Sebaran Penarikan Contoh

Muhammad Arif Rahman
arifelzain@ub.ac.id

Populasi vs Sampel

- Populasi

  Keseluruhan objek penelitian atau keseluruhan elemen yang akan diteliti.
- Sampel

  Sebagian dari populasi

Syarat sampel yang baik

- Representatif → dapat memberi gambaran yang tepat mengenai parameter populasi serta mewakili sebanyak mungkin karakteristik populasi
- Valid, ditentukan oleh dua pertimbangan :

1. Akurasi atau Ketepatan

   semakin sedikit tingkat kesalahan dalam sampel, maka sampel semakin akurat → sampel harus mempunyai selengkap mungkin kriteria populasi

2. Presisi

   sedekat mana estimasi sampel dengan karakteristik populasi → semakin kecil perbedaan antara rata-rata populasi dengan rata-rata sampel, maka semakin tinggi tingkat presisi sampel

Ukuran sampel

- Berapa banyak jumlah sampel yang akan diambil.
- Dalam penelitian kualitatif, ukuran sampel bukan hal utama, karena yang dipentingkan adalah kekayaan informasi dari sampel.
- Sebaliknya, pada penelitian kuantitatif, ukuran sampel menjadi hal yang utama.
- Sampel besar bila $n \geq 30$
- Sampel kecil bila $n < 30$

Daftar jumlah sampel berdasarkan jumlah populasinya

Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384

Krejcie dan Morgan (1970) dalam Uma Sekaran (1992)

Teknik Pengambilan Sampel

- Random Sampling

  teknik pengambilan sampel, dimana setiap anggota populasi memiliki peluang yang sama untuk dijadikan sampel
- Non Random Sampling

  teknik pengambilan sampel, dimana anggota populasi tidak semuanya memiliki peluang untuk dijadikan sampel.

Random Sampling

- <u>**Simple Random Sampling**</u>

  dilakukan bila analisis penelitiannya cenderung deskriptif dan bersifat umum.

  menggunakan cara undian, tabel bilangan random
- Macam :
- Pemulihan

  Sampel yang sudah dipilih masih mempunyai kesempatan untuk dipilih kembali
- Tanpa Pemulihan

  sampel yang sudah dipilih tidak mempunyai kesempatan untuk dipilih kembali

Random Sampling

- <u>**Systematic Sampling**</u>

  dilakukan apabila ukuran populasi banyak dan peneliti tidak memiliki alat pengambil data secara random
- Misal :

  Terdapat 5000 kapal di pelabuhan, kita ingin mengambil sampel sebanyak 250 kapal
- Cara :

  Tentukan intervalnya, misal 20

  Pilih secara acak sampel pertama, misal 7

  berarti 7 merupakan anggota ke-1 dalam sampel

  kemudian 27 merupakan anggota ke-2

  kemudian 47 merupakan anggota ke-3 dst.

Random Sampling

- <u>**Stratified Random Sampling**</u>

  dilakukan bila unsur populasi berkarakteristik heterogen, dan heterogenitas tersebut mempunyai arti yang signifikan pada pencapaian tujuan penelitian.

  dapat dilakukan secara proporsional maupun non proporsional
- Antar Kelas cenderung bersifat heterogen, tetapi anggota dalam suatu kelas cenderung sama (homogen)

- Proporsional

  jumlah sampel dalam setiap stratum sebanding dengan jumlah unsur populasi dalam stratum tersebut (besar mendapatkan besar, kecil mendapatkan kecil)
- Non proporsional

  jumlah sampel dalam setiap stratum tidak sebanding dengan jumlah unsur populasi dalam stratum tersebut (besar belum tentu mendapatkan besar, kecil belum tentu mendapatkan kecil)

Random Sampling

- **Cluster Sampling**

  digunakan untuk mengambil sampel dalam kelompok, bukan individu
- Antar Kelas cenderung bersifat homogen, tetapi anggota dalam suatu kelas cenderung berbeda (heterogen).
- Misal : Dari 40 Rukun Nelayan (RN) yang ada di Kecamatan Brondong diambil 6 RN secara acak untuk dijadikan sampel penelitian.

Random Sampling

- **<u>Area Sampling</u>**

  dipakai ketika populasi penelitian tersebar di berbagai wilayah

  Pengelompokan ditentukan oleh letak geografis atau administratif
- Misal

  seorang marketing manajer sebuah perusahaan perikanan ingin mengetahui tingkat penerimaan masyarakat Jawa Timur atas sebuah product. Maka dapat diambil sampel dengan menentukan wilayah yang akan dijadikan sampel (Kabupaten?, Kotamadya?, Kecamatan?, Desa?).

Non Random Sampling

- **Convenience Sampling**

  Dalam memilih sampel, peneliti tidak mempunyai pertimbangan lain kecuali berdasarkan kemudahan saja.

  Jenis sampel ini biasanya dimanfaatkan untuk penelitian penjajagan, yang kemudian diikuti oleh penelitian lanjutan yang sampelnya diambil secara acak (*random*).

  Beberapa kasus penelitian yang menggunakan jenis sampel ini, hasilnya ternyata kurang obyektif

Non Random Sampling

- **Purposive Sampling**

  sampel diambil dengan maksud atau tujuan tertentu
- Dibedakan dalam 2 jenis
- *Judgement Sampling*

  Sampel dipilih dinilai bahwa dia adalah pihak yang paling baik untuk dijadikan sampel penelitiannya. umumnya memilih sesuatu atau seseorang menjadi sampel karena mereka mempunyai kekayaan informasi.

- *Quota Sampling*

  sampel ini adalah bentuk dari sampel distratifikasikan secara proposional, namun tidak dipilih secara acak melainkan secara kebetulan saja

Non Random Sampling

- <u>**Snowball Sampling**</u>

  Cara ini banyak dipakai ketika peneliti tidak banyak tahu tentang populasi penelitiannya. Dia hanya tahu satu atau dua orang yang berdasarkan penilaiannya bisa dijadikan sampel, kemudian sampel pertama memberitahu orang lain yang bisa dijadikan sampel.
- Hal ini bisa juga dilakukan pada pencandu narkotik, para gay/lesbi, atau kelompok-kelompok sosial lain yang eksklusif (tertutup)

Sebaran Penarikan Sampel

- Sebaran statistik yang diperoleh melalui pemilihan peluang sampel dari suatu populasi dengan ukuran tertentu.

- Sebaran penarikan sampel suatu statistik akan bergantung pada ukuran populasi, ukuran sampel, dan metode pengambilan sampelnya.

Sebaran Penarikan Sampel bagi Nilai Tengah

- Sebuah populasi seragam diskret yang terdiri atas nilai-nilai 0,1,2,dan 3.
- Keempat pengamatan itu menyusun populasi nilai-nilai sebuah peubah acak *X* yang memiliki sebaran peluang

  f(x) = ¼ , untuk *x* = 0,1,2,3,
- Dengan mean

$$\mu = E(X) = \sum_{x=0}^{3} x f(x)$$

$$= \frac{0+1+2+3}{4} = \frac{3}{2}$$

- Dengan ragam

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_{x=0}^{3} (x-\mu)^2 f(x)$$

$$= \frac{\left(0-\frac{3}{2}\right)^2 + \left(1-\frac{3}{2}\right)^2 + \left(2-\frac{3}{2}\right)^2 + \left(3-\frac{3}{2}\right)^2}{4} = \frac{5}{4}$$

- Dari nilai diatas kita ambil 2 nilai secara acak dengan cara pemulihan (sampel yang sudah dipilih, berpeluang dipilih lagi) dan dihitung rata-ratanya.
-

No.	Sampel	Rata-rata
1	0,0	0
2	0,1	0,5
3	0,2	1
4	0,3	1,5
5	1,0	0,5
6	1,1	1
7	1,2	1,5
8	1,3	2

No.	Sampel	Rata-rata
9	2,0	1
10	2,1	1,5
11	2,2	2
12	2,3	2,5
13	3,0	1,5
14	3,1	2
15	3,2	2,5
16	3,3	3

- Kita kumpulkan rata-rata sampelnya dalam tabel sebaran penarikan contoh bagi nilai tengah seperti berikut

Rata-rata	f	Distribusi
0	1	1/16
0,5	2	2/16
1	3	3/16
1,5	4	4/16
2	3	3/16
2,5	2	2/16
3	1	1/16
	16	

- Dari tabel diatas kita bisa melihat bahwa sebaran nilai tengah sampel adalah normal dengan titik kulminasinya adalah 1,5. Jika kita menghitung rata-rata populasi maka kitapun akan mendapatkan hasil 1,5. berarti $\mu = \bar{x}$
- Nilai tengah $\bar{x}$ tidak tergantung pada ukuran contoh. Tetapi ragam $\bar{x}$ bergantung pada ukuran contoh dan nilainya sama dengan ragam populasi asalnya σ^2 dibagi dengan n, sehingga semakin besar ukuran contoh, semakin kecil galat bakunya dan kemungkinannya $\bar{x}$ akan mendekati μ

Sampel Besar

DALIL 1

JIKA

Sampel:
berukuran = $n \geq 30$
rata-rata = $\bar{x}$

diambil DENGAN PEMULIHAN dari

Populasi berukuran = N terhingga
Rata-rata = μ
Simpangan baku = σ

MAKA

Distribusi Rata-rata akan mendekati distribusi Normal dengan :

$$\mu_{\bar{x}} = \mu \qquad \text{dan} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{dan nilai } z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Contoh

- Perusahaan perikanan dalam satu hari rata-rata mampu memproduksi 1.000.000 bungkus bakso ikan. Perusahaan ini menyatakan bahwa rata-rata berat sebungkus bakso ikan adalah 300 g dengan standart deviasi 15 g. Rata-rata populasi dianggap menyebar normal.
- Jika diambil sampel secara acak sebanyak 100 bungkus DENGAN PEMULIHAN, hitunglah peluang isi sebungkus bakso ikan kurang dari 303 g.

Jawab

- N = 1.000.000 $\quad\quad \sigma = 15$
- $\mu_{\bar{x}} = \mu = 300 \quad\quad n = 100$
- $P(\bar{x} < 303) = P(z < ?)$
- Simpangan baku $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{15}{\sqrt{100}} = 1{,}5$

$$z = \frac{303 - 300}{1{,}5} = \frac{3}{1{,}5} = 2.0$$

- $P(\bar{x} < 303) = P(z < 2.0)$ (lihat tabel z)

$= 0.9772 =$ **97.72%**

DALIL 2

JIKA

Sampel:
berukuran = n ≥ 30
rata-rata = $\bar{x}$

} diambil TANPA PEMULIHAN dari

{ Populasi berukuran = N terhingga
Rata-rata = μ
Simpangan baku = σ

MAKA

Distribusi Rata-rata akan mendekati distribusi Normal dengan :

$\mu_{\bar{x}} = \mu$ dan $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ dan nilai $z = \frac{\bar{x}-\mu}{(\sigma/\sqrt{n})\sqrt{\frac{N-n}{N-1}}}$

$\sqrt{\frac{N-n}{N-1}}$ disebut sebagai **FAKTOR KOREKSI** populasi terhingga

- Faktor Koreksi (FK) akan menjadi penting jika sampel berukuran n diambil dari populasi berukuran N yang terhingga/terbatas besarnya
- Jika sampel berukuran n diambil dari populasi berukuran N yang sangat besar maka FK akan mendekati 1.

Contoh soal

- Diketahui rata–rata hasil tangkapan 500 nelayan adalah 165 kg dengan standart deviasi 12 kg. Kemudian diambil 40 nelayan secara acak sebagai sampel. Rata–rata populasi dianggap normal.
- Jika penarikan sampel dilakukan TANPA PEMULIHAN, hitunglah peluang hasil tangkapan nelayan lebih dari 167 kg.

Jawab

- N = 500 σ = 12
- $\mu_{\bar{x}} = \mu$ = 165 n = 40
- P($\bar{x}$ > 167) = P(z > ?)
- Simpangan baku

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{12}{\sqrt{40}}\sqrt{\frac{500-40}{500-1}} = 1{,}75$$

$$z = \frac{167-165}{1{,}75} = \frac{2}{1{,}75} = 1{,}14$$

- P($\bar{x}$ > 167) = 1– P(z > 1,14) (lihat tabel z)

  = 1– 0,8729 = 0,1271 = **1,27%**

Dalil 3 (DALIL LIMIT PUSAT)

JIKA

Sampel:
berukuran = n
rata-rata = $\bar{x}$

} diambil dari

Populasi berukuran = N yang BESAR atau tak hingga
Rata-rata = μ ; simpangan baku = σ

MAKA

Distribusi Rata-rata akan mendekati distribusi Normal dengan :

$$\mu_{\bar{x}} = \mu \qquad \text{dan} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \qquad \text{dan nilai } z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

- Dalil Limit Pusat berlaku untuk penarikan sampel dari populasi yang sangat besar atau takhingga
- Beberapa buku mencatat hal berikut : Populasi dianggap BESAR jika ukuran sampel KURANG DARI 5% ukuran populasi atau $\frac{n}{N} < 5\%$

Contoh

- Sebuah perusahaan lampu mengklaim bahwa rata-rata umur lampu adalah 600 jam dengan standart deviasi 25 jam. Hitunglah peluang bahwa suatu contoh acak 15 bohlam akan mempunyai rata-rata umur lampu kurang dari 590 jam.

Jawab

- σ = 25
- $\mu_{\bar{x}} = \mu = 600$ n = 15

- P($\bar{x}$ < 590) = P(z < ?)
- Simpangan baku $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{25}{\sqrt{15}} = 6{,}5$

$$z = \frac{590 - 600}{6{,}5} = \frac{-10}{6{,}5} = -1{,}54$$

- P($\bar{x}$ < 590) = P(z < -1,54) (lihat tabel z)

  = 0,0618 = **6,18%**

Sebaran *t*

- Bila ukuran contohnya kecil ($n < 30$), nilai-nilai s^2 berfluktuasi cukup besar dari contoh satu ke contoh lainnya, dan sebaran nilai-nilai $(\bar{x} - \mu)/(s/\sqrt{n})$ tidak lagi normal baku. Bila demikian halnya, kita dihadapkan dengan sebaran suatu statistik yang disebut T, yang nilai-nilainya adalah

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Definisi

- Bila $\bar{x}$ dan s^2 masing-masing adalah nilai tengah ragam suatu contoh acak berukuran n yang diambil dari suatu populasi normal dengan nilai tengah μ dan ragam σ^2, maka

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

merupakan sebuah nilai peubah acat T yang mempunyai sebaran t dengan derajat bebas (db) atau *degree of freedom (df)* = v = n-1

- Distribusi-t pada prinsipnya adalah pendekatan distribusi sampel kecil dengan distribusi normal.
- Dua hal yang perlu diperhatikan dalam Tabel t adalah

  1. derajat bebas (db)

  2. nilai α
- Perbedaan Tabel z dan Tabel t

  Tabel z $\rightarrow$ nilai z menentukan nilai α

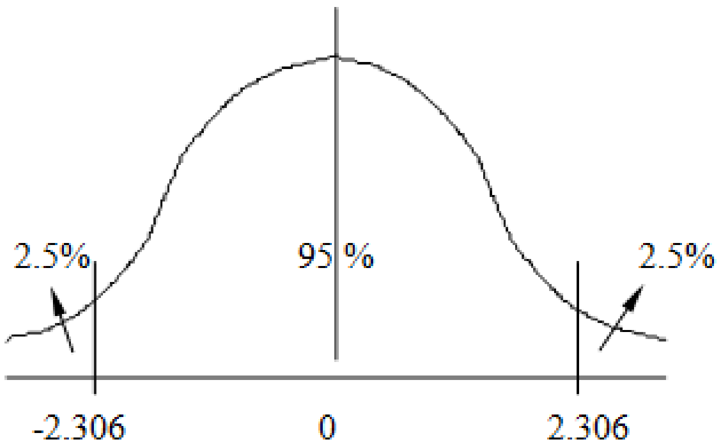
  Tabel t $\rightarrow$ nilai α dan db menentukan nilai t
- Nilai α adalah luas daerah kurva di kanan nilai t atau luas daerah kurva di kiri nilai -t
- Nilai α $\rightarrow$ 0.1 (10%) ; 0.05 (5%) ; 0.025(2.5%) ; 0.01 (1%) ; 0.005(0.5%)
- Nilai α terbatas karena banyak kombinasi db yang harus disusun

Contoh

- PT Cigar menyatakan bahwa 95% rokok produksinya rata-rata mengandung nikotin 1.80 mg, data tersebar normal. Yayasan Konsumen melakukan pengujian nikotin terhadap 9 batang rokok dan didapatkan rata-rata kandungan nikotin = 1.95 mg dengan standar deviasi = 0.24 mg. Apakah hasil penelitian Yayasan Konsumen mendukung pernyataan yang dibuat perusahaan?

Jawab

- 95 % berada dalam selang $\rightarrow$ berarti 5 % berada di luar selang; 2.5 % di kiri t dan 2.5% di kanan t
- $\alpha = 2.5 \% = 0.025$
- $n = 9 \rightarrow db = n - 1 = 8$
- t tabel (db, α) = t tabel(8; 0.025) = 2.306
- Jadi 95 % berada dalam selang $-2.306 < t < 2.306$



Arti Gambar di samping adalah nilai t sampel berukuran n = 9, berpeluang 95% jatuh dalam selang

$-2,306 < t < 2,306$.

Peluang $t > 2,306 = 2.5\%$

dan

Peluang $t < -2.306 = 2.5\%$

- Nilai t

μ = 1.80 n = 9

$\bar{x}$ = 1.95 s = 0.24

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{1{,}95 - 1{,}8}{0{,}24 / \sqrt{9}} = \frac{0{,}15}{0{,}08} = 1{,}875$$

- Nilai t hitung = 1.875 berada dalam selang −2.306 < t < 2.306
- jadi hasil penelitian Yayasan Konsumen masih sesuai dengan pernyataan manajemen PT Cigar

Sebaran Penarikan Sampel Bagi Beda Dua Mean

- Bila sampel-sampel berukuran n_1 dan n_2 diambil dari dua populasi BESAR atau TAKHINGGA dengan μ_1 dan μ_2 serta σ_1^2 dan σ_2^2 , maka beda kedua nilai tengah contoh ($\bar{x}_1$ dan $\bar{x}_2$) akan menyebar menghampiri sebaran normal dengan nilai tengah dan simpangan baku

$$\mu_{\bar{x}_1-\bar{x}_2} = |\mu_1 - \mu_2|$$ dan $$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

dengan demikian
$$z = \frac{|\bar{x}_1 - \bar{x}_2| - |\mu_1 - \mu_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- Beda selisih nilai tengah memakai nilai mutlak

Contoh

- Diketahui rata-rata panjang ikan tengiri yang ditangkap dengan pancing = 125 cm dengan ragam = 119 cm sedangkan rata-rata panjang ikan tengiri yang ditangkap dengan jaring = 128 cm dengan ragam = 181 cm. diasumsikan kedua populasi berukuran besar.

  Jika diambil secara acak 100 ekor ikan tengiri pancing dan 100 ikan tengiri jaring sebagai sampel, berapa peluang terdapat perbedaan panjang kedua kelompok akan kurang dari 2 cm?

Jawab

$$\mu_{\bar{x}_1-\bar{x}_2} = |\mu_1 - \mu_2| = |125-128| = 3$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}} = \sqrt{\frac{119}{100}+\frac{181}{100}} = \sqrt{3}$$

$$z = \frac{|\bar{x}_1 - \bar{x}_2| - |\mu_1 - \mu_2|}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} = \frac{2-3}{\sqrt{3}} = -0{,}58$$

- P ($\bar{x}_1 - \bar{x}_2 < 2$) = P($z < -0{,}58$)
  = 0,2810 = **28,1 %**

TERIMA KASIH