Pokok Bahasan Bab 9
Estimasi Parameter Populasi

Probabilitas & Statistika

Materi yang Dibahas:

1. Penaksiran dengan Metode Klasik
2. Menaksir Rataan dari Sampel Tunggal
3. Interval Kepercayaan μ
4. Menaksir Variansi Sampel Tunggal
5. Interval Kepercayaan σ^2
6. Menaksir Rasio Dua Variansi Dua Sampel

Tim Penyusun

Judhi Santoso
Harlili
Dwi H. Widyantoro

Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung

Global Development Learning Network

Pendahuluan

- Inferensi statistik: proses menggunakan analisis data untuk menyimpulkan properti dari distribusi probabilitas-nya.
- Analisis statistik inferensial menyimpulkan sifat-sifat suatu populasi, misalnya dengan menguji hipotesis dan memperoleh taksiran.
- Diasumsikan bahwa kumpulan data yang diamati diambil sampelnya dari populasi yang lebih besar.
- Bab 9 ini membahas tentang **estimasi/taksiran parameter populasi**.

Metode Estimasi Klasik

Global Development Learning Network

Penaksiran Parameter Populasi Metode Klasik

- Menaksir parameter populasi, seperti rataan, proporsi, variansi diperoleh dari perhitungan besaran statistik sampel random dan teori distribusi sampel.
- Menaksir parameter ada 2 cara:
- 1. Taksiran titik
- 2. Taksiran Interval

Distribusi dan Parameternya

Distribusi	Parameter
Uniform Diskrit/ Kontinu	A = nilai terkecil, B = nilai terbesar
Binomial	N = banyak item, p = proposi sukses
Hypergeometrik	N = banyak item, n=banyak sampel, k= banyak sukses
Poisson	$\beta = \lambda t$ = rataan = variansi; λ = rata item per unit, t=interval unit (waktu, daerah ,dll)
Normal	μ = rataan , σ^2 = variansi
Gamma	α = banyak kejadian, $\beta = 1/\lambda$
Weibull	α = banyak kejadian, $\beta = 1/\lambda$
Eksponensial	β = rataan
Lognormal	μ = rataan , σ^2 = variansi

Penaksiran dengan Metode Klasik

- Sebuah nilai taksiran dari parameter populasi ϑ adalah sebuah nilai tunggal dari $\hat{\theta}$ statistik $\hat{\Theta}$
- Sebuah nilai penaksir tidak diharapkan dapat menaksir parameter populasi tanpa kesalahan, misalkan tidak perlu $\bar{X}$ dapat menaksir μ secara tepat, tetapi diharapkan tidak terlalu jauh dari parameter yang ditaksir.

Penaksir Tak Bias (1)

Penaksir yang baik memiliki sifat **tak bias (*unbiased*)**.

- Definisi:

Sebuah statistik $\hat{\Theta}$ dikatakan penaksir tak bias dari parameter ϑ jika:

$$\mu_{\hat{\Theta}} = E(\hat{\Theta}) = \theta$$

Artinya nilai Ekspektasi, rataan penaksir $\hat{\Theta}$ sama dengan parameter yang ditaksir.

Penaksir Tak Bias (2)

- Contoh:
  Tunjukkan bahwa S^2 adalah penaksir tak bias dari parameter σ^2!

Penaksir Tak Bias (3)

- Jawab: Kita tuliskan

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

Sekarang tentukan

$$E(S^2) = E\left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i - \mu)^2 - nE(\bar{X} - \mu)^2\right]$$
$$= \frac{1}{n-1}\left(\sum_{i=1}^{n}\sigma^2{}_{X_i} - n\sigma^2{}_{\bar{X}}\right)$$

Penaksir Tak Bias (4)

Tetapi

$$\sigma^2{}_{X_i} = \sigma^2 \;\; untuk \;\; i = 1,2,\dots,n \; dan \; \sigma^2{}_{\bar{X}} = \frac{\sigma^2}{n}$$

Sehingga

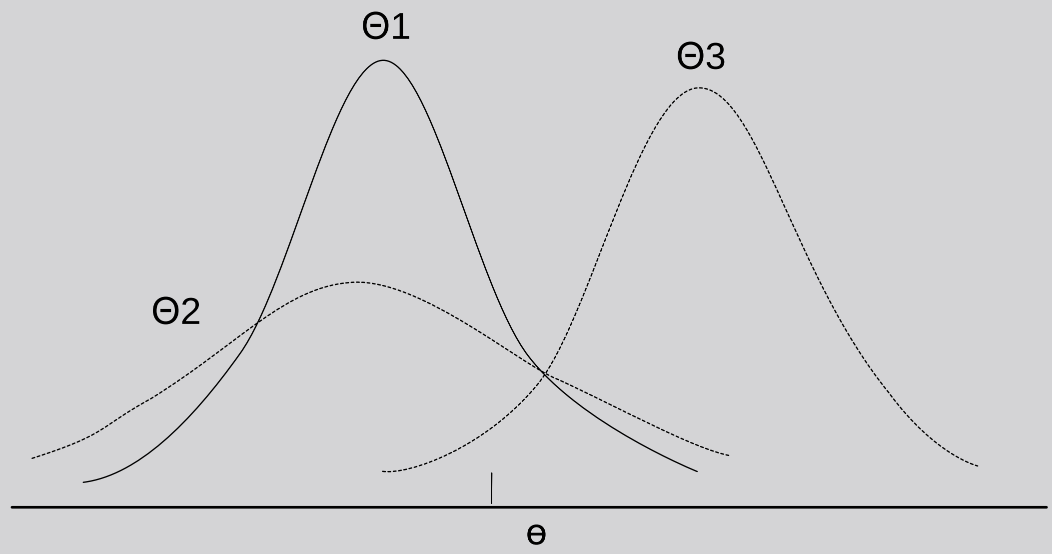
$$E(S^2) = \frac{1}{n-1}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) = \sigma^2$$

Variansi Nilai Penaksir Titik

Penaksir yang baik memiliki variansi terkecil .

- Jika kita mengumpulkan semua penaksir tak bias yang mungkin dari parameter ϑ, maka salah satu penaksir yang memiliki variansi terkecil dikatakan penaksir yang paling efisien dari ϑ.

Kurva Distribusi Sampling



Gambar 7.1.1 Distribusi *Sampling* dari Penaksir θ yang Berbeda

Sampel Tunggal

Taksiran Interval Kepercayaan
Interval Prediksi
Limit Toleransi

Penaksiran Interval Rataan

- Sebuah penaksiran interval dari parameter populasi ϑ adalah sebuah selang $\hat{\theta}_L < \theta < \hat{\theta}_U$ yang bergantung pada nilai statistik untuk sampel tertentu dan distribusi *sampling*-nya.
- Sampel-sampel yang berbeda secara umum akan menghasilkan nilai yang berbeda dari $\widehat{\Theta}$, atau $\widehat{\theta}_L$ dan $\ddot{\theta}_U$, sehingga kita dapat menentukan suatu nilai *α* = error yang terletak antara 0 dan 1 demikian sehingga:

$$P\left(\widehat{\Theta}_L < \theta < \widehat{\Theta}_U\right) = 1 - \alpha$$

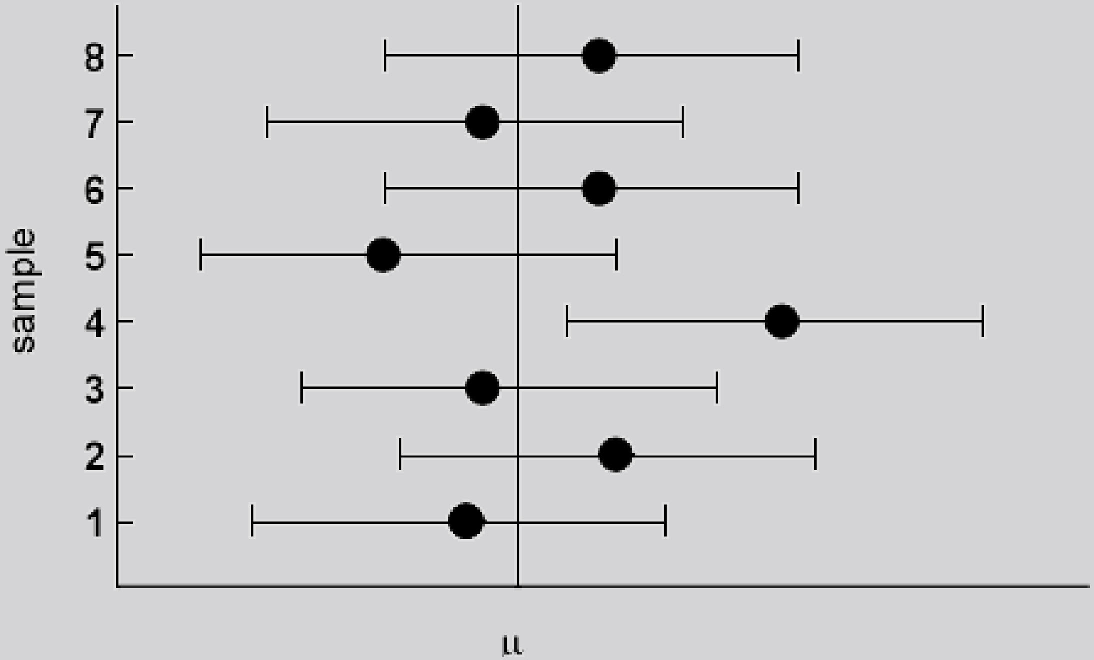
- Interval $\ddot{\theta}_L < \theta < \ddot{\theta}_U$ dihitung dari sampel dengan **interval kepercayaan** , *confidence interval*, (1-*α*) 100%

Arti Interval Kepercayaan (Confidence Interval)

- Jika error α = 1 % =0,01 maka interval kepercayaan = 1- α = 99 % = 0,99
- Jika error α = 5 % =0,05 maka interval kepercayaan = 1- α = 95 % = 0,95
- Interval kepercayaan 95 % untuk taksiran rataan adalah interval ($\hat{\theta}_L$, $\ddot{\theta}_U$) artinya 95 % dari seluruh data terletak pada interval tersebut.
- Idealnya interval pendek dengan selang kepercayaan tinggi.

Taksiran Interval μ

- Kurva:



Gambar 7.2.1 Interval Kepercayaan μ

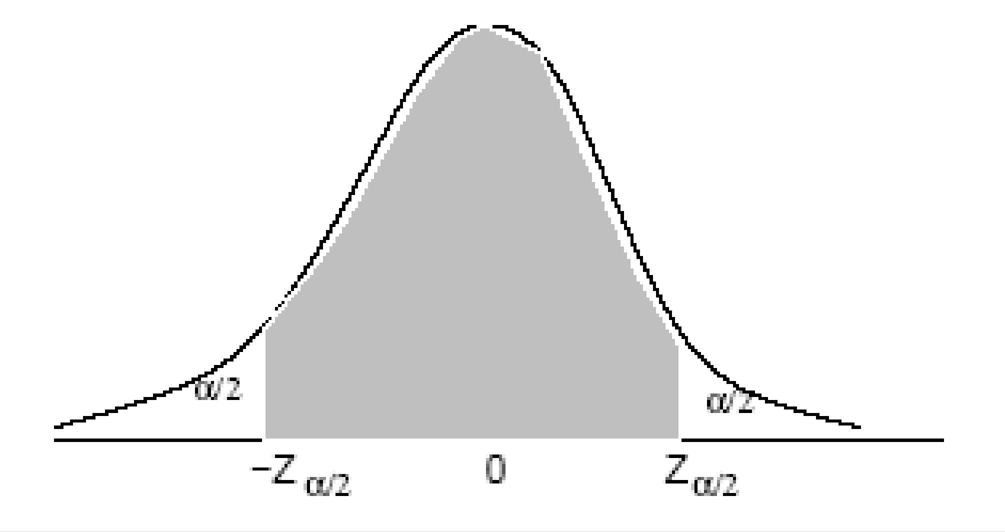
Interval taksiran μ untuk selang berbeda menghasilkan taksiran yang berbeda pula.

Penaksiran Rataan Dari Sampel Tunggal (1)

- Akan ditentukan taksiran interval dari μ. Misalkan sampel diambil dari populasi normal, atau jika tidak normal sampel mempunyai ukuran yang besar. Sesuai dengan teorema limit pusat, diharapkan distribusi sampel $\bar{X}$ akan mendekati normal dengan rataan $\mu_{\bar{X}} = \mu$ dan simpangan baku $\sigma_{\bar{X}} = \sigma/\sqrt{n}$.

Penaksiran Rataan Dari Sampel Tunggal (2)

- Selanjutnya peluang Z yang terletak antara $-z_{\alpha/2}$ *dan* $z_{\alpha/2}$ ditunjukkan pada kurva berikut:



Gambar 7.1.2 $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$

Penaksiran Rataan Dari Sampel Tunggal (3)

- Dari Gambar 7.1.2 dapat dilihat:

$$P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1-\alpha$$

di mana:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

Sehingga:

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

Atau dapat dituliskan:

$$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

Penaksir Interval μ (1)

- Definisi Penaksir interval μ untuk σ diketahui:

  Jika $\bar{x}$ adalah rataan dari sampel random dengan ukuran *n* dari sebuah populasi dan variansi σ^2, penaksir interval μ dengan tingkat kepercayaan $(1-\alpha)100\%$ adalah:

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

  di mana $z_{\alpha/2}$ adalah nilai z yang memberikan luas $\frac{\alpha}{2}$ sebelah kanan nilai tersebut.

Nilai kesalahan , e <= $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Ukuran sampel , dengan nilai kesalahan e adalah n, $n = \left(\frac{z_{\alpha/2}\sigma}{e}\right)^2$

Contoh 9.2:

Rataan konsentrasi zinc dari pengukuran di 36 lokasi diperoleh 2.6 gram per mililiter. Hitunglah Interval penaksiran rataan dengan kepercayaan 95% .

Jawab

Nilai taksiran dari μ adalah $\bar{x}$ = 2.6. Nilai z yang memberikan luas 0.025 sebelah kanan atau 0.975 sebelah kiri adalah sehingga selang kepercayaan 95 % adalah

$$z_{0.025} = 1.96$$

atau

$$2.6 - (1.96)\frac{0.3}{\sqrt{36}} < \mu < 2.6 + (1.96)\frac{0.3}{\sqrt{36}}$$

$$2.50 < \mu < 2.70$$

Banyak sampel μ

If $\bar{x}$ is used as an estimate of μ, we can be $100(1-\alpha)\%$ confident that the error will not exceed a specified amount e when the sample size is

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

Banyak sampel μ

- Contoh 9.3:

  Berapa jumlah sampel yang diperlukan pada contoh 9.2 sebelumnya agar kita memiliki tingkat kepercayaan 95% bahwa taksiran μ memiliki kesalahan kurang dari 0.05?

  Jawab:

  Simpangan baku populasi adalah σ = 0.3. Dengan teorema sebelumnya,

$$n = \left(\frac{(1.96)(0.3)}{0.05}\right)^2 = 138.3 \approx 139$$

One-sided Confidence Bounds - μ dan σ diketahui

- Jika $\bar{x}$ adalah rataan dari sampel random dengan ukuran *n* dari sebuah populasi dan variansi σ^2, batas atas/bawah μ dengan tingkat kepercayaan $(1-\alpha)100\%$ adalah:

- Batas atas penaksiran satu sisi = $\bar{x} + z_\alpha \sigma/\sqrt{n};$
- Batas bawah penaksiran satu sisi = $\bar{x} - z_\alpha \sigma/\sqrt{n}.$

One-sided Confidence Bounds - σ tidak diketahui

- Gunakan distribusi T untuk melakukan penaksiran (Lihat kembali Bab 8).

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal population with unknown variance σ^2, a $100(1-\alpha)\%$ confidence interval for μ is

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

where $t_{\alpha/2}$ is the t-value with $v = n - 1$ degrees of freedom, leaving an area of $\alpha/2$ to the right.

Interval Prediksi (Prediction Interval) σ diketahui

- Prediksi interval dengan kepercayaan $(1-\alpha)100\%$ untuk nilai observasi yang akan datang (future) adalah x_0 dari populasi normal dengan rataan = μ tidak diketahui dan variansi = σ^2 diketahui

$$\bar{x} - z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} < x_0 < \bar{x} + z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}}$$

Interval Prediksi (Prediction Interval) - σ tidak diketahui

- Prediksi interval dengan kepercayaan $(1-\alpha)100\%$ untuk nilai observasi yang akan datang (future) adalah x_0 dari populasi normal dengan rataan = μ dan variansi = σ^2 tidak diketahui:

$$\bar{x} - t_{\alpha/2} s\sqrt{1+1/n} < x_0 < \bar{x} + t_{\alpha/2} s\sqrt{1+1/n},$$

Menggunakan distribusi T, dengan v = n-1 derajat kebebasan.

Limit Toleransi

- Untuk populasi berdistribusi normal dengan rataan = μ dan variansi = σ² tidak diketahui, limit toleransi = $\bar{x} \pm ks$
- Interval toleransi ($\bar{x} - ks$,$\bar{x} + ks$)
- Nilai k ditentukan dari table A7 dengan kepercayaan (1-α) memuat 100(1-γ)% pengukuran.

Overview: Taksiran Interval μ

σ DIKETAHUI

- Taksiran interval kepercayaan μ (Two-sided confidence bounds):

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

  One-sided confidence bounds:

$$\mu > \bar{x} - z_{\alpha}\frac{\sigma}{\sqrt{n}}$$

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

- Taksiran interval Prediksi future observation:

$$\bar{x} - z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}} < x_0 < \bar{x} + z_{\alpha/2}\sigma\sqrt{1+\frac{1}{n}}$$

σ TIDAK DIKETAHUI

- Taksiran interval kepercayaan μ:

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

- Taksiran Prediksi interval Prediksi future observation:

$$\bar{x} - t_{\alpha/2}s\sqrt{1+\frac{1}{n}} < x_0 < \bar{x} + t_{\alpha/2}s\sqrt{1+\frac{1}{n}}$$

- Taksiran Tolerance limit:

$$\bar{x} \pm ks$$

Case Study 1

- A machine produces metal pieces that are cylindrical in shape. A sample of these pieces is taken and the diameters are found to be 1.01, 0.97, 1.03, 1.04, 0.99, 0.98, 0.99,1.01 and 1,03. Use these data to calculate three interval types and draw interpretations that illustrate the distinction between them in the context of the system. For all computations, assume an approximately normal distribution. The sample mean and standard deviation for the given data = 1.0056 dan S = 0.0246.

Case Study 1 (Lanjutan)

- A) Find a 99 % confidence interval on the mean diameter.
- B) Compute a 99 % prediction interval on a measured diameter of a single metal piece taken from the machine.
- C) Find the 99% tolerance limit that will contain 95 % of the metal pieces produce by this machine.

Case Study 1: Solusi

(a) The 99% confidence interval for the mean diameter is given by

$$\bar{x} \pm t_{0.005} s/\sqrt{n} = 1.0056 \pm (3.355)(0.0246/3) = 1.0056 \pm 0.0275.$$

Thus, the 99% confidence bounds are 0.9781 and 1.0331.

(b) The 99% prediction interval for a future observation is given by

$$\bar{x} \pm t_{0.005} s \sqrt{1 + 1/n} = 1.0056 \pm (3.355)(0.0246)\sqrt{1 + 1/9},$$

with the bounds being 0.9186 and 1.0926.

(c) From Table A.7, for $n = 9$, $1 - \gamma = 0.99$, and $1 - \alpha = 0.95$, we find $k = 4.550$ for two-sided limits. Hence, the 99% tolerance limits are given by

$$\bar{x} + ks = 1.0056 \pm (4.550)(0.0246),$$

with the bounds being 0.8937 and 1.1175. We are 99% confident that the tolerance interval from 0.8937 to 1.1175 will contain the central 95% of the distribution of diameters produced.

Dua Sampel

Taksiran 2 Rataan
Taksiran Paired Observations

Penaksir Interval Kepercayaan μ_1-μ_2

- Jika kita memiliki 2 populasi dengan rataan μ_1 dan μ_2 dan variansi σ_1^2 dan σ_2^2, taksiran selisih μ_1-μ_2 dapat dilakukan dengan statistic X_1-X_2 (lihat Bab 8).
- Untuk σ_1^2 dan σ_2^2 diketahui:

Confidence Interval for $\mu_1 - \mu_2$, σ_1^2 and σ_2^2 Known

If $\bar{x}_1$ and $\bar{x}_2$ are means of independent random samples of sizes n_1 and n_2 from populations with known variances σ_1^2 and σ_2^2, respectively, a $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ is given by

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

where $z_{\alpha/2}$ is the z-value leaving an area of $\alpha/2$ to the right.

Penaksir Interval Kepercayaan μ_1-μ_2

- Untuk $\sigma_1^2 = \sigma_2^2$, keduanya tidak diketahui nilainya, menggunakan nilai-t:

Confidence Interval for $\mu_1 - \mu_2$, $\sigma_1^2 = \sigma_2^2$ but Both Unknown	If $\bar{x}_1$ and $\bar{x}_2$ are the means of independent random samples of sizes n_1 and n_2, respectively, from approximately normal populations with unknown but equal variances, a $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ is given by $$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$ where s_p is the pooled estimate of the population standard deviation and $t_{\alpha/2}$ is the t-value with $v = n_1 + n_2 - 2$ degrees of freedom, leaving an area of $\alpha/2$ to the right.

Penaksir Interval Kepercayaan μ_1-μ_2

- Untuk $\sigma_1^2 \neq \sigma_2^2$, keduanya tidak diketahui nilainya, menggunakan nilai-t:

Confidence Interval for $\mu_1 - \mu_2$, $\sigma_1^2 \neq \sigma_2^2$ and Both Unknown

If $\bar{x}_1$ and s_1^2 and $\bar{x}_2$ and s_2^2 are the means and variances of independent random samples of sizes n_1 and n_2, respectively, from approximately normal populations with unknown and unequal variances, an approximate $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$ is given by

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

where $t_{\alpha/2}$ is the t-value with

$$v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{[(s_1^2/n_1)^2/(n_1-1)] + [(s_2^2/n_2)^2/(n_2-1)]}$$

degrees of freedom, leaving an area of $\alpha/2$ to the right.

Penaksir Interval Kepercayaan μ_1-μ_2 untuk Paired Observations

- Paired observations: 2 populasi untuk eksperimen homogen, tidak acak, setiap eksperimen memiliki pasangan observasi, 1 untuk setiap populasi . Contoh: Populasi berat badan sebelum dan sesudah diet.

Confidence Interval for $\mu_D = \mu_1 - \mu_2$ for Paired Observations

If $\bar{d}$ and s_d are the mean and standard deviation, respectively, of the normally distributed differences of n random pairs of measurements, a $100(1-\alpha)\%$ confidence interval for $\mu_D = \mu_1 - \mu_2$ is

$$\bar{d} - t_{\alpha/2}\frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\alpha/2}\frac{s_d}{\sqrt{n}},$$

where $t_{\alpha/2}$ is the t-value with $v = n - 1$ degrees of freedom, leaving an area of $\alpha/2$ to the right.

Taksiran Proporsi dan Variansi

Proporsi Sampel Tunggal
Proporsi Dua Sampel
Variansi Sampel Tunggal
Variansi Dua Sampel

Institut Teknologi Bandung

Global Development Learning Network

Interval Kepercayaan proporsi

- Proporsi *p*: mengikuti prinsip eksperimen binomial
- Banyak sukses *x*: jumlah nilai *n*, yang isinya hanya 0 atau 1
- *p*: rataan sampel dari nilai-*n* tsb
- Untuk sampel tunggal:

Large-Sample Confidence Intervals for p

If $\hat{p}$ is the proportion of successes in a random sample of size n and $\hat{q} = 1 - \hat{p}$, an approximate $100(1-\alpha)\%$ confidence interval, for the binomial parameter p is given by (method 1)

$$\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

or by (method 2)

$$\frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n}}{1 + \frac{z^2_{\alpha/2}}{n}} - \frac{z_{\alpha/2}}{1 + \frac{z^2_{\alpha/2}}{n}}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}} < p < \frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n}}{1 + \frac{z^2_{\alpha/2}}{n}} + \frac{z_{\alpha/2}}{1 + \frac{z^2_{\alpha/2}}{n}}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}},$$

where $z_{\alpha/2}$ is the z-value leaving an area of $\alpha/2$ to the right.

Interval Kepercayaan proporsi

- Taksiran selisih dua proporsi:

Large-Sample Confidence Interval for $p_1 - p_2$

If $\hat{p}_1$ and $\hat{p}_2$ are the proportions of successes in random samples of sizes n_1 and n_2, respectively, $\hat{q}_1 = 1 - \hat{p}_1$, and $\hat{q}_2 = 1 - \hat{p}_2$, an approximate $100(1-\alpha)\%$ confidence interval for the difference of two binomial parameters, $p_1 - p_2$, is given by

$$(\hat{p}_1 - \hat{p}_2) - z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}},$$

where $z_{\alpha/2}$ is the z-value leaving an area of $\alpha/2$ to the right.

Penaksir Interval Kepercayaan Variansi σ^2

- Definisi:

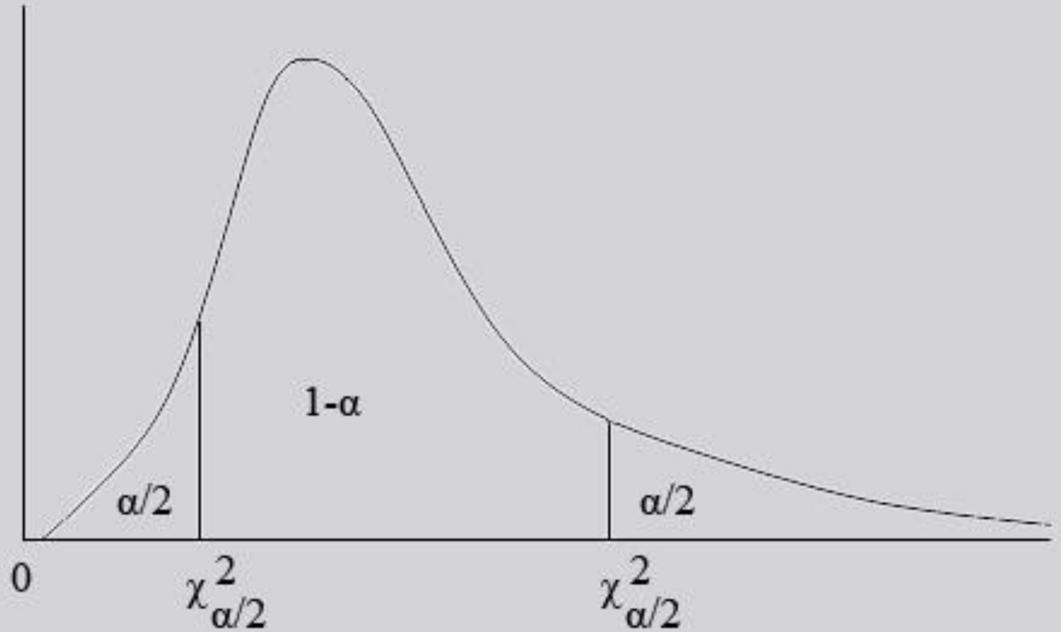
Jika s^2 adalah variansi sampel random berukuran n dari populasi normal, interval kepercayaan $(1-\alpha)100\%$ dari σ^2 adalah:

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$$

dengan $\chi^2_{\alpha/2}$ dan $\chi^2_{1-\alpha/2}$ adalah nilai *chi-squared* dengan n-1 derajat kebebasan yang mempunyai luas di sebelah kanan $\alpha/2$ dan $1-\alpha/2$

Penjelasan Penaksir Variansi σ^2

- Kurva:



$$P(\chi^2_{1-\alpha/2} < X^2 < \chi^2_{\alpha/2}) = 1 - \alpha$$

Gambar 7.2.2 Interval Penaksiran

Contoh : Example 9.18

Berat 10 paket biji rumput yang didistribusikan oleh perusahaan tertentu adalah 46.4; 46.1; 45.8; 47.0; 46.1; 45.9; 45.8; 46.9; 45.2; 46.0.
Hitunglah selang kepercayaan 95% dari variansinya, asumsi distribusi normal.

Analisis Soal

Berat 10 paket biji rumput yang didistribusikan oleh perusahaan tertentu adalah 46.4; 46.1; 45.8; 47.0; 46.1; 45.9; 45.8; 46.9; 45.2; 46.0.
Hitunglah selang kepercayaan 95% dari variansinya, asumsi distribusi normal.

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$$

n=10 → v=9

$\alpha=0.05 \rightarrow \chi^2_{\alpha/2} = \chi^2_{0.025} = 19.023$ dan $\chi^2_{0.975} = 2.7$

s^2 dihitung, lalu tentukan intervalnya.

Institut Teknologi Bandung

Solusi Interval Kepercayaan σ²

Hitung dulu $$s^2 = \frac{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2}{n(n-1)}$$
$$= \frac{(10)(21{,}273.12) - (461.2)^2}{(10)(9)}$$
$$= 0.286$$

Untuk selang 95%, pilih α=0.05, dengan tabel nilai χ^2 dengan v=9 maka $\chi^2_{0.025}$ =19.023 dan $\chi^2_{0.975}$ =2.7

$$\frac{(9)(0.286)}{19.023} < \sigma^2 < \frac{(9)(0.286)}{2.700}$$

Dengan demikian intervalterval kepercayaan 95% adalah: $0.135 < \sigma^2 < 0.953$

Contoh

Perusahaan baterai mobil mengklaim bahwa produknya secara rata-rata berumur 3 tahun dengan simpangan 1 tahun. Jika 5 baterai mempunyai umur 1.9; 2.4; 3.0; 3.5; dan 4.2 tahun, tentukan selang kepercayaan 95% untuk σ^2 dan berilah pendapat apakah klaim perusahaan yang menyatakan bahwa $\sigma^2 = 1$ adalah valid?
Asumsi distribusi umur baterai adalah normal.

Analisis Soal

Perusahaan baterai mobil mengklaim bahwa produknya secara rata-rata berumur 3 tahun dengan simpangan 1 tahun. Jika 5 baterai mempunyai umur 1.9; 2.4; 3.0; 3.5; dan 4.2 tahun, tentukan selang kepercayaan 95% untuk σ^2 dan berilah pendapat apakah klaim perusahaan yang menyatakan bahwa $\sigma^2 = 1$ adalah valid? Asumsi distribusi umur baterai adalah normal.

Klaim: $\mu=3$; $\sigma=1$

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}$$

$n=5 \rightarrow v=4$

$\alpha=0.05 \rightarrow \chi^2_{\alpha/2} = \chi^2_{0.025} = 11.143$ dan $\chi^2_{0.975} = 0.484$

s^2 dihitung, lalu tentukan intervalnya.

Solusi

Hitung dulu $$s^2 = \frac{n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$$

$$= \frac{(5)(48.26) - (15)^2}{(5)(4)} = 0.815$$

Untuk selang 95%, pilih α=0.05, dengan tabel nilai χ^2 dengan v=4 maka $\chi^2_{0.025}$ =11.143 dan $\chi^2_{0.975}$ =0.484

Dengan demikian selang kepercayaan 95% adalah:

$$\frac{(4)(0.815)}{11.143} < \sigma^2 < \frac{(4)(0.815)}{0.484}$$ atau $$0.293 < \sigma^2 < 6.736$$

Kesimpulan: klaim perusahaan bisa diterima karena nilai 1 masih terletak pada selang tersebut.

Penaksir Rasio Dua Variansi Dua Sampel

- Definisi:

Taksiran rasio dua variansi populasi σ_1^2/σ_2^2 adalah rasio dari variansi sampel s_1^2/s_2^2.

Jadi, **statistik S_1^2/S_2^2 adalah penaksir dari σ_1^2/σ_2^2**

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

Penaksir Rasio Dua Variansi Dua Sampel

- Definisi (2):

Jika s_1^2 dan s_2^2 adalah variansi dari dua sampel saling bebas berukuran n_1 dan n_2 dari populasi normal, maka interval kepercayaan $(1-\alpha)100\%$ untuk σ_1^2/σ_2^2 adalah:

$$\frac{s_1^2}{s_2^2}\frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\frac{\alpha}{2}}(v_2, v_1)$$

dengan $f_{\alpha/2}(v1,v2)$ adalah nilai f dengan derajat kebebasan $v_1=n_1-1$ dan $v_2=n_2-1$ yang mempunyai luas sebelah kanan $\alpha/2$, dan $f_{\alpha/2}(v1,v2)$ bernilai-f yang sama dengan derajat kebebasan $v_1=n_1-1$ dan $v_2=n_2-1$.

Penjelasan

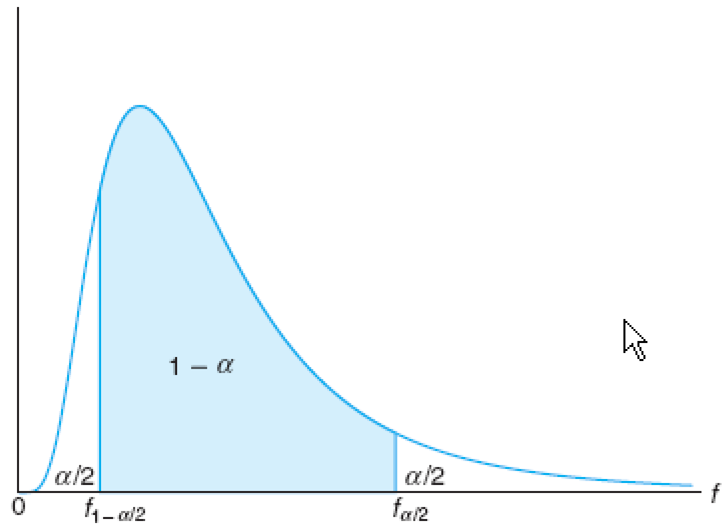


Figure 9.8: $P[f_{1-\alpha/2}(v_1, v_2) < F < f_{\alpha/2}(v_1, v_2)] = 1 - \alpha$.

$$P[f_{1-\alpha/2}(v_1, v_2) < F < f_{\alpha/2}(v_1, v_2)] = 1 - \alpha$$

$$P\left[f_{1-\alpha/2}(v_1, v_2) < \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} < f_{\alpha/2}(v_1, v_2)\right] = 1 - \alpha$$

$$P\left[\frac{S_1^2}{S_2^2}\frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2}\frac{1}{f_{1-\alpha/2}(v_1, v_2)}\right] = 1 - \alpha$$

$$P\left[\frac{S_1^2}{S_2^2}\frac{1}{f_{\alpha/2}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\alpha/2}(v_2, v_1)\right] = 1 - \alpha$$

$$f_{1-\alpha}(v1, v2) = \frac{1}{f_{\alpha}(v2, v1)}$$

Contoh 9.19

- A confidence interval for the difference in the mean orthophosphorus H_3PO_4 contents, measured in milligrams per liter, at two stations on the James River was constructed by **assuming the normal population variance to be unequal.**
  Fifteen samples were collected from station 1, and 12 samples were obtained from station 2. The 15 samples from station 1 had an average orthophosphorus content of 3.84 milligrams per liter and a standard deviation of 3.07 milligrams per liter, while the 12 samples from station 2 had an average content of 1.49 milligrams per liter and a standard deviation of 0.80 milligram per liter.
  Justify this assumption by constructing 98% confidence intervals for σ_1^2/σ_2^2 and for σ_1/σ_2, *where* σ_1^2 and σ_2^2 are the variances of the populations of orthophosphorus contents at station 1 and station 2, respectively.

Solusi

- Asumsi: the normal population variance to be unequal.
- $n_1 = 15$; $\bar{X}_1 = 3.84$; $s_1 = 3.07$
- $n_2 = 12$; $\bar{X}_2 = 1.49$; $s_1 = 0.80$
- $\alpha = 0.02 \rightarrow f_{0.01}(14,11) \approx 4.30$, $f_{0.01}(11,14) \approx 3.87$
- 98% confidence intervals for σ_1^2/σ_2^2 and for σ_1/σ_2

$$\frac{s_1^2}{s_2^2}\frac{1}{f_{\frac{\alpha}{2}}(v_1, v_2)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2} f_{\frac{\alpha}{2}}(v_2, v_1)$$

$$\left(\frac{3.07^2}{0.80^2}\right)\left(\frac{1}{4.30}\right) < \frac{\sigma_1^2}{\sigma_2^2} < \left(\frac{3.07^2}{0.80^2}\right)(3.87)$$

$3.425 < \sigma_1^2/\sigma_2^2 < 56.991$

- 98% confidence intervals for σ_1/σ_2:

$1.851 < \sigma_1/\sigma_2 < 7.549$

Nilai f

Table A.6 (continued) Critical Values of the F-Distribution

v_2	$f_{0.01}(v_1, v_2)$					
	v_1					
	10	12	15	20	24	30
1	6055.85	6106.32	6157.28	6208.73	6234.63	6260.65
2	99.40	99.42	99.43	99.45	99.46	99.47
3	27.23	27.05	26.87	26.69	26.60	26.50
4	14.55	14.37	14.20	14.02	13.93	13.84
5	10.05	9.89	9.72	9.55	9.47	9.38
6	7.87	7.72	7.56	7.40	7.31	7.23
7	6.62	6.47	6.31	6.16	6.07	5.99
8	5.81	5.67	5.52	5.36	5.28	5.20
9	5.26	5.11	4.96	4.81	4.73	4.65
10	4.85	4.71	4.56	4.41	4.33	4.25
11	4.54	4.40	4.25	4.10	4.02	3.94
12	4.30	4.16	4.01	3.86	3.78	3.70
13	4.10	3.96	3.82	3.66	3.59	3.51
14	3.94	3.80	3.66	3.51	3.43	3.35
15	3.80	3.67	3.52	3.37	3.29	3.21

- $f_{0.01}(14,11) \approx 4.30$,
- $f_{0.01}(11,14) \approx 3.87$

PR

- Bab 9 : #41, 53