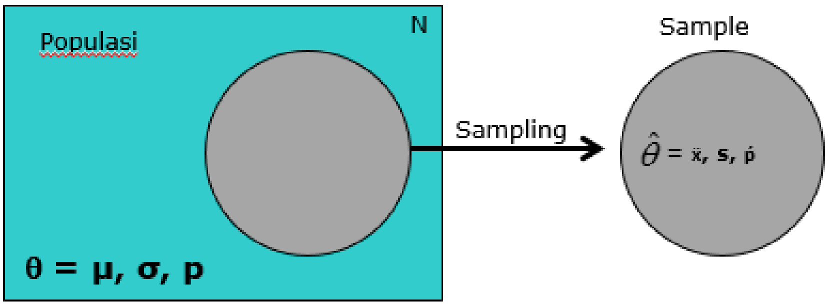
BAB 6. DISTRIBUSI PENCUPLIKAN / DISTRIBUSI SAMPEL

A. Pengertian dan Konsep Dasar

Istilah-istiah yang sering digunakan dalam distribusi meliputi informasi berikut. Populasi merupakan kumpulan/himpunan dari keseluruhan elemen yang menjadi obyek penelitian/penyelidikan, sedangkan sampel adalah sebagian obyek dari populasi. Elemen merupakan sesuatu yang menjadi obyek penyelidikan. Sementara itu pengertian dari sensus adalah cara mengumpulkan data dari populasi, dan sampling adalah cara mengumpulkan data dari sampel. Kelemahan dari metode sensus dalam pengumpulan data yaitu dibutuhkan biaya mahal, waktu lama, tenaga yang besar. Kelemahan sensus diatasi dengan teknik sampel (sampling). Sedangkan keuntungan teknik sampel adalah biaya yang rendah serta waktu yang pendek tanpa mengurangi keakuratan. Karateristik populasi merupakan sifat-sifat atau ciri-ciri yang diamati dalam suatu populasi. Selanjutnya pengertian dari parameter adalah karakteristik yang dihitung dari populasi, sedangkan statistik adalah karakteristik yang dihitung dari sampel.



Gambar 6.1 Grafik hubungan antara populasi dan sampel

Terdapat gap antara populasi dan sampel yang disebut sebagai kesalahan (penyimpangan). Sebab kesalahan sampel: kesalahan pemilihan sampel, kesalahan hitung, dan lain lain. Sampel yang representatif memiliki ciri: ukuran tertentu yang memakai syarat, kesalahan terkecil, dan dipilih dengan prosedur yang benar berdasarkan teknik sampel tertentu.

Tabel 6.1 Perbedaan karakteristik populasi dan sampel

Karakteristik populasi	**Karakteristik Sampel**
Ukuran N	Ukuran n
Parameter	Statistik
Mean μ	Mean X
Simpangan baku σ	Simpangan baku S
Populasi berhingga dan tak berhingga	Sampel besar dan kecil

Sampel merupakan potret dari populasi, dimana dengan adanya sampel diharapkan bisa memperoleh gambaran yang sesungguhnya mengenai karakteristik populasi. Hasil perhitungan dari data sampel akan dipakai untuk menyimpulkan parameter populasi. Sebagai contoh jika jumlah mahasiswa di UMY N= 20.000 (populasi mahasiswa) dan hanya n=100 yang pilih dari 100 sampel mahasiswa yg terpilih diharapkan mampu mewakili 20.000 populasi dalam penelitian.

Parameter populasi ditulis dengan huruf latin θ, di mana θ bisa berupa rata-rata populasi μ, simpangan baku populasi σ, proporsi populasi p. Sedangkan statistik dari sampel ditulis $\hat{\theta}$ (θ topi), bisa berupa rata-rata sampel **X**, simpangan baku sampel S, proporsi sampel $\hat{p}$.

B. Teknik Pengambilan Sampel

Teknik pengambilan sampel menurut proses memilihnya terbagi menjadi dua bagian yaitu pengambilan sampel dengan pengembalian, dimana ukuran populasi adalah tetap. Sesuai untuk ukuran populasi terbatas serta selanjutnya adalah pengambilan sampel tanpa pengembalian, dimana ukuran populasi akan berkurang, sesuai untuk populasi tak terbatas. Sementara itu, teknik pengambilan sampel menurut probabilitasnya terbagi menjadi empat yaitu teknik pengambilan sampel acak sederhana, sistematik, stratifikasi dan kluster.

Teknik pengambilan sampel acak sederhana diartikan sebagai pengambilan sampel sebanyak n sedemikian rupa sehingga setiap unit dalam populasi mempunyai kesempatan yang sama untuk dipilih menjadi sampel dan setiap ukuran sampel n juga mempunyai kesempatan yang sama untuk terambil. Metode yang bisa digunakan pada pengambilan ini adalah dengan cara undian (digoncang seperti arisan) atau tabel bilangan acak. Sedangkan teknik pengambilan sampel acak sistematik dilakukan dengan mengambil setiap unsur ke-k dalam populasi untuk dijadikan sampel, dengan titik awal ditentukan secara acak diantara k unsur yang pertama. Tenik ini sering digunakan karena dapat menarik kesimpulan yang tepat mengenai parameter populasi sebab sampelnya menyebar secara merata di seluruh populasi. Sementara itu, teknik pengambilan sampel acak statifikasi dilakukan dengan membagi populasi menjadi beberapa kelompok / sub populasi (strata), sehingga setiap kelompok akan memiliki nilai variabel yang tidak terlalu bervariasi (relatif homogen). Selanjutnya dari setiap strata dipilih sampel melalui proses pengambilan secara sederhana. Sedangkan pada teknik pengambilan sampel acak stratifikasi dilakukan dengan membagi populasi menjadi beberapa strata (tingkatan) kemudian sampel diambil secara acak dari setiap tingkatan. Teknik ini dilakukan bila populasinya heterogen. Cara pengambilan sampel untuk setiap tingkatan tidak sama, harus sebanding dengan jumlah anggota setiap tingkatan (proporsional). Rumusan umum teknik pengambilan sampel acak stratifikasi adalah $n_i = \frac{N_i}{N} n.$ Terakhir, teknik pengambilan sampel acak kluster dilakukan dengan mengambil beberapa kluster (kelompok) secara acak kemudian semua atau sebagian dari anggota masing-masing kelompok diambil secara acak sebagai sampel.

C. Distribusi Sampel

Statistik yang dihitung dari sampel digunakan untuk menyimpulkan parameter populasi. Dalam hal ini berarti dilakukan generalisasi data. Generasilasi akan lebih baik dan meyakinkan jika sampel diambil secara berulang-ulang dan acak sehingga diperoleh banyak sampel acak yang berlainan dari populasi yang sama. Data dari sampel dihitung karakteristiknya dan karena sampel bersifat acak, maka statistik dari sampel juga bersifat acak sehingga secara keseluruhan statistik yang diperoleh akan membentuk suatu distribusi yaitu distribusi statistik. Distribusi probabilitas dari suatu statistik disebut distribusi sampel.

Pada distribusi sampel, statistik sampel yang diperoleh bersifat acak (variabel acak) yang mengikuti suatu distribusi tertentu. Bila statistik yang dihitung dari sampel adalah rata-rata maka akan diperoleh distribusi sampel rata-rata. Bila statistik yang dihitung dari sampel adalah proporsi maka akan diperoleh distribusi sampel proporsi.

D. Distribusi Sampel Rata-rata

Pada distribusi sampel rata-rata, bila populasi **terbatas** berukuran N dengan rata-rata μx dan simpangan baku σx diambil sampel berukuran n secara berulang **tanpa pengembalian**, maka diperoleh rumusan sebagai berikut.

Rata-rata untuk distribusi sampel rata rata $\mu_{\bar{x}} = \mu_x$ ……………………………….…. (6.1)

Simpangan baku pada sampel rata rata $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ ………………………………. (6.2)

dimana $\sqrt{\frac{N-n}{N-1}}$ disebut sebagai faktor koreksi.

Pada distribusi sampl rata-rata, bila populasi **tak terbatas** dengan rata-rata μx dan simpangan baku σx diambil sampel berukuran n **tanpa pengembalian**, maka diperoleh rumusan sebagai berikut.

Rata-rata untuk distribusi sampel rata rata $\mu_{\bar{x}} = \mu_x$ ……………………………….…. (6.3)

Simpangan baku pada sampel rata rata $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$ ………………..…………………….... (6.4)

Karena N besar sekali (populasi tak terbatas) maka perbandingan $\frac{N-n}{N-1} = 1$ dan $\sqrt{\frac{N-n}{N-1}} = 1$

Pada distribusi sampl rata-rata, bila sampel diambil cukup besar, n>=30, maka **distribusi sampelnya akan mendekati distribusi normal** sehingga variabel random Z dapat dihitung dengan rumus sebagai berikut.

$$Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu_x}{\sigma_{\bar{x}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \ (6.5)$$

Contoh soal 1. Suatu sampel acak sebesar n = 10 diambil tanpa pengembalian dari populasi **tak terbatas** yang mempunyai rata-rata $\mu x = 5{,}5$ dan simpangan baku $\sigma x = 2{,}92$. Tentukan rata-rata dan simpangan baku dari sampel rata-rata tersebut.

Jawab 1.

Cara menjawab pertanyaan ini gunakan rumus 6.3 dan 6.4 sebagaimana berikut.

Rata-rata untuk distribusi sampel rata rata $\mu_{\bar{x}} = \mu_x$

Simpangan baku pada sampel rata rata $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{2{,}92}{\sqrt{10}} = 0{,}92$

Contoh soal 2. Kecepatan maksimum 2000 mobil mempunyai rata-rata 135,5 km/jam dengan simpangan baku 5,2 km/jam. Jika sampel sebesar 150 mobil dipilih secara acak tanpa pengembalian, hitung probabilitas kecepatan maksimum rata-rata dari 150 mobil tersebut yang lebih besar dari 136,1 km/jam!

Jawab 2.

- Populasi terbatas N= 2000
- Sampel besar (n>30), n= 150
- Kecepatan rata-rata $\mu x = 135{,}5$ Km/Jam
- Simpangan baku $\sigma x = 5{,}2$ Km/jam
- Distribusi sampel kecepatan maksimum rata-rata merupakan distribusi normal sehingga statistik Z mempunyai distribusi normal standar $Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu_x}{\sigma_{\bar{x}}}$

  Lihat tabel distribusi normal 1,46 ➔ 0,9279

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{5{,}2}{\sqrt{150}}.\sqrt{\frac{2000-150}{2000-1}} = 0{,}41$$

$$Z = \frac{\bar{X} - \mu_X}{\sigma_{\bar{X}}} = \frac{136{,}1 - 135{,}5}{0{,}41} = 1{,}46$$

Jadi probabilitas kecepatan maksimum rata-rata mobil yang lebih besar dari 136,1 km/jam adalah P(X>136,1) = P(Z>1,46) = 1- 0,9279 = 0,0721.

E. Distribusi Sampel Proporsi

Pada distribusi sampel proporsi, bila populasi berukuran N mengandung jenis p sebanyak X, maka proporsi p adalah X/N. Jika dari populasi tersebut diambil sampel berukuran n yang juga mengandung proporsi x/n dan sampel diambil berulang maka statistik yang bersifat acak. Sehingga mempunyai distribusi sampel proporsi dengan rata-rata dan simpangan baku sebagai berikut.

Rumus umum proporsi $\hat{p} = \frac{x}{n}$

Rata-rata pada distribusi sampel proporsi $\mu_{\hat{p}} = \mu_p = \frac{X}{N}$ …………………..……………... (6.6)

Simpangan baku sampel proporsi bila populasi terbatas $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}$ ……... (6.7)

Simpangan baku sampel proporsi bila populasi tak terbatas $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ …….......... (6.8)

Variabel random Z pada distribusi sampel proporsi $Z = \frac{\hat{p}-p}{\sigma_{\hat{p}}}$ ……………….................. (6.9)

Contoh soal 3. Diketahui sebanyak 10% dari ibu-ibu rumah tangga di Bandung memakai detergen A untuk mencuci pakaiannya. Jika dari populasi tersebut diambil sampel berukuran 100. Tentukan rata-rata dan simpangan baku dari populasi ibu-ibu rumah tangga yang memakai detergen A! Bila dari sampel tersebut ternyata terdapat paling sedikit 15 ibu rumah tangga yang memakai detergen A, tentukan probabilitasnya!

Jawab 3.

- p = 10% = 0,1
- Populasi ibu-ibu rumah tangga di Bandung ➔ N populasi tak terbatas
- Sampel besar (n>30) ➔ n= 100
- Rata-rata µp(topi) dan Simpangan baku σp(topi) [$\hat{p}$ => $p(topi)$] ibu-ibu rumah tangga yang memakai detergen A ➔ Rata-rata µp(topi) dan Simpangan baku σp(topi) distribusi sampel proporsi
- Bila x Paling sedikit 15 (x≥15) dalam sampel n = 100 maka p(topi) = x/n = 15/100 = 0,15

Rata-rata = 0,1

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0,1.0,9}{100}} = 0,03$$

Proporsi yang memakai detergen A adalah 15/100 = 0,15

$$Z = \frac{\hat{p}-p}{\sigma_{\hat{p}}} = \frac{0,15-0,1}{0,03} = 1,67$$

- Dari tabel distribusi normal untuk nilai Z=1,67 ➔ 0.9525
- Sehingga P(Z>1,67) = 1- 0.9525 = 0,0475

F. Distribusi Sampel Beda Dua Rata-rata

Pada distribusi sampel beda dua rata-rata, jika terdapat 2 populasi dimana populasi 1 sebanyak N1, mempunyai rata-rata μ1 dan simpangan baku σ1 kemudian Populasi 2 sebanyak N2 mempunyai rata-rata μ2 serta simpangan baku σ2. Dari populasi 1 diambil sampel acak sebanyak n1 dengan rata-rata X1 dan dari populasi 2 sampel acak sebanyak n2 dengan rata-rata X2 dimana kedua sampel tersebut dianggap saling bebas. Dari sampel X1 dan X2 dapat dibuat sampel baru yang juga bersifat acak, yaitu sampel beda dua rata-rata. Rumus umum untuk nilai rata-rata dan simpangan baku dari distribusi sampel beda dua rata-rata adalah sebagai berikut.

Rata-rata pada distribusi sampel beda dua rata-rata $\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2$ ……………... (6.10)

Simpangan baku pada distribui sampel beda dua rata-rata bila populasi terbatas adalah

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} . \sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{(N_1+N_2)-1}} \quad \text{……...…………………………...…. (6.11)}$$

Simpangan baku bila populasi tak terbatas $\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$ ………….......... (6.12)

Variabel random Z pada distribui sampel beda dua rata-rata yaitu

$$Z = \frac{(\bar{X}1-\bar{X}2)-(\mu_1-\mu_2)}{\sigma_{\bar{x}_1-\bar{x}_2}} \quad \text{…………….. (6.13)}$$

Contoh soal 4. Di suatu universitas diketahui rata-rata tinggi badan mahasiswa laki-laki adalah 164 cm dengan simpangan baku 5,3 cm. Sedangkan mahasiswa perempuan tinggi badannya rata-rata 153 cm dengan simpangan baku 5,1 cm. Dari dua populasi tersebut diambil sampel acak yang saling bebas masing-masing 150 orang. Berapa probabilitas rata-rata tinggi mahasiswa laki-laki paling sedikit 12 cm lebihnya daripada rata-rata tinggi mahasiswa perempuan?

Jawab 4.

Diketahui populasi 1: $\mu_1 = 164$ cm, $\sigma_1 = 5{,}3$ cm dan sampel 1 : $n_1 = 150$ orang

populasi 2: $\mu_2 = 153$ cm, $\sigma_2 = 5{,}1$ cm dan sampel 2 : $n_2 = 150$ orang

Misal $\overline{X}_1$ = rata-rata tinggi badan mahasiswa laki-laki

$\overline{X}_2$ = rata-rata tinggi badan mahasiswa perempuan

Rata-rata: $\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2 = 164 - 153 = 11$ cm

Simpangan baku: $\sigma_{\overline{X}_1-\overline{X}_2} = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} = \sqrt{\frac{5{,}3^2}{150} + \frac{5{,}1^2}{150}} = 0{,}6$

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\overline{X}_1-\overline{X}_2}} = \frac{(\overline{X}_1 - \overline{X}_2) - 11}{0{,}6}$$

Karena rata- rata tinggi badan mahasiswa laki-laki **paling sedikit 12 cm lebihnya** dari pada rata-rata tinggi mahasiswa perempuan, maka $(\bar{X}1-\bar{X}2)\geq 12$ sehingga $Z=\frac{12-11}{0,6}=1{,}67$ Dari tabel untuk nilai Z=1,67 ➔ 0.9525 sehingga P(Z≥1,67) = 1- 0.9525 = 0,0475.

G. Distribusi Sampel Beda Dua Proporsi

Pada distribusi sampel beda dua proporsi, terdapat dua populasi dimana populasi 1 berukuran N1 terdapat jenis X1 dengan proporsi X1/N1 dan populasi 2 berukuran N2 terdapat jenis X2 dengan proporsi X2/N2. Bila populasi 1 diambil sampel acak berukuran n1 maka sampel ini akan mengandung jenis x1 dengan proporsi x1/n1. Demikian juga dengan populasi 2 diambil sampel acak berukuran n2 maka sampel ini akan mengandung jenis x2 dengan proporsi x2/n2. Sampel 1 dan 2 dapat membentuk sampel acak baru yaitu sampel beda dua proporsi. Rumus umum distribusi sampel beda dua proporsi yaitu sebagai berikut.

Rata-rata pada distribusi sampel beda dua proporsi $\mu_{\hat{p}_1-\hat{p}_2}=p_1-p_2$ ………….... (6.14)

Simpangan baku pada distribui sampel beda dua proporsi bila populasi terbatas adalah

$$\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{p_1\,(1-p_1)}{n_1}+\frac{p_2\,(1-p_2)}{n_2}}.\sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{(N_1+N_2)-1}} \quad \text{……..…………..……… (6.15)}$$

Simpangan baku bila populasi tak terbatas $\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{p_1\,(1-p_1)}{n_1}+\frac{p_2\,(1-p_2)}{n_2}}$ ……. (6.16)

Variabel random Z pada distribui sampel beda dua rata-rata yaitu

$$Z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sigma_{\hat{p}_1-\hat{p}_2}} \quad \text{………………... (6.17)}$$

Contoh soal 5. Terdapat 5% barang di gudang timur cacat, sedangkan barang yang cacat di gudang barat sebanyak 10%. Bila diambil sampel acak sebanyak 200 barang dari gudang timur dan 300 barang dari gudang barat. Tentukan probabilitas persentase barang yang cacat dalam gudang barat 2% lebih banyak dibanding gudang timur!

Jawab 5.

$$\text{Gudang barat}: n_1=300, p_1=0{,}1$$
$$\text{Gudang timur}: n_2=200, p_2=0{,}05$$
$$\hat{p}_1=\text{proporsi barang yang cacat di gudang barat dalam sampel}$$
$$\hat{p}_2=\text{proporsi barang yang cacat di gudang timur dalam sampel}$$
$$\sigma_{\hat{p}_1-\hat{p}_2}=\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}=\sqrt{\frac{0{,}1(0{,}9)}{300}+\frac{0{,}05(0{,}95)}{200}}=0{,}023$$
$$Z=\frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sigma_{\hat{p}_1-\hat{p}_2}}=\frac{(\hat{p}_1-\hat{p}_2)-(0{,}1-0{,}05)}{0{,}023}$$

Karena barang cacat di gudang barat 2% lebih banyak daripada di gudang timur, maka $(\hat{p}_1 - \hat{p}_2) > 0{,}02$ sehingg a diperoleh $Z = \frac{0{,}02-0{,}05}{0{,}023} = -1{,}3$ dari tabel distribusi normal untuk nilai Z= -1,3 ➔ 0.0968 sehingga $(\hat{p}_1 - \hat{p}_2) > 0{,}02$= P(Z>-1,3) =1- 0.0968 = 0,9032 = 90,32%.

Soal Latihan.

1. Pada suatu pengiriman barang yang terdiri dari 2000 tube elektronika telah diketahui terdapat 600 unit tube yang tidak memenuhi standar mutu. Jika sampel acak sebanyak 500 unit dipilih dari populasi tersebut tanpa pengembalian, berapakah probabilitas sampel populasi yang tidak memenuhi standar mutu :
   a. akan kurang dari 150/500
   b. antara 144/500 sampai dengan 145/500
   c. lebih besar dari 164/500

2. Besi baja yang diproduksi perusahaan A mempunyai rata-rata daya regang sebesar 4500 lbs dan variansi sebesar 40000 lbs, sedangkan yang diproduksi perusahaan B mempunyai ratarata daya regang sebesar 4000 lbs dan variansi sebesar 90000 lbs. Misalkan sampel random sebanyak 50 diambil dari perusahaan A dan sampel random sebanyak 100 diambil dari perusahaan B, berapakah probabilitas rata-rata daya regang beda dua rata-rata dari dua sampel itu yang lebih besar dari 600 lbs?