Bab 5
Distribusi Sampling

Pendahuluan

Untuk mempelajari populasi kita memerlukan sampel yang diambil dari populasi yang bersangkutan. Meskipun kita dapat mengambil lebih dari sebuah sampel berukuran n dari populasi berukuran N, pada prakteknya hanya sebuah sampel yang biasa diambil dan digunakan untuk hal itu.

Sampel yang diambil ialah sampel acak dan dari sampel tersebut nilai-nilai statistiknya dihitung untuk digunakan seperlunya. Untuk itu diperlukan sebuah teori yang dikenal dengan nama distribusi sampling.

Distribusi sampling biasanya diberi nama tergantung pada nama statistik yang digunakan.

Misalnya distribusi sampling rata-rata, distribusi sampling proporsi, distribusi simpangan baku dan sebagainya.

Distribusi Sampling Rata-rata

Jika sebuah populasi mempunyai rata-rata μ dan simpangan baku σ yang besarnya terhingga, maka untuk ukuran sampel acak n cukup besar, distribusi rata-rata sampel mendekati distribusi normal dengan rata-rata $\mu_{\bar{x}} = \mu$ dan simpangan baku

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Jika ukuran populasi kecil dibandingkan dengan ukuran sampel, yaitu (n/N) > 5%

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Jika ukuran populasi besar dibandingkan dengan ukuran sampel, yaitu (n/N) $\leq$ 5%

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Distribusi normal yang didapat dari distribusi rata-rata perlu distandarkan agar daftar distribusi normal baku dapat digunakan. Ini perlu untuk perhitungan-perhitungan. Untuk ini digunakan transformasi :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

Contoh Soal :

Tinggi badan mahasiswa rata-rata mencapai 165 cm dan simpangan baku 8,4 cm. Untuk itu diambil sampel acak terdiri atas 45 mahasiswa. Tentukanlah berapa peluang tinggi rata-rata ke-45 mahasiswa tersebut :

a. Antara 160 cm dan 168 cm
b. Paling sedikit 166 cm

Jawab:

Jika ukuran populasi tidak dikatakan besarnya, selalu dianggap cukup besar untuk berlakunya teori. Ukuran sampel n = 45 tergolong sampel besar sehingga dalil limit pusat berlaku. Jadi rata-rata $\overline{\mathrm{x}}$ untuk tinggi mahasiswa akan mendekati distribusi normal dengan :

$$\text{Rata} - \text{rata} \qquad \mu_{\bar{x}} = 165\ \text{cm}$$

$$\text{Simpangan Baku} \qquad \sigma_{\bar{x}} = \frac{8{,}4}{\sqrt{45}} = 1{,}252\ \text{cm}$$

a. Dengan $\bar{x}$ = 160 cm dan $\bar{x}$ = 168 cm didapat :

$$z_1 = \frac{160 - 165}{1,252} = -3,99 \text{ dan } z_2 = \frac{168 - 165}{1,252} = 2,40$$

Dari daftar distribusi normal baku memberikan luas kurva = 0,5 + 0,4918 = 0,9918

Peluang rata-rata tinggi ke-45 mahasiswa antara 160 cm dan 168 cm adalah 0,9918

b. Rata-rata tinggi paling sedikit 166 cm memberikan angka z paling sedikit

$$z = \frac{166 - 165}{1,252} = 0,80$$

Dari daftar normal baku, luas kurva 0,5 – 0,2881 = 0,2119.

Peluang yang dicari = 0,2119

Apabila dari populasi diketahui variannya dan perbedaan antara rata-rata dari sampel ke sampel diharapkan tidak lebih dari sebuah harga d yang ditentukan, maka berlaku hubungan :

$$\sigma_{\bar{x}} \leq \mathrm{d}$$

Dari rumus ini ukuran sampel yang paling kecil sehubungan dengan distribusi rata-rata dapat ditentukan

Contoh:

Untuk contoh diatas, misalkan harga-harga $\bar{x}$ dari sampel yang satu dengan sampel lainnya diharapkan tidak ingin lebih dari 1 cm. Jika populasi cukup besar, maka :

$$\frac{\sigma}{\sqrt{n}} \leq d \text{ menghasilkan } \frac{8,4}{\sqrt{n}} \leq 1 = \frac{70,56}{n} \leq 1$$

$$\text{atau } n >= 70,56$$

Paling sedikit perlu diambil sampel terdiri atas 71 mahasiswa

Distribusi Sampling Proporsi

Misalkan populasi diketahui berukuran N yang didalamnya terdapat peristiwa A sebanyak Y diantara N. Maka didapat parameter proporsi peristiwa A sebesar µ = (Y/N)

Jika dari suatu populasi diambil sampel acak berukuran n dan dimisalkan didalamnya ada peristiwa A sebanyak x. Sampel ini memberikan statistik proporsi peristiwa A = p = x/n.

Jika semua sampel yang mungkin diambil dari populasi itu maka didapat sekumpulan harga-harga statistik proporsi. Dari kumpulan ini kita dapat menghitung rata-ratanya,diberi symbol

μ_p dan simpangan bakunya diberi simbol σ_p

Jika ukuran populasi kecil dibandingkan dengan ukuran sampel, yaitu (n/N) > 5% , maka :

$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}\sqrt{\frac{N-n}{N-1}}$$

dan jika ukuran populasi besar dibandingkan dengan ukuran sampel, yaitu (n/N) $\leq$ 5%, maka :

$$\mu_p = \pi$$

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

σ_p dinamakan kekeliruan baku proporsi atau galat baku proporsi.

Contoh

Ada petunjuk kuat bahwa 10 % anggota masyarakat tergolong dalam golongan A. Sebuah sampel acak terdiri atas 100 orang telah diambil.

a. Tentukan peluangnya bahwa dari 100 orang itu akan ada paling sedikit 15 orang dari golongan A.

b. Berapa orang harus diselidiki agar persentase golongan A dari sampel yang satu dengan yang lainnya diharapkan berbeda paling besar dengan 2 % ?

Jawab :

Populasi yang dihadapi berukuran cukup besar dengan π = 0,10 dan $1 - \pi = 0,90$

a. Untuk ukuran sampel 100, diantaranya paling sedikit 15 tergolong kategori A, maka paling sedikit p = 0,15.

Kekeliruan bakunya adalah :

$$\sigma_p = \sqrt{\frac{\pi\,(1-\pi)}{n}} = \sqrt{\frac{0{,}10 \text{ x } 0{,}90}{100}} = 0{,}03$$

Bilangan z paling sedikit $= \dfrac{0{,}15 - 0{,}10}{0{,}03} = 1{,}67$

Dari daftar normal baku luasnya = 0,5 – 0,4525 = 0,0475

Peluang dalam sampel itu akan ada paling sedikit 15 kategori A adalah 0,0475.

b. π = 0,10 dan $1 - \pi$ = 0,90 sedangkan d = 0,02

maka :

$$\sqrt{\frac{0{,}1 \text{ x } 0{,}9}{n}} \leq 0{,}02 \text{ yang menghasilkan } n \geq 225$$

Paling sedikit sampel harus berukuran 225

Distribusi Sampling Selisih Rata-rata

Misalkan kita punya dua populasi masing-masing berukuran N_1 dan N_2 . Populasi kesatu mempunyai rata-rata μ_1 dan simpangan baku σ_1 , sedangkan populasi kedua mempunyai rata-rata μ_2 dan simpangan baku σ_2. Dari setiap populasi secara independen kita ambil sampel-sampel acak berukuran n_1 dari populasi kesatu dan berukuran n_2 dari populasi kedua. Untuk membedakan, populasi kesatu dimisalkan mempunyai variabel X dan populasi kedua mempunyai variabel Y .

Dari sampel-sampel ini, kemudian dihitung rata-ratanya. Dan didapat rata-rata sampel

$\bar{x}_1$, $\bar{x}_2$, ..., $\bar{x}_k$ dan $\bar{y}_1$, $\bar{y}_2$, ..., $\bar{y}_r$ dengan k = banyak sampel yang dapat diambil dari populasi kesatu dan r = banyak sampel yang dapat diambil dari populasi kedua.

Bentuklah sekarang semua selisih antara rata-rata dari sampel-sampel dalam kumpulan kesatu dan rata-rata dari sampel-sampel dalam kumpulan kedua . Didapat kumpulan selisih rata-rata yang bentuk umumnya :

$\bar{x}_i - \bar{y}_j$ dengan i = 1, 2, 3, ….,k dan j = 1,2,3, …, r

Kumpulan selisih rata-rata sampel akan membentuk distribusi selisih rata-rata. Dari kumpulan ini kita dapat menghitung rata-ratanya, diberi simbol

$\mu_{\bar{x}-\bar{y}}$, dan menghitung simpangan bakunya, diberi simbol $\sigma_{\bar{x}-\bar{y}}$

Ternyata bahwa, untuk N_1 dan N_2 cukup besar dan sampel-sampel acak diambil secara independen satu sama lain, didapat hubungan :

$$\mu_{\bar{x}-\bar{y}} = \mu_1 - \mu_2$$

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}$$

Untuk ukuran-ukuran sampel cukup besar, maka selisih rata-rata $\bar{x} - \bar{y}$ akan mendekati distribusi normal. Untuk membuat distribusi normal ini menjadi distribusi normal baku digunakan transformasi

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}-\bar{y}}}$$

Contoh :

Rata-rata tinggi mahasiswa laki-laki 163 cm dan simpangan bakunya 5,2 cm ; sedangkan untuk mahasiswa perempuan, parameter tersebut berturut-turut 152 cm dan 4,9 cm.

Dari kedua kelompok mahasiswa itu masing-masing diambil sampel acak, secara independen, berukuran sama, ialah 140 orang.

Berapa peluang rata-rata tinggi mahasiswa laki-laki paling sedikit 10 cm lebihnya dari rata-rata tinggi mahasiswa perempuan?

Jawab:

Misalkan $\bar{x}$ dan $\bar{y}$ masing-masing menyatakan rata-rata tinggi dari sampel untuk mahasiswa laki-laki dan perempuan. Yang ditanyakan adalah peluang $\bar{x}$ - $\bar{y}$ paling sedikit 10 cm. Dari yang diketahui didapat μ_1 = 163 cm, σ_1 = 5,2 cm, μ_2 = 152 cm, σ_2 = 4,9 cm, dan n_1 = n_2 = 140 , maka :

$$\text{Rata-rata} \qquad \mu_{\bar{x}-\bar{y}} = (163-152)\ \text{cm} = 11\ \text{cm}$$

$$\text{Simpangan Baku} \quad \sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{(5,2)^2}{140} + \frac{(4,9)^2}{140}}\ \text{cm} = 0,6038\ \text{cm}$$

$$z = \frac{10-11}{0,6038} = -1,66$$

Luas daerah normal baku yang diperlukan adalah 0,5 + 0,4515 = 0,9515.

Jadi peluang yang dicari = 0,9515

Distribusi Sampling Selisih Proporsi

Misalkan ada dua populasi masing-masing berdistribusi Binom, kedua-duanya berukuran cukup besar. Didalam kedua populasi itu ada peristiwa A dengan proporsi π_1 untuk populasi kesatu dan π_2 untuk populasi kedua. Dari kedua populasi itu, secara independen diambil sampel-sampel acak berukuran n_1 dari populasi kesatu dan berukuran n_2 dari populasi kedua. Untuk peristiwa A, didapat kumpulan proporsi

$$p_1 = \frac{x_i}{n_1}, i = 1, 2, ..., k \text{ dan } p_2 = \frac{y_i}{n_2}, j = 1, 2, ..., r$$

Dengan x_i = adanya peristiwa A dalam sampel yang diambil dari populasi kesatu, y_j = adanya peristiwa A dalam sampel yang diambil dari populasi kedua, k dan r masing-masing banyak sampel yang mungkin diambil dari populasi kesatu dan kedua.

Selisih proporsi $(p_1 - p_2)$ dapat dibentuk sehingga terdapat kumpulan selisih proporsi. Dari kumpulan ini dapat dihitung rata-ratanya dan diberi simbol μ_{sp} dan simpangan bakunya diberi simbol σ_{sp}, dengan sp = $(p_1 - p_2)$ = selisih antara proporsi sampel kesatu dan proporsi sampel kedua. Ternyata untuk ini berlaku :

$$\mu_{sp} = \pi_1 - \pi_2$$

$$\sigma_{sp} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

Untuk ukuran sampel cukup besar distribusi selisih proporsi akan mendekati distribusi normal. Agar supaya distribusi normal ini menjadi distribusi normal baku maka diperlukan transformasi :

$$z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sigma_{sp}}$$

Contoh :

Ada petunjuk kuat bahwa calon A akan mendapat suara 60 % dalam pemilihan. Dua buah sampel acak secara independen telah diambil masing-masing terdiri atas 300 orang.

Tentukan peluangnya akan terjadi perbedaan persentase tidak lebih dari 10 % akan memilih A.

Jawab :

Kedua sampel diambil dari sebuah populasi, jadi kita anggap dua populasi yang sama, sehingga $\pi_1 = \pi_2 = 0,6$. Jika x = banyak orang yang memilh A dalam sampel kesatu, dan y = banyak orang yang memilih B dalam sampel kedua, maka yang dicari adalah peluang $p_1 - p_2$ < 10 % atau $p_2 - p_1$ < 10 %

Setelah digabungkan menjadi -10 % < $p_1 - p_2$ < 10 %

$$\mu_{sp} = 0,6 - 0,6 = 0$$

$$\sigma_{sp} = \sqrt{\frac{0,6 \text{ x } 0,4}{300} + \frac{0,6 \text{ x } 0,4}{300}} = 0,04$$

Bilangan z yang perlu adalah :

$$z_1 = \frac{-0,1 - 0}{0,04} = -2,50 \quad \text{dan} \quad z_2 = \frac{0,1 - 0}{0,04} = 2,50$$

Luas daerah normal baku yang diperlukan = 2 (0,4938) = 0,9876

Latihan Soal

1. Suatu populasi terdiri dari 5 angka yaitu : 2, 3, 6, 8, dan 11. Dari populasi ini kemudian ditarik semua sampel yang beranggotakan 2 dengan pengembalian yang mungkin dapat diambil dari populasi tersebut. Hitunglah :
   a. Mean populasi
   b. Deviasi standar populasi
   c. Mean dari distribusi sampling harga mean
   d. Deviasi standar dari distribusi sampling harga deaviasi standar.
2. Diketahui 500 bola yang mempunyai berat rata-rata 5,02 ons dan deviasi standar 0,3 ons. Hitunglah probabilitasnya bahwa dari sampel random 100 bola mempunyai berat rata-rata keseluruhan antara :
   a. Antara 4,96 dan 5,00 ons
   b. Lebih dari 5,10 ons

3. Tinggi badan mahasiswa suatu Perguruan Tinggi berdistribusi normal dengan mean 68 cm dan deviasi standar 3 cm. Apabila terdapat 25 mahasiswa suatu PT, tentukan peluang rata-rata tinggi antara :

   a. Terletak antara 66,8 dan 68,3 cm
   b. Kurang dari 66,4

4. Bola lampu hasil produksi Pabrik A mempunyai umur rata-rata 1400 jam dengan deviasi standar 200 jam, sedangkan bola lampu hasil produksi Pabrik B mempunyai umur rata-rata 1200 jam dengan deviasi standar 100 jam. Jika sampel random sebanyak 125 bola lampu diambil dari masing-masing merk diuji, berapa probabilitasnya bahwa merk A mempunyai umur rata-rata paling sedikit :

   a. 160 jam lebih lama daripada merk B.
   b. 250 jam lebih lama daripada merk B.

5. Misalkan bahwa rata-rata mahasiswa Indonesia 162 cm dengan simpangan baku 6,5 cm. Jika ada 49 mahasiswa, tentukan peluang rata-rata tinggi mahasiswa :

a. Paling rendah 155 cm

b. Paling tinggi 175 cm

c. Antara 158 cm dan 172 cm

d. Kurang dari 160 cm

6. Dari pengalaman memperlihatkan bahwa 10% anggota masyarakat menderita penyakit A. Penelitian dilakukan terhadap 900 orang, tentukan :

a. Rata-rata dan simpangan baku untuk proporsi yang menderita penyakit A

b. Peluang sampel itu akan berisikan anggota yang menderita penyakit A

   1. Antara 80 dan 95 orang
   2. Lebih dari 98 orang
   3. Paling banyak 75 orang

7. Merk lampu A rata-rata menyala 1400 jam dan merk lampu B rata-rata menyala 1300 jam. Simpangan bakunya masing-masing 160 jam dan 125 jam. Dari tiap populasi diambil sebuah sampel acak berukuran 85 dari lampu A dan 100 dari lampu B. Tentukan peluang rata-rata menyala lampu dalam sampel dari A paling sedikit 50 jam lebihnya dari rata-rata menyala lampu dalam sampel dari B.
8. Pengalaman mencatat bahwa 65% dari penduduk ternyata menyenangi pemimpin A. Dua buah sampel acak telah diambil masing-masing berukuran 250. Tentukan bagaimana peluangnya bahwa kedua sampel itu akan memperlihatkan perbedaan presentase lebih dari 12% yang menyenangi pemimpin A.
9. Diketahui bahwa produk yang dihasilkan mesin tertentu 2%-nya rusak. Berapa probabilitasnya bahwa dari pengiriman sebanyak 400 produk itu.
   a. 3% atau lebih ternyata rusak
   b. 2% atau kurang ternyata rusak