MODUL III

DISTRIBUSI *SAMPLING*

MODUL III

DISTRIBUSI *SAMPLING*

1. TUJUAN PRAKTIKUM

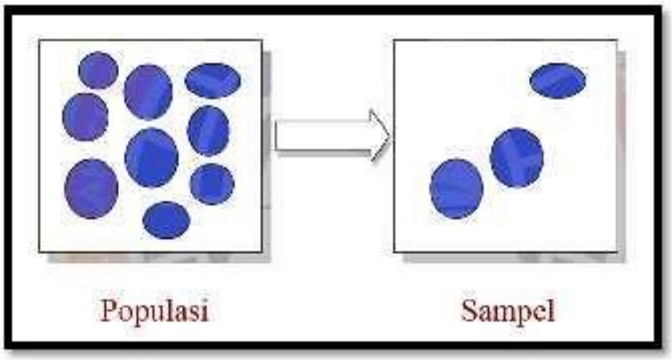
Tujuan dari praktikum ini adalah sebagai berikut:

1. Mahasiswa mampu menentukan:
   a. *Mean* sampel
   b. Variansi sampel
   c. Perbedaan *mean* dua populasi
   d. Rasio variansi dua populasi
   e. Proporsi
2. Mahasiswa mengetahui pengaruh ukuran sampel terhadap ketelitian pada penentuan nilai rata-rata populasi.
3. Mahasiswa mampu membuktikan teorema limit pusat (*central limit theorem*).

2. LANDASAN TEORI

Sampling sering digunakan dalam kehidupan sehari-hari. Teknik *sampling* sangat berguna dalam upaya penarikan kesimpulan (*inference*) yang valid dan dapat dipercaya. Hal ini disebabkan karena informasi yang diperoleh dari data sampel tidak mungkin lebih baik dari pada informasi yang sesungguhnya pada populasi.

Sampel adalah suatu himpunan bagian dari populasi, yang dianggap bias mewakili populasi. Populasi merupakan kumpulan dari keseluruhan elemen-elemen suatu objek yang menjadi perhatianya memiliki kuantitas dan karakteristik tertentu. Berikut ini hubungan antara populasi dengan sampel yang digambarkan sebagai berikut :



Gambar 1. Hubungan Populasi dengan sampel

Distribusi Sampling merupakan distribusi teoritis (distribusi kemungkinan) dari semua hasil sampel yang mungkin, dengan ukuran sampel yang tetap N, pada statistik (karakteristik sampel) yang digeneralisasikan ke populasi. Cara- cara pengambilan sampel antara lain:

- Daftar pertanyaan (*questionnaire*).
- Wawancara.
- Observasi atau pengamatan langsung.
- Melalui pos, telepon, atau alat komunikasi lainnya.

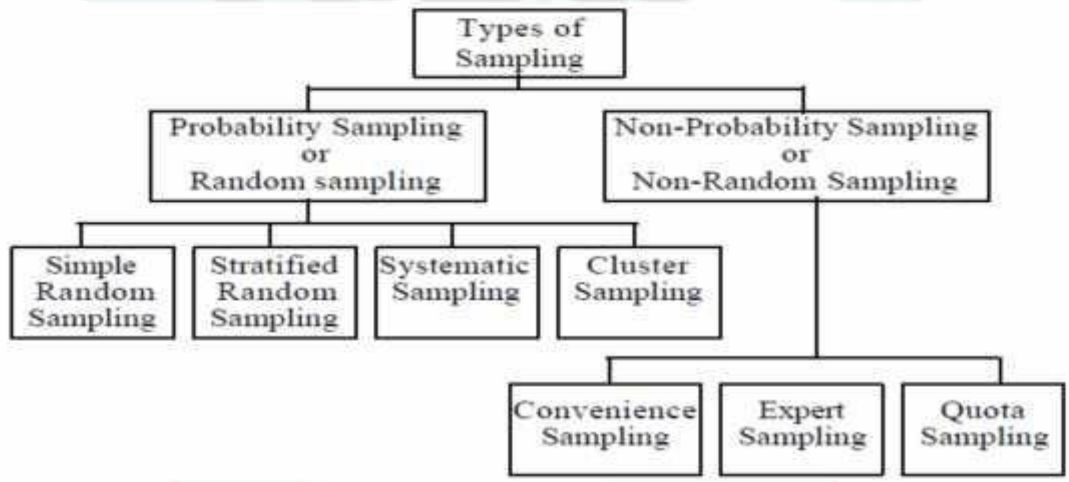
2.1 Parameter dan Statistik

Konstanta statistik dari populasi seperti *mean*/rataan (μ), variansi(s^2), proporsi (p) disebut dengan **parameter.** Jadi, **parameter** adalah bilangan/angka yang menggambarkan karakteristik dari populasi. Sedangkan ukuran statistik seperti *mean*/rataan ($\bar{X}$), variansi (s^2), proporsi (*p*) yang dihitung dari pengamatan sampel dikenal dengan **statistik. Statistik** adalah bilangan/angka yang menggambarkan karakteristik suatu sampel. Statistik merupakan perkiraan/taksiran dasar pada sampel data untuk menggambarkan perbedaan tentang parameter populasi.

Tabel 1. Parameter dan Statistik

Ukuran/Ciri	Parameter Populasi	Statistik Sampel
Rata-Rata	μ : (myu)	$\bar{x}$: (x bar)
Selisih 2 Rata-rata	$\lvert\mu_1 - \mu_2\rvert$: (nilai mutlak)	$\lvert\bar{x}_1 - \bar{x}_2\rvert$: (nilai mutlak)
Standar Deviasi = Simpangan Baku	σ : (sigma)	S
Varians = Ragam	σ^2	s^2
Proporsi	π : (phi atau p)	$\bar{p}$ atau $\hat{p}$
Selisih 2 proporsi	$\lvert\pi_1 - \pi_2\rvert$:(nilai mutlak)	$\lvert p_1 - p_2\rvert$: (nilai mutlak)

2.2 Metode Penarikan Sampel



Gambar 2. Metode Penarikan Sampel

(i) *Probabilitas Sampling* atau *Random Sampling*

a. *Probability Sampling*

Simple random sampling merupakan dasar dari probabilitas *sampling*. Pada kasus khusus probabilitas *sampling* dalam setiap unit pada populasi memiliki kemungkinan yang sama menjadi sampel. *Sampling* bisa dilakukan dengan atau tanpa penggantian.

b. *Stratified Random Sampling*

Stratified random sampling meliputi pembagian populasi menjadi kelompok yang disebut dengan **strata** dimana anggota di dalam satu strata cenderung sama (homogen) dan antara strata cenderung berbeda (heterogen). Langkah selanjutnya adalah mengambil sampel secara acak pada masing-masing strata. Kemudian sampel pada stratum atau *subgroup* tersebut digabung menjadi satu.

c. *Systematic Sampling*

Systematic sampling merupakan teknik *sampling* yang membutuhkan waktu yang sedikit dan biaya yang murah dibandingkan dengan simple *random sampling*.

d. *Cluster Sampling*

Cluster sampling merupakan teknik *sampling* dimana populasi dibagi menjadi beberapa *group*/gerombol (*cluster*) yang masing-masingnya dapat memrepresentasikan populasi tersebut.

(ii) *Non-probability Sampling* atau *Non-Random Sampling*

a. *Purposive Sampling*

Sampel diseleksi dengan tujuan yang jelas berdasarkan sudut pandang dan pemilihan unit sampel bergantung secara menyeluruh pada pertimbangan dan kebijakan dari pengamat.

b. *Quota Sampling*

Merupakan tipe pembatas dari *purpose sampling*. *Sampling* ini terdiri dari kuota sampel yang spesifik yang digambarkan dari kelompok-kelompok yang berbeda dan kemudian menggambarkan kebutuhan sampel dari kelompok tersebut dengan *purposive sampling*. *Quota sampling* ini sangat berguna sekali dalam penyelidikan/ *survey* pasar.

c. *Expert Opinion Sampling or Expert Sampling*

Expert opinion sampling melibatkan kumpulan dari beberapa orang yang memiliki pengetahuan dan keahlian dalam pengambilan keputusan terhadap suatu permasalahan yang sangat penting.

2.3 Distribusi *Sampling* Rata- Rata

Distribusi *sampling* rata- rata adalah distribusi probabilitas untuk nilai- nilai yang dapat terjadi dari rata- rata sampel yang didasarkan pada sejumlah sampel tertentu.

Mean dan standar deviasinya:

- Jika sampling tanpa pergantian dari suatu populasi terhingga berukuran N :

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

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

- Jika sampling dengan pergantian, yang berarti populasi tak terhingga :

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

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

Keterangan:

$\mu_{\bar{x}} =$ *Mean* daridistribusi *mean* sampel

$\mu =$ *Mean* populasi

$\sigma_{\bar{x}} =$ Deviasi standar dari distribusi *mean* populasi

$s =$ Deviasi standar sampel

$N =$ Ukuran populasi

$n =$ Ukuran sampel

Contoh Soal:

Dalam suatu pengujian kelelahan (*fatigue test*), material titanium diberi pembebanan berulang sampai deteksi timbulnya retak (*crack initiation*). Siklus pembebanan rata-rata sampai mulai retak adalah 25000 kali dengan deviasi standar 5000. Jika diuji 25 spesimen material titanium yang dipilih secara acak, berapakah :

- *Mean* dari sampel tersebut?
- Standar deviasi dari sampel tersebut? Jawab:
- *Mean* dari sampel

  $\mu_{\bar{x}} = \mu = 25000$
- Standar deviasi dari sampel

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{5000}{\sqrt{25}} = 1000$$

2.4 *Distribusi Proporsi Sampling*

Distribusi proporsi sampling adalah distribusi proporsi-proporsi dari seluruh sampel acak berukuran *n* yang mungkin dipilih dari sebuah populasi. Jika populasinya tak berhingga dan probabilitas terjadinya suatu kejadian atau *event* dikatakan sukses adalah P . Dan Q = 1 – P menunjukkan probabilitas gagal. Anggap semua kemungkinan ukuran sampel *n* digambarkan dari populasi. Sebagai contoh, tentukan proporsi *p* sukses. Dengan menggunakan Teorema Limit Pusat**,** jika ukuran sampel besar, distribusi proporsi sampel *p* mengikuti distribusi normal dengan rataan/*mean* $\mu_p = P$ dan S.D $\sigma_p = \sqrt{\frac{PQ}{n}}$.

Contoh Soal:

Divisi pengendalian mutu pabrik perkakas mesin mencatat bahwa 1,5% dari bearing mengalami cacat. Jika dalam pengiriman satu kotak produk terdiri dari 100 bearing, tentukan probabilitas banyaknya bearing yang cacat sebanyak 2% atau lebih!

Jawab:

Mean dan standar deviasi:

$\mu_p = P = 0{,}015$

$$\sigma_p = \sqrt{\frac{PQ}{n}} = \sqrt{\frac{0{,}015(1-0{,}015)}{100}} = 0{,}0122$$

Faktor koreksi variabel diskrit = 1/2n = 1/200 = 0,005

Proporsi (2%) setelah dikoreksi, p = 0,02 – 0,005 = 0,015

Maka,

$$
\begin{aligned}
P(p > 0{,}01) &= 1 - P(p \le 0{,}01) \\
&= 1 - P\left(Z_p \le \frac{0{,}015 - 0{,}015}{0{,}0122} \right) \\
&= 1 - P(Z_p \le 0) = 1 - 0{,}5 = 50\%
\end{aligned}
$$

2.5 Distribusi *Sampling* Beda Rata- Rata

Merupakan distribusi dari perbedaan dari besaran rata-rata yang muncul dari sampel-sampel dua populasi.

- Rata- rata

$$\mu_{X-X} = \mu_1 - \mu$$

- Simpangan Baku

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

2.6 Distribusi *Sampling* Beda Proporsi

Merupakan distribusi dari perbedaan dua besaran proporsi yang muncul dari sampel dua populasi.

- Rata- rata

$$\mu_{p-p} = P_1 - P_2$$

- Simpangan Baku

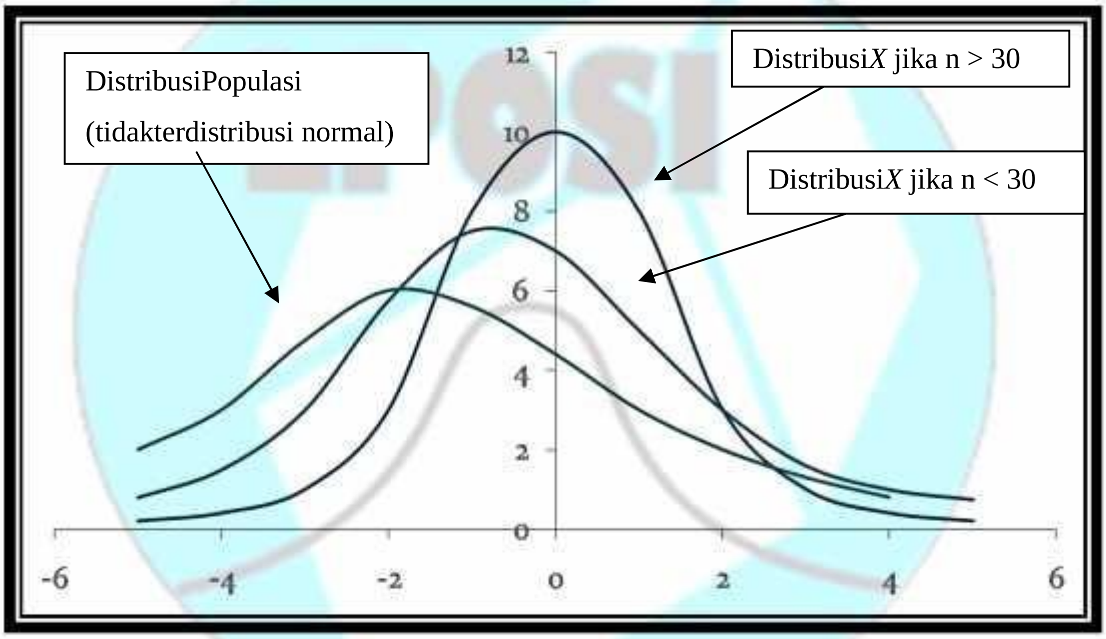
$$\sigma_{p_1-p_2} = \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

2.7 Teorema Limit Pusat (*Central Limit Theorem*)

Suatu populasi yang memiliki distribusi normal, distribusi *mean* sampling juga terdistribusi normal untuk nilai *n* berapapun (tidak tergantung ukuran sampel). Dengan kata lain, jika dimisalkan X_1, X_2,X_3, ...,X_n adalah suatu sampel acak dari suatu populasi yang terdistribusi normal dengan *mean* μ dan standar deviasi σ maka untuk sembarang nilai *n*.

Sementara itu dari suatu populasi yang tidak terdistribusi secara normal, jika ukuran sampel cukup besar (n > 30), distribusi *mean* sampling akan mendekati suatu distribusi normal (*gaussian*) apapun bentuk asli distribusi probabilitasnya. Pernyataan ini dikenal sebagai *Teorema Limit Pusat*). Dengan kata lain, seandainya X_1, X_2,X_3, ...,$X_n$adalah suatu sampel acak dari suatu populasi tidak terdistribusi secara normal dengan *mean* μ dan standar deviasi σ, maka untuk nilai n yang cukup besar (n > 30).

$$\overline{X} \approx N\left[\mu, \frac{\sigma^2}{n}\right]$$



Gambar 14. Ilustrasi Teorema Limit Pusat

Contoh:

1) Suatu perusahaan memproduksi bola lampu yang umurnya berdistribusi hampir normal dengan rataan 800 jam dan simpangan baku 40 jam. Hitunglah peluangnya bahwa suatu sampel acak dengan 16 bola lampu akan mempunyai umur rata- rata kurang dari 775 jam.

   Jawab:

   Misalkan: X = bola lampu.

   X ~ N[800, 40].

$\overline{X}$ = rataan/*mean mean* sampel. Kemudian

$$\overline{X} \sim \mathrm{N}\left[\mu, \frac{\sigma^2}{n}\right] = \mathrm{N}\left[800, \frac{1600}{16}\right] = \mathrm{N}[800{,}100]$$

Maka, P($\overline{X}$ < 775) = P(Z < -2.5)

= 0.0062

2) Suatu pabrik dapat memproduksi voltmeter dengan kemampuan pengukuran tegangan, rataan 40 volt dan standar deviasi 2 volt.

Misalkan tegangan tersebut berdistribusi normal. Dari 1000 voltmeter yang diproduksi, berapa voltmeter yang tegangannya melebihi 43 volt?

Jawab:

Misalkan: X = Tegangan voltmeter.

X ~ N[40, 2].

Dengan transformasi $Z = \frac{x-\mu}{\sigma}$

P(X > 43) = P (Z > $\frac{X-4}{2}$)

= P (Z > 1,5)

= 1 – P (Z ≤ 1,5)

= 1- 0,9332

= 0,0668

2.8 *Standard Error*

Standar deviasi dari distribusi *sampling* pada statistik disebut ***standard error*.** Standar deviasi dari distribusi *mean* sampel disebut ***standard error of the mean***. Begitu guja dengan standar deviasi pada distribusi proporsi sampel disebut ***standard error of the proportion***. *Standard errors* dari ratan sampel/*mean* sampel $\overline{X}$ dan proporsi sampel *p* digunakan untuk memperoleh limit kepercayaan untuk rataan populasi μ and proporsi populasi P masing-masing.

Tabel 2. Tabel Perbedaan *Mean* Sampel dan Proporsi Sampel

Statistik	Standar Error	Keterangan
Mean sampel $\bar{X}$	$\frac{\sigma}{\sqrt{n}}$	Ukuran populasi tak terhingga atau sampel dengan penggantian
	$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$	Ukuran populasi N terhingga atau sampel tanpa penggantian
Proporsi sampel p	$\sqrt{\frac{PQ}{n}}$	Ukuran populasi tak terhingga atau sampel dengan penggantian
	$\sqrt{\frac{PQ}{n}}\sqrt{\frac{N-n}{N-1}}$	Ukuran populasi N terhingga atau sampel tanpa penggantian