BAB VII
PENDUGAAN SELANG KEPERCAYAAN

Capaian Pembelajaran

- Mahasiswa mengenai jenis pendugaan (titik dan selang)
- Mahasiswa memahami pendugaan titi dari mean dan proporsi populasi
- Mahasiswa memahami pendugaan keragaman populasi
- Mahasiswa memahami pendugaan selang untuk proporsi

7.1 Selang Kepercayaan

Distribusi yang terkait dengan populasi sering diketahui kecuali untuk satu atau lebih parameter. Masalah estimasi memiliki hubungan yang erat dengan cara terbaik untuk menduga nilai parameter, bagaimana memberikan ukuran kepercayaan pada estimasi tersebut. Dengan kata lain, berapa nilai kepastian dalam perkiraan itu. Hal yang paling menjadi perhatian yaitu: rata-rata, varian, dan proporsi.

Notasi 1: θ menjelaskan parameter. Statistik yang berfungsi untuk memperkirakan θ ditunjukkan dengan kapital $\hat{\Theta}$ dan merupakan variabel acak. Nilai spesifik dari variabel acak tersebut ditunjukkan oleh $\hat{\theta}$. Estimator adalah cara untuk memperoleh nilai estimasi titik berdasarkan pengamatan $X_1,\ X_2,\ \cdots, X_n$

Notasi 2: Rata-rata populasi dilambangkan dengan μ, variansi populasi dengan σ^2, dan proporsi populasi dengan p .

Definisi : Suatu statistik $\hat{\Theta}$ dikatakan sebagai estimator tak bias dari suatu parameter θ jika

$$E\left(\hat{\Theta}\right)=\theta$$

Estimator dengan varians terkecil di antara semua estimator tak bias disebut sebagai estimator θ yang paling efisien.

7.2 Estimator titik pada varian

Kadang-kadang kita memiliki penaksir lebih dari titik untuk parameter yang sama. Untuk membuat pilihan, dibutuhkan batasan tertentu. Salah satunya nilai estimator harus tidak boleh bias. Nilai penaksir lainnya harus memiliki varians yang lebih kecil. Di antara semua penaksir tak bias, penaksir dengan varian terkecil disebut penaksir paling efisien. Sebagai contoh, untuk sampel ukuran 3, baik $\overline{X}_3$ dan $\hat{\Theta} = \frac{X_1 + 2X_2 + 3X_3}{6}$ Θˆ =X1+2X2+3X36

tidak bias. Tapi $\overline{X}_3$ lebih efisien jika,

$$Var\left(\overline{X}_3\right) = \frac{\sigma^2}{3}, Var\left(\hat{\Theta}\right) = \frac{14}{36}\sigma^2$$

7.3 Estimasi Selang/interval

Selain penaksir titik untuk parameter, akan dijelaskan pula penaksiran interval yang memberikan beberapa ukuran ketidakpastian. Untuk membuat perkiraan interval, kita perlu menemukan dua variabel acak, $\hat{\Theta}_L, \hat{\Theta}_U$ sehingga

$$P(\hat{\Theta}_L < \theta < \hat{\Theta}_U) = 1 - \alpha, 0 < \alpha < 1.$$

Dengan tujuan parameter dimuat dalam interval. Interval yang dihitung dari sampel, $\hat{\Theta}_L < \theta < \hat{\Theta}_U$ disebut 100 (1 − α) % selang kepercayaan. Titik akhir interval disebut batas kepercayaan. Secara umum, interval yang baik adalah interval kepercayaan yang pendek dengan tingkat kepercayaan yang tinggi.

7.4 Mengestimasi nilai rata-rata pada populasi tunggal dimana σ diketahui

Distribusi pengambilan sampel $\overline{X}$ berpusat pada μ, dan di sebagian besar aplikasi varians lebih kecil daripada estimator μ lainnya. Dengan demikian, sampel rata-rata $\overline{X}$ akan digunakan sebagai estimasi titik untuk rata-rata populasi μ. Ditulis $\sigma^2 X = \sigma^2 / n$ sehingga sampel yang besar akan menghasilkan nilai $\overline{X}$ yang berasal dari sampel distribusi dengan varian yang kecil. Oleh karena itu, $\overline{X}$ kemungkinan merupakan perkiraan yang sangat akurat dari μ ketika n besar. Mari kita pertimbangkan estimasi interval μ. Jika sampel kami dipilih dari populasi normal atau, jika gagal, jika n cukup besar, kita dapat menetapkan interval kepercayaan untuk μ dengan mempertimbangkan distribusi sampel $\overline{X}$. Menurut Teorema Limit Pusat, kita dapat mengharapkan distribusi sampling dari kira-kira terdistribusi normal dengan rata-rata μ dan standard deviation $\overline{X} = \sigma \sqrt{n}$. Menulis zα/2 untuk nilai z di atas yang kita temukan luas α/2 di bawah kurva normal, kita dapat melihat dari Gambar berikut:

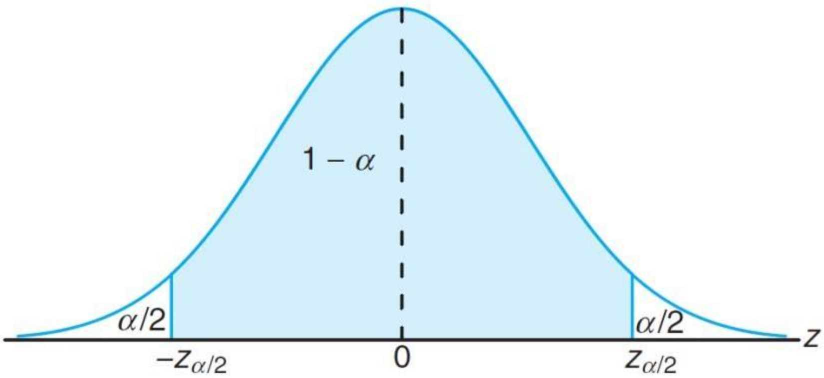
$$P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1 - \alpha$$

dimana

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

sehingga,

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



Gambar 7.1. $P\left(-z_{\alpha/2} < Z < z_{\alpha/2}\right) = 1 - \alpha$

Mengalikan setiap suku dalam pertidaksamaan dengan $\sigma/\sqrt{n}$ dan kemudian mengurangkan $\bar{X}$ dari setiap suku dan mengalikannya dengan −1 (membalikkan rasa pertidaksamaan), kita memperoleh:

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

Contoh:

Jumlah lemak mentega dalam pound yang dihasilkan oleh seekor sapi selama periode produksi susu 305 hari berdistribusi normal dengan rata-rata μ dan standar deviasi 89,75. Hitunglah selang kepercayaan 95% untuk rata-rata μ berdasarkan sampel acak ukuran 20 ketika rata-rata sampel adalah 507,5.

Diketahui, $\bar{X}_{20} = 507.5$, $\sigma = 89.75$, α = 0,05. Karena $z_{0.025} = 1.96$ z0,025 = 1,96. Selang kepercayaan menjadi:

$$507.5 \pm 1.96 \frac{89.75}{\sqrt{20}} = 507.5 \pm 39.34$$

Dapat dilihat bahwa ukuran sampel terlalu besar. Jika tingkat kepercayaan diturunkan menjadi 90%, dan margin of error meningkat menjadi e = 45, berapa ukuran sampel yang diperlukan?

$$n = \left(\frac{1.645(89.75)}{45} \right)^2 = 10.76 \cong 11$$

7.5 Mengestimasi nilai rata-rata pada populasi tunggal dimana σ tidak diketahui

Jika varian dari populasi normal tidak diketahui, kita dapat membuat selang kepercayaan berdasarkan distribusi variabel acak

$$T = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{S}$$

yang merupakan Student t dengan n−1 derajat kebebasan. Dalam hal ini, kita memiliki probabilitas (1 − α), rata-rata μ akan dimasukkan dalam interval $\bar{X}_n \pm t_{\alpha/2} \frac{S}{\sqrt{n}}$ sehingga menjadi persamaan berikut:

$$\hat{\Theta}_L = \bar{X}_n - t_{\alpha/2} \frac{S}{\sqrt{n}}, \hat{\Theta}_U = \bar{X}_n + t_{\alpha/2} \frac{S}{\sqrt{n}}$$

Contoh:

Jumlah lemak mentega dalam pound yang dihasilkan oleh seekor sapi selama periode produksi susu 305 hari berdistribusi normal dengan rata-rata μ. Berdasarkan sampel acak berukuran 20, rata-rata sampel adalah 507,5. dan standar deviasi sampel adalah 89,75. Buatlah kepercayaan 95% untuk rata-rata μ.

Diketahui, $\bar{X}_{20} = 507.5$, $s = 89.75$, α = 0,05. Karena $z_{0.025} = 2.093$ dengan derajat bebas n-1. Selang kepercayaan menjadi:

$$507.5 \pm 2.093 \frac{89.75}{\sqrt{20}} = 507.5 \pm 42.00$$

Dapat dilihat bahwa intervalnya lebih lebar ketika varians tidak diketahui. Hal ini disebabkan pada $\sigma/\sqrt{n}$ sebagai kesalahan standar dan $S/\sqrt{n}$ sebagai perkiraan kesalahan standar.

7.6 Estimasi perbedaan antara dua rata-rata

Misalkan satu lahan peternakan memiliki dua program pemberian pakan yang berbeda untuk penggemukan sapi potong. Peneliti tertarik untuk melihat apakah ada perbedaan yang signifikan di antara mereka. Misalkan kita memiliki dua populasi normal dengan rata-rata μ1, μ2 dan varian σ_1^2, σ_2^2 berturut-turut. Dihitung selang kepercayaan untuk selisih μ1 −μ2. Beberapa kasus muncul dengan sendirinya yang dapat kita tabulasikan sebagai berikut.

Misalkan $\bar{X}_1, \bar{X}_2$ mewakili rata-rata sampel dari dua populasi berdasarkan sampel n_1 n_2 secara berurutan.

Varians	Selang Kepercayaan	Definisi	d.f.
σ_1^2, σ_2^2 diketahui	$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$		
$\sigma_1^2 = \sigma_2^2 = \sigma^2$	$\bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	$n_1 + n_2 - 2$
σ_1^2, σ_2^2 tidak diketahui	$\bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} S_p \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$		

Penggunaan R untuk Selang Kepercayaan

Package R Basic tidak terdapat fasilitas untuk menghitung selang kepercayaan (*Confidence Interval*/ CI), diperlukan *package* `epicalc` yang bisa diunduh dan kemudian dimuat dalam sistem *R*.

```
> library(epicalc)
```

7.7.1 Jika σ diketahui

Diketahui data tinggi badan dari 34 mahasiswa (diberikan pada tabel) mempunyai σ^2 sebesar 64 atau σ sebesar 8.

171 173 160 173 162 173 173 173 162 173
161 171 175 167 175 167 155 160 165 169
151 153 150 163 161 159 159 150 151 160
153 153 152 155

Maka cara mendapatkan selang kepercayaan (CI) adalah:

Masukkan data ke dalam *R*

Hitung selang kepercayaan 95% dari data TB berikut:

```
> library(confintr)
> TB = c(171, 173, 160, 173, 162, 173, 173, 173, 162, 173,
161,171, 175, 167, 175, 167, 155, 160, 165, 169, 151, 153, 150,
163,161, 159, 159, 150, 151, 160, 153, 153, 152, 155)
> ci_mean(TB, probs = c(0.025,0.975))

	Two-sided 95% t confidence interval for the population mean

Sample estimate: 162.5882
Confidence interval:
     2.5%     97.5%
159.6587 165.5177
```

Didapatkan 95% selang kepercayaan adalah (159.7969 ,165.3796). Untuk merubah besaran selang kepercayaan dengan α = 10%, sintaks diatas harus digantikan dengan `qnorm(0.95)`.