MODUL 1

Sampel dan Distribusi Sampling

Prof. Dr. Zanzawi Soejoeti

PENDAHULUAN

Pada modul pertama ini, akan dipelajari terlebih dahulu mengenai sampel dan sifat-sifatnya serta *sampling*-nya. Materi ini sebenarnya telah banyak disajikan pada mata kuliah metode statistika, materi ini sangat penting dan perlu dipelajari lebih mendalam karena pemahaman mengenai sampel dan distribusi *sampling*-nya merupakan dasar untuk mempelajari materi lanjutan yang berkaitan dengan inferensi statistik yang akan dipelajari pada modul-modul berikutnya.

Pada Modul 1 ini, ada dua subpokok bahasan yang akan disajikan, yaitu tentang sampel dan sifat-sifatnya dan beberapa distribusi sampling khusus.

Pada Kegiatan Belajar 1 akan dibahas terlebih dahulu mengenai pengertian statistik, distribusi sampling statistik, sifat distribusi normal, serta distribusi lainnya yang berkaitan dengan distribusi normal, sedangkan pada kegiatan belajar 2 akan dibahas mengenai distribusi t , F , dan Beta.

Setelah mempelajari modul ini Anda diharapkan dapat menerapkan distribusi sampling berbagai statistik dan sifat-sifatnya dalam analisis statistik.

KEGIATAN BELAJAR 1

Sampel dan Sifat-sifatnya

Pertama-tama marilah kita ingat kembali definisi sampel random sebagai berikut.

Definisi 1.1.1

Himpunan variabel random $X_1, X_2, ..., X_n$ dikatakan sebagai sampel random berukuran n dari suatu populasi dengan fungsi kepadatan $f(x)$ jika fungsi kepadatan peluang bersamanya berbentuk

$f\left(X_1, X_2, ..., X_n\right) = f(x_1).f(x_2).....f(x_n)$.

Fungsi distribusi empiris digunakan untuk memberikan nilai-nilai mean dan variansi sampel sebagai taksiran bagi mean dan variansi distribusi populasinya. Dalam modul ini akan dikenalkan konsep tentang statistik yang termasuk mean sampel dan variansi sampel sebagai kasus khususnya, dan akan diturunkan sifat-sifat statistik tertentu yang memainkan peranan penting dalam modul-modul mendatang.

A. STATISTIK

Kita pandang himpunan variabel random teramati $X_1, X_2, ..., X_n$; misalnya sampel random berukuran n dari suatu populasi.

Definisi 1.1.2

Suatu fungsi variabel random teramati $T(\underset{\sim}{x}) = t(x_1, x_2, ..., x_n)$ yang tidak bergantung pada sesuatu parameter yang tidak diketahui dinamakan ***statistik***.

Huruf t menunjukkan fungsi yang kita terapkan pada X_1, ..., X_n guna mendefinisikan nilai statistik, yang selanjutnya ditulis dengan T.

Di sini variabel-variabel itu harus dapat diamati karena akan menggunakan statistik itu. Kita ingin melakukan inferensi tentang distribusi himpunan variabel random; jika variabel-variabel itu tidak dapat diamati atau jika fungsi $t\left(x_1, x_2, ..., x_n\right)$ bergantung pada parameter yang tidak diketahui, maka T tidak akan bermanfaat guna melakukan inferensi seperti itu. Sebagai contoh kita pandang data tahan hidup yang diamati (dalam bulan) suatu sampel random 40 benda elektris, yang telah kita urutkan dari kecil ke besar.

0,15	2,37	2,90	7,39	7,99	12,05	15,17	17,56
22,40	34,84	35,39	36,38	39,52	41,07	46,50	50,52
52,54	58,91	58,93	66,71	71,48	71,84	77,66	79,31
80,90	90,87	91,22	96,35	108,92	112,26	122,71	126,87
127,05	137,96	167,59	183,53	282,49	335,33	341,19	409,97

Akan kita gunakan k = sembilan interval dengan lebar 50, yakni I_1 = (0; 50), I_2 = (50; 100), dan seterusnya. Maka, kita peroleh distribusi frekuensi sebagai berikut.

Tabel 1.1
Distribusi Frekuensi Tahan Hidup 40 Benda Elektris.

Interval	frekuensi (f_i)
0 – 50	15
50 – 100	13
100 – 150	6
150 – 200	2
200 – 250	0
250 – 300	1
300 – 350	2
350 – 400	0
400 – 450	1

Dapat kita anggap bahwa tahan hidup itu merupakan nilai-nilai yang diamati suatu sampel random berukuran 40 dari suatu populasi benda-benda elektris seperti itu. Biasanya, populasi seperti itu akan mempunyai satu parameter atau lebih yang tidak diketahui, misalnya mean populasi μ yang tidak diketahui. Guna melakukan inferensi tentang populasinya, andaikan kita perlu menghitung suatu fungsi data yang juga bergantung pada parameter yang tidak diketahui, misalnya $t(x_1, x_2, ..., x_{40}) \; = (x_1, x_2, ..., x_{40}) / \mu$. Tentu saja hitungan semacam itu tidak mungkin kita selesaikan karena μ tidak diketahui. Oleh karena itu, fungsi semacam itu tidak sesuai untuk kita gunakan sebagai definisi statistik. Perlu juga dicatat bahwa, pada umumnya, himpunan variabel random tidak selalu merupakan sampel random. Misalnya, himpunan variabel random berurut $X_{(1)}, X_{(2)}, \ldots, X_{(10)}$ bukan suatu sampel random. Tetapi, fungsi variabel-variabel ini yang tidak bergantung pada parameter yang tidak diketahui misalnya $t(x_{(1)}, x_{(2)}, \ldots, x_{(10)}) =$ $(x_{(1)}, x_{(2)}, \ldots, x_{(10)}) + 3.x_{(10)}$ adalah suatu statistik.

Kebanyakan pembicaraan kita dalam modul ini akan melibatkan sampel random.

Contoh 1.1.1

Misalkan, $X_1, X_2, \ldots, X_n$ adalah sampel random dari suatu populasi dengan fungsi peluang *f*(*x*). *Mean sampel* merupakan suatu contoh statistik dengan fungsi $t(x_1, x_2, \ldots, x_n) =$ $t(x_1 + x_2 + \ldots + x_n)/n$. Statistik ini biasanya ditulis dengan:

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$$

Jika sampel random itu diamati, nilai $\bar{X}$ yang dihitung dari data biasanya ditulis dengan huruf kecil $\bar{x}$. Sebagaimana telah kita pelajari $\bar{x}$ berguna sebagai nilai taksiran mean populasinya ($\mu = E(X)$).

Teorema berikut memberikan sifat-sifat penting mean sampel.

Teorema 1.1.1

Jika $X_1, X_2, \ldots, X_n$ merupakan sampel random dari f (x) dengan E(X) = μ, dan Var (X) $= \sigma^2$ maka:

$$E(\bar{X}) = \mu \qquad (1.1.1)$$

$$Var(\bar{X}) = \frac{\sigma^2}{n} \qquad (1.1.2)$$

Sifat 1.1.1 menunjukkan bahwa jika mean suatu sampel digunakan untuk menaksir mean populasinya, maka nilai-nilai taksiran sampel itu rata-ratanya akan sama dengan mean populasi μ. Tentu saja untuk sesuatu sampel nilai $\bar{x}$ mungkin berbeda cukup besar dengan μ. Statistik dengan sifat, seperti dalam Contoh 1.1.1 ini dikatakan tak bias untuk parameter yang ditaksirnya. Sifat ini akan kita pelajari lebih jauh dalam model mendatang. Satu kasus khusus yang penting teorema ini terjadi jika populasinya berdistribusi Bernoulli.

Contoh 1.1.2

Kita pandang variabel random $X_1, X_2, ..., X_n$ yang merupakan sampel random berukuran n dari distribusi Bernoulli, $X_i \sim BIN(1; p)$. Distribusi Bernoulli merupakan satu model untuk suatu populasi bernilai dua atau dikotomi. Mean dan variansi populasi seperti itu adalah $\mu = p$ dan $\sigma^2 = pq$, dengan $q = 1 - p$. Dalam hal ini mean sampel $\bar{X} = Y/n$ dengan Y adalah variabel binomial, biasanya dinamakan proporsi sampel dan ditulis $\hat{p} = Y/n$. Mudah dilihat bahwa $\hat{p}$ adalah penaksir tak bias untuk p, yakni:

$$E(\hat{p}) = p \qquad (1.1.3)$$

dan

$$Var(\hat{p}) = \frac{pq}{n} \qquad (1.1.4)$$

Sebagaimana telah kita pelajari, distribusi binomial memberikan suatu model untuk keadaan pengambilan sampel *dengan pengembalian*. Telah pernah kita pelajari pula bedanya dengan pengambilan sampel *tanpa pengembalian* yang menghasilkan distribusi hipergeometrik $Y \sim Hip(n, M, N)$. Misalkan, kita ingin menaksir M/n, yakni proporsi benda yang cacat dalam populasi, berdasarkan proporsi sampel Y/n. Kita tahu bahwa

$$E\left(\frac{Y}{n}\right) = \frac{M}{N} = p \qquad (1.1.5)$$

yang berarti bahwa Y/n tak bias untuk p, dan

$$Var\left(\frac{Y}{n}\right) = \frac{1}{n}\, p\,(1 - p)\left(\frac{N-n}{N-1}\right) \qquad (1.1.6)$$

yang mendekati nol untuk n menjadi besar. Sebenarnya di sini dapat kita lihat bahwa $Var(Y/n)$ akan sama dengan nol jika $n = N$, yakni semua elemen populasi diamati.

Contoh 1.1.3

Telah kita pelajari pula bahwa variansi sampel diberikan dengan rumus :

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} \qquad (1.1.7)$$

Bentuk rumus berikut ini dapat kita jabarkan dengan mudah

$$S^2 = \frac{\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2 / n}{n-1} \tag{1.1.8}$$

$$= \frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n-1} \tag{1.1.9}$$

Teorema berikut memberikan sifat penting variansi sampel.

Teorema 1.1.2

Misalkan, $X_1, X_2, ..., X_n$ suatu sampel random berukuran n dari f(x) dengan E(X) = μ dan Var(X) = σ^2. Maka,

$$E(S^2) = \sigma^2 \tag{1.1.10}$$

$$Var(S^2) = \left(\mu_4' - \frac{n-3}{n-1}\sigma^4\right)/n \; ; \; n > 1 \tag{1.1.11}$$

Bukti **:**

$$\begin{aligned} E(S^2) &= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right) \\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n} E\left(X_i^2\right) - nE\left(\bar{X}^2\right)\right] \\ &= \frac{1}{n-1}\left[n(\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right] \\ &= \frac{1}{n-1} E\left[(n-1)\sigma^2\right] \\ &= \sigma^2. \end{aligned}$$

Bukti untuk (1.1.11) tidak diberikan di sini.

Menurut Sifat (1.1.10), variansi sampel memberikan contoh lain statistik tak bias, dan inilah pula alasan utama kita menggunakan pembagian (*n* – 1), dan bukannya *n*.

B. DISTRIBUSI SAMPLING STATISTIK

Suatu statistik adalah juga variabel random, yang distribusinya bergantung pada distribusi populasinya dan pada bentuk fungsi $t(x_1, x_2, ..., x_n)$. Distribusi suatu statistik sering dikenal sebagai *distribusi turunan* atau *distribusi sampling*.

Banyak statistik yang penting dapat ditulis sebagai kombinasi linear variabel random normal independen.

Kombinasi linear variabel normal

Teorema 1.1.3

Jika $X_i \sim N(\mu_i; \sigma_i^2),\ i = 1, 2, ..., n$ merupakan variabel normal independen, maka

$$Y = \sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i \mu_i ; \sum_{i=1}^{n} a_i^2 \sigma_i^2\right) \quad (1.1.12)$$

Bukti **:**

$$M_Y(t) = \prod_{i=1}^{n} M_{X_i}(a_i t)$$

$$= \prod_{i=1}^{n} e^{a_i \mu_i t + a_i^2 \sigma_i^2 t^2 / 2}$$

$$= \exp\left[t \sum_{i=1}^{n} a_i \mu_i + t^2 \sum_{i=1}^{n} a_i^2 \sigma_i^2 / 2\right]$$

yang merupakan fungsi pembentuk momen variabel normal dengan mean $\sum_{i=1}^{n} a_i \mu_i$ dan variansi $\sum_{i=1}^{n} a_i^2 \sigma_i^2$.

Korolari 1.1.1

Jika $X_1, X_2, ..., X_n$ merupakan sampel random dari $N(\mu, \sigma^2)$ maka $\overline{X} \sim N(\mu ; \sigma^2/n)$.

Bukti gunakan Teorema 1.1.3 dengan $\mu_i = \mu$; $\sigma_i^2 = \sigma^2$ dan $a_i = \dfrac{1}{n}$

60 bulan

Contoh 1.1.4

Kita ingin menyelidiki pernyataan bahwa $X \sim (60;\,36)$, yakni X yang merupakan tahan hidup (dalam bulan) suatu baterai berdistribusi normal dengan mean 60 dan variansi 36. Untuk ini 25 baterai semacam itu di uji hidupnya, dan rata-rata tahan hidup 25 baterai tersebut dihitung. Jika pernyataan itu benar, rata-rata tahan hidup 25 baterai itu harus lebih besar dari nilai berapa, sebanyak 95% kali ? Kita punyai $E(\bar{X}) = 60$ dan $Var(\bar{X}) = 36/25$ dan

$$P(\bar{X} > c) = 1 - \Phi\left(\frac{c-60}{\sqrt{36}/5}\right) = 1 - \Phi\left(\frac{c-60}{6/5}\right) = 0{,}95\,.$$

Jadi, $\dfrac{c-60}{6/5} = Z_{0,05} = 1{,}645$ dan $c = 58{,}026$ bulan

Pada umumnya, untuk suatu tingkat peluang tertentu $(1-\alpha)$, akan kita punyai:

$$c = \mu + \frac{Z_\alpha \sigma}{\sqrt{n}}$$

Jadi, prosedur yang masuk akal adalah menerima pernyataan jika nilai pengamatan $\bar{x} \geq 58{,}026$, tetapi menolaknya jika $\bar{x} < 58{,}026$ karena hal ini akan terjadi dengan peluang yang sangat kecil (kurang dari 0,05) jika pernyataan itu benar. Jika kita ingin lebih yakin sebelum menolaknya, maka nilai α yang lebih kecil, misalnya α = 0,01, harus kita gunakan untuk menentukan nilai kritis c. Prosedur uji ini menguntungkan konsumen karena dengan prosedur ini kita tidak menolaknya jika diperoleh mean tahan hidup yang besar. Tentu saja, dapat juga kita bentuk prosedur uji yang sesuai untuk arah yang lain (atau dua arah).

Contoh 1. 1.5

Kita pandang dua sampel random independen $X_1, X_2, \ldots, X_{n_1}$ dan $Y_1, Y_2, \ldots, Y_{n_2}$ dengan ukuran sampel masing-masing n_1 dan n_2, dari populasi-populasi berdistribusi normal $X_i \sim N(\mu_1;\sigma_1^2)$ dan $Y_i \sim N(\mu_2;\,\sigma_2^2)$. Kita tulis mean sampel masing-masing $\bar{X}$ dan $\bar{Y}$. Menurut Teorema 1.1.3, selisihnya juga berdistribusi normal, yakni:

$$\boxed{\left(\bar{X} - \bar{Y}\right) \sim N(\mu_1 - \mu_2\,;\,\sigma_1^2/n_1 + \sigma_2^2/n_2).}$$

Jelas bahwa n_1 suku yang pertama selisih itu mempunyai koefisien $a_i = 1/n_1$, dan n_2 suku terakhirnya mempunyai koefisien $a_i = -1/n_2$. Sehingga mean selisih itu sama dengan:

$n_1(1/n_1)\mu_1 + n_2(-1/n_2)\mu_2 = \mu_1 - \mu_2$, dan variansinya

$$n_1(1/n_1)^2\sigma_1^2 + n_2(-1/n_2)^2\sigma_2^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Sifat-sifat selanjutnya yang akan kita pelajari melibatkan kasus khusus dalam distribusi gamma, yakni distribusi Khi-kuadrat.

Distribusi Khi-kuadrat

Pertama-tama marilah kita tulis kembali fungsi peluang distribusi gamma ($X \sim GAM(\theta; \beta)$):

$$f(x) = \frac{1}{\theta^\beta \Gamma(\beta)} x^{\beta-1} e^{-x/\theta};\ x > 0,\ \theta > 0,\ \beta > 0. \qquad (1.1.13)$$

Jika $\theta = 2$ dan $\beta = v/2$, maka distribusi gamma itu akan menjadi distribusi khi-kuadrat dengan derajat bebas v, yang distribusi peluangnya adalah:

$$f(x) = \frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-x/2}\ ;\ x > 0 \qquad (1.1.14)$$

Notasi yang sering kita pakai untuk ini adalah $X \sim GAM(2\ ;\ v/2)$ atau $X \sim \chi^2(v)$.

Teorema 1.1.4

Jika X ~ $\chi^2(v)$, maka fungsi pembentuk momennya adalah:

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

$$E(X^k) = 2^k \frac{\Gamma(v/2+k)}{\Gamma(v/2)}$$

sehingga

E(X) = v

Var(X) = 2v.

Bukti: (sebagai latihan).

Distribusi khi-kuadrat kumulatif telah ditabelkan secara luas dalam literatur. Kebanyakan memberikan nilai persentil $\chi^2_\gamma(v)$ untuk γ yang kita inginkan dan nilai-nilai v yang berbeda-beda. Khususnya, jika $X \sim \chi^2(v)$, maka $\chi^2_\gamma(v)$ adalah nilai yang memenuhi persamaan:

$$P\left[X \le \chi^2_\gamma(v)\right] = \gamma \qquad (1.1.15)$$

Nilai-nilai $\chi^2_\gamma(v)$ diberikan dalam Tabel 1.4 (lampiran) untuk berbagai nilai γ dan v. Nilai-nilai ini dapat juga digunakan untuk mendapatkan nilai-nilai persentil bagi distribusi gamma.

Teorema 1.1.5

Jika X ~ GAM (θ ; β), maka Y = 2X / θ ~ $\chi^2(2\beta)$.

Bukti:

$$\begin{aligned} M_Y(t) &= M_{2X/\theta}(t) \\ &= M_X(2t/\theta) \\ &= (1-2t)^{-2\beta/2} \end{aligned}$$

yang tak lain adalah fungsi pembentuk momen distribusi khi-kuadrat dengan derajat bebas 2β.

Fungsi distribusi gamma dapat juga dinyatakan dalam bentuk notasi khi-kuadrat. Jika $X \sim GAM(\theta\ ;\ \beta)$ dan jika $H(y\ ;\ v)$ menunjukkan fungsi distribusi khi-kuadrat dengan derajat bebas v, maka

$$F_X(x) = H(2x/\theta\ ;\ 2\beta) \qquad (1.1.16)$$

Peluang khi-kuadrat kumulatif $H(c\ ;\ v)$ diberikan dalam Tabel 1.5 (lampiran) untuk berbagai nilai c dan v.

Contoh 1.1.6

Tahan hidup (dalam tahun) suatu jenis komponen tertentu berdistribusi gamma dengan $\theta = 3$ dan $\beta = 2$. Diinginkan untuk menentukan periode garansi yang 90% komponen akan tetap hidup, yakni kita inginkan persentil ke 10, kita tulis $x_{0,10}$, sedemikian hingga $P(X \le x_{0,10}) = 0,10$. Kita peroleh:

$P[X < x_{0,10}] = H(2x_{0,10}/\theta; 2\beta) = 0{,}10.$

Dengan menuliskan persamaan

$$\frac{2x_{0,10}}{\theta} = \chi^2_{0,10}; (2\beta)$$

kita dapatkan

$$x_{0,10} = \frac{\theta.\chi^2_{0,10};(2\beta)}{2}$$

Untuk $\theta = 3$ dan $\beta = 2$,

$$x_{0,10} = \frac{3\chi^2_{0,10}(4)}{2} = \frac{(3.1{,}06)}{2} = 1{,}59 \text{ tahun}$$

Jelas bahwa pada umumnya persentil ke p distribusi gamma dapat dinyatakan sebagai

$$x_p = \frac{\theta.\chi^2_p;(2\beta)}{2} \qquad (1.1.17)$$

Teorema berikut menyajikan sifat bahwa jumlah variabel khi-kuadrat yang independen adalah juga variabel yang berdistribusi khi-kuadrat.

Teorema 1.1.6

Jika $Y_i \sim \chi^2(v_i)$; i = 1, 2, ..., n, adalah variabel khi-kuadrat yang independen maka:

$$V = \sum_{i=1}^{n} Y_i \sim \chi^2\left(\sum_{i=1}^{n} v_i\right).$$

*Bukti***:**

$$M_V(t) = (1-2t)^{-v_1/2}.(1-2t)^{-v_2/2}(1-2t)^{-v_n/2}$$

$$= (1-2t)^{-\sum_{i=1}^{n} v_i/2}$$

yang tidak lain adalah fungsi pembentuk momen $\chi^2_{\left(\sum_{i=1}^{n} v_i\right)}$.

Teorema berikut menegaskan hubungan antara variabel normal standar dan variabel khi-kuadrat.

Teorema 1.1.7

Jika Z ~ N(0 ; 1), maka $Z^2 \sim \chi^2(1)$

Bukti:

$$
\begin{aligned}
M_{Z^2}(t) &= E(e^{tZ^2}) \\
&= \int_{-\infty}^{\infty} \frac{e^{tz^2} e^{-z^2/2}}{\sqrt{2\pi}} dz \\
&= \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \frac{\sqrt{1-2t}}{\sqrt{2\pi}} e^{-z^2(1-2t)/2} dz \\
&= (1-2t)^{-\frac{1}{2}}
\end{aligned}
$$

yang adalah fungsi pembentuk momen distribusi khi-kuadrat dengan derajat bebas satu.

Korolari 1.1.2

Jika $X_1, X_2, ..., X_n$ merupakan suatu sampel random dari $N(\mu; \sigma^2)$ maka:

$$\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi^2(n) \qquad (1.1.18)$$

$$\frac{n(\bar{X} - \mu)^2}{\sigma^2} \sim \chi^2(1) \qquad (1.1.19)$$

Variansi sampel telah kita bicarakan di atas, dan untuk sampel random dari suatu populasi normal distribusinya dapat dikaitkan dengan distribusi khi-kuadrat. Distribusi sampling S^2 tidak mengikuti secara langsung Korolari 1.1.2 di atas, sebab suku-suku $(X_i - \bar{X})$ tidak independen. Kita tahu suku-suku itu dependen karena $\sum_{i=1}^{n} (X_i - \bar{X}) = 0$.

Teorema 1.1.8

Jika $X_1, X_2, ..., X_n$ merupakan sampel random dari $N(\mu;\sigma^2)$, maka

1. $\overline{X}$ dan suku-suku $(X_i - \overline{X})$, i = 1, 2, …, n, independen
2. $\overline{X}$ dan S^2 independen
3. $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.

Bukti **:**

Guna membuktikan bagian 1, pertama-tama kita perhatikan bahwa dengan menambahkan dan mengurangkan $\overline{x}$ dan menjabarkannya dapat kita peroleh hubungan:

$$\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2} = \sum_{i=1}^{n}\frac{(x_i-\overline{x})^2}{\sigma^2} + \frac{n(\overline{x}-\mu)^2}{\sigma^2} \tag{1.1.20}$$

Maka, fungsi peluang bersama $X_1, X_2, ..., X_n$ dapat ditulis sebagai berikut.

$$f(x_1,...,x_n) = (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right]$$

$$= (2\pi)^{-\frac{n}{2}}\sigma^{-n}\exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i=1}^{n}(x_i-\overline{x})^2 + n(\overline{x}-\mu)^2\right\}\right]$$

Sekarang kita pandang transformasi bersama:

$y_1 = \overline{x}$, $y_i = x_i - \overline{x}$, $i = 2, 3, \ldots, n$.

Kita tahu bahwa:

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

sehingga

$$x_1 - \overline{x} = -\sum_{i=2}^{n}(x_i - \overline{x}) = -\sum_{i=2}^{n} y_i$$

dan

$$\sum_{i=1}^{n}(x_i - \overline{x})^2 = \left(-\sum_{i=2}^{n} y_i\right)^2 + \sum_{i=2}^{n} y_i^2$$

Maka,

$$g(y_1,\ldots,y_n)=\frac{|J|}{(2\pi)^{\frac{n}{2}}\sigma^n}\exp\left[-\frac{1}{2\sigma^2}\left\{\left(-\sum_{i=2}^{n}y_i\right)^2+\sum_{i=2}^{n}y_i^2+n(y_1-\mu)^2\right\}\right].$$

Dengan mudah dapat kita lihat bahwa Jacobian J adalah suatu konstan, dan di sini dapat kita tunjukkan bahwa $|J| = n$. Maka, fungsi peluang bersama itu dapat kita uraikan menjadi fungsi peluang marginal y_1 kali suatu fungsi $y_2, y_3, \ldots, y_n$ saja. Kenyataan ini menunjukkan bahwa $Y_1=\bar{X}$ dan suku-suku $Y_i = X_i - \bar{X}$, i =2, 3, …, n adalah independen. Oleh karena $X_1-\bar{X}=-\sum_{i=2}^{n}(X_i-\bar{X})$ maka $\bar{X}$ dan $X_1-\bar{X}$ juga independen.

Bagian 2 mengikuti bagian 1 karena S^2 adalah fungsi $(X_i-\bar{X})$ saja.

Untuk membuktikan bagian 3. kita pandang lagi persamaan (1.1.20) yang diterapkan bagi sampel random itu

$$V_1=\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}=\frac{(n-1)S^2}{\sigma^2}+\frac{n(\bar{X}-\mu)^2}{\sigma^2}$$
$$=V_2+V_3.$$

Dari Korolari 1.1.2; $V_1\sim\chi^2(n)$ dan $V_3\sim\chi^2(1)$. Juga, dikarenakan V_2 dan V_3 independen maka:

$$M_{v_1}(t)=M_{v_2}(t).M_{v_3}(t)$$

dan

$$M_{v_2}(t)=\frac{M_{v_1}(t)}{M_{v_3}(t)}=\frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}}=(1-2t)^{-(n-1)/2}$$

Jadi, $V_2=\left[(n-1)S^2\right]/\sigma^2\sim\chi^2(n-1)$

Sehingga jika c_γ adalah persentil ke γ distribusi S^2 maka:

$$c_\gamma=\frac{\sigma^2\chi_\gamma^2(n-1)}{n-1} \qquad (1.1.21)$$

Kita pandang kembali Contoh 1.1.6, di mana kita menganggap bahwa $X \sim N(60; 36)$. Andaikan kita inginkan untuk mengambil sampel 25 baterai, dan menolak pernyataan bahwa $\sigma^2 = 36$ jika $s^2 \geq 54{,}63$, dan kita tidak menolak pernyataan itu jika $s^2 < 54{,}63$. Dengan prosedur ini, berapakah peluang akan menolak pernyataan itu jika kenyataannya $\sigma^2 = 36$? Kita lihat bahwa:

$$\begin{aligned} P(S^2 \geq 54{,}63) &= P(24\ S^2/36 \geq 36{,}42) \\ &= 1 - H(36{,}42; 24) \\ &= 0{,}05 \end{aligned}$$

Jika sebaliknya, kita hanya ingin salah 1% kali jika menolak pernyataan itu maka prosedur yang kita tempuh adalah akan menolaknya jika $s^2 \geq c_{0,99}$, dengan:

$$c_{0,99} = \frac{\sigma^2 \chi^2_{0,99(n-1)}}{n-1} = \frac{36(42{,}98)}{24} = 64{,}47.$$

LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

1) Misalkan, X menunjukkan berat dalam pound satu kotak benda, dengan $X \sim N(101; 4)$. Berapakah peluang bahwa 20 kotak benda itu akan mempunyai paling sedikit 100 pon?
2) Misalkan, $X_1, X_2, \ldots, X_n$ adalah sampel random berukuran n dari suatu distribusi normal. $X_i \sim N(\mu; \sigma^2)$. Kita definisikan $U = \sum_{i=1}^{n} X_i$ dan $W = \sum_{i=1}^{n} X_i^2$.
   a. Carilah suatu statistik yang merupakan fungsi U dan W serta tak bias untuk parameter
      $\theta = 2\mu - 5\sigma^2$.
   b. Carilah suatu statistik yang tak bias untuk $\sigma^2 + \mu^2$.

c. Misalkan, c suatu konstan, kita definisikan $Y_i = 1$ jika $X_i \leq c$ dan nol untuk nilai-nilai X_i yang lain. Carilah suatu statistik yang merupakan fungsi $Y_1, Y_2, \ldots, Y_n$ dan juga tak bias untuk

$$F_X(c) = \Phi\left(\frac{c-\mu}{\sigma}\right).$$

3) Kita anggap bahwa X_1 dan X_2 adalah variabel random normal independen, $X_i \sim N(\mu;\sigma^2)$ Misalkan, pula $Y_1 = X_1 + X_2$ dan $Y_2 = X_1 - X_2$. Tunjukkan bahwa Y_1 dan Y_2 independen dan berdistribusi normal.

4) Misalkan, $X \sim \chi^2(m)$, $Y \sim \chi^2(n)$, $S = (X + Y) \sim \chi^2(m+n)$, dan X dan Y independen. Gunakan fungsi pembentuk momen untuk menunjukkan bahwa $(S - X) \sim \chi^2(n)$

RANGKUMAN

Kita pelajari sifat-sifat distribusi normal dan kita jabarkan distribusi-distribusi lain yang berkaitan dengan distribusi normal yang timbul dalam analisis statistik untuk data yang diambil dari populasi yang berdistribusi normal.

Satu sifat penting distribusi normal bahwa kombinasi linear variabel-variabel random normal independen juga berdistribusi normal. Sebagai contoh, mean sampel adalah berdistribusi normal karena mean sampel merupakan kombinasi linear variabel-variabel random normal independen.

Suatu fungsi tertentu variansi sampel berdistribusi khi-kuadrat; serta telah kita tunjukkan pula bahwa mean sampel dan variansi sampel adalah variabel-variabel random yang independen.

TES FORMATIF 1

Pilihlah satu jawaban yang paling tepat!

1) Misalkan, S menunjukkan diameter suatu corong (shaft) dan B diameter suatu peluru (*bearing*), dengan S dan B independen, serta $S \sim N(1;$ $0{,}0004)$ dan $B \sim N(1{,}01;\ 0{,}0009)$

a. Jika satu corong dan satu peluru dipilih secara random, maka peluang bahwa diameter corong akan lebih besar dari diameter peluru adalah ….
   A. 0,31
   B. 0,39
   C. 0,42
   D. 0,47

b. Kita anggap variansinya sama $\sigma_1^2 = \sigma_2^2 = \sigma^2$ maka nilai σ yang akan menghasilkan peluang adanya gangguan *(interference)* sebesar 0,95 adalah ….
   A. 0,0043
   B. 0,043
   C. 0,43
   D. 4,3.

2) Satu komponen baru dipasang untuk beroperasi dan masih tersedia sembilan cadangannya. Tahan hidup-tahan hidup (dalam hari) adalah variabel independen berdistribusi eksponensial, $T_i \sim Exp(100)$

a. Maka, $\sum_{i=1}^{10} T_i$ berdistribusi ….
   A. *Gamma*(100; 10)
   B. *Gamma*(90; 20)
   B. *Gamma* (80; 30)
   D. *Gamma* (70; 40)

b. Peluang bahwa operasi yang berhasil dapat dipelihara untuk paling sedikit 1,5 tahun adalah ….
   A. 0,80
   B. 0,85
   C. 0,90
   D. 0,95

c. Berapa banyak suku cadang diperlukan supaya 95% yakin bahwa operasi yang berhasil untuk paling sedikit 2 tahun?
   A. 10.
   B. 12.
   C. 14.
   D. 16.

3) Kita ulangi soal 2 di atas, dengan menganggap bahwa $T_i \sim Gamma\ (100;\ 12)$ maka ….
   a. A. *Gamma* (70; 42)
      B. *Gamma* (80; 32)
      C. *Gamma* (90; 22)
      D. *Gamma* (100; 12)
   b. A. 0,99
      B. 0,95
      C. 0,90
      D. 0,87
   c. A. 8
      B. 10
      C. 12
      D. 14.

4) Misalkan, $Z \sim N(0;\ 1)$.
   a. Maka, dengan menggunakan nilai-nilai tabel distribusi normal, $P(Z^2 < 3{,}84)$ sama dengan ….
      A. 0,90
      B. 0,93
      C. 0,95
      D. 0,99
   b. Maka, dengan menggunakan nilai-nilai tabel distribusi khi-kuadrat, $P(Z^2 < 3{,}84)$ sama dengan ….
      A. 0,90
      B. 0,93
      C. 0,95
      D. 0,99

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 1 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 1.

$$\text{Tingkat penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100\%$$

Arti tingkat penguasaan: 90 - 100% = baik sekali
80 - 89% = baik
70 - 79% = cukup
< 70% = kurang

Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan Kegiatan Belajar 2. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 1, terutama bagian yang belum dikuasai.

KEGIATAN BELAJAR 2

Beberapa Distribusi Sampling Khusus

Akan kita bicarakan beberapa distribusi sampling yang diturunkan dari distribusi normal. Distribusi-distribusi ini sangat penting dalam analisis statistik.

A. DISTRIBUSI *t*, *F* DAN BETA

Penurunan Rumus-rumus t, F, dan Beta dari Distribusi Normal

Distribusi *t*

Kita tahu bahwa S^2 dapat digunakan untuk melakukan inferensi tentang parameter σ^2 dalam suatu distribusi normal. Demikian juga $\bar{X}$ bermanfaat bagi inferensi parameter μ ; tetapi distribusi $\bar{X}$ juga bergantung pada parameter σ^2 ini menyebabkan tidak dimungkinkannya kita menggunakan $\bar{X}$ untuk melakukan inferensi bagi μ, jika σ^2 tidak diketahui dengan menggunakan prosedur berdasarkan distribusi itu. Oleh karena itu kita harus mencari prosedur lain, yakni dengan mengganti dengan S dalam kuantitas $\sqrt{n}(\bar{X}-\mu)/\sigma$, menjadi $\boxed{\sqrt{n}(\bar{X}-\mu)/S}$. Kuantitas terakhir ini tidak lagi berdistribusi normal standar tetapi tidak lagi bergantung σ. Distribusinya dapat dijabarkan menggunakan metode transformasi. Untuk ini pelajari dahulu beberapa teorema sebagai berikut.

Teorema 1.2.1

Jika Z berdistribusi normal standar, Z ~ N(0; 1), dan W berdistribusi khi-kuadrat dengan derajat bebas v, W ~ $\chi^2(v)$, serta Z dan W independen maka distribusi variabel random:

$$\boxed{T = \frac{Z}{\sqrt{\frac{W}{v}}}}$$

dikenal sebagai distribusi t dengan derajat bebas v, T ~ t(v). Fungsi peluangnya adalah:

PDF t distribution

$$\boxed{f(t)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)}\frac{1}{\sqrt{v\pi}}\left(1+\frac{t^2}{v}\right)^{-(v+1)/2} ;-\infty<t<\infty} \qquad (1.2.1)$$

Bukti:

Oleh karena Z dan W independen, maka fungsi peluang bersama Z dan W adalah $f(z\ ,\ w)=g(z).h(w)$.

$$=\frac{1}{\sqrt{2\pi}}e^{-z^2/2}.\frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)}w^{\frac{v}{2}-1}e^{-w/2},\quad 0<w<\infty\ ;\ -\infty<z<\infty.$$

Pandang transformasi:

$$T=\frac{Z}{\sqrt{\frac{W}{v}}}\quad \text{dan}\quad Y=W$$

maka $w=y$ dan $z=t\sqrt{\frac{y}{v}}$, dan Jacobian transformasinya adalah:

$$J=\begin{vmatrix}\frac{\partial w}{\partial y} & \frac{\partial w}{\partial t}\\ \frac{\partial z}{\partial y} & \frac{\partial z}{\partial t}\end{vmatrix}=\begin{vmatrix}1 & 0\\ \frac{t}{\sqrt{yv}} & \sqrt{\frac{y}{v}}\end{vmatrix}=\sqrt{\frac{y}{v}}$$

Maka, fungsi peluang bersama T dan Y adalah:

$$f(t,y)=\frac{1}{\sqrt{2\pi}}e^{-t^2y/2v}.\frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)}y^{\frac{v}{2}-1}e^{-y/2}.\sqrt{\frac{y}{v}}$$

$$=\frac{1}{\sqrt{2\pi}\,2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)\sqrt{v}}y^{\left(\frac{v}{2}+\frac{1}{2}\right)-1}e^{-\left(\frac{t^2}{2v}+\frac{1}{2}\right)}\ ;\quad y>0\ ;-\infty<t<\infty.$$

Maka, fungsi peluang kita peroleh sebagai fungsi peluang marginal $f(t,y)$, yakni:

$$f(t) = \int_0^\infty f(t, y)dy.$$

Setelah dilakukan hitungan-hitungan dan penyederhanaan seperlunya, maka kita peroleh $f(t)$ seperti tertuang dalam (1.2.1).

Catatan:

Sering digunakan notasi $Z \sim N(0, 1)$ yang dibaca: Z berdistribusi $N(0, 1)$

Teorema 1.2.2 Estimasi dan Variansi Distribusi t

Jika T ~ t(v), maka untuk v > 2r berlaku:

$$E(T^{2r}) = \frac{\Gamma\left((2r+1)/2\right)\left[\Gamma\left((v-2r)/2\right)\right].v^r}{\Gamma(1/2)\Gamma(v/2)}$$

Untuk setiap T pangkat genap

$$E(T^{2r-1}) = 0, \quad r = 1,\ 2,\ \ldots$$

Untuk setiap T pangkat ganjil

$$Var(T) = \frac{v}{v-2}; \quad v > 2$$

*Bukti***:**

$$T = \frac{Z}{\sqrt{\frac{W}{v}}} = Z\left(\frac{W}{v}\right)^{-\frac{1}{2}}.$$

Oleh karena Z dan W independen maka:

$$E(T) = E(Z).E\left(\frac{W}{v}\right)^{-\frac{1}{2}}.$$

Momen ke-2r untuk T adalah:

$$E(T^{2r}) = E(Z^{2r}).E\left(\frac{W}{v}\right)^{-r}.$$

Oleh karena $Z \sim N\ (0;1)$, maka $E(Z^{2r}) = \dfrac{(2r)!}{r!2^r}$; dan karena $W \sim \chi^2_{(v)}$ maka:

$$E(W^{-r}) = \int_0^\infty W^{-r}.\frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)} w^{\frac{v}{2}-1} e^{-w/2} dw$$

$$= \int_0^\infty \frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)} w^{\left(\frac{v}{2}-r\right)^{-1}} e^{-w/2} dw$$

$$= \frac{2^{\frac{v}{2}-r}\Gamma\left(\frac{v}{2}-r\right)}{2^{\frac{v}{2}}\Gamma\left(\frac{v}{2}\right)} = \frac{\Gamma\left(\frac{v}{2}-r\right)}{2^r\Gamma\left(\frac{v}{2}\right)}.$$

Maka,

$$E\left(\frac{W}{v}\right)^{-r} = v^r E(W^{-r}) = \frac{v^r\Gamma\left(\frac{v}{2}-r\right)}{2^r\Gamma\left(\frac{v}{2}\right)}$$

Jadi,

$$E(T^{2r}) = \frac{(2r)!}{r!2^r}.\frac{v^r\Gamma\left(\frac{v}{2}-r\right)}{2^r\Gamma\left(\frac{v}{2}\right)}$$

$$= \frac{(2r)!}{r!2^r}.\frac{\Gamma\left[(v-2r)/2\right].v^r}{\Gamma\left(\frac{v}{2}\right)} \qquad (1.2.2)$$

Selanjutnya dapat ditunjukkan bahwa:

$$\frac{(2r)!}{r!2^{2r}} = \frac{\Gamma\left[(2r+1)/2\right]}{\Gamma\left[\frac{1}{2}\right]}$$

Jadi, $E(T^{2r}) = \dfrac{\Gamma\left[(2r+1)/2\right]\Gamma\left[(v-2r)/2\right]v^r}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{v}{2}\right)}$

Selanjutnya, $E(T^{2r-1}) = 0$ karena $E(Z^{2r-1})$ selalu sama dengan 0. Akhirnya kita hitung $Var(T)$.

$Var(T) = E(T^2) - \left[E(T)\right]^2$.

Oleh karena $E(T) = 0$, maka $Var(T) = E(T^2)$, dan dari (1.2.2):

$$E(T^2) = \frac{2!}{1!2^2}\frac{\Gamma\left(\frac{v}{2}-1\right).v}{\Gamma\left(\frac{v}{2}\right)}$$

$$= \frac{1}{2}.\frac{\Gamma\left(\frac{v}{2}-1\right).v}{\left(\frac{v}{2}-1\right)\Gamma\left(\frac{v}{2}-1\right)}$$

$$= \frac{v}{v-2}$$

Teorema 1.2.3 Statistik Uji t

Jika $X_1, X_2, ..., X_n$ sampel random dari $N(\mu;\sigma^2)$ maka:

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)}.$$

Bukti:

Kita tulis:

$$\frac{\sqrt{n}(\bar{X}-\mu)/\sigma}{\sqrt{\frac{(n-1)S^2}{\sigma^2}\Big/(n-1)}} = \frac{Z}{\sqrt{\frac{W}{v}}}$$

dengan $Z = \dfrac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0;\ 1)$

$$W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

Dan $v = (n-1)$

Maka, menurut Teorema 1.2.1

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)}$$

atau berdistribusi t dengan derajat bebas $n - 1$ karena $\bar{X}$ dan S independen.

Distribusi *F*

Distribusi lain, yang diturunkan dari distribusi normal yang sangat penting dalam statistika adalah distribusi F. Kita pelajari hal-hal sebagai berikut.

Teorema 1.2.4

Jika $W_1 \sim \chi^2_{(v_1)}$ dan $W_2 \sim \chi^2_{(v_2)}$ independen maka variabel random

$X = \dfrac{W_1/v_1}{W_2/v_2}$ mempunyai fungsi peluang, untuk $x > 0$.

PDF dari distribusi F

$$g(x) = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{v_2}\right)\Gamma\left(\frac{v_2}{2}\right)}\left(\frac{v_1}{v_2}\right)^{v_1/2} x^{\left(\frac{v_1}{2}\right)-1}\left(1+\frac{v_1}{v_2}x\right)^{-(v_1+v_2)/2}$$

Distribusi ini dikenal sebagai distribusi F dengan derajat bebas pembilang v_1 dan derajat penyebut v_2, dan ditulis dengan $X \sim F_{(v_1;v_2)}$.

Bukti:

Fungsi peluang bersama W_1 dan W_2 adalah:

$$f(w_1,w_2) = \frac{1}{2^{v_1/2}\Gamma\left(\frac{v_1}{2}\right)} w_1^{\frac{v_1}{2}-1} e^{-w_1/2} . \frac{1}{2^{v_2/2}\Gamma\left(\frac{v_2}{2}\right)} w_2^{\frac{v_2}{2}-1} e^{-w_2/2}$$

$X = \dfrac{W_1/v_1}{W_2/v_2}$, dan misalkan $Y = W_2$ dari transformasi itu kita punyai

$$w_1 = \frac{v_1 xy}{v_2} \text{ dan } w_2 = y$$

Jacobian transformasi itu adalah:

$$J = \begin{vmatrix} \frac{v_1}{v_2} y & \frac{v_1}{2} x \\ 0 & 1 \end{vmatrix} = \frac{v_1}{v_2} y.$$

Maka, fungsi peluang bersama X dan Y adalah:

$$h(x,y) = \frac{1}{2^{\frac{v_1}{2}} \Gamma\left(\frac{v_2}{2}\right)} \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}-1} x^{\frac{v_1}{2}-1} y^{\frac{v_1}{2}-1} e^{\frac{v_1}{v_2} xy/2} \cdot \frac{1}{2^{\frac{v_2}{2}} \Gamma\left(\frac{v_2}{2}\right)} y^{\frac{v_2}{2}-1} e^{-y/2} . \frac{v_1}{v_2} y$$

$$= \frac{1}{2^{(v_1+v_2)/2} \Gamma\left(\frac{v_1}{2}\right) \Gamma\left(\frac{v_2}{2}\right)} \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} x^{\frac{v_1}{2}-1} y^{\frac{v_1+v_2}{2}-1} e^{-\left(\frac{v_1}{v_2}x+1\right)\frac{y}{2}} \; ; x > 0 \; ; y > 0$$

Maka, fungsi peluang X adalah fungsi peluang marginal dari $h(x, y)$.

Maka, karena $\int_0^\infty y^{\frac{v_1+v_2}{2}-1} e^{-\left(1+\frac{v_1}{v_2}x\right)y/2} dy = \left(\frac{1+\frac{v_1}{v_2}x}{2}\right)^{-\left(\frac{v_1+v_2}{2}\right)} . \Gamma\left(\frac{v_1+v_2}{2}\right)$

kita peroleh fungsi peluang X sebagai:

$$g(x) = \frac{\Gamma\left(\frac{v_1+v_2}{2}\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \left(\frac{v_1}{v_2}\right)^{\frac{v_1}{2}} x^{\frac{v_1}{2}-1} \left(1+\frac{v_1}{v_2}x\right)^{-\left(\frac{v_1+v_2}{2}\right)} .$$

Teorema 1.2.5 Estimasi dan Variansi dari Distribusi F

Jika X ~ $F_{(v_1;v_2)}$ maka

$$E(X^r) = \frac{\left(\frac{v_2}{v_1}\right)^r \Gamma\left(\frac{v_1}{2}+r\right)\Gamma\left(\frac{v_2}{2}-r\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)} \; ; \quad v_2 > 2r$$

$$E(X) = \frac{v_2}{v_2 - 2} \quad ; \; v_2 > 2$$

$$Var(X) = \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)} \quad ; \; v_2 > 4.$$

*Bukti***:**

Dari $X = \dfrac{W_1 / v_1}{W_2 / v_2}$ maka $X^r = \left(\dfrac{W_1}{v_1}\right)^r . \left(\dfrac{W_2}{v_2}\right)^{-r}$

sehingga $E(X^r) = \dfrac{1}{v_1^r} E(W_1^r).v_2^r E(W_2^{-r})$, karena W_1 dan W_2 independen

$$E(W_1^r) = 2^r \frac{\Gamma\left(\frac{v_1}{2} + r\right)}{\Gamma\left(\frac{v_1}{2}\right)}, \quad \text{dan}$$

$$E(W_2^{-r}) = \frac{\Gamma\left(\frac{v_2}{2} - r\right)}{2^r \Gamma\left(\frac{v_2}{2}\right)}$$

Jadi,

$$E(X^r) = \frac{\left(\frac{v_2}{v_1}\right)^r \Gamma\left(\frac{v_1}{2} + r\right)\Gamma\left(\frac{v_2}{2} - r\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}.$$

Untuk $r = 1$maka:

$$E(X) = \frac{\frac{v_2}{v_1}\Gamma\left(\frac{v_1}{2} + 1\right)\Gamma\left(\frac{v_2}{2} - 1\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}.$$

$$= \frac{\frac{v_2}{v_1}\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_1}{2}\right).\Gamma\left(\frac{v_2}{2}\right)\left(\frac{v_2}{2} - 1\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}$$

$$= \frac{\frac{v_2}{2}}{\frac{v_2}{2}-1} = \frac{v_2}{v_2-2}.$$

$$Var(X) = E(X^2) - [E(X)]^2.$$

Kita hitung :

$$E(X^2) = \frac{\left(\frac{v_2}{v_1}\right)^2 \Gamma\left(\frac{v_1}{2}+2\right)\Gamma\left(\frac{v_2}{2}-2\right)}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}$$

$$= \frac{\left(\frac{v_2}{v_1}\right)^2 \left(\frac{v_1}{2}+1\right)\frac{v_1}{2}\Gamma\left(\frac{v_1}{2}\right).\Gamma\left(\frac{v_2}{2}\right)\left[\left(\frac{v_2}{2}-1\right)\right]\left[\frac{v_2}{2}-2\right]}{\Gamma\left(\frac{v_1}{2}\right)\Gamma\left(\frac{v_2}{2}\right)}$$

$$= \frac{v_2^2\left(\frac{v_1}{2}+1\right).\frac{v_1}{2}}{v_1^2\left(\frac{v_2}{2}-1\right)\left(\frac{v_2}{2}-2\right)}.$$

Maka,

$$Var(X) = \frac{v_2^2(v_1+2)}{v_1(v_2-2)(v_2-4)} - \frac{v_2^2}{(v_2-2)^2}$$

$$= \frac{v_2^2(v_1+2)(v_2-2) - v_1 v_2^2 (v_2-4)}{v_1(v_2-4)(v_2-2)^2}$$

$$= \frac{2v_2^3 - 4v_2^2 + 2v_1 v_2^2}{v_1(v_2-2)^2(v_2-4)} = \frac{2v_2^2(v_1+v_2-2)}{v_1(v_2-2)^2(v_2-4)}.$$

Nilai-nilai persentil $f_\gamma(v_1,v_2)$ untuk variabel random $X \sim F(v_1,v_2)$ sehingga $P\left[X \le f_\gamma(v_1,v_2)\right] = \gamma$

Telah ditabelkan untuk nilai-nilai γ, v_1 dan v_2 yang dipilih. Persentil untuk nilai-nilai γ yang kecil dapat diperoleh dengan mengingat $X \sim F(v_1,v_2)$ maka

persentil

$$Y = \frac{1}{X} \sim F(v_1, v_2)$$

Nilai persentil distribusi F

Sehingga

$$\begin{aligned} 1-\gamma &= P\left[X < f_{1-\gamma}(v_1, v_2)\right] \\ &= P\left[Y > \frac{1}{f_{1-\gamma}(v_1, v_2)}\right] \\ &= 1 - P\left[Y \le \frac{1}{f_{1-\gamma}(v_1, v_2)}\right]. \end{aligned}$$

Maka, $\dfrac{1}{f_{1-\gamma}(v_1, v_2)} = f_\gamma(v_2, v_1)$

atau

$$\boxed{f_{1-\gamma}(v_1, v_2) = \frac{1}{f_\gamma(v_2, v_1)}.}$$

Contoh 1.2.1

Misalkan, $X_1, X_2, ..., X_n$ dan $Y_1, Y_2, ..., Y_n$ adalah dua sampel random independen masing-masing dari distribusi $X_i \sim N(\mu_1; \sigma_1^2)$ dan $Y_j \sim N(\mu_2; \sigma_2^2)$

Jika $v_1 = n_1 - 1$ dan $v_2 = n_2 - 1$ maka

$$\frac{v_1 S_1^2}{\sigma_1^2} \sim \chi^2(v_1) \text{ dan } \frac{v_2 S_2^2}{\sigma_2^2} \sim \chi^2(v_2)$$

sehingga

$$\frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2} \sim F(v_1, v_2).$$

Jadi,

$$\boxed{P\left[\frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_2^2} \le f_{0,95}(v_1, v_2)\right] = 0,95}$$

Perbandingan variansi kedua variabel

dan

$$\boxed{P\left[\frac{S_1^2}{S_2^2 f_{0,95}(v_1, v_2)} \le \frac{\sigma_1^2}{\sigma_2^2}\right] = 0,95.}$$

Jika $n_1 = 16$ dan $n_2 = 20$ maka menurut tabel distribusi F yang ada $f_{0,95}$ (15,20) = 2,20 dan untuk keadaan itu kita mengatakan bahwa kita *yakin* 95% bahwa perbandingan σ_1^2 / σ_2^2 lebih besar dari $S_1^2 / \left[S_2^2 f_{0,95}(15,20) \right]$. Pengertian-pengertian serupa ini akan kita kembangkan lebih lanjut dalam modul- modul mendatang.

Distribusi Beta

Suatu variabel random berdistribusi F dapat ditransformasi untuk mendapatkan distribusi beta.

Jika $X \sim F(v_1, v_2)$ maka

$$Y = \frac{(v_1 / v_2)/X}{1 + (v_1 / v_2)X}$$

berdistribusi beta dengan fungsi peluang

PDF

$$f(y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} \;;\; 0 < y < 1$$

Parameter Beta diambil dari parameter DF pada distribusi F

dengan $a = v_1 / 2$, dan $b = v_2 / 2$. Distribusi beta ini mempunyai parameter $a > 0$ dan $b > 0$, dan ditulis sebagai $Y \sim Beta(a\,;\,b)$.

Mean dan variansi Y dengan mudah dapat dihitung dan kita peroleh

$$E(Y) = \frac{a}{a+b}$$

dan

$$Var(Y) = \frac{ab}{(a+b+1)(a+b)^2}.$$

Persentil ke γ suatu distribusi beta dapat dinyatakan dalam bentuk persentil distribusi F sebagai berikut.

$$y_\gamma(a,b) = \frac{af_\gamma(2a;2b)}{b + af_\gamma(2a;2b)}.$$

Jika a dan b bilangan bulat positif, maka pengintegralan bagian demi bagian berturut-turut menghasilkan hubungan antara fungsi distribusi beta dengan distribusi binomial. Jika X berdistribusi binomial ditulis $X \sim bin\ (n; p)$ dan $Y \sim beta\ (n - i + 1; i)$ maka $F_x\ (i - 1) = F_y\ (1 - p)$. Distribusi beta timbul dalam kaitannya dengan distribusi statistik berurut. Untuk suatu variabel random kontinu $X \sim f(x)$, fungsi peluang statistik berurut ke k dari suatu sampel random berukuran n, ditulis $X_{(k)}$, diberikan oleh:

$$g_k(X_{(k)}) = \frac{n!}{(k-1)!(n-k)!}\left[F(X_{(k)})\right]^{k-1}\left[1-F(X_{(k)})\right]^{n-k} f(x_{(k)}) .$$

Dengan membuat perubahan variabel $U_{(k)} = F(X_{(k)})$ kita peroleh: $U_{(k)} \sim Beta\,(k; n-k+1)$.

Oleh karena $U = F(X)$ berdistribusi uniform (0, 1) maka $U_{(k)}$ adalah statistik berurut ke k variabel random uniform. Fungsi distribusi $X_{(k)}$ dapat dinyatakan dalam bentuk fungsi distribusi beta. Karena:

$$\begin{aligned} G_k(X_{(k)}) &= P\left[X_{(k)} \le x_{(k)}\right] \\ &= P\left[F(X_{(k)}) \le F(x_{(k)})\right] \\ &= H\left[F(x_{(k)}); k, n-k+1\right] \end{aligned}$$

dengan $H(y\,; a,b)$ menunjukkan fungsi distribusi $Y \sim Beta(a, b)$.

Contoh 1.2.2

Misalkan, X berdistribusi eksponensial dengan mean θ. Kita ingin menghitung peluang tentang $X_{(k)}$. Kita punyai:

$$F(x) = 1 - e^{-x/\theta}$$

$$U_{(k)} \sim F(X_{(k)}) \sim Beta(k, n-k+1)$$

dan

$$\begin{aligned} P(X_{(k)} \le c) &= P\left[F(X_{(k)} \le F(c)\right] \\ &= P\left[U_{(k)} \le F(c)\right] \\ &= P\left[\frac{kU_{(k)}}{(n-k+1)(1-U_{(k)})} \le \frac{kF(c)}{(n-k+1)(1-F(c))}\right] \end{aligned}$$

dengan peluang terakhir ini memuat variabel yang berdistribusi $F\,[2k, 2\,(n - k + 1)]$. Jadi untuk nilai-nilai θ, c, k, dan n tertentu, peluang ini dapat diperoleh dari suatu tabel beta kumulatif atau dari suatu tabel F kumulatif jika tersedia tingkat α tertentu. Sebagai contoh, kita ingin menghitung c_γ sehingga $P\left[X_{(k)} \le c_\gamma\right] = \gamma$.

Maka, $f_\gamma\left[2k, 2(n-k+1)\right] = \dfrac{kF(c_\gamma)}{(n-k+1)\left[1-F(c_\gamma)\right]}$

$$= \frac{k\left[1-\exp(-c_\gamma/\theta\right]}{(n-k+1)\exp(-c_\gamma/\theta)}$$

dan

$$c_\gamma = \theta \ln\left[1+\frac{(n-k+1)f_\gamma\left[2k, 2(n-k+1)\right]}{k}\right]$$

Jika n = 11, k = 6, dan γ = 0,95 maka

$$c_\gamma = \theta \ln\left[1+\frac{6(2,69)}{6}\right] = 1,31\theta \text{ dan } P\left[X_{(6)} \le 1,31\ \theta\right] = 0,95$$

atau

$$P\left[\frac{X_{(6)}}{1,31} \le \theta\right] = 0,95$$

Di sini $X_{(6)}$ adalah median (dari sampel berukuran 11) dan θ adalah mean populasinya.

Kita telah mempelajari distribusi beta dan kita telah melihat hubungannya dengan distribusi F dan distribusi kumulatif binomial. Kita juga telah mempelajari aplikasinya untuk distribusi variabel random uniform berurut. Distribusi beta merupakan model generalisasi distribusi uniform dan merupakan model dua parameter yang agak luwes untuk berbagai jenis variabel yang harus terletak antara 0 dan 1.

B. PENDEKATAN SAMPEL BESAR

Distribusi sampling yang telah kita bicarakan di atas mempunyai bentuk pendekatan yang berlaku bagi ukuran-ukuran sampel besar.

Teorema 1.2.6

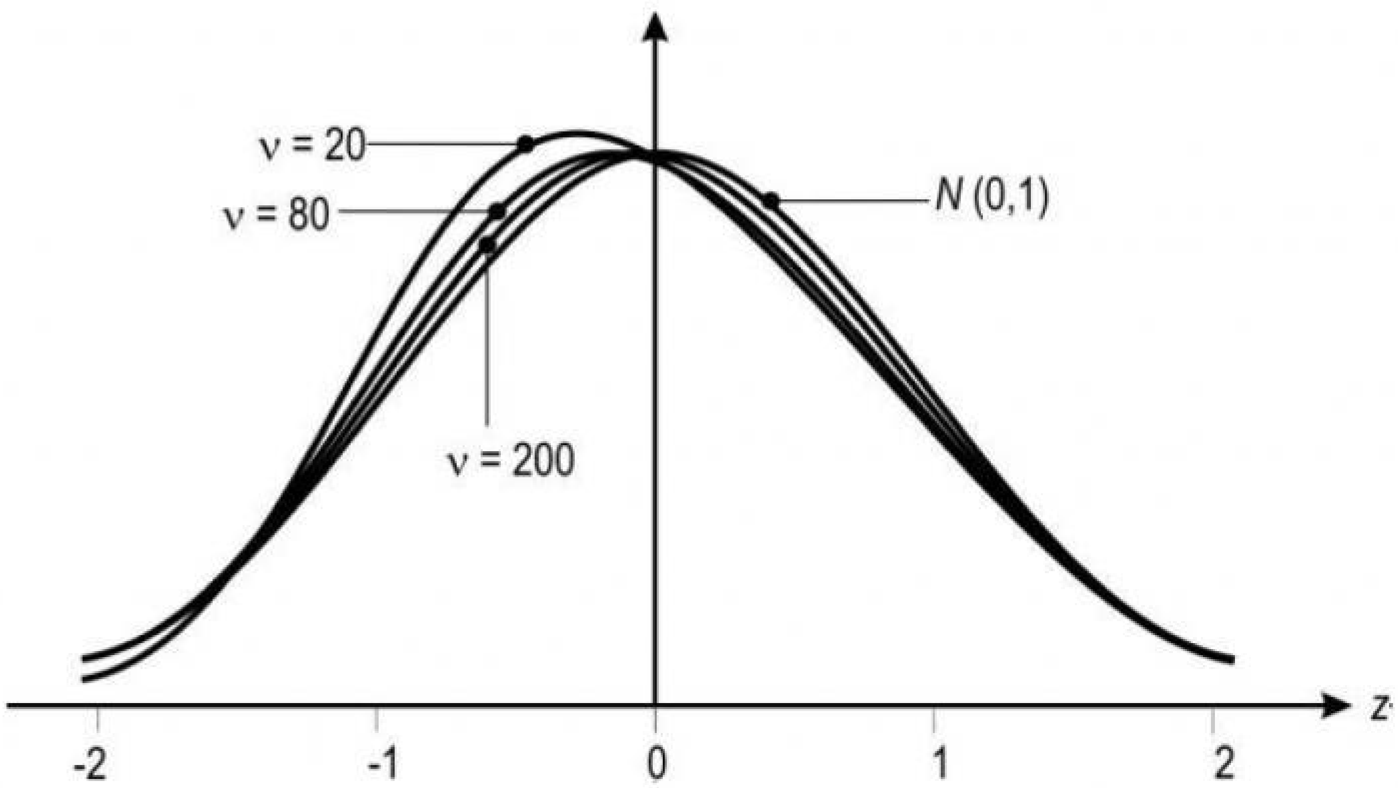
Jika $Y_v \sim \chi^2(v)$ maka

$$Z_v = \frac{Y_v - v}{\sqrt{2v}} \xrightarrow{d} Z \sim N(0;1) \text{ untuk } v \to \infty.$$

Bukti:

Kita tahu bahwa $E(Y_v) = v$ dan $Var\ (Y_v\ ) = 2v$. Ini berarti dapat kita pandang Y_v berdistribusi seperti jumlah $\sum_{i=1}^{v} X_i$ dengan $X_1, X_2, ..., X_v$ independen dan $X_i \sim \chi^2(1)$ sehingga $E(X_i) = 1$ dan $var\ (X_i) = 2$. Maka, menurut teorema limit pusat, Z_v berdistribusi asimtotik normal standar.

Kita juga mengharap fungsi peluang Z_v akan sangat dekat dengan fungsi peluang Z untuk nilai-nilai v yang besar. Hal ini dilukiskan dalam Gambar 1.2.1 yang menunjukkan fungsi peluang untuk v = 20; 80; dan 200.



Gambar 1.2.1.
Pembandingan fungsi peluang distribusi khi-kuadrat standar dan normal standar

Ini berarti bahwa nilai-nilai persentil khi-kuadrat dapat didekati oleh nilai-nilai persentil normal standar untuk v besar. Khususnya:

$$\gamma = P\left[Y_v \le \chi^2_\gamma(v)\right]$$

$$= \Phi\left[\frac{\chi^2_\gamma(v) - v}{\sqrt{2v}}\right]$$

Sehingga

$$\frac{\chi^2_\gamma(v) - v}{\sqrt{2v}} = Z_\gamma$$

$$\chi^2_\gamma(v) = v + z_\gamma\sqrt{2v}.$$

$$\chi^2_{0.95}(30) = 30\left[1 - \frac{2}{9\times 30} + 1.645\sqrt{\frac{2}{9\times 30}}\right]^3$$

Misalnya, untuk v = 30 dan γ = 0,95

$\chi^2_{0,95}(30) = 30 + 1,645\sqrt{60} = 42,74$ dibandingkan dengan nilai yang eksak $\chi^2_{0,95}(30) = 43,77$. Pendekatan yang lebih akurat, yang dikenal sebagai pendekatan **Wilson Hilferty**, diberikan oleh:

$$\boxed{\chi^2_\gamma(v) = v\left[1 - \frac{2}{9v} + Z_\gamma\sqrt{\frac{2}{9v}}\right]^3}$$

Wilson-Hilferty adalah suatu pendekatan untuk menghitung persentil chi-kuadrat yang melibatkan pendekatan asimtotik terhadap distribusi normal standar.

Ini akan memberikan nilai $\chi^2_\gamma(v)/v$ dalam jarak 0,01 dari nilai yang sebenarnya untuk $v \geq 3$, dan $0,01 \leq \gamma \leq 0,99$. Misalnya, jika v = 30 dan γ = 0,95 nilai pendekatan **Wilson-Hilferty** memberikan $\chi^2_{0,95}(30) = 43,77$ yang sama dengan nilai yang sebenarnya sampai dua angka di belakang koma.

Dimungkinkan juga untuk menjabarkan distribusi normal asimtotik secara langsung bagi S_n^2 dan S_n.

Contoh 1.2.3

Misalkan S_n^2 merupakan variansi sampel suatu sampel random berukuran n dari $N(\mu,\sigma^2)$. Kita tahu bahwa dan menurut Teorema 1.2.6, yaitu:

$$\frac{V_n - (n-1)}{\sqrt{2(n-1)}} \xrightarrow{d} Z \sim N(0;1)$$

yakni

$$\frac{\sqrt{n-1}\left[S_n - \sigma^2\right]}{\sigma^2\sqrt{2}} \xrightarrow{d} Z$$

Atau kira-kira

$$S_n^2 \sim N\left[\sigma^2; \frac{2\sigma^4}{n-1}\right].$$

Jika $Y_n = S_n^2$ dan $g(y) = \sqrt{y}$ maka $g'(y) = \dfrac{1}{2\sqrt{y}}$, $g'(\sigma^2) = \dfrac{1}{2\sigma}$ dan secara pendekatan

$$S_n \sim N\left[\sigma; \frac{\sigma^2}{2(n-1)}\right].$$

Artinya, dengan meningkatnya ukuran sampel, distribusi dari S_n mendekati distribusi normal dengan rata-rata σ dan varians $\frac{\sigma^2}{2(n-1)}$.

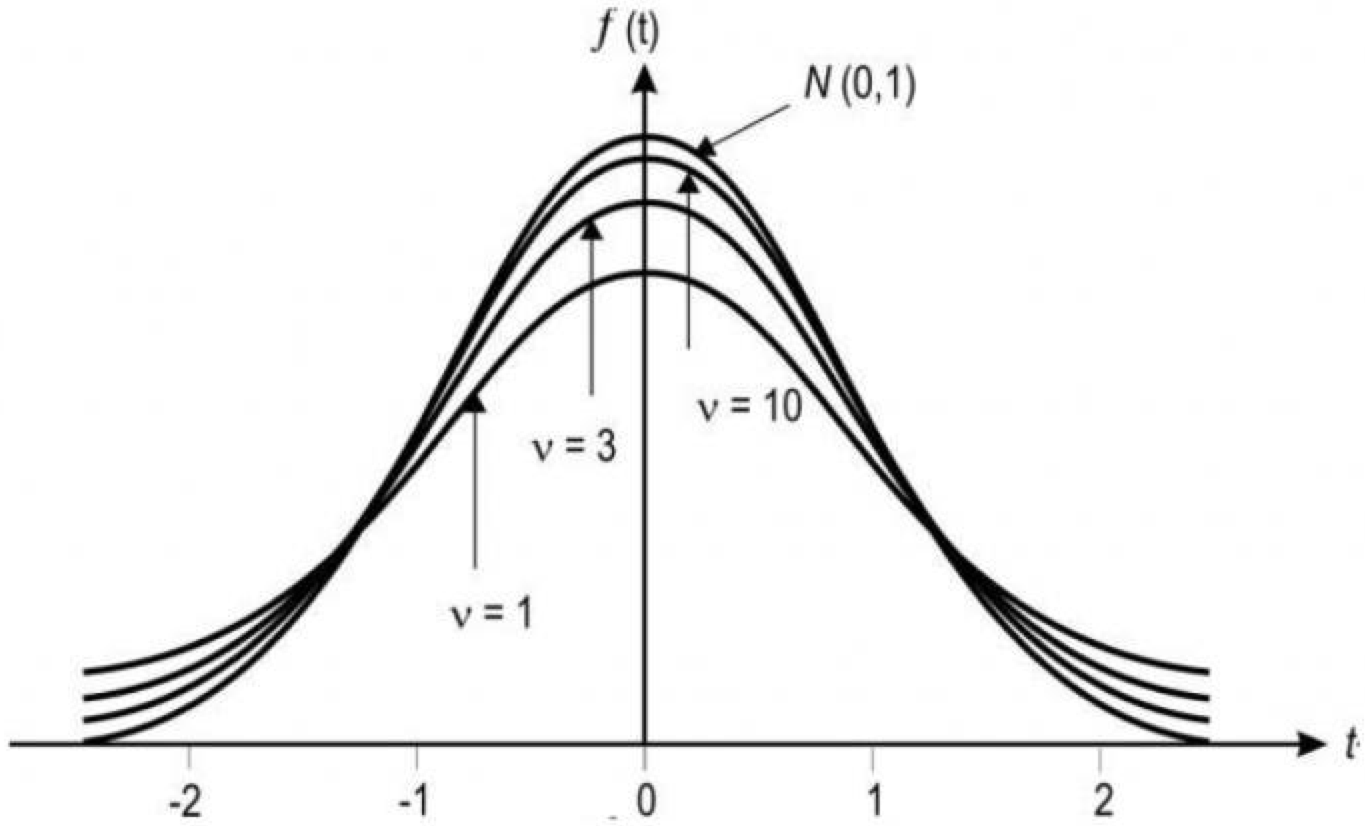
Dapat juga ditunjukkan bahwa variabel random berdistribusi t mempunyai distribusi limit normal standar jika derajat bebas v bertambah besar. Untuk melihat kenyataan ini, kita pandang variabel $T_v \sim t(v)$ dengan:

$$T_v = \frac{Z}{\sqrt{\chi_v^2 / v}}.$$

Kita tahu bahwa $E\left[\chi_v^2 / v\right] = 1$, v a χ_v^2 (v =/ v) dan dengan pertidaksamaan Chebyshev:

$$P\left[\left|\chi_v^2 / v - 1\right| < \varepsilon\right] \geq 1 - \frac{2}{v\varepsilon^2} \text{ sehingga } \frac{\chi_v^2}{v} \xrightarrow{p} 1, \text{ jika } v \to \infty.$$

Jadi, distribusi student's t mempunyai distribusi limit normal standar (*Teorema Slutsky*), yaitu $T_v = \dfrac{Z}{\sqrt{\chi_v^2 / v}} \xrightarrow{d} Z \sim N(0,1)$,



Gambar 1.2.2.

Pembandingan fungsi peluang distribusi t dan normal standar

Hal ini dilukiskan dalam Gambar 1.2.2 yang menunjukkan fungsi peluang $N(0,1)$ dan $t(v)$ untuk $v = 1$, 3, dan 10.

Ini mengakibatkan bahwa persentil t, $(t_\gamma(v))$ mendekati Z_γ untuk v besar, sehingga kita lihat bahwa pada Tabel distribusi t, nilai-nilai pada baris terakhir (bersesuaian dengan $v \to \infty$) sama dengan persentil distribusi normal standar.

Dengan jalan pikiran yang sama dapat kita peroleh nilai pendekatan persentil untuk distribusi F. Misalnya $X_{v_1,v_2} = (W_1/v_1)/(W_2/v_2)$. Telah kita pelajari di atas bahwa $W_2/v_2 \to 1$ jika $v_2 \to \infty$. Jadi, jika v_1 kita ambil tetap $X_{v_1,v_2} \xrightarrow{d} W_1/v_1$, jika $v_2 \to \infty$. Hasil pendekatannya untuk persentil F adalah $f_\gamma(v_1,v_2) = \chi^2_\gamma(v_1)/v_1$ untuk v_2 besar. Argumen yang serupa dapat kita gunakan untuk memperoleh persentil pendekatan bagi v_1 besar, yakni $f_\gamma(v_1,v_2) = v_2/\chi^2_{1-\gamma}(v_2)$.

LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

1) Misalkan, $T \sim t(v)$. Hitunglah distribusi T^2.
2) Misalkan, $X_i \sim N(\mu\,;\,\sigma^2)$, $i = 1, 2, \ldots, n$, dan $Z_i \sim N(0\,;\,1)$, $i = 1, 2, \ldots, k$, dan semua variabel adalah independen. Sebutkan distribusi dari tiap-tiap variabel berikut jika distribusi itu punya nama, jika demikian katakan saja "tak diketahui"

   A. $\dfrac{X_1 - X_2}{\sigma S_z \sqrt{2}}$ B. Z_i^2

   C. $\dfrac{\sqrt{nk}\,(\bar{x} - \mu)}{\sigma \sqrt{\sum_{i=1}^{k} Z_i^2}}$ D. $\dfrac{\bar{X}}{\sigma^2} + \dfrac{\sum_{i=1}^{k} Z_i}{k}$.

3) Gunakan nilai-nilai dalam tabel distribusi yang sesuai untuk memperoleh nilai

   A. $P(7{,}26 < Y < 22{,}31)$ jika $Y \sim \chi^2_{(15)}$

   B. $P\left(\dfrac{Y}{1+Y} > \dfrac{11}{16}\right)$ jika $Y \sim \chi^2_{(6)}$

   C. nilai b sehingga $P(T < b) = 0{,}6$ jika $T \sim t(26)$

   D. $P(2{,}91 < X < 5{,}52)$ jika $X \sim F(7\,;\,12)$.

4) Misalkan, $Y_n \sim \chi^2_{(n)}$. Hitunglah distribusi limit dari $(Y_n - n)/\sqrt{2n}$ jika $n \to \infty$ dengan menggunakan fungsi pembentuk momen.

5) Misalkan, $X_1, X_2, \ldots, X_n$ suatu sampel random dari distribusi yang empat momennya yang pertama ada. Misalkan, pula $S_n^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 /(n-1)$. Tunjukkan bahwa $S_n^2 \xrightarrow{P} \sigma^2$ jika $n \to \infty$

RANGKUMAN

Kita kembangkan metode statistik guna menganalisis mean populasi jika variansi populasi normal tidak diketahui. Ini berkaitan dengan distribusi (*student's*) *t* yang kita peroleh sebagai distribusi variabel random normal standar dibagi dengan akar dari variabel khi-kuadrat yang independen dibagi dengan derajat bebasnya. Atau dapat dituliskan sebagai:

Misalkan, $Z \sim N\ (0\ ;\ 1)$ dan $W \sim \chi^2(n)$ yang independen satu dengan lain maka:

$$T = \frac{Z}{\sqrt{\frac{W}{v}}} \sim t(v).$$

Kita jabarkan juga distribusi *F* dengan yang kita peroleh sebagai distribusi dari perbandingan dua variabel khi-kuadrat yang independen masing-masing dibagi dengan derajat bebasnya. Misalkan, $X \sim \chi^2(v_1)$ dan $Y \sim \chi^2(v_2)$ independen maka:

$$F = \frac{X / v_1}{Y / v_2} \sim F(v_1, v_2).$$

Distribusi variabel ini penting dalam analisis statistik guna membandingkan variansi dua populasi normal yang independen.

Kita pelajari juga variabel random yang merupakan transformasi variabel random F.

Misalkan, X ~ $F(v_1, v_2)$, maka:

$$Y = \frac{(v_1 / v_2) / X}{1 + (v_1 / v_2) X} \sim Beta(a, b)$$

dengan $a = v_1 / 2$ dan $b = v_2 / 2$

TES FORMATIF 2

Pilihlah satu jawaban yang paling tepat!

1) Misalkan $X_i \sim N(\mu;\sigma^2)$; $i = 1, 2, \ldots, n$, dan $Z_i \sim N(0;1)$; $i = 1, 2, \ldots, k$, dan semua variabel adalah independen, maka

a. $\dfrac{Z_1}{\sqrt{Z_2^2}}$ berdistribusi ….

A. $t_{(1)}$
B. $t_{(2)}$
C. $Z_{(1)}$
D. $Z_{(2)}$

Kita tahu bahwa $|Z_2|$ memiliki distribusi chi-kuadrat dengan satu derajat kebebasan (χ_1^2).

Jadi, $\frac{Z_1}{|Z_2|}$ adalah distribusi t-student dengan satu derajat kebebasan (t_1).

b. $\dfrac{Z_1^2}{Z_2^2}$ berdistribusi ….

A. $F_{(0;\,1)}$
B. $F_{(1;\,1)}$
C. $F_{(2;\,0)}$
D. $F_{(2;\,2)}$.

kebebasan tertentu. Dalam hal ini, kita memiliki $\frac{Z_1^2}{Z_2^2}$, di mana $Z_1^2 \sim \chi_1^2$ dan $Z_2^2 \sim \chi_1^2$.

Rasio dua variabel chi-kuadrat yang independen dengan derajat kebebasan v_1 dan v_2 memiliki distribusi F dengan parameter $F_{(v_1, v_2)}$.

Sehingga, $\frac{Z_1^2}{Z_2^2}$ memiliki distribusi F dengan parameter $F_{(1,1)}$.

c. $\dfrac{\bar{X}}{\bar{Z}}$ berdistribusi ….

A. $t_{(1)}$
B. $F_{(1\,;\,1)}$
C. $Z_{(2)}$
D. tidak diketahui.

$$\frac{\bar{X}}{\bar{Z}} = \frac{\frac{\bar{X}}{\mu}}{\frac{\bar{Z}}{0}}$$

Namun, informasi lebih lanjut diperlukan untuk memastikan distribusi ini. Jadi, jawaban yang benar adalah D. tidak diketahui, karena kita tidak memiliki informasi lebih lanjut mengenai distribusi

d. $\dfrac{(k-1)\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)\ \sigma^2 \sum_{i=1}^{k}(Z_i - \bar{Z})^2}$ berdistribusi ….

A. $Z_{(k-1)}$
B. $t_{(n-1)}$
C. $F_{(n-1;\,k-1)}$
D. tidak tahu.

Mengingat bahwa $\sum_{i=1}^{n}(X_i - \bar{X})^2$ dan $\sum_{i=1}^{k}(Z_i - \bar{Z})^2$ adalah variabel chi-kuadrat dengan derajat kebebasan $n-1$ dan $k-1$ masing-masing, dan keduanya bersifat independen, kita dapat menyederhanakan ekspresi ini menjadi rasio dua variabel chi-kuadrat independen.

Sehingga, ekspresi ini memiliki distribusi F dengan parameter $F_{(n-1,k-1)}$.

2) Gunakan tabel distribusi yang sesuai.

a. Maka, nilai b sehingga $P(Y \le b) = 0{,}75$ jika $Y \sim \chi^2\,(23)$ adalah ….

A. 25,48
B. 26,16

Cari nilai Chi Square dengan df=23 dengan alpha 1-0.75

k	0.995	0.99	0.975	0.95	0.9	0.75	0.5	0.25	0.1
	Probability Content, p, between χ								
1	3.927e-5	1.570e-4	9.820e-4	0.00393	0.0157	0.102	0.455	1.323	2.706
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.386	2.773	4.605
3	0.0717	0.115	0.216	0.352	0.584	1.213	2.366	4.108	6.251
4	0.207	0.297	0.484	0.711	1.064	1.923	3.357	5.385	7.779
5	0.412	0.554	0.831	1.145	1.610	2.675	4.351	6.626	9.236
6	0.676	0.872	1.237	1.635	2.204	3.455	5.348	7.841	10.645
7	0.989	1.239	1.690	2.167	2.833	4.255	6.346	9.037	12.017
8	1.344	1.646	2.180	2.733	3.490	5.071	7.344	10.219	13.362
9	1.735	2.088	2.700	3.325	4.168	5.899	8.343	11.389	14.684
10	2.156	2.558	3.247	3.940	4.865	6.737	9.342	12.549	15.987
11	2.603	3.053	3.816	4.575	5.578	7.584	10.341	13.701	17.275
12	3.074	3.571	4.404	5.226	6.304	8.438	11.340	14.845	18.549
13	3.565	4.107	5.009	5.892	7.042	9.299	12.340	15.984	19.812
14	4.075	4.660	5.629	6.571	7.790	10.165	13.339	17.117	21.064
15	4.601	5.229	6.262	7.261	8.547	11.037	14.339	18.245	22.307
16	5.142	5.812	6.908	7.962	9.312	11.912	15.338	19.369	23.542
17	5.697	6.408	7.564	8.672	10.085	12.792	16.338	20.489	24.769
18	6.265	7.015	8.231	9.390	10.865	13.675	17.338	21.605	25.989
19	6.844	7.633	8.907	10.117	11.651	14.562	18.338	22.718	27.204
20	7.434	8.260	9.591	10.851	12.443	15.452	19.337	23.828	28.412
21	8.034	8.897	10.283	11.591	13.240	16.344	20.337	24.935	29.615
22	8.643	9.542	10.982	12.338	14.041	17.240	21.337	26.039	30.813
23	9.260	10.196	11.689	13.091	14.848	18.137	22.337	27.141	32.007
24	9.886	10.856	12.401	13.848	15.659	19.037	23.337	28.241	33.196

C. 27,14
D. 28,71

b. Jika $T \sim t(13)$, maka $P[0{,}87 < T < 2{,}65]$ sama dengan ….
A. 0,16
B. 0,17
C. 0,18
D. 0,19.

We need to compute $\Pr(0.87 \leq T \leq 2.65)$, where T has a t-distribution with $df = 13$ degrees of freedom. Therefore, the probability is computed as:

$$\begin{aligned}\Pr(0.87 \leq T \leq 2.65) &= \Pr(T \leq 2.65) - \Pr(T \leq 0.87) \\ &= 0.99 - 0.8 \\ &= 0.19\end{aligned}$$

c. Jika $T \sim t\ (23)$, maka nilai k sehingga $P(|T| \geq k) = 0{,}02$ adalah ….
A. 0,001
B. 0,01
C. 0,05
D. 0,1

Seharusnya 2.5 (tidak ada jawaban)

d. Jika $X \sim F_{(20;\ 8)}$, maka $P\left[\dfrac{1}{X} > 0{,}25\right]$ sama dengan ….
A. 0,975
B. 0,950
C. 0,925
D. 0,900

Diketahui:
X ~ F_{(20,8)}
Artinya, X berdistribusi F dengan derajat kebebasan pembilang (df1) = 20 dan derajat kebebasan penyebut (df2) = 8.

Ditanyakan:
P[1/X > 0.25]

Langkah-langkah:
1. 1/X ~ F_{(8,20)} karena sifat distribusi F bahwa 1/F ~ F_{(df2,df1)}
2. P(1/X > 0.25) = P(X < 4)
3. Gunakan tabel distribusi F untuk mencari peluang X < 4 pada df1 = 8 dan df2 = 20
4. Didapat P(X < 4) = 0,951

Jadi, P[1/X > 0.25] = 0,951

Sehingga, jawabannya adalah 0,951

3) Pandang kembali soal nomor 2 Tes Formatif 1. Maka, dengan menggunakan distribusi normal.
a. Soal 2 (b) memberi jawaban ….
A. 0,6666
B. 0,7712
C. 0,8362
D. 0,9236.

b. Soal nomor 2 (c) memberi jawaban ….
A. 10
B. 12
C. 14
D. 16.

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 2 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 2.

$$\text{Tingkat penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100\%$$

Arti tingkat penguasaan: 90 - 100% = baik sekali
80 - 89% = baik
70 - 79% = cukup
< 70% = kurang

Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan modul selanjutnya. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 2, terutama bagian yang belum dikuasai.

Kunci Jawaban Tes Formatif

Tes Formatif 1

1)	a. B	b. A	
2)	b. A	b. D	c. B
3)	c. D	b. C	c. C
40	a. C	b. C	

Tes Formatif 2

1)	a. A	b. B	c. D	d. C
2)	a. C	b. D	c. B	d. A
30	a. D	b. B		

Daftar Pustaka

Bain, L.J. & Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics* 2^{nd}. California: Duxbury Press.

Hogg, R.V. & Craig, A.T. (1995). *Introduction to Mathematical Statistics* 5^{th}. Englewood Cliffs, New Jersey: Prentice Hall.