elemennya tersebar dalam k sel. Maka untuk n besar, variabel random positif :

$$W = \sum_{i=1}^{k} \left(\frac{(n_i - np_i)}{np_i} \right)^2$$

berdistribusi Chi Kuadrat dengan derajat bebas $k-1$, dan p_i adalah probabilitas suatu observasi sampel berada pada sel ke-i, n_i adalah frekuensi observasi, np_i adalah harga harapan sel ke-i, $\sum_{i=1}^{k} p_i = 1$.

Bukti :

Karena $n_1, n_2, ..., n_k$ berdistribusi multinomial dengan probabilitas $p_1, p_2, ..., p_k$, maka dengan menggunakan pendekatan stirling untuk faktorial diperoleh :

$$P = \frac{(\sqrt{2\pi})e^{-n}n^{n+\frac{1}{2}}P_1^{n_1}P_2^{n_2}\ldots P_k^{n_k}}{(\sqrt{2\pi})^k e^{-(n_1+n_2+\ldots+n_k)}n_1^{n_1+\frac{1}{2}}n_2^{n_2+\frac{1}{2}}\ldots n_k^{n_k+\frac{1}{2}}}$$

$$= \frac{e^{-n}n^{n+\frac{1}{2}}\left(\frac{np_1}{n_1}\right)^{n_1+\frac{1}{2}}\left(\frac{np_2}{n_2}\right)^{n_2+\frac{1}{2}}\ldots\left(\frac{np_k}{n_k}\right)^{n_k+\frac{1}{2}}}{(\sqrt{2\pi})^{k-1}e^{-n}n^{n_1+n_2+\ldots+n_k+(\frac{k}{2})}(p_1\ldots p_k)^{(\frac{1}{2})}}$$

$$= C\prod_{i=1}^{k}\left(\frac{np_i}{n_i}\right)^{n_i+\frac{1}{2}} \quad \text{dengan } C = \frac{1}{(\sqrt{2\pi})^{k-1}n^{\frac{(k-1)}{2}}(p_1\ldots p_k)^{\left(\frac{1}{2}\right)}}$$

dan P adalah peluang terjadinya C_i sebanyak n_i kali dalam n trial dengan i = 1,...,k adalah konstanta yang independen dengan n_i.

$$\log P = \log C + \sum_{i=1}^{k}\left(n_i + \frac{1}{2}\right)\log\left(\frac{np_i}{n_i}\right)$$

$$\log(P/C) = \sum_{i=1}^{k}\left(n_i + \frac{1}{2}\right)\log\left(\frac{np_i}{n_i}\right) \quad \text{.......(3.1)}$$

dengan λ_i adalah harga harapan pada sel ke-i, yaitu $E(n_i) = np_i = \lambda_i (i = 1,2,\ldots,k)$

Didefinisikan : $\varphi_i = \frac{n_i - \lambda_i}{\sqrt{\lambda_i}}$

sehingga $n_i - \lambda_i = \varphi_i\sqrt{\lambda_i}$ dan $n_i = \lambda_i + \varphi_i\sqrt{\lambda_i}$ ....(3.2)

Persamaan (3.2) disubstitusikan ke persamaan (3.1) diperoleh

$$\log(P/C) = -\sum_{i=1}^{k}(\lambda_i + \varphi_i\sqrt{\lambda_i} + 1/2\log(1 + \varphi_i/\sqrt{\lambda_i}) \quad \text{.......(3.3)}$$

Diasumsikan φ_i kecil dibandingkan dengan λ_i, maka persamaan (3.3) menjadi :

$$\log(P/C) = -\sum_{i=1}^{k}(\lambda_i + \varphi_i\sqrt{\lambda_i} + 1/2)\left(\frac{\varphi_i}{\sqrt{\lambda_i}} - \frac{1\varphi_i^{2}}{2\lambda_i} + O\left(\frac{1}{\lambda_i^{3/2}}\right)\right)$$

$$= -\left(\sum_{i=1}^{k}\left(\varphi_i\sqrt{\lambda_i} + \right) + (1/2)\varphi_i^2 + O\left(\frac{1}{\lambda_i^{3/2}}\right)\right) \quad \text{..(3.4)}$$

Karena n besar, maka $\lambda_i = np_i$ juga besar sehingga $O\left(\frac{1}{\lambda_i^{1/2}}\right)$ mendekati 0 dan juga $\sum_{i=1}^{k}\varphi_i\sqrt{\lambda_i} = \sum_{i=1}^{k}(n_i - \lambda_i) = 0$ sehingga persamaan (3.4) dapat disederhanakan menjadi :

$$\log(P/C) = -\frac{1}{2}\sum_{i=1}^{k}\varphi_i^2$$

$$P = C\exp\left(-\frac{1}{2}\sum_{i=1}^{k}\varphi_i^2\right)$$

yang menunjukkan bahwa $\varphi_i(1,2,...,k)$ merupakan variabel random independen yang berdistribusi Normal standar. Sehingga :

$$W = \sum_{i=1}^{k}\varphi_i^2 = \sum_{i=1}^{k}\left(\frac{(n_i - np_i)^2}{np_i}\right)$$

adalah jumlah kuadrat variabel random normal standar, sehingga W berdistribusi Chi Kuadrat dengan derajat bebas $(k-1)$. Satu derajat bebas berkurang karena ada satu konstrain linear yaitu $\sum_{i=1}^{k}\varphi_i\sqrt{\lambda_i} = \sum_{i=1}^{k}(n_i - \lambda_i) = 0$ atau $\sum_{i=1}^{k}n_i = \sum_{i=1}^{k}\lambda_i$.

3.1 UJI *GOODNESS OF FIT*

Diambil N observasi yang independen dari suatu variabel random x. Ke-N observasi dikelompokkan dalam c kalas atau sel, dan banyaknya observasi pada masing-masing kalas disajikan dalam tabel seperti berikut :

Tabel 3.1 N observasi dalam kelas

Kelas	1	2	...	c	Jumlah
Frekuensi	O_1	O_2	...	O_c	N

Dengan O_i adalah jumlah observasi pada kelas ke-i untuk i =1, 2, ..., c.

Akan dilakukan uji hipotesis :

H_0 = Fungsi distribusi dari observasi variabel random adalah fungsi distribusi tertentu.

H_1 = Fungsi distribusi dari observasi variabel random berbeda dengan fungsi distribusi tertentu.

Misal P_i adalah probabilitas suatu observasi berada pada kelas ke-i, dengan asumsi bahwa fungsi distribusi tertentu adalah fungsi distribusi dari x didefinisikan.

$$E_i = P_i N \qquad i = 1,2,..,c$$

dengan E_i adalah harga harapan observasi berada pada kelas ke-i jika H_0 benar.

Statistik penguji yang digunakan adalah

$$W = \sum_{i=1}^{c} \frac{(O_i - E_i)^2}{E_i}$$

Dari teorema 3.1 statistik W berdistribusi Chi Kuadrat dengan derajat bebas c – 1. Harga W yang besar umumnya disebabkan oleh besar selisih antara frekuensi observasi O_i dengan harga harapan E_i. Nilai Chi Kuadrat (W) yang kecil (mendekati nol) menunjukkan perbedaan antara frekuensi amatan dan frekuensi harapan yang kecil atau tidak signifikan dan sebaliknya dengan nilai Chi Kuadrat yang besar menunjukkan bahwa frekuensi harapan dan frekuensi amatan berbeda secara signifikan. Sehingga H_0 ditolak jika $W > \chi^2_{(c-1)}$. Dalam melakukan pengujian harus mengukuti peraturan yaitu diperlukan frekuensi harapan paling sedikit 5. Jika frekuensi harapan lebih kecil dari 5, maka sel tersebut harus digabungkan dengan sel yang lain atau sampel harus diperbesar.

➢ **Menghitung Nilai Derajat Kebebasan**

Nilai derajat kebebasan sama dengan jumlah sel (k) dikurangi 1 derajat kebebasan.

Formula untuk derajat kebebasan :

Jumlah derajat kebebasan n dalam uji *Goodness of Fit* diberikan :

a) $n = k - 1$, jika frekuensi harapan dihitung tanpa memperkirakan beberapa parameter populasi dari sampel data.

b) $n = k - 1 - r$, jika frekuensi harapan dihitung dengan memperkirakan parameter populasi r dari sampel data. Seringkali dijumpai kasus dengan probabilitas $p_1, p_2, \ldots, p_c$ tidak sepenuhnya diketahui, tetapi tergantung pada parameter-parameter $\theta_1, \theta_2, \ldots, \theta_r$ yang tidak diketahui $(r < k - 1)$. Jika terjadi yang demikian, maka parameter yang tidak diketahui harus diestimasi dari sampel. Sehingga statistik W berdistibusi $\chi^2_{(k-r-1)}$.

Dalam perhitungan derajat kebebasan selalu mengurangi paling sedikit 1 dari k karena keharusan bahwa frekuensi harapan jumlahnya n. Jadi jika $k-1$ frekuensi harapan diketahui, maka frekuensi yang lainnya dapat diketahui.

3.1.1 UJI *GOODNESS OF FIT* DISTRIBUSI NORMAL

Pada beberapa uji hipotesis sampel yang diambil dianggap barasal dari distribusi normal. Anggapan normalitas kadang-kadang didukung olah teori tertentu namun kadang-kadang dapat juga didukung oleh hasil uji hipotesis tertentu yang dikenal sebagai uji *Goodness of Fit* distribusi normal. Ada beberapa cara untuk Uji Normalitas, salah satu diantaranya adalah dengan metode Chi Kuadrat. Diambil sampel berukuran n kemudian dilakukan uji hipotesis:

H_0 = Sampel diambil dari populasi berdistribusi Normal.
H_1 = Sampel diambil bukan dari populasi berdistribusi Normal.

Digunakan statistik uji :

$$W = \sum_{i=1}^{c} \frac{(O_i - E_i)^2}{E_i}$$

Misal dari observasi diperoleh data sebagai berikut :

Tabel 3.2 Frekuensi amatan dan kelas interval

Kelas interval	Frekuensi
k_1	f_1
...	...
k_c	f_c

Dari data tersebut dicari mean dan deviasi standarnya untuk menghitung harga harapan dari masing-masing kelas interval. Kemudian dihitung harga statistik W. Statistik W ini berdistribusi Chi Kuadrat dengan derajat bebas $c-3$, karena selain $\sum_{i=1}^{c} O_i = \sum_{i=1}^{c} E_i = n$, ada 2 kuantitas lagi yang dihitung dari observasi yaitu mean dan standar deviasi.

3.2 Tabel Kategorik 2×2

Tabel kategorik 2×2 merupakan bentuk khusus tabel kategorik $a \times w$ dengan $a = 2$ dan $w = 2$. Dimana n observasi dan masing-masing observasi diklasifikasikan dalam 2 kategorik, dengan masing-masing kategorik terdiri dari 2 peristiwa. Kategorik yang pertama terdiri dari kejadian A dan A^c dan kategorik kedua terdiri dari kejadian B dan B^c. Tabel kategorik disajikan dalam bentuk :

Tabel 3.3 Tabel Kategorik 2×2

Kejadian	B	B^c	Jumlah
A	a	b	n_1
A^c	c	d	n_2
Jumlah	m_1	m_2	n

Baik untuk Uji Homogenitas maupun Uji Independensi statistik penguji yang digunakan adalah :

$$W = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

yang berdistribusi Chi Kuadrat dengan derajat bebas 1. Statistik W tersebut dapat disajikan dalam bentuk yang lebih sederhana :

$$W = \frac{n(ad - bc)^2}{n_1 n_2 m_1 m_2}$$

Bukti :

Harga harapan untuk sel $(i, j)(i, j = 1.2)$ adalah $E(a) = \frac{m_1 n_1}{n}$; $E(b) = \frac{m_2 n_1}{n}$;

$E(c) = \frac{m_1 n_2}{n}$ dan $E(d) = \frac{m_2 n_2}{n}$.

$$a - E(a) = a - \frac{m_1 n_1}{n} = \frac{ad - bc}{n}$$

Dengan cara yang sama diperoleh :

$$b - E(b) = c - E(c) = d - E(d) = \frac{ad - bc}{n}$$

Hasil yang diperoleh disubstitusikan ke statistik W, maka :

$$W = \frac{(ad - bc)^2}{n^2}\left(\frac{1}{E(a)} + \frac{1}{E(b)} + \frac{1}{E(c)} + \frac{1}{E(d)}\right)$$

Setelah masing-masing harga harapan dimasukkan dan kemudian disederhanakan maka diperoleh $W = \frac{n(ad - bc)^2}{n_1 n_2 m_1 m_2}$ (perhitungan secara lengkap terdapat pada lampiran 4).

- **Koreksi Yate**

Dalam tabel kategorik 2×2 banyaknya derajat bebas adalah 1.jika ada satu nilai E_i yang kecil, maka diadakan penggabungan yang berakhibat statistik W berdistribusi Chi Kuadrat dengan derajat bebas 0. Dalam mengatasi masalah ini digunakan koreksi Yate, dengan rumus sebagai berikut :

$$W = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{\left(| O_{ij} - E_{ij} | - \frac{1}{2} \right)}{E_{ij}}$$

yaitu menambahkan 0,5 pada sel yang mempunyai E_i yang kecil dan frekuensi sel yang lain menyesuaikan sehingga jumlah observasi dari masing-masing kejadian tetap. Kemudian uji hipotesis dilakukan tanpa penggabungan.

3.3 Tabel Kategorik $a \times w$

Tabel kategorik atau tabel kontingensi $a \times w$ adalah susunan bilangan asli yang diatur dalam a baris dan w kolom dan mempunyai $a \times w$ sel, dimana bilangan-bilangan tersebut menunjukkan jumlah atau frekuensi observasi.

Diambil n observasi dan masing-masing observasi diklasifikasikan menurut dua variabel kategorik. Variabel pertama mempunyai a tingkat yaitu $A_1, A_2, \ldots, A_a$ dan variabel kedua mempunyai w tingkat yaitu $B_1, B_2, \ldots, B_w$. Data yang diperoleh disusun sebagai berikut :

Tabel 3.4 Tabel Kategorik $a \times w$

Variabel 2 / Variabel 1	B_1	B_j	B_w	Jumlah
A_1	O_{11}	O_{1j}	O_{1w}	n_1
A_i	O_{i1}	O_{ij}	O_{iw}	n_i
A_a	O_{a1}	O_{aj}	O_{aw}	n_a
Jumlah	m_1	m_j	m_w	n

Dengan O_{ij} adalah banyaknya observasi dengan sifat A_iB_j dimana $i = 1,2,\ldots,a$; $j = 1,2,\ldots,w$ dan $n = \sum_{i=1}^{a} n_i = \sum_{j=1}^{w} m_j$

➢ **Menentukan Derajat Kebebasan**

Jumlah derajat kebebasan dari Chi Kuadrat digambarkan atau ditunjukkan oleh frekuensi harapan yang dapat diambil secara bebas, menghasilkan baris dan kolom total dari tabel frekuensi harapan yang identik dengan baris dan kolom total dari tabel frekuensi amatan. Seandainya diisi semua sel pada baris $a-1$ yang pertama dan kolom $w-1$ dan membiarkan kosong semua sel pada baris dan kolom yang terakhir dari tabel $a \times w$. Proses ini terdiri dari memasukkan nilai $(a-1)(w-1)$ kedalam tabel. Ini akan jelas jika telah mengetahui frekuensi $(a-1)(w-1)$ ini dan juga jika total baris dan kolom diketahui, sehingga sel yang kosong dalam tabel itulah yang satu-satunya dapat ditentukan.Derajat kebebasan untuk tabel kategorik $a \times w$ adalah $(a-1)(w-1)$.

3.3.1 Uji Homogenitas

Akan diuji homogenitas a populasi yang masng-masing berdistribusi binomial. Diambil sampel random dari masing-masing populasi dengan ukuran $n_1, n_2, \ldots, a$. Tabel probabilitas untuk populasi adalah sebagai berikut :

Tabel 3.5 Tabel Probabilitas untuk Populasi $a \times w$

Kejadian / Populasi	B_1	B_j	B_w	Jumlah
1	p_{11}	p_{1j}	p_{1w}	1
i	p_{i1}	p_{ij}	p_{iw}	1
a	p_{a1}	p_{aj}	p_{aw}	1

Untuk menguji bahwa a populasi homogen, maka :

$H_0 : p_{i1} = p_1 ; \ldots ; p_{iw} = p_w$ untuk $i = 1,2,\ldots,a$ dengan $p_1, p_2, \ldots, p_W$ tidak diketahui
$H_1 : p_{i1} \neq p_1 ; \ldots ; p_{iw} \neq p_w$ untuk $i = 1,2,\ldots,a$ dengan $p_1, p_2, \ldots, p_W$ tidak diketahui

Misal E_{ij} adalah harga harapan untuk masing-masing sel, jika H_0 benar

$$E_{ij} = n_i \frac{m_j}{n} \quad ; i = 1,2,\ldots,a ; j = 1,2,\ldots,w$$

Statistik yang digunakan sebagai uji hipotesis adalah :

$$W = \sum_{i=1}^{a} \sum_{j=1}^{w} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Sesuai dengan teorema 3.1 statistik W berdistribusi Chi Kuadrat. Pada tabel kategorik di muka :

$\sum_{i=1}^{a} O_{ij} = n_i ; i = 1, 2, ..., a$

$\sum_{j=1}^{w} O_{ij} = m_j ; j = 1, 2, ..., w$

Sehingga ada $a \times w$ konstrain yang harus dipenuhi. Tetapi karena $\sum_{i=1}^{a} n_i = \sum_{j=1}^{w} m_j = n$ maka jumlah konstrain yang harus dipenuhi berkurang 1 menjadi $a + w - 1$. Dan karena jumlah sel adalah $a \times w$, maka derajat bebasnya adalah $aw - (a + w - 1) = (a - 1)(w - 1)$. Sehingga statistik W berdistribusi $\chi^2_{(a-1)(w-1)}$.

3.3.2 Uji Independensi

Diambil sampel random yang beranggotakan n. Setiap anggota sampel doklasifikasikan dalam 2 kategorik, kejadian $(A_1, A_2, ..., A_a)$ dan $(B_1, B_2, ..., B_w)$. Probabilitas kejadian A_i dan B_j dalam populasi adalah P_{ij}.

Tabel 3.6 Tabel Probabilitas kejadian A_i dan B_j dalam populasi adalah P_{ij}.

Kejadian 1 / Kejadian 2	B_1	B_j	B_w	Jumlah
A_1	P_{11}	P_{1j}	P_{1w}	$P(A_1)$
A_i	P_{i1}	P_{ij}	P_{iw}	$P(A_i)$
A_a	P_{a1}	P_{aj}	P_{aw}	$P(A_a)$
Jumlah	$P(B_1)$	$P(B_j)$	$P(B_w)$	1

Dilakukan uji hipotesis :

$$H_0 = P_{ij} = P(A_i \cap B_j) = P(A_i)\,P(B_j)$$
$$H_1 = P_{ij} \neq P(A_i \cap B_j) \neq P(A_i)\,P(B_j)$$

Harga harapan dari masing-masing sel adalah :

$$E_{ij} = \frac{n_i m_j}{n}\,;\, i = 1,2,...,a \text{ dan } j = 1,2,...,w$$

Statistik uji yang digunakan adalah :

$$W = \sum_{i=1}^{a} \sum_{j=1}^{w} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Seperti juga pada uji Homogenitas, statistik W berdistribusi $\chi^2_{(a-1)(w-1)}$

BAB IV

CONTOH KASUS APLIKASI DISTRIBUSI CHI KUADRAT

Pada bab ini akan diberikan beberapa contoh kasus yang dapat diselesaikan dengan menggunakan bentuk aplikasi dari distribusi Chi Kuadrat. Yang dalam penyelesaiannya kali ini menggunakan bantuan program komputer yaitu SPSS versi 10. Keunggulan dari program SPSS dalam pengolahan data statistik adalah dalam hal kecepatan dan ketepatan khususnya dalam hal perhitungan. Contoh kasus kali ini diambil dari soal–soal yang dimuat dalam buku " Ilmu Peluang Dan Statistika Untuk Insinyur Dan Ilmuwan " pengarang Ronald E Walpole dan Raymond H Myers.

4.1 UJI *GOODNESS OF FIT* DISTRIBUSI NORMAL

Kasus :

Terdapat frekuensi umur baterai mobil yang serupa sebanyak 40 buah, yang terbagi dalam kelas interval (satuan dalam tahun) seperti pada tabel 4.1 berikut :

Test Statistics

	Umur Baterai Mobil
Chi-Square[a]	8.763
df	6
Asymp. Sig.	.187

a. 3 cells (42.9%) have expected frequencies less than 5. The minimum expected cell frequency is .5.

Analisis Output 4.1 :

Karena dari hasil output menunjukkan harga frekuensi harapan ada 3 sel yang kurang dari 5, sehingga sebelum analisis dilanjutkan sesuai dengan ketentuan yang ada dilakukan penggabungan terlebih dahulu. Dan hasilnya seperti pada tabel berikut :

Tabel 4.2 Frekuensi amatan setelah dilakukan penggabungan

Kelas Interval	Frekuensi
1,5 – 2,9	7
3,0 – 3,4	15
3,5 – 3,9	10
4,0 – 4,9	8

Tabel 4.1 Kelas interval dan frekuensi amatan umur baterai

Kelas Interval	Frekuensi
1,5 – 1,9	2
2,0 – 2,4	1
2,5 – 2,9	4
3,0 – 3,4	15
3,5 – 3,9	10
4,0 – 4,4	5
4,5 – 4,9	3

Sumber : Buku "Ilmu Peluang Dan Statistika Untuk Insinyur Dan Ilmuwan", edisi-4.

Dari data tersebut uji hipotesis bahwa distribusi frekuensi umur baterai berdistribusi normal dengan $\mu = 3,5$ dan $\sigma = 0,7$.

Penyelesaian :

Output 4.1 :

Kelas Interval Umur Baterai Mobil

	Observed N	Expected N	Residual
1,5 - 1,9	2	.5	1.5
2,0 - 2,4	1	2.2	-1.2
2,5 - 2,9	4	6.0	-2.0
3,0 - 3,4	15	10.5	4.5
3,5 - 3,9	10	10.9	-.9
4,0 - 4,4	5	7.1	-2.1
4,5 - 4,9	3	2.8	.2
Total	40		

Penyelesaian :

Output 4.2 :

Kelas Interval Umur Baterai Mobil

	Observed N	Expected N	Residual
1,5 - 2,9	7	8.7	-1.7
3,0 - 3,4	15	10.5	4.5
3,5 - 3,9	10	10.9	-.9
4,5 - 4,9	8	9.9	-1.9
Total	40		

Test Statistics

	Umur Baterai Mobil
Chi-Square[a]	2.668
df	3
Asymp. Sig.	.446

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 8.7.

Analisis Output 4.2 :

Tahap-tahap uji hipotesis :

1. H_0 = Umur baterai mobil berdistribusi normal dengan $\mu = 3{,}5$ dan $\sigma = 0{,}7$.
   H_1 = Umur baterai mobil tidak berdistribusi normal dengan $\mu = 3{,}5$ dan $\sigma = 0{,}7$.

2. Dengan tingkat signifikan 0,05.

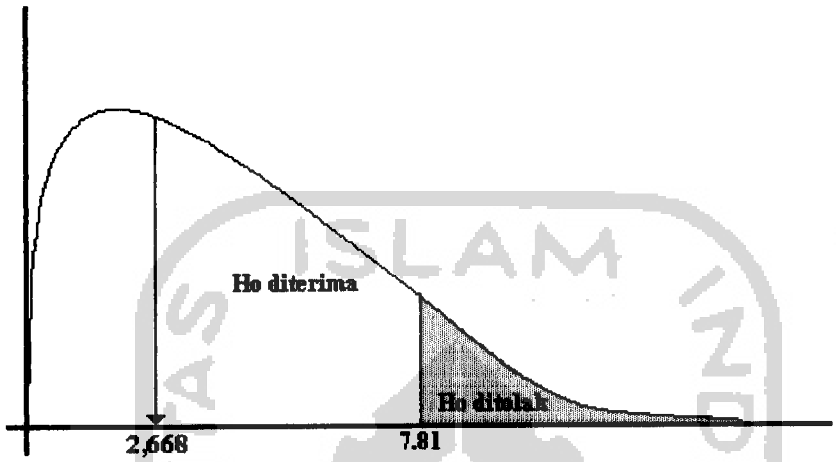
3. Dasar pengambilan keputusan :

   a. Berdasarkan perbandingan Uji Chi Kuadrat dan tabel

   - Jika Chi Kuadrat Hitung < Chi Kuadrat Tabel, maka H_0 diterima.
   - Jika Chi Kuadrat Hitung > Chi Kuadrat Tabel, maka H_0 ditolak.

   Chi Kuadrat Hitung (lihat pada output SPSS) adalah 2,668.

Chi Kuadrat Tabel bisa dihitung pada Tabel Chi Kuadrat, dengan $\alpha = 5\%$ dan derajat bebas 3. Diperoleh Chi Kuadrat Tabel adalah 7,81.



Gambar 4.1 Kurva Chi Kuadrat dengan $\alpha = 5\%$ dan derajat bebas 3

Keputusan : Karena Chi Kuadrat Hitung < Chi Kuadrat Tabel (2,668 < 7,81), maka H_0 ditolak.

b. Berdasarkan probabilitas

- Jika probabilitas > 0,05 maka H_0 diterima.
- Jika probabilitas < 0,05 maka H_0 ditolak.
- Keputusan : Terlihat bahwa pada kolom *Asymp. Sig/Asymptotic significance* adalah 0,446, atau probabilitas di atas 0,05. Maka H_0 diterima.

4. Kesimpulan : Dengan menggunakan tingkat signifikan 0,05 dari kedua analisis di atas, dapat diambil kesimpulan yang sama, yaitu H_0 diterima, atau umur baterai mobil berdistribusi normal dengan $\mu = 3,5$ dan $\sigma = 0,7$.

4.2 UJI HOMOGENITAS

Kasus :

Suatu pusat kesehatan universitas melakukan percobaan untuk menentukan tingkat kesembuhan yang diberikan oleh 3 jenis obat batuk. Tiap obat batuk dicobakan pada 50 mahasiswa dan diperoleh data tabel 4.3 berikut ini :

Tabel 4.3 Frekuensi amatan mahasiswa yang menggunakan obat batuk

	Obat Batuk		
	Nyquil	Robitussin	Triaminic
Tidak Sembuh	11	13	9
Agak Tertolong	32	28	27
Sembuh	7	9	14

Sumber : Buku "Ilmu Peluang Dan Statistika Untuk Insinyur Dan Ilmuwan", edisi-4.

Uji hipotesis, pada tingkat signifikan 0,05, bahwa ketiga obat batuk tersebut sama baiknya (homogen).

Penyelesaian :

Output 4.3

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Efek minum obat * Obat	150	100.0%	0	.0%	150	100.0%

Efek minum obat * Obat Crosstabulation

Count

		OBAT			Total
		Nyquil	Robitussin	Triaminic	
Efek minum obat	Tidak Sembuh	11	13	9	33
	Agak Sembuh	32	28	27	87
	Sembuh	7	9	14	30
Total		50	50	50	150

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.810[a]	4	.432
Likelihood Ratio	3.739	4	.442
Linear-by-Linear Association	1.918	1	.166
N of Valid Cases	150		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.00.

Analisis Output 4.3:

Tahap-tahap uji hipotesis :

1. H_0 = Ketiga jenis obat batuk sama baiknya.

   H_1 = Ketiga jenis obat batuk tidak sama baiknya atau minimal ada satu yang tidak baik.

2. Dengan tingkat signifikan 0,05.

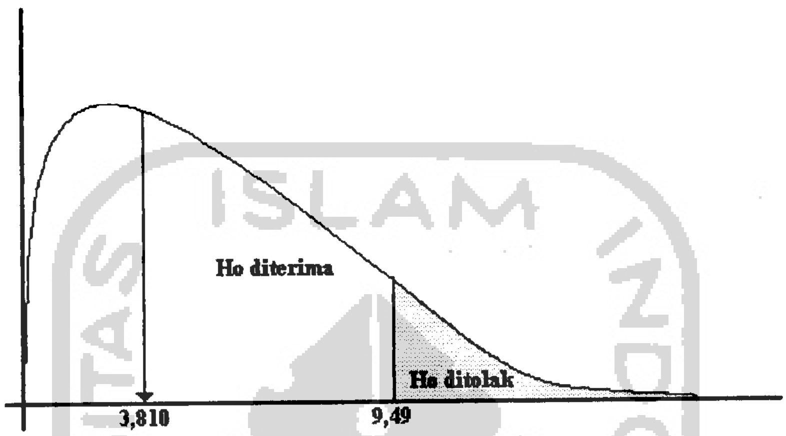
3. Dasar pengambilan keputusan :

   a. Berdasarkan perbandingan Uji Chi Kuadrat dan tabel

      - Jika Chi Kuadrat Hitung < Chi Kuadrat Tabel, maka H_0 diterima.
      - Jika Chi Kuadrat Hitung > Chi Kuadrat Tabel, maka H_0 ditolak.

Chi Kuadrat Hitung (lihat pada output SPSS) adalah 3,810.

Chi Kuadrat Tabel bisa dihitung pada Tabel Chi Kuadrat, dengan $\alpha = 5\%$ dan derajat bebas 4. Diperoleh Chi Kuadrat Tabel adalah 9,49.



Gambar 4.2 Kurva Chi Kuadrat t, dengan $\alpha = 5\%$ dan derajat bebas 4

Keputusan : Karena Chi Kuadrat Hitung < Chi Kuadrat Tabel (3,810 < 9,49), maka H_0 diterima.

c. Berdasarkan probabilitas

- Jika probabilitas > 0,05maka H_0 diterima.
- Jika probabilitas < 0,05 maka H_0 ditolak.
- Keputusan : Terlihat bahwa pada kolom *Asymp. Sig/Asymptotic significance* adalah 0,432, atau probabilitas di atas 0,05. Maka H_0 diterima.

5. Kesimpulan : Dengan menggunakan tingkat signifikan 0,05 dari kedua analisis di atas, dapat diambil kesimpulan yang sama, yaitu H_0 diterima, atau terjadinya ketiga jenis obat batuk sama baiknya.

4.3 UJI INDEPENDENSI

Kasus :

Seorang Kriminolog melakukan sigi untuk menentukan apakah terjadinya berbagai kejahatan tertentu berbeda dari satu bagian ke bagian lain dalam suatu kota besar. Kejahatan yang ingin diselidiki adalah penodongan, pembongkaran, pencurian dan pembunuhan. Tabel 4.2 menunjukkan banyaknya kejahatan yang terjadi di 4 bagian kota tahun lalu.

Tabel 4.4 Frekuensi amatan berbagai jenis kejahatan pada empat daerah dalam suatu kota

Daerah	Jenis Kejahatan			
	Penodongan	Pembongkaran	Pencurian	Pembunuhan
1	162	118	451	18
2	310	196	996	25
3	258	193	458	10
4	280	175	390	19

Sumber : Buku "Ilmu Peluang Dan Statistika Untuk Insinyur Dan Ilmuwan", edisi-4.

Dapatkah disimpulkan dari data tersebut dengan tingkat signifikan 0,01 bahwa dari tiap jenis kejahatan tidak ada hubungannya dengan daerah di kota itu ?

Penyelesaian :

Output 4.4

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Daerah * Jenis Kejahatan	4059	100.0%	0	.0%	4059	100.0%

Daerah * Jenis Kejahatan Crosstabulation

Count

		Jenis Kejahatan				Total
		penodongan	pembongkaran	pencurian	pembunuhan	
Daerah	1	162	118	451	18	749
	2	310	196	996	25	1527
	3	258	193	458	10	919
	4	280	175	390	19	864
Total		1010	682	2295	72	4059

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	124.530[a]	9	.000
Likelihood Ratio	124.822	9	.000
Linear-by-Linear Association	70.594	1	.000
N of Valid Cases	4059		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.29.

Analisis Output 4.4 :

Tahap-tahap uji hipotesis :

1. H_0 = Kejahatan tidak ada hubungannya dengan daerah di kota itu.

   H_1 = Kejahatan tergantung dengan daerah di kota itu.

2. Dengan tingkat signifikan 0,01.

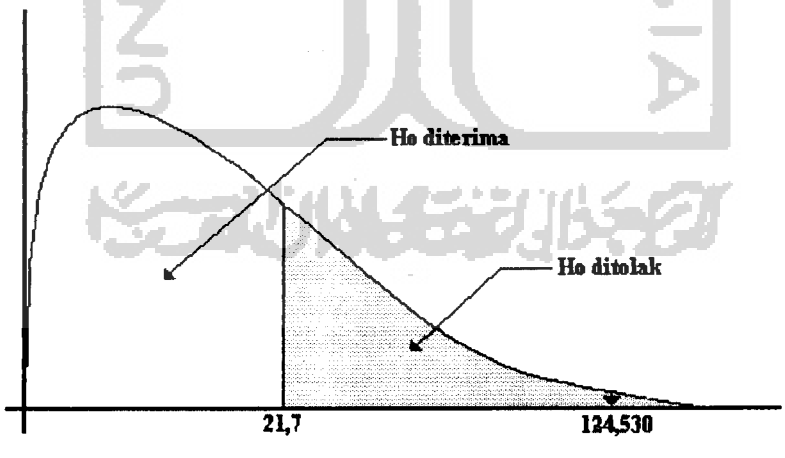
3. Dasar pengambilan keputusan :

   a. Berdasarkan perbandingan Uji Chi Kuadrat dan tabel

      - Jika Chi Kuadrat Hitung < Chi Kuadrat Tabel, maka H_0 diterima.
      - Jika Chi Kuadrat Hitung > Chi Kuadrat Tabel, maka H_0 ditolak.

      Chi Kuadrat Hitung (lihat pada output SPSS) adalah 124,530.

      Chi Kuadrat Tabel bisa dihitung pada Tabel Chi Kuadrat, dengan $\alpha = 1\%$ dan derajat bebas 9. Diperoleh Chi Kuadrat Tabel adalah 21,7.



Gambar 4.3 Grafik Chi Kuadrat dengan $\alpha = 1\%$ dan derajat bebas 9

Keputusan : Karena Chi Kuadrat Hitung > Chi Kuadrat Tabel (124,530 > 21,7), maka H_0 ditolak.

b. Berdasarkan probabilitas

- Jika probabilitas > 0,01 maka H_0 diterima.
- Jika probabilitas < 0,01 maka H_0 ditolak.

Keputusan : Terlihat bahwa pada kolom *Asymp. Sig/Asymptotic significance* adalah 0,000, atau probabilitas di bawah 0,01. Maka H_0 ditolak.

4. Kesimpulan : Dengan menggunakan tingkat signifikan 0,01 dari kedua analisis di atas, dapat diambil kesimpulan yang sama, yaitu H_0 ditolak, atau terjadinya kejahatan tergantung dengan daerah di kota itu.

BAB V

KESIMPULAN DAN SARAN

Sesuai dengan tujuan dari tugas akhir ini yang telah diuraikan pada Bab I yaitu untuk membahas tentang teori Distribusi Chi Kuadrat dan beberapa bentuk aplikasinya, maka untuk selanjutnya pada akhir bab ini penulis akan mengambil suatu kesimpulan dan sedikit saran-saran sehubungan dengan kesimpulan yang diambil.

5.1 KESIMPULAN

Dari hasil pembahasan dapat diambil beberapa kesimpulan :

1. Distribusi Chi Kuadrat dapat dibagi dalam dua bentuk yaitu Distribusi Chi Kuadrat Terpusat dan Distribusi Chi Kuadrat Tidak Terpusat. Dimana kedua bentuk distribusi tersebut sama-sama berasal dari kuadrat distribusi normal, letak perbedaannya adalah jika Distribusi Chi Kuadrat Terpusat berasal dari distribusi normal standar (memiliki $\mu = 0$ dan $\sigma = 1$) sedangkan Distribusi Chi Kuadrat Tidak Terpusat berasal dari distribusi normal dengan $\mu = \mu_i$ (untuk $i = 1, 2, 3, ...$) dan $\sigma = 1$.
2. Pada dasarnya bentuk aplikasi dari Distribusi Chi Kuadrat untuk Uji Independensi hampir sama bentuknya dengan Uji Homogenitas. Perbedaannya adalah jika Uji Independensi didasarkan pada satu sampel dari suatu populasi

sedangkan untuk Uji Homogenitas didasarkan pada a sampel yang diambil dari r populasi.

5.2 SARAN

Saran ini diberikan mengingat keterbatasan penulis dalam menyampaikan materi tentang Distribusi Chi Kuadrat yang belum semua tercakup, sehingga bagi penulis lain yang berminat tugas akhir ini dapat dikembangkan dengan pembahasan bentuk aplikasi dari Distribusi Chi Kuadrat yang lain serta pembahasan hubungan Distribusi Chi Kuadrat dengan distribusi yang lain. Sehingga tugas akhir ini menjadi suatu bentuk tulisan yang sempurna.

DAFTAR PUSTAKA

James, L. Kenkel, 1984, "*Introductory Statistics for Mamagement and Economics Second Edition*", Duxbury Press.

Kapur, J.N. dan H. C. Sawena, 1981, "*Mathematical Statistics*", S. Chand and Company Ltd, New Delhi.

Kendall, M.G, 1958, "*The Advanced Theory of Statistics Vol.1*", Hafner Publishing and Co, New York.

Kenney J.F. dan E.S. Keeping, 1951, "*Matematic of Statistic Part Two*", D. Van Nostrnd Company, Inc.

Rohatgi, V.K, 1984, "*Statistical Inference*", Jonh Willey and Sons, New York.

Santoso, S. , 2000, "*Buku Latihan SPSS Statistik Parametrik*", Penerbit PT Elex Komputindo Kelompok – Jakarta.

Walpole, R.E. and Myers, R.H, "*Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan, Edisi-4*" ,Penerbit ITB.

Lampiran 1

Fungsi Pembangkit Momen jika x adalah suatu variabel random berdistribusi Chi Kuadrat dengan derajat bebas n adalah :

$$M_x(t) = E\left(e^{tx}\right)$$

$$= \frac{1}{2^{n/2}\Gamma(n/2)}\int_0^\infty e^{tx}e^{-(1/2)x}x^{(1/2)n-1}dx$$

$$= \frac{1}{2^{n/2}\Gamma(n/2)}\int_0^\infty x^{(1/2)n-1}e^{-\frac{(1-2t)x}{2}}dx$$

$$\text{misal}: y = \frac{(1-2t)x}{2} \quad ; \quad dx = \frac{2}{(1-2t)x}$$

$$M_x(t) = \frac{1}{2^{n/2}\Gamma(n/2)}\int_0^\infty \left[\frac{2y}{1-2t}\right]^{(1/2)n-1} e^{-y}\frac{2}{(1-2t)}dy$$

$$= \frac{1}{2^{n/2}\Gamma(n/2)}\left[\frac{2}{1-2t}\right]^{(1/2)n-1}\frac{2}{(1-2t)}\int_0^\infty y^{(1/2)n-1}e^{-y}dy$$

$$= \frac{1}{2^{n/2}\Gamma(n/2)}\left[\frac{2}{1-2t}\right]^{(1/2)n}\frac{(1-2t)}{2}\frac{2}{(1-2t)}\int_0^\infty y^{(1/2)n-1}e^{-y}dy$$

$$= \frac{1}{2^{n/2}\Gamma(n/2)^{(1/2)n}}\int_0^\infty y^{(1/2)n-1}e^{-y}dy$$

$$= (1-2t)^{-1/2n}$$

Lampiran 2

Kumulan $k_1, k_2, ..., k_r$ didefinisikan :

$$\exp\left[k_1 t + \frac{k_2 t^2}{2!} + ... + \frac{k_r t^r}{r!} + ...\right] = 1 + \mu_1' + \frac{\mu_2' t^2}{2!} + ... + \frac{\mu_r' t^r}{r!} + ... \qquad ....(1)$$

Dari definisi momen diperoleh :

$$M_x(t) = E(e^{tx}) = E\left[1 + tx + \frac{t^2 x^2}{2!} + ... + \frac{t^r x^r}{r!} + ...\right]$$

$$= 1 + t(E(x)) + \frac{t^2}{2!}(E(x^2)) + ... + \frac{t^r}{r!}(E(x^r)) + ... \qquad ....(2)$$

$$= 1 + \mu_1' t + \frac{\mu_2' t^2}{2!} + ... + \frac{\mu_r' t^r}{r!} + ...$$

sehingga dari (1) dan (2) diperoleh :

$$k_1 t + \frac{k_2 t^2}{2!} + ... + \frac{k_r t^r}{r!} + ... = \ln M_x(t)$$

$$1 + \mu_1' t + \frac{\mu_2' t^2}{2!} + ... + \frac{\mu_r' t^r}{r!} + ... = \exp\left[k_1 t + \frac{k_2 t^2}{2!} + ... + \frac{k_r t^r}{r!} + ...\right]$$

$$= \exp(k_1 t)\exp\left(\frac{k_2 t^2}{2!}\right)...\exp\left(\frac{k_r t^r}{r!}\right)...$$

$$= \left(1 + \frac{k_1 t}{1!} + \frac{k_1^2 t^2}{2!} + ...\right)\left(1 + \frac{k_2 t^2}{2!} + \frac{1}{2!}\left(\frac{k_2^2 t^4}{2!}\right) + ...\right)...$$

$$\left(1 + \frac{k_r t^r}{r!} + \frac{1}{2!}\left(\frac{k_r t^r}{r!}\right) + ...\right)...$$

sehingga :

$\mu_1' = k_1$

$\mu_2' = k_1^2 + k_2$

$\mu_3' = k_1^3 + 3k_1 k_2 + k_3$

dan seterusnya

Lampiran 3

Tabel Chi Kuadrat

	α					
db	0.25	0.1	0.05	0.025	0.01	0.005
1	1.3233	2.7055	3.8415	5.0239	6.6349	7.8794
2	2.7726	4.6052	5.9915	7.3778	9.2104	10.5965
3	4.1083	6.2514	7.8147	9.3484	11.3449	12.8381
4	5.3853	7.7794	9.4877	11.1433	13.2767	14.8602
5	6.6257	9.2363	11.0705	12.8325	15.0863	16.7496
6	7.8408	10.6446	12.5916	14.4494	16.8119	18.5475
7	9.0371	12.0170	14.0671	16.0128	18.4753	20.2777
8	10.2189	13.3616	15.5073	17.5345	20.0902	21.9549
9	11.3887	14.6837	16.9190	19.0228	21.6660	23.5893
10	12.5489	15.9872	18.3070	20.4832	23.2093	25.1881
11	13.7007	17.2750	19.6752	21.9200	24.7250	26.7569
12	14.8454	18.5493	21.0261	23.3367	26.2170	28.2997
13	15.9839	19.8119	22.3620	24.7356	27.6882	29.8193
14	17.1169	21.0641	23.6848	26.1189	29.1412	31.3194
15	18.2451	22.3071	24.9958	27.4884	30.5780	32.8015
16	19.3689	23.5418	26.2962	28.8453	31.9999	34.2671
17	20.4887	24.7690	27.5871	30.1910	33.4087	35.7184
18	21.6049	25.9894	28.8693	31.5264	34.8052	37.1564
19	22.7178	27.2036	30.1435	32.8523	36.1908	38.5821
20	23.8277	28.4120	31.4104	34.1696	37.5663	39.9969
21	24.9348	29.6151	32.6706	35.4789	38.9322	41.4009
22	26.0393	30.8133	33.9245	36.7807	40.2894	42.7957
23	27.1413	32.0069	35.1725	38.0756	41.6383	44.1814
24	28.2412	33.1962	36.4150	39.3641	42.9798	45.5584
25	29.3388	34.3816	37.6525	40.6465	44.3140	46.9280
26	30.4346	35.5632	38.8851	41.9231	45.6416	48.2898
	31.5284	36.7412	40.1133	43.1945	46.9628	49.6450
28	32.6205	37.9159	41.3372	44.4608	48.2782	50.9936
29	33.7109	39.0875	42.5569	45.7223	49.5878	52.3355
30	34.7997	40.2560	43.7730	46.9792	50.8922	53.6719
31	35.8871	41.4217	44.9853	48.2319	52.1914	55.0025
32	36.9730	42.5847	46.1942	49.4804	53.4857	56.3280
33	38.0575	43.7452	47.3999	50.7251	54.7754	57.6483
34	39.1408	44.9032	48.6024	51.9660	56.0609	58.9637
35	40.2228	46.0588	49.8018	53.2033	57.3420	60.2746
36	41.3036	47.2122	50.9985	54.4373	58.6192	61.5811
37	42.3833	48.3634	52.1923	55.6680	59.8926	62.8832
38	43.4619	49.5126	53.3835	56.8955	61.1620	64.1812

Lampiran 4.

Pembuktian $W = \sum_{i=0}^{2}\sum_{j=0}^{2}\left(\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}\right) = \frac{n(ad - bc)^2}{m_1 m_2 n_1 n_2}$

$$E(a) = \frac{m_1 n_1}{n} \quad ; \quad E(b) = \frac{m_2 n_1}{n} \quad ; \quad E(c) = \frac{m_1 n_2}{n} \quad ; \quad E(d) = \frac{m_2 n_2}{n}$$

$$a - E(a) = b - E(b) = c - E(c) = d - E(d) = \frac{ad - bd}{n}$$

$$W = \sum_{i=0}^{2}\sum_{j=0}^{2}\left(\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}\right)$$

$$= \left(\frac{ad - bc}{n}\right)^2\left(\frac{1}{E(a)} + \frac{1}{E(b)} + \frac{1}{E(c)} + \frac{1}{E(d)}\right)$$

$$= \left(\frac{ad - bc}{n}\right)^2\left(\frac{1}{\frac{m_1 n_1}{n}} + \frac{1}{\frac{m_2 n_1}{n}} + \frac{1}{\frac{m_1 n_2}{n}} + \frac{1}{\frac{m_2 n_2}{n}}\right)$$

$$= \left(\frac{ad - bc}{n}\right)^2\left(\frac{n}{m_1 n_1} + \frac{n}{m_2 n_1} + \frac{n}{m_1 n_2} + \frac{n}{m_2 n_2}\right)$$

$$= \frac{(ad - bc)^2}{n^2} n\left(\frac{m_2 n_2 + m_1 n_2 + m_2 n_1 + m_1 n_1}{m_1 m_2 n_1 n_2}\right)$$

$$= \frac{(ad - bc)^2}{n^2} n\left(\frac{n_2(m_1 + m_2) + n_1(m_1 + m_2)}{m_1 m_2 n_1 n_2}\right)$$

$$= \frac{(ad - bc)^2}{n^2} n^2\left(\frac{n_2 + n_1}{m_1 m_2 n_1 n_2}\right)$$

$$= \frac{(ad - bc)^2}{n^2} n^3\left(\frac{1}{m_1 m_2 n_1 n_2}\right)$$

$$= \frac{n(ad - bc)^2}{m_1 m_2 n_1 n_2}$$