

Metode Sampling Penelitian Kedokteran dan Kesehatan

dr. I Ketut Tangking Widarsa, MPH
dr. Putu Ayu Swandewi Astuti, MPH, Ph.D
Ni Made Dian Kurniasari, S.KM., MPH

Metode Sampling Penelitian Kedokteran dan Kesehatan

Penulis:

dr. I Ketut Tangking Widarsa, MPH

dr. Putu Ayu Swandewi Astuti, MPH, PhD

Ni Made Dian Kurniasari, S.KM., MPH



PENERBIT

BASWARA PRESS

Jl. ByPass Ngurah Rai Nomor: 888 xx, Denpasar, Bali – Indonesia

(0361) 6209990; (+62)85858962311

info@baswarapress.com

baswarapress@gmail.com

Metode Sampling Penelitian Kedokteran dan Kesehatan

Penulis:

dr. I Ketut Tangking Widarsa, MPH

dr. Putu Ayu Swandewi Astuti, MPH, PhD

Ni Made Dian Kurniasari, S.KM., MPH

ISBN (PDF) : 978-623-99689-0-8

Copyright © Januari 2022

Hak Cipta dilindungi oleh Undang-Undang. Pertama kali diterbitkan di Indonesia dalam Bahasa Indonesia oleh BASWARA PRESS. Penggunaan material dalam karya ini diatur dalam Lisensi Publik Creative Commons 4.0 Atribusi-NonKomersial-TanpaTurunan (CC-BY-ND). Pengguna dapat mengutip konten dalam karya ini dengan memberi kredit yang sepatutnya kepada penulis asli karya ini. Dilarang menggunakan karya ini untuk kepentingan komersial dalam bentuk apapun. Dilarang mengubah dan menyebarkan isi karya ini dalam bentuk apa pun tanpa izin tertulis dari penulis.

PENERBIT

BASWARA PRESS

Jl. ByPass Ngurah Rai Nomor: 888 xx, Denpasar, Bali – Indonesia

(0361) 6209990; (+62)85858962311

info@baswarapress.com

baswarapress@gmail.com

Kata Pengantar

Puji syukur kehadapan Tuhan Yang Maha Esa atas rahmat dan karunia-NYA sehingga buku “Metode Sampling Penelitian Kedokteran dan Kesehatan” dapat diselesaikan. Penulisan buku ini dilatarbelakangi oleh banyaknya pertanyaan dari mahasiswa kedokteran dan kesehatan tentang metode sampling. Banyak yang bertanya apa itu tingkat kemaknaan (α), apa itu tingkat reliabilitas, apa itu power penelitian, apa itu presisi, apa itu efek desain dan bagaimana hubungannya dengan besar sampel penelitian kedokteran dan kesehatan. Banyak juga mahasiswa yang masih bingung mengenai bagaimana cara menentukan besar sampel untuk penelitian trial klinik, kohort, case-control, cross-sectional, uji diagnostic, dan survei. Mereka juga belum memahami kapan menggunakan metode random dan kapan menggunakan metode non-random. Mereka juga banyak berkonsultasi tentang cara pemilihan sampel pada penelitian trial klinik, kohort, case-control, cross-sectional, dan uji diagnostic. Pada buku ini dijelaskan tentang metode sampling penelitian kedokteran dan kesehatan yang mencakup cara penentuan besar sampel minimal dan cara pemilihan sampel untuk berbagai jenis penelitian kedokteran dan kesehatan.

Buku ini terdiri dari 7 Bagian yang meliputi Bagian 1: Pengantar Metode Sampling, Bagian 2: Metode Sampling Penelitian Survei, Bagian 3: Metode Sampling Penelitian Trial, Bagian 4: Metode Sampling Penelitian Uji Diagnostik, Bagian 5: Metode Sampling Penelitian Kohort, Bagian 6: Metode Sampling Penelitian Case-Control, dan Bagian 7: Metode Sampling Penelitian Korelasi. Pada setiap Bagian dijelaskan tentang sasaran belajar, topik bahasan, dan bahan diskusi.

Pada kesempatan ini, penulis menyampaikan ucapan terima kasih kepada semua pihak yang telah berkontribusi. Secara khusus, penulis menyampaikan penghargaan dan ucapan terima kasih kepada Ibu Aniek Pratiwi, SE, MPH dan dr. I Made Indra Wijaya, Ph.D. yang telah mereview secara mendalam naskah buku ajar ini.

Penulis berharap agar buku ini dapat membantu mahasiswa untuk memahami metode sampling penelitian kedokteran dan kesehatan. Penulis menyadari bahwa buku ini masih jauh dari sempurna, oleh karena itu, dengan kerendaha hati, penulis mengharapkan masukan yang berarti yang dapat dipakai untuk menyempurnakan buku ajar ini.

Denpasar, 2 Mei 2021

Penulis

Daftar Isi

Kata Pengantar	iv
Daftar Isi	v
DAFTAR SINGKATAN	ix
BAGIAN 1: PENGANTAR METODE SAMPLING	1
1.1 Pengantar	2
1.2 Pengertian Populasi, Populasi Target, dan Populasi Terjangkau	2
1.3 Sampel	3
1.4 Representativitas sampel	7
1.5 Metode pemilihan sampel	7
1.6 Diskusi Kelompok	8
1.7 Praktikum	9
BAGIAN 2. METODE SAMPLING PENELITIAN SURVEI	10
2.1 Pengantar	11
2.2 Metode Random	11
2.3 Simple Random Sampling (Acak Sederhana)	13
2.3.1 Pengertian	13
2.3.2 Indikasi	13
2.3.3 Kelebihan dan kekurangan	14
2.3.4 Cara pemilihan sampel	14
2.3.5 Cara penentuan besar sampel	18
2.3.6 Estimasi Parameter Populasi	22
2.3.6 Latihan	24
2.4 Systematic Random Sampling (Acak Sistematis)	25
2.4.1 Pengertian	25
2.4.2 Asumsi	25
2.4.3 Kelebihan dan kekurangan	25
2.4.4 Cara pemilihan sampel	25
2.4.5 Cara penentuan besar sampel	27
2.4.6. Estimasi Parameter Populasi	28
2.5 Stratified Random Sampling (Acak Berlapis)	30

2.5.1 Pengertian	30
2.5.2 Asumsi:	30
2.5.3 Kelebihan dan kekurangan:	30
2.5.4 Cara pemilihan sampel:	30
2.5.5 Cara penentuan besar sampel:	31
2.5.6 Estimasi Parameter Populasi	34
2.5.7 Latihan	36
2.6 <i>Cluster sampling</i> (Acak Kelompok).....	38
2.6.1 Pengertian	38
2.6.2 Indikasi	38
2.6.3 Kelebihan dan kekurangan	39
2.6.4 Cara pemilihan sampel	39
2.6.5 Cara penentuan besar sampel:	44
2.6.6 Estimasi Parameter Populasi	45
2.6.7 Latihan	46
2.7 Multistage Random Sampling (Acak Bertingkat)	47
2.7.1 Pengertian	47
2.7.2 Indikasi:	47
2.7.3 Kelebihan dan kekurangan:	47
2.7.4 Langkah pemilihan sampel.....	47
2.7.5 Penentuan <i>sampling fraction</i>	48
2.7.6 Estimasi parameter populasi.....	49
2.7.7 Penentuan besar sampel	50
2.7.8 Latihan	50
2.8 Metode Non-Random.....	51
2.8.1 Pengertian <i>non-random sampling</i>	51
2.8.2 Jenis <i>non-random sampling</i>	51
2.8.3 Indikasi <i>non-random sampling</i>	54
2.8.4 Kelebihan metode <i>non-random sampling</i>	54
2.8.5 Perbedaan metode <i>non-random</i> dengan metode <i>random</i>	54
2.8.6 Latihan	55
2.9 Survei Cepat (<i>Rapid survey</i>)	56
2.9.1 Pengertian	56
2.9.2 Indikasi	56

2.9.3 Kelebihan dan kekurangan	56
2.9.4 Cara penentuan besar sampel	57
2.9.5 Cara pemilihan sampel	59
2.9.6 Latihan	60
BAGIAN 3. METODE SAMPLING PENELITIAN TRIAL	62
3.1 Pengantar	63
3.2 Trial klinik paralel.....	63
3.3 Trial klinik menyilang (Cross-over design).....	64
3.4 Cara penentuan besar sampel	64
3.4.1 Parameter penentuan besar sampel.....	64
3.4.2 Penghitungan besar sampel untuk hipotesis perbedaan <i>mean</i> (untuk data numerik).....	66
3.4.3 Penghitungan besar sampel untuk hipotesis perbedaan insiden (variabel nominal).....	67
3.4.3 Penghitungan besar sampel Trial Negatif.....	68
3.5 Alokasi Sampel Secara <i>Random</i>	70
3.5.1 Alokasi <i>Random</i> dengan Bilangan <i>Random</i>	70
3.5.2 Alokasi <i>Random</i> dengan Blok <i>Random</i>	71
3.6 Latihan	72
BAGIAN 4. METODE SAMPLING PENELITIAN UJI DIAGNOSTIK	73
4.1 Pengantar	74
4.2 Parameter penentuan besar sampel.....	74
4.3 Penentuan besar sampel	75
4.4 Cara Pemilihan Sampel.....	76
4.4 Latihan	78
BAGIAN 5. METODE SAMPLING PENELITIAN KOHORT	80
5.1 Pengantar	81
5.2 Penelitian Kohort	81
5.3 Parameter penentuan besar sampel.....	82
5.4 Cara penentuan besar sampel	84
5.4 Metode pemilihan sampel	87
5.5 Latihan	94
BAGIAN 6. METODE SAMPLING PENELITIAN CASE-CONTROL	95
6.1 Pengantar	95
6.2 Rancangan Penelitian Case-Control.....	96
6.3 Parameter penentuan besar sampel.....	97

6.4	Cara penentuan besar sampel	99
6.4.1	Penentuan Besar Sampel <i>Independent Case-Control</i>	99
6.4.2	Penentuan besar sampel <i>Case-Control</i> dengan C kontrol per kasus	102
6.4.3	Penentuan besar sampel <i>Matched Case-Control</i>	103
6.5	Cara Pemilihan Sampel	104
6.6	Latihan	107
BAGIAN 7. METODE SAMPLING PENELITIAN KORELASI		109
7.1	Pengantar	109
7.2	Penelitian korelasi	110
7.3	Parameter penentuan besar sampel.....	111
7.4	Cara Penentuan Besar Sampel.....	112
7.8	Cara pemilihan sampel.....	113
7.9	LATIHAN.....	114
LAMPIRAN		115
Informasi Tentang Penulis		126

DAFTAR SINGKATAN

AKDR	: Alat kontrasepsi dalam rahim
Balita	: Bawah lima tahun
Batita	: Bawah tiga tahun
BBLR	: Berat badan lahir rendah
BCG	: <i>Bacillus Calmette–Guérin</i>
BKKBN	: Badan Koordinasi Keluarga Berencana Nasional
Bumil	: Ibu hamil
<i>CI</i>	: <i>Confidence interval</i>
DE	: <i>Design Effect</i>
DM	: Diabetes melitus
DO	: <i>Drop out</i>
EPI	: <i>Extended program for immunization</i>
Germas	: Gerakan masyarakat sehat
IK	: interval kepercayaan
Lansia	: Lanjut usia
LR	: <i>Likelihood ratio</i>
<i>OR</i>	: <i>Odd Ratio</i>
PA	: patologi anatomi
PJK	: Penyakit jantung koroner
PPS	: <i>Probability proportional to size</i>
<i>PSU</i>	: <i>Primary Sampling Unit</i>
PUS	: Pasangan Usia Subur
RR	: <i>Relative Risk</i>
RT	: Rumah tangga
RW	: Rukun warga
SD	: Standar deviasi

SMA : Sekolah Menengah Atas
SKRT : Survei Kesehatan Rumah Tangga
S.Ked : Sarjana Kedokteran
TBC : Tuberkulosis
WHO : World Health Organization
WUS : Wanita usia subur

BAGIAN 1: PENGANTAR METODE SAMPLING

dr. I Ketut Tangking Widarsa, MPH

Dosen Biostatistik Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com

Tujuan Pembelajaran

- Memahami pengertian populasi, populasi target, populasi terjangkau
- Memahami pengertian sampel
- Memahami kriteria eligibilitas sampel
- Memahami pengertian representativitas sampel dan faktor yang memengaruhi
- Memahami jenis metode pemilihan sampel

Pokok Bahasan:

- 1.1 Pengantar
- 1.2 Pengertian Populasi, Populasi Target, dan Populasi Terjangkau
- 1.3 Sampel
- 1.4 Faktor yang memengaruhi representativitas sampel
- 1.5 Jenis metode pemilihan sampel

Bahan Bacaan:

1. Cochran W.G. 1977. Sampling Techniques, Third edition, John Wiley&Sons, New York.
2. Levy P.S. and Lemeshow S. 1980. Sampling for Health Professionals, Lifetime Learning Publication, California.
3. Lwanga S.K. and Lemeshow S. 1991. Sample Size Determination in Health Studies, World Health Organization, Geneva.

1.1 Pengantar

Penelitian di bidang kedokteran dan kesehatan dalam dekade terakhir berkembang dengan sangat pesat sejalan dengan perkembangan ilmu pengetahuan dan teknologi. Pesatnya perkembangan di bidang ilmu kedokteran dan kesehatan tidak terlepas dari perkembangan dan kemajuan riset di bidang kedokteran dan kesehatan. Hal ini tercermin dari jumlah publikasi hasil penelitian di bidang kedokteran dan kesehatan yang terus bertambah setiap tahunnya.

Pada umumnya penelitian kedokteran dan kesehatan dilakukan pada sebagian dari populasi yang disebut sampel. Banyak keuntungan penggunaan sampel di dalam penelitian, seperti biaya lebih murah, waktu yang dibutuhkan lebih singkat, pelaksanaan lebih simpel, dan validitas data lebih terkendali. Selain beberapa keuntungan dari penggunaan sampel dalam penelitian kedokteran dan kesehatan, juga ada beberapa kelemahannya, seperti ada kemungkinan hasil yang didapat dari sampel berbeda dengan gambaran yang seharusnya di populasi. Perbedaan hasil tersebut bisa terjadi karena *sampling error* dan juga bisa sebagai akibat dari *systematic error*. *Sampling error* berkaitan dengan penetapan kriteria populasi, sampel, dan metode pemilihan sampel yang dipakai. Pada bagian ini akan dibahas tentang populasi dan sampel dalam penelitian kedokteran dan kesehatan, sedangkan metode pemilihan sampel akan dibahas pada sesi selanjutnya.

1.2 Pengertian Populasi, Populasi Target, dan Populasi Terjangkau

Unit analisis merupakan satuan analisis yang dipakai di dalam suatu penelitian. Unit analisis juga disebut sebagai unit observasi. Kumpulan dari unit analisis disebut populasi. Untuk menentukan subyek penelitian, peneliti terlebih dahulu harus menentukan siapa yang menjadi unit analisis dari penelitian tersebut. Unit analisis tersebut bisa individu, keluarga atau rumah tangga, organisasi atau lembaga, dan geografis (desa, kecamatan, kabupaten, provinsi dan lainnya). Setiap unit analisis harus memiliki kriteria tertentu. Misalnya pada penelitian kejadian stunting pada anak balita di Kabupaten Gianyar tahun 2021, maka anak balita yang berdomisili di Kabupaten Gianyar pada tahun 2021 menjadi unit analisis penelitian tersebut. Pada penelitian ketersediaan pangan keluarga di Kabupaten Gianyar, keluarga atau rumah tangga di Kabupaten Gianyar menjadi unit analisisnya. Populasi penelitian dibedakan menjadi populasi target dan populasi terjangkau.

Populasi Target:

Populasi target dalam suatu penelitian merupakan kelompok individu yang akan diteliti atau merupakan kelompok individu kemana hasil penelitian akan digeneralisasikan. Deskripsi ciri-ciri dari populasi target suatu penelitian sangat penting karena menjadi dasar penentuan kriteria sampel. Karakteristik yang menjadi ciri dari populasi target umumnya terdiri dari karakteristik demografi seperti jenis kelamin, umur, dan lainnya serta karakteristik klinis seperti penderita diare, stroke, hipertensi dan lainnya. Misalnya akan dilakukan trial klinik tahap III tentang efektivitas dan keamanan vaksin XX untuk pencegahan COVID-19. Vaksin tersebut diharapkan akan diterapkan pada semua penduduk yang sehat. Oleh karena itu, yang menjadi populasi target dari uji klinik tahap III vaksin tersebut adalah semua penduduk yang sehat. Dari contoh ini, semua penduduk merupakan ciri demografi dan sehat secara fisik adalah ciri klinis. Contoh lain, misalnya akan diteliti kejadian diare pada balita, maka pada penelitian ini balita sebagai ciri demografi dan diare sebagai ciri klinis.

Populasi Terjangkau:

Populasi terjangkau adalah bagian dari populasi target yang akan dijangkau pada penelitian tersebut. Pembatasan tersebut dilakukan karena adanya keterbatasan baik waktu, tenaga, dan biaya dan juga didasari atas pertimbangan efisiensi, efektivitas, dan validitas hasil penelitian. Supaya populasi tersebut bisa dijangkau, maka ciri populasi perlu ditambahkan dengan ciri administrasi seperti tempat dan waktu. Misalnya pada penelitian trial klinik tahap III vaksin XX untuk COVID-19, hasilnya akan diterapkan pada semua penduduk yang sehat sebagai populasi target. Karena adanya keterbatasan pada trial tersebut, trial hanya dilakukan pada penduduk dewasa usia 15-50 tahun yang sehat dan berdomisili di Kota Bandung pada tahun 2020.

1.3 Sampel

Sampel adalah bagian dari suatu populasi. Sampel, pada hakekatnya merupakan representasi dari populasi target yang benar-benar diteliti yang menjadi sumber data penelitian. Agar sampel benar-benar merupakan representasi dari populasi target, maka perlu ditetapkan kriteria eligibilitas seseorang yang bisa dipilih sebagai sampel. Sebelum melakukan pemilihan sampel tersebut maka harus terlebih dahulu ditetapkan unit sampelnya. **Unit sampel** adalah **unit** yang dijadikan dasar penarikan **sampel** baik berupa individu maupun kumpulan individu (klaster). Unit sampel bisa sama dengan unit analisis atau merupakan kumpulan dari unit analisis. Sebagai contoh, rumah tangga dapat dijadikan sebagai

unit sampel dan atau kumpulan rumah tangga pada wilayah tertentu dapat dijadikan **unit sampel** seperti blok sensus, RT/RW, bahkan desa.

Banyak penelitian di bidang kedokteran seperti penelitian trial, kohort, case-control atau uji diagnostik mempunyai unit analisis dan unit sampel yang sama, yaitu pasien. Sebaliknya, penelitian survei, umumnya memiliki lebih dari satu unit analisis, maka unit sampelnya bisa berbeda dengan unit analisisnya. Misalnya pada Survei Kesehatan Rumah Tangga (SKRT), unit sampelnya adalah blok sensus, sedangkan unit analisisnya adalah bayi, balita, ibu hamil, lansia, rumah tangga (RT), desa, kecamatan, kabupaten/kota.

Kriteria Eligibilitas:

Terdapat beberapa pertimbangan di dalam pemilihan sampel, yaitu 1) representativitas sampel, 2) validitas data, 3) keamanan sampel, dan 4) validitas penelitian. Dari ke empat pertimbangan tersebut, selanjutnya ditetapkan kriteria eligibilitas yang umumnya dibedakan menjadi kriteria inklusi, kriteria eksklusi, dan kriteria *drop out*.

1) Kriteria inklusi:

Kriteria inklusi merupakan batasan yang dipakai menentukan agar sampel merupakan representasi populasi. Oleh karena itu kriteria inklusi harus identik dengan kriteria populasi terjangkau. Selain itu, hak individu harus juga diakomodasikan dalam pemilihan sampel penelitian. Oleh Karena itu, kesediaan dengan sukarela di dalam penelitian agar dijadikan kriteria inklusi dengan menanda tangani formulir *informed consent* apabila merupakan prasyarat untuk mendapatkan *ethical clearance*.

Contoh Kriteria Inklusi:

Kriteria inklusi:

1. Penduduk usia 15-50 tahun yang berdomisili di Kota Bandung pada tahun 2020.
2. Bersedia terlibat dalam penelitian sampai akhir penelitian dengan menandatangani Formulir *Informed consent*.

2) Kriteria eksklusi:

Kriteria eksklusi merupakan batasan yang ditetapkan untuk menyaring atau mengeluarkan individu yang sudah memenuhi kriteria inklusi dari sampel dengan tujuan menjaga validitas data, menjamin keselamatan calon sampel dan mengendalikan variabel perancu yang kuat.

Kriteria eksklusi yang berkaitan dengan validitas data:

Kriteria eksklusi dapat ditetapkan untuk menjamin kualitas dan validitas data. Sebagai contoh, bila data dikumpulkan dengan wawancara, maka sampel (atau responden) harus mampu memberikan respon yang adekuat selama proses wawancara. Oleh karena itu, individu yang tidak bisa memberikan respon dengan baik misalnya memiliki tingkat kesadaran menurun sampai koma, hilang ingatan atau mengalami kelainan jiwa berat tidak mungkin diwawancara karena data yang diperoleh dari mereka kemungkinan tidak valid. Kriteria-kriteria seperti ini dipakai untuk mengeluarkan individu yang telah lolos kriteria inklusi sebelumnya untuk meningkatkan validitas data.

Kriteria eksklusi yang berkaitan dengan keselamatan sampel (*sample safety*):

Dasar pertimbangan lain untuk mengeluarkan individu yang sudah memenuhi kriteria inklusi adalah untuk menjamin keselamatan dan memenuhi kaidah etika penelitian. Sebagai contoh, pada penelitian uji coba suatu obat, individu yang memiliki kontra indikasi atau riwayat alergi terhadap obat yang akan diujicobakan, harus dikeluarkan dari sampel demi keselamatan mereka. Misalnya pada trial klinik tahap III vaksin COVID-19, penduduk yang demam dikeluarkan, anak-anak dan orang dengan tekanan darah $> 180/110$ mmHg tidak diikutsertakan sebagai sampel karena pertimbangan faktor keamanan sampel dari risiko yang mungkin terjadi.

Kriteria eksklusi yang berkaitan dengan pengendalian perancu:

Kriteria eksklusi juga bisa dipakai untuk mengendalikan satu atau beberapa variabel perancu yang kuat untuk meningkatkan validitas dari penelitian. Individu yang telah memenuhi kriteria inklusi dikeluarkan berdasarkan keberadaan variabel perancu ini. Penetapan kriteria eksklusi untuk mengendalikan variabel perancu ini merupakan salah satu upaya pengendalian perancu pada tahap merancang penelitian (*control by design*). Sebagai contoh, pada penelitian tentang kejadian infeksi saluran reproduksi pada wanita usia subur (WUS), WUS yang menggunakan alat kontrasepsi dalam Rahim (AKDR) dikeluarkan karena AKDR merupakan salah satu faktor kuat yang sudah terbukti dari literatur memengaruhi terjadinya infeksi saluran reproduksi.

Berikut adalah rangkuman beberapa contoh kriteria eksklusi dan tujuannya:

1. Penduduk dengan kesadaran menurun → validitas data
2. Penduduk yang mengalami gangguan kejiwaan berat → validitas data
3. Penduduk yang tidak sehat atau mempunyai penyakit penyerta → *patient safety*
4. Pasien yang memiliki riwayat alergi obat → *patient safety*
5. Usia di bawah 15 atau usia di atas 50 tahun → *patient safety*
6. WUS yang menggunakan AKDR → kontrol perancu

3) Kriteria *drop out*:

Kriteria *drop out* merupakan kriteria yang ditetapkan untuk mengoptimalkan kelengkapan data baik selama masa pengamatan maupun saat analisis. Kriteria *drop out* yang ditetapkan pada masa pengamatan umumnya ditentukan pada penelitian longitudinal (studi kohort atau trial klinik) Tujuan dari penetapan kriteria *drop out* (DO) adalah untuk mencegah adanya *missing data* yang dapat memengaruhi hasil analisis. Kriteria ini juga dikaitkan dengan kelengkapan data, misalnya pasien yang *loss to follow up* apapun penyebabnya seperti misalnya sampel menolak untuk dilibatkan lagi, sampel tidak ditemukan, begitu pula sampel yang datanya tidak lengkap bisa dikeluarkan pada tahap analisis.

Contoh Kriteria *drop out*:

1. Sampel tidak ditemukan (*loss to follow up*)
2. Sampel menolak melanjutkan penelitian
3. Sampel dengan data yang tidak lengkap (hasil lab, rontgen, dll tidak ada)

NOTE:

Agar dicermati apakah dengan penerapan kriteria tersebut sampelnya masih memiliki ciri seperti populasinya atau malah sangat berbeda, terutama kriteria yang ditujukan untuk pengendalian perancu. Bila ya, sebaiknya pengendalian perancu dilakukan *by analysis* dengan metode *multivariate*.

1.4 Representativitas sampel

Representativitas atau keterwakilan dapat digambarkan dari tingkat kesamaan ciri-ciri sampel terhadap populasi. Bila dilihat dari hasil pengukuran, representativitas digambarkan dari ketepatan atau tingkat presisi hasil pengukuran sampel dengan parameter populasi yang diwakili. Hasil yang didapatkan dari sampel akan dipakai menggambarkan kondisi populasi atau akan digeneralisasikan ke populasi, oleh karena itu representativitas menjadi indikator penting untuk menentukan sampel yang baik.

Perbedaan atau bias antara statistik sampel dengan parameter populasi niscaya tidak bisa dihindari pada penelitian sampel, tetapi diusahakan agar bias tersebut sekecil mungkin. Adanya perbedaan antara hasil sampel dengan keadaan di populasi tersebut disebabkan oleh dua faktor yaitu *sampling error* dan *systematic error*. Bias yang disebabkan oleh *sampling error* terjadi karena: 1) risiko atau probabilitas kejadian yang diteliti pada setiap individu di populasi bervariasi, 2) cara pemilihan sampel yang tidak memberikan kesempatan yang sama untuk terpilih bagi semua individu, dan 3) jumlah sampel yang dipakai kurang mencukupi. Sedangkan, *systematic error* berkaitan dengan instrumen, cara pengukuran, cara pengumpulan data, dan cara analisis data. Salah satu faktor yang memengaruhi besarnya *sampling error* adalah jumlah dan metode pemilihan sampel.

1.5 Metode pemilihan sampel

Metode pemilihan sampel dibedakan menjadi dua kelompok, yaitu metode pemilihan sampel secara acak (*random sampling*) dan metode pemilihan sampel secara tidak acak (*non-random sampling*).

Random Sampling

Metode pemilihan sampel secara acak (*random*) adalah pemilihan sampel dimana setiap individu dari populasi memiliki kesempatan yang sama terpilih sebagai sampel. Salah satu prasyarat penting untuk pemilihan sampel secara acak adalah adanya daftar unit sampling yang akan menjadi kerangka sampel (*sampling frame*) dalam pemilihan sampel tersebut. Metode *random sampling* dibedakan menjadi: 1) *simple random sampling* (acak sederhana), 2) *systematic random sampling* (acak sistematis), 3) *stratified random sampling* (acak berlapis), 4) *cluster sampling* (acak kelompok), dan 5) *multiple stages sampling* (acak berjenjang).

Non-Random Sampling

Metode pemilihan sampel tidak acak (*non random*) adalah metode pemilihan sampel dimana tidak setiap individu mempunyai kesempatan yang sama terpilih sebagai sampel. Ada beberapa pertimbangan penggunaan metode *non random* yaitu: tidak ada atau tidak mungkin membuat kerangka sampel, misalnya populasi penelitiannya adalah kelompok *under grown* seperti pengguna narkoba, dan lainnya. Pertimbangan lain adalah tidak semua individu di satu populasi relevan dijadikan sampel dikaitkan dengan informasi yang ingin dikumpulkan. Misalnya ingin meneliti sikap ibu terhadap keberadaan ruang laktasi di tempat kerja. Tidak semua ibu tepat untuk dijadikan sampel, sehingga sampelnya ditentukan berdasarkan pertimbangan peneliti. Metode *non random* yang sering dipakai, seperti: 1) *purposive sampling*, 2) *incidental sampling*, 3) *convenience sampling*, 4) *consecutive sampling*, dan 5) *snowball sampling*. Secara ringkas, metode *random* dan *non-random sampling* disajikan pada matrik di bawah (Table 1.1) ini.

Tabel 1.1. Metode *random sampling* dan *non random sampling*

<i>Random Sampling</i>	<i>Non Random Sampling</i>
1. <i>Simple Random Sampling</i>	1. <i>Purposive Sampling</i>
2. <i>Systematic Random Sampling</i>	2. <i>Incidental Sampling</i>
3. <i>Stratified Random Sampling</i>	3. <i>Convenience Sampling</i>
4. <i>Cluster Random Sampling</i>	4. <i>Consecutive Sampling</i>
5. <i>Multistage Random Sampling</i>	5. <i>Snowball Sampling</i>

1.6 Diskusi Kelompok

1. Apakah batasan dan perbedaan ciri antara populasi target dengan populasi terjangkau? Sebutkan dengan contoh!
2. Apakah pengertian unit sampel dan unit analisis serta kapan unit sampel dikatakan berbeda dengan unit analisis dan kapan bisa dikatakan sama? Sebutkan dengan contoh!
3. Apakah kelebihan dan kekurangan dari penggunaan sampel?
4. Sebutkan jenis, indikasi, dan prinsip metode *random* dan *non-random sampling*!

1.7 Praktikum

Kasus 1.

Survei rumah tangga di Bali pada tahun 2017 dilakukan pada 1250 RT dengan tujuan untuk mengetahui prevalensi anemia pada wanita usia subur (usia 15-45 th) atau disebut WUS.

Jelaskan:

- Siapa menjadi populasi target
- Siapa populasi terjangkau
- Siapa unit sampel
- Siapa unit analisis
- Apa metode sampling yang dipakai (*random* atau *non-random*)

Kasus 2.

Yayasan Hipokrates meneliti perilaku penggunaan jarum suntik dan perilaku seksual beresiko pada pemakai obat narkotika suntik di Bali.

Jelaskan:

- Siapa menjadi populasi target
- Siapa populasi terjangkau
- Siapa unit sampel
- Siapa unit analisis
- Apa metode sampling yang dipakai (*random* atau *non-random*)

BAGIAN 2. METODE SAMPLING PENELITIAN SURVEI

dr. I Ketut Tangking Widarsa, MPH.*

dr. Ayu Swandewi, MPH, Ph.D.**

*) *Dosen Biostatistik Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com*

***) *Dosen Biostatistik Fakultas Kedokteran Universitas Udayana. E-mail: ayuswandewi@unud.ac.id*

Tujuan Pembelajaran

- Memahami pengertian survei.
- Memahami berbagai metode pemilihan sampel survei mencakup indikasi, kelebihan dan kekurangan, dan langkah-langkah pemilihan.
- Memahami cara penentuan besar sampel survei menurut jenis metode sampling yang dipakai.

Pokok Bahasan:

- 2.1 Pengantar
- 2.2 Metode *Random*
- 2.3 *Simple Random Sampling* (Acak Sederhana)
- 2.4 *Systematic Random Sampling* (Acak Sistematis)
- 2.5 *Stratified Random Sampling* (Acak Berlapis)
- 2.6 *Cluster Sampling* (Acak Kelompok)
- 2.7 *Multistage Sampling* (Acak Bertingkat)
- 2.8 Metode *Non Random*
- 2.9 Survei Cepat (*Rapid Survei*)

Bahan Bacaan:

1. Cochran W.G. 1977. *Sampling Techniques*, Third edition, John Wiley&Sons, New York.
2. Levy P.S. and Lemeshow S. 1980. *Sampling for Health Professionals*, Lifetime Learning Publication, California.

3. Lwanga S.K. and Lemeshow S. 1991. *Sample Size Determination in Health Studies*, World Health Organization, Geneva.
4. Sopiudin Dahlan M. 2009. *Besar Sampel dan Cara Pengambilan Sampel Dalam Penelitian Kedokteran dan Kesehatan*, Edisi 2, Salemba Medika, Jakarta.

2.1 Pengantar

Penelitian kedokteran dan kesehatan dibedakan menjadi penelitian eksperimental dan observasional berdasarkan ada atau tidaknya perlakuan. Penelitian eksperimental adalah penelitian yang ditandai oleh adanya perlakuan, randomisasi, replikasi, dan adanya kelompok kontrol. Sebaliknya, pada penelitian observasional tidak ada perlakuan yang diteliti dan peneliti hanya mengobservasi fenomena yang terjadi pada subyek penelitian. Penelitian observasional dibedakan lagi menjadi penelitian analitik dan deskriptif berdasarkan kepada ada atau tidak adanya kontrol.

Survei adalah penelitian observasional yang tidak ada kelompok kontrolnya. Survei dipakai untuk meneliti fenomena yang terjadi di populasi. Survei dapat dipakai mempelajari banyak fenomena pada lebih dari satu kelompok populasi. Misalnya pada survei kesehatan rumah tangga (SKRT) mempunyai banyak tujuan dan banyak unit analisis seperti rumah tangga, anak umur dibawah lima tahun (balita), ibu hamil (bumil), kelompok lanjut usia (lansia), dan lainnya. Survei pada umumnya dilakukan pada sampel sehingga sering disebut sebagai *sample survey*. Pada bagian ini akan dibahas tentang metode sampling untuk survei dan cara penentuan besar sampel untuk setiap metode sampling yang dipakai. Secara garis besar metode sampling pada survei dibedakan menjadi metode *random* dan *non random*.

2.2 Metode Random

Pengertian:

Metode pemilihan sampel secara *random* adalah cara pemilihan sampel dimana pada prinsipnya setiap individu yang menjadi populasi penelitian mempunyai kesempatan atau probabilitas yang sama untuk terpilih sebagai sampel.

Persyaratan:

Metode pemilihan sampel secara *random* bisa diterapkan apabila kerangka sampel yang berisikan daftar nama dari semua individu yang menjadi populasi terjangkau ada. Daftar semua individu dari populasi terjangkau akan menjadi kerangka sampel atau *sampling frame* pada pemilihan sampel secara acak. Setiap individu pada kerangka sampel tersebut diberi nomor urut yang unik dan selanjutnya sampel secara acak akan dipilih dari kerangka sampel tersebut. Misalnya pada survei cakupan imunisasi dasar anak usia di bawah 3 tahun (batita) di Kota Denpasar, maka semua batita yang berdomisili di Kota Denpasar pada saat survei dilakukan menjadi populasi terjangkau. Kerangka sampel survei tersebut dibuat berdasarkan data balita yang terdapat di register posyandu di semua wilayah di Kota Denpasar. Karena kerangka sampel bisa dibuat maka metode *random sampling* bisa diterapkan pada survei tersebut. Berbeda halnya dengan survei prevalensi penyakit jantung koroner pada perokok di Bali. Data perokok di Bali tidak tersedia, maka pada survei ini tidak bisa diterapkan metode pemilihan sampel secara *random*, karena tidak dimungkinkan peneliti membuat kerangka sampelnya. Metode pemilihan sampel secara acak (*random*) terdiri dari beberapa jenis.

Jenis:

Metode pemilihan sampel secara *random* dibagi menjadi *Single Stage Sampling* dan *multi-stage random sampling*. *Single Stage Sampling* terdiri dari beberapa metode sebagai berikut.

- 1) Acak sederhana (*simple random sampling*)
- 2) Acak sistematis (*systematic random sampling*)
- 3) Acak berlapis (*stratified random sampling*)
- 4) Acak kelompok (*cluster sampling*) yang dibedakan lagi berdasarkan cara pemilihan sampel klasternya menjadi acak kelompok secerhana (*simple cluster sampling*) dan *probability proportional to size (PPS) cluster sampling*.

Metode sampling secara bertingkat (*multi-stage random sampling*) terdiri dari:

- 5) Acak berjenjang (*multi-stage random sampling*)
- 6) Acak kelompok dua tingkat (*two stage cluster sampling*) dengan metode PPS.

2.3 Simple Random Sampling (Acak Sederhana)

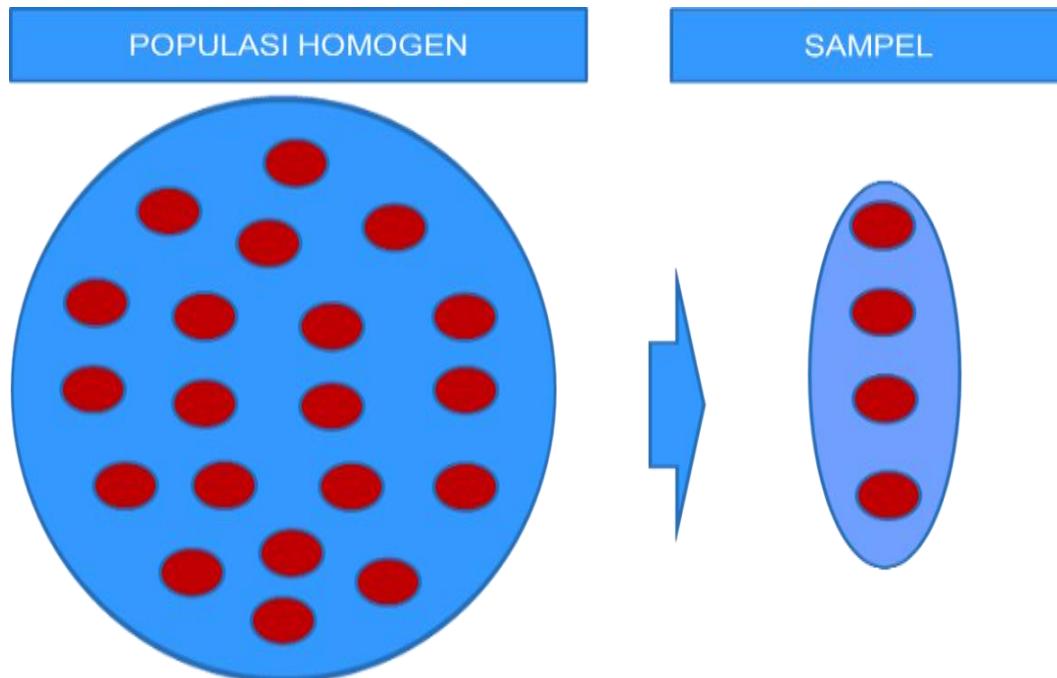
2.3.1 Pengertian

Metode acak sederhana adalah cara pemilihan sampel dimana setiap individu mempunyai probabilitas yang sama untuk terpilih sebagai sampel. Misalnya akan dipilih sampel sebanyak n individu dari N populasi secara acak, maka sampel pertama akan dipilih 1 orang secara acak dari N populasi tersebut. Pemilihan sampel kedua dan seterusnya juga dipilih 1 orang dari N populasi tersebut, sehingga probabilitas sampel ke 1 sampai ke n adalah sama. Pemilihan sampel secara acak sederhana bisa menggunakan tabel bilangan *random* atau menggunakan bantuan komputer. Mahasiswa bisa mengakses website yang menyediakan metode *random sampling* dan menggunakan piranti tersebut sebagai alat bantu dalam pemilihan sampel secara acak.

2.3.2 Indikasi

Metode ini bisa dipakai apabila:

- 1) Kerangka sampel yang berisikan daftar nama semua individu dari populasi terjangkau penelitian ada atau bisa dibuat.
- 2) Populasi penelitian homogen atau setiap individu di populasi penelitian memiliki risiko yang sama terhadap kejadian fenomena yang akan diteliti. Cara mengetahui apakah populasi penelitian homogen atau tidak dapat dilihat dari risiko kejadian yang diteliti pada semua populasi terjangkau. Apabila risikonya sama, maka populasi tersebut dinyatakan homogen, sebaliknya bila risikonya tidak sama maka populasi tersebut dinyatakan heterogen. Misalnya akan diteliti kejadian diare pada penduduk di Kota Denpasar dengan Rumah Tangga sebagai unit sampel. Pada penelitian ini diasumsikan bahwa setiap rumah tangga memiliki kejadian diare yang sama, maka unit sampel RT dianggap homogen. Sebaliknya, bila akan diteliti kejadian merokok pada mahasiswa, populasi mahasiswa diasumsikan tidak homogen karena risiko mahasiswa dan mahasiswi merokok tidak sama. Ilustrasi pemilihan sampel secara acak sederhana, dimana risiko kejadian yang diteliti sama pada semua anggota populasi seperti pada Gambar 2.1.



Gambar 2.1. Bagan pemilihan sampel secara acak sederhana (*simple random sampling*)

2.3.3 Kelebihan dan kekurangan

1) Kelebihan:

Kelebihan dari metode acak sederhana adalah hasil surveinya akan memiliki varians terkecil, lebih mudah dan lebih cepat pelaksanaannya bila dibandingkan dengan metode sampling lainnya.

2) Kelemahan:

Kelemahan utama dari metode ini adalah distribusi sampel yang dihasilkan bisa tidak merata di semua wilayah survei. Misalnya survei kejadian diare pada penduduk Kota Denpasar dilakukan dengan rumah tangga (RT) sebagai unit sampelnya dan sampel dipilih secara acak sederhana. Pada studi tersebut terdapat kemungkinan semua sampel hanya berasal dari satu kecamatan saja dan apabila hal tersebut terjadi, maka representativitas sampel akan diragukan.

2.3.4 Cara pemilihan sampel

Cara pemilihan sampel secara acak sederhana dapat dilakukan dengan cara undian, menggunakan tabel bilangan *random* atau menggunakan bantuan komputer atau aplikasi *online*. Berikut akan dijelaskan cara pemilihan sampel secara acak sederhana dengan menggunakan tabel bilangan *random* dan bantuan program Excel.

1) Menggunakan Tabel Bilangan *Random*

Langkah 1: Membuat kerangka sampel

Semua unit sampel didaftar dalam kerangka sampel dan setiap unit sampel harus mempunyai nomor urut masing-masing yang unik.

Misalnya jumlah populasi $N = 899$ dan jumlah sampel $n = 30$, maka kerangka sampel berisi unit sampel dari no 1 s/d 899. Dari kerangka sampel tersebut nanti dipilih sebanyak 30 sampel.

Langkah 2: Menentukan batasan bilangan *random*

Tabel bilangan *random* ada terdiri dari 1 digit, 2 digit, 4 digit atau 5 digit. Sebelum menggunakan tabel bilangan *random*, peneliti harus menetapkan bilangan *random* yang bisa dipilih (*eligible*) sebagai sampel, yaitu bilangan *random* yang besarnya dari 1 sampai jumlah populasi N . Jumlah digit dari bilangan *random* yang dibaca disesuaikan dengan jumlah digit populasi. Untuk $N = 899$, maka jumlah digit bilangan *random* yang dicoba adalah 3 digit dari depan.

Dari contoh di atas, bilangan *random* yang besarnya antara 001 s/d 899 bisa dipakai sampel dan yang di luar itu tidak bisa. Misalnya bilangan *random* yang didapatkan adalah 01245, tiga digit pertama bilangan *random* tersebut adalah 012, nomor ini memenuhi syarat dan sampelnya adalah individu nomor urut 12. Bila bilangan *random* yang didapat adalah 56990, tiga digit pertama adalah 569, maka sampelnya adalah individu nomor urut 569. Sebaliknya, bila bilangan *random* yang terpilih adalah 97860, maka tiga digit pertamanya adalah 978. Nomor tersebut tidak ada di kerangka sampel, maka nomor tersebut tidak bisa dipilih.

Langkah 3: Menentukan *starting point*

Pemilih bilangan *random* pada tabel didahului dengan penetapan *starting point*. Penentuan *starting point* pada tabel bilangan *random* dapat dilakukan dengan cara menjatuhkan pensil pada tabel bilangan *random* tersebut. Misalnya ujung pensil jatuh pada bilangan *random* 19184, maka *starting point* pemilihan sampel dimulai dari bilangan *random* 19184, seperti Gambar 2.2 di bawah. Sebagai sampel pertama adalah unit sampel nomor urut 191 (tiga digit pertama).

Langkah 4: Memilih sampel selanjutnya

Pemilihan sampel selanjutnya dilakukan dengan cara mencari bilangan *random* pada tabel dari *starting point* ke arah atas, ke samping kanan, ke bawah, dan ke samping kiri.

Bila mendapatkan nomor bilangan *random* yang *eligible* selanjutnya dipilih sebagai sampel. Cara ini diulang berkali-kali sampai jumlah sampel yang dibutuhkan tercukupi.

	00000 12345	00001 67890	11111 12345	11112 67890	22222 12345	22223 67890	33333 12345	33334 67890	44444 12345	44445 67890
01	85967	73152	14511	85285	36009	95892	36962	67835	63314	50162
02	07483	51453	11649	86348	76431	81594	95848	36738	25014	15460
03	96283	01898	61414	83525	04231	13604	75339	11730	85423	60698
04	49174	12074	98551	37895	93547	24769	09404	76548	05393	96770
05	97366	39941	21225	93629	19574	71565	33413	56087	40875	13351
06	90474	41469	16812	81542	81652	45554	27931	93994	22375	00953
07	28599	64109	09497	76235	41383	31555	12639	00619	22909	29563
08	25254	16210	89717	65997	82667	74624	36348	44018	64732	93589
09	28785	02760	24359	99410	77319	73408	58993	61098	04393	48245
10	84725	86576	86944	93296	10081	82454	76810	52975	10324	15457
11	41059	66456	47679	66810	15941	84602	14493	65515	19251	41642
12	67434	41045	82830	47617	36932	46728	71183	36345	41404	81110
13	72766	68816	37643	19959	57550	49620	98480	25640	67257	18671
14	92079	46784	66125	94932	64451	29275	57669	66658	30818	58353
15	29187	40350	62533	73603	34075	16451	42885	03448	37390	96328
16	74220	17612	65522	80607	19184	64164	66962	82310	18163	63495
17	03786	02407	06098	92917	40434	60602	82175	04470	78754	90775
18	75085	55558	15520	27038	25471	76107	90832	10819	56797	33751
19	09161	33015	19155	11715	00551	24909	31894	37774	37953	78837
20	75707	48992	64998	87080	39333	00767	45637	12538	67439	94914
21	21333	48660	31288	00086	79889	75532	28704	62844	92337	99695
22	65626	50061	42539	14812	48895	11196	34335	60492	70650	51108
23	84380	07389	87891	76255	89604	41372	10837	66992	93183	56920
24	46479	32072	80083	63868	70930	89654	05359	47196	12452	38234
25	59847	97197	55147	76639	76971	55928	36441	95141	42333	67483
26	31416	11231	27904	57383	31852	69137	96667	14315	01007	31929
27	82066	83436	67914	21465	99605	83114	97885	74440	99622	87912
28	01850	42782	39202	18582	46214	99228	79541	78298	75404	63648
29	32315	89276	89582	87138	16165	15984	21466	63830	30475	74729
30	59388	42703	55198	80380	67067	97155	34160	85019	03527	78140
31	58089	27632	50987	91373	07736	20436	96130	73483	85332	24384
32	61705	57285	30392	23660	75841	21931	04295	00875	09114	32101
33	18914	98982	60199	99275	41967	35208	30357	76772	92656	62318
34	11965	94089	34803	48941	69709	16784	44642	89761	66864	62803
35	85251	48111	80936	81781	93248	67877	16498	31924	51315	79921
36	66121	96986	84844	93873	46352	92183	51152	85878	30490	15974
37	53972	96642	24199	58080	35450	03482	66953	49521	63719	57615
38	14509	16594	78883	43222	23093	58645	60257	89250	63266	90858
39	37700	07688	65533	72126	23611	93993	01848	03910	38552	17472
40	85466	59392	72722	15473	73295	49759	56157	60477	83284	56367
41	52969	55863	42312	67842	05673	91878	82738	36563	79540	61935
42	42744	68315	17514	02878	97291	74851	42725	57894	81434	62041
43	26140	13336	67726	61876	29971	99294	96664	52817	90039	53211
44	95589	56319	14563	24071	06916	59555	18195	32280	79357	04224
45	39113	13217	59999	49952	83021	47709	53105	19295	88318	41626
46	41392	17622	18994	98283	07249	52289	24209	91139	30715	06604
47	54684	53645	79246	70183	87731	19185	08541	33519	07223	97413
48	89442	61001	36658	57444	95388	36682	38052	46719	09428	94012
49	36751	16778	54888	15357	68003	43564	90976	58904	40512	07725
50	98159	02564	21416	74944	53049	88749	02865	25772	89853	88714

Gambar 2.2. Tabel bilangan *random* dan ilustrasi pemilihan sampel

Contoh:

Starting point: 19184, sampel pertama adalah no. 191

Ke samping kanan didapatkan: 64164 (641), 66962 (669), 82310 (823), 18163 (181), 63495 (634)

Ke bawah didapatkan: 40434 (404), 25471 (254), 00551 (005), 39333 (393), 79689 (796), 48895 (488) dan seterusnya.

Ke samping kiri didapatkan: 80607 (806), 65522 (655), 17612 (176), 74220 (742)

Ke atas didapatkan: 34015 (340), 64451 (644), 57550 (575), 36932 (369), dan seterusnya.

Bila sampel yang terpilih belum cukup, pindah *starting point* dan selanjutnya dilakukan hal yang sama seperti di atas tadi, sampai jumlah sampel yang dibutuhkan tercukupi.

Langkah 5: Membuat *list* sampel terpilih.

List nomor identitas (ID) sampel yang terpilih dibuat kemudian sampel terpilih diurut dari nomor ID terkecil sampai terbesar.

2) Bantuan komputer

Sudah banyak tersedia *computer tools* untuk pemilihan sampel, misalnya Excel, dan lainnya. Berikut adalah langkah pemilihan sampel acak sederhana dengan program Excel.

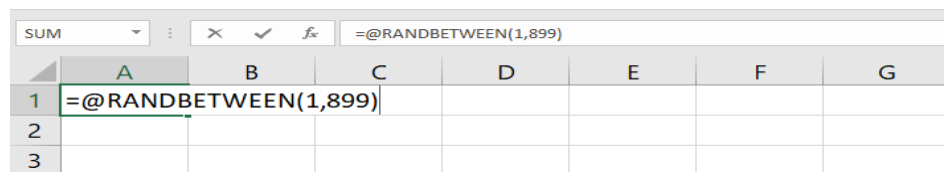
Langkah 1: Buka program Excel

Langkah 2: Menggunakan *command* `@RANDBETWEEN()`

Ketik pada kolom A, baris 1 sbb: “=RANDBETWEEN (1,899)” Enter, maka muncul bilangan *random* antara 1 sampai 899, misalnya muncul bilangan *random* 380.

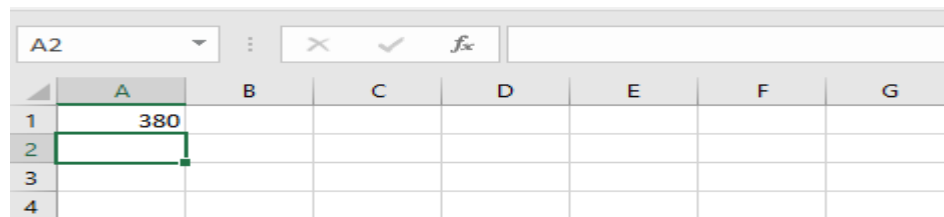
Range Between (1,899) adalah rentang nomor urut unit sampel terkecil sampai terbesar, yaitu yang terkecil adalah nomor 1 dan yang terbesar adalah nomor 899.

Contoh:



	A	B	C	D	E	F	G
1	=@RANDBETWEEN(1,899)						
2							
3							

Tampak hasil sbb:



	A	B	C	D	E	F	G
1	380						
2							
3							
4							

Langkah 3: *Copy* rumus

Copy rumus pada kolom A baris 1, ke bawah sampai baris 3 dan ke kanan sampai Kolom J (30 sel) yaitu banyak sampel yang dibutuhkan. Nomor sampel yang terpilih adalah sebagai berikut.

	A	B	C	D	E	F	G	H	I	J
1	320	559	170	425	310	205	429	316	468	671
2	292	137	151	127	28	502	141	902	9	729
3	305	404	479	716	138	235	849	225	702	419

Setelah diurut, sampel menjadi sebagai berikut.

	A	B	C	D	E	F	G	H	I	J
1	9	28	127	137	138	141	151	170	205	225
2	235	292	305	310	316	320	404	419	425	429
3	468	479	502	559	671	702	716	729	849	902

2.3.5 Cara penentuan besar sampel

Sampel survei pada umumnya bertujuan untuk memperkirakan kejadian di populasi dengan hasil yang diperoleh dari sampel. Kejadian fenomena di populasi disebut parameter yang diberi simbol μ untuk means atau π proporsi. Sedangkan kejadian yang didapat dari sampel disebut statistik dengan simbol \bar{x} untuk rerata sampel dan p untuk prevalensi. Pada kenyataannya sering dijumpai bahwa parameter populasi tidak sama dengan statistik sampel ($\mu \neq \bar{x}$) atau ($\pi \neq p$). Perbedaan tersebut merupakan bias atau deviasi (d) antara kejadian di populasi dengan hasil sampel. Bias atau deviasi tersebut terjadi sebagai akibat dari *sampling error*. Lebarnya deviasi tersebut ditentukan oleh besar sampel yang dipakai. Oleh karena itu, presisi atau lebar deviasi yang diinginkan akan menentukan berapa besar sampel yang dibutuhkan. Semakin tinggi presisi atau semakin kecil deviasi (d) yang diinginkan, maka semakin besar sampel yang diperlukan.

Seperti diuraikan di atas, statistik sampel akan dipakai memprediksi parameter populasi. Tingkat kepercayaan atau reliabilitas hasil prediksi ditentukan dari level kepercayaan yang dipakai. Apabila level kepercayaan yang dipakai 95%, maka ada kemungkinan terjadi kesalahan estimasi sebesar 5% dan bila level kepercayaan yang dipakai 99%, maka ada kemungkinan kesalahan estimasi sebesar 1%. Selain tingkat presisi, besar sampel survei juga ditentukan oleh level kepercayaan dari sampel yang diinginkan. Makin tinggi level kepercayaan sampel yang diinginkan, makin besar jumlah sampel yang dibutuhkan. Dalam penghitungan besar sampel, level kepercayaan ditentukan dengan nilai *Z score* untuk level kepercayaan yang dipakai yang diberi simbol Z_{α} . Apabila level kepercayaan sebesar 95%, maka nilai $Z_{\alpha} = 1,96$ dan bila level kepercayaan 99% maka nilai $Z_{\alpha} = 2,575$.

Variasi kejadian di populasi juga berpengaruh terhadap jumlah sampel yang dibutuhkan. Apabila kejadian di populasi memiliki rentang nilai kejadian yang lebar, maka sampel yang dibutuhkan akan

lebih banyak dibandingkan bila rentang nilai kejadian di populasi yang sempit. Ukuran yang dipakai menggambarkan variasi data adalah standar deviasi (SD) untuk data numerikal dan prevalensi (p) untuk data nominal.

Berdasarkan uraian di atas, maka besar sampel survei ditentukan dari nilai parameter berikut:

- 1) Variasi data di populasi yang ditentukan dari perkiraan kejadian (P) atau variasi (SD) dari peristiwa yang akan diteliti
- 2) Presisi atau besarnya *margin of error* atau lebar deviasi dari sampel yang diinginkan (D)
- 3) Level kepercayaan atau tingkat reliabilitas yang diinginkan (Z_α)
- 4) Besar sampel dihitung dengan rumus di bawah ini.

Rumus Besar Sampel Survei dengan metode <i>Simple Random</i>	
Data Numerikal	Data Kategorikal
$n = \left[\frac{Z_\alpha SD}{D} \right]^2$	$n = \frac{Z_\alpha^2 P(1-P)}{D^2}$
<p>Keterangan:</p> <ul style="list-style-type: none"> • n = jumlah sampel minimal • Z_α = nilai <i>Z score</i> untuk kepercayaan (1-α) yang besarnya 1,96 untuk kepercayaan 95% dan 2,57 untuk kepercayaan 99% • SD = standar deviasi • D = presisi atau <i>margin of error</i> 	<p>Keterangan:</p> <ul style="list-style-type: none"> • n = jumlah sampel minimal • Z_α = nilai <i>Z score</i> untuk kepercayaan (1-α) yang besarnya 1,96 untuk kepercayaan 95% dan 2,57 untuk kepercayaan 99% • P = estimasi kejadian • D = presisi atau <i>margin of error</i>
<p>Rumus jumlah sampel dengan koreksi, apa bila jumlah Populasi (N) < 10.000</p> $n_k = \frac{n}{1 + n/N}$ <p>Keterangan: n_k = jumlah sampel koreksi n = jumlah sampel sebelum dikoreksi N = jumlah populasi</p>	

1) Penentuan Besar Sampel secara Manual

Contoh kasus 1: variabel yang disurvei berskala numerik

Survei rumah tangga (RT) dengan tujuan untuk mengetahui pengeluaran rumah tangga untuk kesehatan di Desa Bunga dengan jumlah populasi rumah tangga sebanyak 1500 RT.

Besar sampel yang dibutuhkan berdasarkan perkiraan pengeluaran kesehatan rumah tangga 2 juta rupiah dengan SD sebesar 0,75 juta rupiah, *margin of error* (D) sebesar 5% dari 2 juta atau 0,1 juta, tingkat kepercayaan sampel 95% ($Z_{\alpha} = 1,96$), bisa dihitung seperti dibawah ini

Jumlah sampel minimal tanpa koreksi dihitung sebagai berikut.

$$n = \left[\frac{Z_{\alpha} SD}{D} \right]^2 = \left[\frac{1,96 \times 0,75}{0,1} \right]^2 = 216,09$$

Jumlah sampel minimal tanpa koreksi adalah 216 RT. Karena jumlah populasi RT < 10.000, maka jumlah sampel perlu dikoreksi, sehingga jumlah sampel setelah dikorekasi menjadi seperti di bawah ini.

$$n_k = \frac{n}{1 + n/N} = \frac{216,09}{(1 + 216,09/1500)} = 188,87 \text{ dibulatkan menjadi } 189 \text{ RT}$$

Contoh kasus 2 variabel yang disurvei berskala nominal

Survei rumah tangga dilakukan dengan tujuan untuk mengetahui proporsi RT yang memiliki akses air bersih di Desa Bunga dengan jumlah populasi RT sebanyak 1500. Besar sampel minimal yang dibutuhkan apabila perkiraan RT yang punya akses air bersih $P = 65\%$, *margin of error* (D) = 5%, dan tingkat kepercayaan sampel 95%, bisa ditentukan seperti di bawah ini.

Jumlah sampel minimal tanpa koreksi adalah sebagai berikut:

$$n = \frac{Z_{\alpha}^2 P(1-P)}{D^2} = \frac{1,96^2 (0,65 \times 0,35)}{0,05^2} = 349,58$$

Jumlah sampel setelah koreksi menjadi seperti di bawah ini:


$$n_k = \frac{n}{1 + n/N} = \frac{349,58}{(1 + 349,58/1500)} = 238,50 \text{ dibulatkan menjadi } 239 \text{ RT}$$

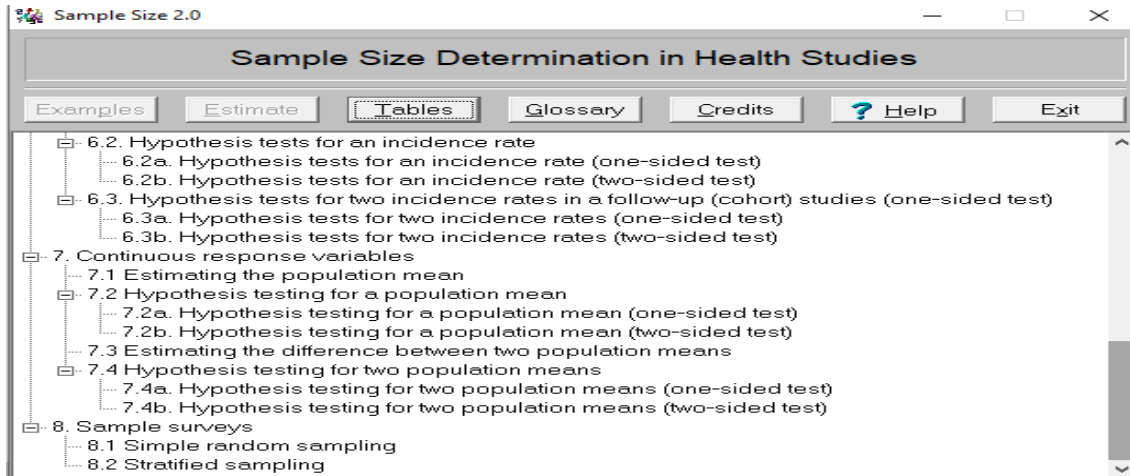
2) Penghitungan besar sampel menggunakan program *Sample Size Determination in Health Studies Versi 2.0 (WHO)*

Kasus:

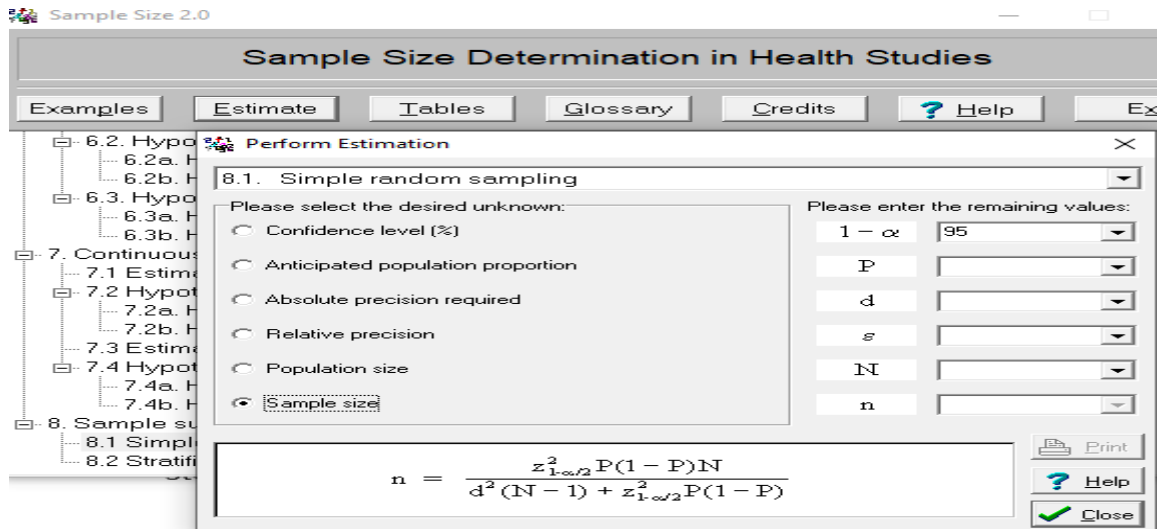
Survei angka akseptor di Kabupaten Badung dengan populasi survei adalah Pasangan Usia Subur (PUS) dengan jumlah $N = 15.000$. Berapa besar sampel minimal, apabila asumsi akseptor rate $P = 65\%$, *margin of error* survei adalah $D = 0,05$, level kepercayaan survei adalah 95%, dan sampel dipilih secara acak sederhana (*simple random sampling*).

Langkah penghitungan besar sampel sbb:

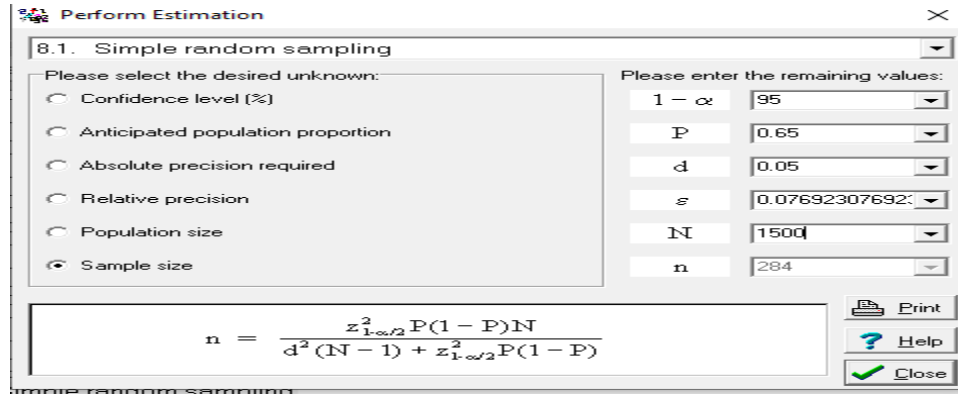
Step 1. Aktifkan program dengan *double click* “ Sample Size 2.0”, maka akan tampak gambar sbb:



Step 2. Klik “8.1 *Simple random sampling* dan *Estimate*, maka akan tampak sbb:



Step 3: masukan data % *Confidence Level* (95), perkiraan kejadian dipopulasi (P) = 0.65, masukan data *margin of error* (d) = 0.05, dan jumlah populasi N = 1500, maka akan tampak jumlah sampel minimal yang dibutuhkan, sbb:



Jumlah sampel minimal untuk survei tersebut $n = 284$.

2.3.6 Estimasi Parameter Populasi

Nilai statistik sampel seperti mean (\bar{x}) dan proporsi (p) dapat dipakai sebagai estimator dari parameter populasi seperti rerata populasi (μ) dan prevalensi di populasi (π), standar deviasi sampel (SD) juga bisa dipakai sebagai estimator standar deviasi populasi (σ), dimana rerata dan standar deviasi populasi dan sampel dihitung dengan rumus berikut :

Data	Penghitungan parameter populasi	Penghitungan statistik sampel	Penghitungan nilai <i>Confidence Interval</i>
Berskala Interval	$\mu = \frac{\sum_{i=1}^N x_i}{N}$ $\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N - 1}}$ <p>Keterangan: μ = rerata populasi σ = standar deviasi x_i = data orang ke i N = jumlah populasi</p>	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ $sd = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$ <p>Keterangan: \bar{x} = rerata sampel sd = standar deviasi x_i = data sampel ke i n = jumlah sampel</p>	$CI_{1-\alpha} = \bar{x} \pm Z_{\alpha} sd / \sqrt{n}$ <p>Keterangan: $CI_{1-\alpha}$ = confidence interval \bar{x} = rerata sampel sd = standar deviasi Z_{α} = Z skor untuk α tertentu n = jumlah sampel</p>
Data berskala Nominal	$\pi = \frac{m}{N}$ $\sigma = \sqrt{\frac{\pi(1-\pi)}{N-1}}$ <p>Keterangan: π = angka kejadian σ = standar deviasi m = jumlah kejadian N = jumlah populasi</p>	$p = \frac{m}{n}$ $sd = \sqrt{\frac{p(1-p)}{n-1}}$ <p>Keterangan: p = kejadian di sampel sd = standar deviasi m = jumlah kejadian n = jumlah sampel</p>	$CI_{1-\alpha} = p \pm Z_{\alpha} \sqrt{pq/n}$ <p>Keterangan: $CI_{1-\alpha}$ = confidence interval p = proporsi sampel Z_{α} = Z skor untuk α tertentu n = jumlah sampel</p>

Estimasi parameter populasi berdasarkan statistik sampel dapat dilakukan dengan *point estimate* yaitu dengan rerata sampel dan juga bisa dengan *interval estimate*. Kemungkinan bias hasil estimasi dengan *point estimate* jauh lebih besar dibandingkan dengan menggunakan *interval estimate*. Rentang nilai estimasi parameter populasi dengan interval pada level kepercayaan 95% atau 99%. Rentang nilai tersebut disebut nilai interval kepercayaan (IK) atau *confidence interval* (CI). CI 95% untuk level kepercayaan 95% dan CI 99% untuk level kepercayaan 99%.

Contoh:

Misalnya dari 30 mahasiswa kedokteran diketahui IP kumulatifnya sebagai berikut:

2.5	2.8	3.0	3.0	3.2	3.5	3.6	3.7	3.9	4.0
2.7	3.0	3.0	3.0	3.1	3.1	3.3	3.4	3.5	3.5
2.8	3.0	3.0	3.1	3.2	3.2	3.5	3.6	3.7	4.0

Dari data populasi tersebut diketahui bahwa *mean* IPK (μ) = 3,263 dengan standar deviasi (σ) = 0,383

Dari populasi tersebut dipilih 10 sampel secara acak sederhana, dan didapatkan data sbb:

No.	22	17	14	1	12	23	20	10	2	30
IPK	3.0	3.3	3.0	2.5	3.0	3.0	3.5	4.0	2.8	4.0

Dari data sampel didapatkan nilai rata-rata (\bar{x}) = 3,21 dengan standar deviasi (SD) = 0,49

Estimasi IPK di populasi dengan kepercayaan 95% adalah:

$$\text{Batas bawah: } CI_{1-\alpha} = \bar{x} - Z_{\alpha}sd / \sqrt{n} = 3,21 - 1,96 \times 0,49 / \sqrt{10} = 2,904$$

$$\text{Batas atas : } CI_{1-\alpha} = \bar{x} + Z_{\alpha}sd / \sqrt{n} = 3,21 + 1,96 \times 0,49 / \sqrt{10} = 3,516$$

Estimasi rerata IPK populasi berdasarkan rerata sampel dengan kepercayaan 95% adalah antara 2,904 – 3,516 dan rerata IPK populasi adalah 3,263 berada di dalam CI. Hal ini membuktikan bahwa rerata sampel yang dipilih dengan metode *simple random* tidak bias.

2.3.6 Latihan

Kasus

Survei Rumah Tangga di Desa Bunga dilakukan dengan tujuan untuk mengetahui proporsi RT sehat berdasarkan 12 indikator GERMAS. RT dikategorikan sehat apabila minimal 80% indikator GERMAS terpenuhi. Terdapat 1500 RT di desa tersebut. Sampel akan dipilih dengan metode *simple random*

Jelaskan:

- 1) Siapa populasi terjangkau dari survei tersebut?
- 2) Kenapa sampel dipilih dengan *simple random*, berikan penjelasannya!
- 3) Apa kelebihan dan kekurangan dari metode *simple random* tersebut?
- 4) Berapa sampel RT minimal yang dibutuhkan apabila diperkirakan proporsi RT sehat $P = 45\%$, *sampling error* $D = 5\%$, kepercayaan sampel 95%, dan total populasi $N = 1500$ RT?
- 5) Demonstrasikan cara pemilihan sampel menggunakan Tabel Bilangan *Random*
- 6) Demonstrasikan cara pemilihan sampel dengan bantuan program Excel.

2.4 Systematic Random Sampling (Acak Sistematis)

2.4.1 Pengertian

Systematic random sampling adalah metode pemilihan sampel dimana hanya sampel pertama yang dipilih secara *random*, sedangkan sampel berikutnya dipilih dengan kelipatan interval tertentu. Oleh karena itu, metode *systematic random sampling* juga disebut metode sampling yang menggabungkan metode *random* dan *non random*.

2.4.2 Asumsi

Metode *systematic random sampling* memiliki asumsi yang sama dengan metode *simple random sampling* yaitu populasi survei homogen untuk aspek yang diteliti dan memungkinkan dibuatnya kerangka sampel.

2.4.3 Kelebihan dan kekurangan

Kelebihan metode ini adalah pelaksanaannya lebih mudah dan sampel terpilih akan terdistribusi secara merata. Sebaliknya, kelemahan dari metode ini adalah ada kemungkinan sampel hanya terdistribusi di satu area tertentu saja terutama apabila distribusi populasi seperti sistem *battery* dimana rumah tertata menjadi blok dan setiap blok terdapat jumlah rumah yang sama seperti pada perkotaan, perumahan atau kompleks tertentu.

2.4.4 Cara pemilihan sampel

Systematic random sampling mempunyai dua metode pemilihan sampel yaitu metode linier dan metode sirkuler. Perbedaan kedua metode tersebut terjadi pada cara pemilihan sampel pertama.

Pemilihan sampel secara linier

Misalnya survei akan dilakukan pada $N = 8500$ RT dan sampel minimal yang dibutuhkan $n = 30$, maka langkah pemilihan sampel secara linier adalah sebagai berikut.

Step 1: Menghitung Interval sampel (I) dengan rumus $I = N/n$

Apabila $N = 8500$ dan $n = 30$, maka $I = N/n = 8500/30 = 283,33$ (jangan dibulatkan!!!)

Step 2: Memilih sampel pertama

Sampel pertama dipilih secara *simple random* dari Interval sampel pertama yaitu dari unit sampel nomor 1 sampai dengan nomor 283.

Misalnya yang terpilih adalah unit sampel nomor 57.

Step 3: Memilih sampel berikutnya yaitu dengan menambah interval I, sbb

No urut sampel	Deskripsi penentuan sampel	No ID Sampel
1	Dipilih secara acak dari 1-283	57
2	57 + 283,33	290,33 (290)
3	290,33 + 283,33	573,66 (574)
4	573,66 + 283,33	856,99 (857)
Sampel seterusnya	dstnya	dstnya

Pemilihan sampel secara sirkuler

Step 1: Menghitung Interval sampel (I) dengan rumus $I = N/n$

Apabila $N = 8500$ dan $n = 30$, maka $I = N/n = 8500/30 = 283,33$ (jangan dibulatkan!!!)

Step 2: Memilih sampel pertama

Sampel pertama dipilih secara *simple random* dari semua populasi yaitu dari unit sampel nomor 1 sampai 8500, misalnya yang terpilih adalah nomor 8450.

Step 3: Memilih sampel berikutnya yaitu dengan menambah interval I, sbb:

No urut sampel	Deskripsi pemilihan	No ID Sampel
1	Dipilih secara acak dari 1-8500	8450
2	8450 + 283,33	8733,33 Lebih besar dari jumlah populasi, maka nomor tersebut dikurang N menjadi: $8733,33 - 8500 = 233,33$ (233)
3	233,33+ 283,33	516,63 (517)
4	516,63 + 283,33	799,96 (800)
Dan seterusnya	dstnya	dstnya

Cara pemilihan sampel dengan Excel

Step 1: Menghitung Interval Sampel $I = N/n$

	A	B	C	D	E	F
1	STEP 1: Membuat Interval Sampel					
2		Jumlah Populasi (N)		8500		
3		Jumlah sampel (n)		30		
4		Interval Sampel (I)		283.3333		

Step 2: memilih sampel pertama

	A	B	C	D	E	F
1	STEP 1: Membuat Interval Sampel					
2		Jumlah Populasi (N)		8500		
3		Jumlah sampel (n)		30		
4		Interval Sampel (I)		283.3333		
5						
6	STEP 2: Pemilihan Sampel					
7		1	secara acak	57		

Step 3: memilih sampel berikutnya

	A	B	C	D	E	F
1	STEP 1: Membuat Interval Sampel					
2		Jumlah Populasi (N)		8500		
3		Jumlah sampel (n)		30		
4		Interval Sampel (I)		283.3333		
5						
6	STEP 2: Pemilihan Sampel					
7		1	secara acak	57		
8		2	$57 + 283.3333$	$=D7+283.3333$		

Selanjutnya copy rumus di kolom D7 ke bawah sampai nomor sampel ke 30 kemudian dilakukan pembulatan dengan meniadakan angka di belakang koma, dengan langkah sbb:

	A	B	C	D	E	F
6	STEP 2: Pemilihan Sampel					Pembulatan
7		1	secara acak	57		57
8		2	$57 + 283.3333$	340.333		340
9		3		623.667		624
10		4		907.000		907
11		5		1190.333		1190
12		6		1473.667		1474
13		7		1757.000		1757
14		8		2040.333		2040
15		9		2323.666		2324
16		10		2607.000		2607
17		11		2890.333		2890
18		12		3173.666		3174
19		13		3457.000		3457
20		14		3740.333		3740
21		15		4023.666		4024

Dan seterusnya!!

2.4.5 Cara penentuan besar sampel

Cara penentuan besar sampel survei dengan pemilihan sampel secara *systematic random sampling* adalah sama dengan cara penentuan besar sampel survei dengan metode *simple random sampling* seperti yang telah diuraikan sebelumnya.

2.4.6. Estimasi Parameter Populasi

Statistik sampel seperti mean (\bar{x}) dan proporsi (p) dari sampel yang dipilih dengan *systematic random sampling* dapat dipakai sebagai estimator dari parameter populasi seperti rerata populasi (μ) dan prevalensi di populasi (π) serta standar deviasi sampel (SD) juga bisa dipakai sebagai estimator standar deviasi populasi (σ).

Cara penghitungan nilai parameter populasi dan statistik sampel sama dengan cara penghitungan pada simple *random sampling* seperti di atas.

Misalnya dari 30 mahasiswa kedokteran diketahui IP kumulatifnya sebagai berikut:

2.5	2.8	3.0	3.0	3.2	3.5	3.6	3.7	3.9	4.0
2.7	3.0	3.0	3.0	3.1	3.1	3.3	3.4	3.5	3.5
2.8	3.0	3.0	3.1	3.2	3.2	3.5	3.6	3.7	4.0

Dari data populasi tersebut diketahui bahwa *mean* IPK (μ) = 3,263 dengan standar deviasi (σ) = 0,383

Dari populasi tersebut dipilih 10 sampel secara acak sistematis, dan didapatkan data sbb:

No.	2	5	8	11	14	17	20	23	26	29
IPK	2.8	3.2	3.7	2.7	3.0	3.3	3.5	3.0	3.2	3.7

Dari data sampel didapatkan nilai rata-rata (\bar{x}) = 3,21 dengan standar deviasi (SD) = 0,347

Estimasi IPK di populasi dengan kepercayaan 95% adalah 2,994 – 3,386. Jadi dengan kepercayaan 95%, estimasi rerata populasi berada di dalam nilai CI rerata sampel yaitu antara 3,033 – 3,424 dimana rerata IPK populasi adalah 3,263. Hal ini membuktikan bahwa rerata sampel dengan *systematic random sampling* tidak bias dan pada umumnya memiliki CI yang lebih sempit dari acak sederhana.

2.4.6 Latihan:

Kasus

Survei Rumah Tangga di Desa Bunga dilakukan dengan tujuan untuk mengetahui proporsi RT sehat berdasarkan 12 indikator GERMAS. RT dikategorikan sehat apabila minimal 80% indikator GERMAS

terpenuhi. Terdapat 8500 RT di desa tersebut. Sampel akan dipilih dengan metode *systematic random sampling*.

Jelaskan:

- 1) Siapa populasi terjangkau dari survei tersebut?
- 2) Kenapa sampel dipilih dengan *systematic random*, berikan penjelasannya!
- 3) Apa kelebihan dan kekurangan dari metode *simple random* tersebut?
- 4) Berapa jumlah sampel RT minimal yang dibutuhkan apabila diperkirakan proporsi RT sehat $P = 45\%$, *sampling error* $D = 5\%$, tingkat kepercayaan sampel 95%?
- 5) Demonstrasikan cara pemilihan sampel menggunakan metode linier!
- 6) Demonstrasikan cara pemilihan sampel dengan metode sirkuler!

2.5 Stratified Random Sampling (Acak Berlapis)

2.5.1 Pengertian

Stratified random sampling atau juga disebut pemilihan sampel secara acak berlapis, merupakan metode pemilihan sampel yang digunakan pada populasi yang heterogen. Pada pemilihan sampel secara acak berlapis, populasi yang heterogen dikelompokkan menjadi sub-sub populasi yang homogen dan sub kelompok tersebut disebut strata. Heterogenitas populasi bisa dilihat dari prevalensi/risiko permasalahan yang diteliti. Apabila risiko setiap individu di populasi tidak sama, maka populasi tersebut disebut heterogen. Misalnya, survei rumah tangga untuk mengetahui proporsi RT dengan akses air bersih dan jamban saniter di Kabupaten Gianyar. Dari laporan Survei Kesehatan Rumah Tangga (SKRT) 2020, diketahui bahwa prevalensi akses air bersih dan jamban saniter pada rumah tangga di kota dua kali lebih tinggi dari rumah tangga di desa. Oleh karena itu populasi survei rumah tangga di Kabupaten Gianyar tersebut adalah heterogen, maka populasi survei dijadikan dua strata, yaitu strata 1 adalah RT di kota dan strata 2 adalah RT di desa. Selanjutnya sampel dipilih dari masing-masing strata dengan metode *simple random sampling*.

2.5.2 Asumsi:

Metode acak berlapis (*stratified random sampling*) dipakai apabila populasi survei heterogen.

2.5.3 Kelebihan dan kekurangan:

Kelebihan dari metode acak berlapis adalah sampel yang terpilih akan memiliki tingkat representativitas yang lebih baik. Selain kelebihan tersebut, metode acak berlapis juga memiliki beberapa kelemahan seperti metode pemilihan sampel ini tidak dapat dilaksanakan apabila tidak tersedia data sekunder yang bisa dipakai acuan pembuatan strata serta ketersediaan data jumlah populasi dan angka prevalensi kejadian di setiap strata.

2.5.4 Cara pemilihan sampel:

Pemilihan sampel dilakukan pada setiap strata dengan metode acak sederhana menggunakan tabel bilangan *random* atau dengan bantuan komputer, dengan langkah sbb:

1) Pembuatan strata

Populasi yang heterogen dipilah menjadi kelompok-kelompok yang homogen, misalnya populasi survei dibagi menjadi dua strata yaitu kelompok kota dan kelompok desa.

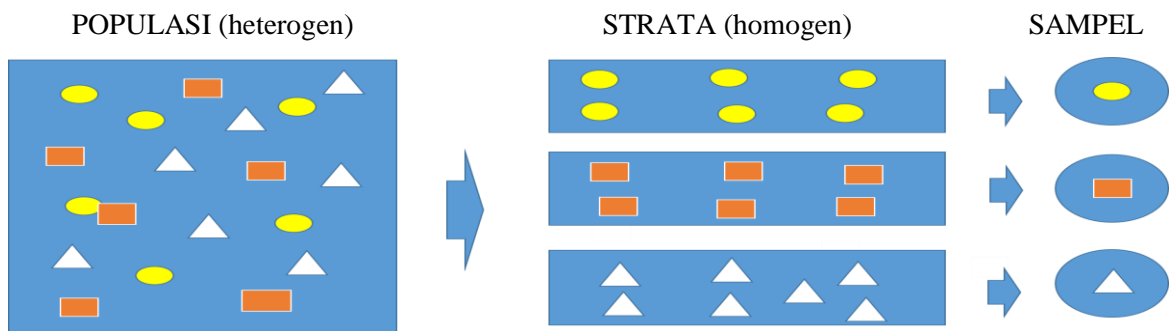
2) Membuat kerangka sampel di setiap strata

Semua unit sampel dari masing-masing strata dibuatkan daftar dan diberi nomor urut yang unik. Selanjutnya daftar tersebut dijadikan sebagai kerangka sampel untuk setiap strata. Misalnya semua RT di kota dan di desa dibuatkan *list*-nya dan diberikan nomor urut yang unik.

3) Menentukan jumlah sampel dari setiap strata

Penentuan jumlah sampel untuk setiap strata dapat dilakukan secara proporsional berdasarkan jumlah unit sampel/populasi di masing-masing strata (*proportional sample size*). Bisa juga ditetapkan jumlahnya sama untuk setiap strata (*equal sample size*). Apabila terdapat strata dengan jumlah populasi yang sangat kecil, maka penentuan besar sampel di masing-masing strata dapat dilakukan dengan pendekatan optimum (*optimum sample size*).

4) Pemilihan sampel dari setiap strata dengan cara acak sederhana.



Gambar 2.3. Bagan pemilihan sampel secara acak berlapis (*stratified random sampling*)

2.5.5 Cara penentuan besar sampel:

Penentuan besar sampel pada rancangan survei dengan acak berlapis ditentukan dengan beberapa parameter seperti, prevalensi atau kejadian peristiwa yang diteliti di masing-masing strata, jumlah populasi dari masing-masing strata, *margin of error*, dan tingkat kepercayaan sampel. Besar sampel dihitung dengan rumus sebagai berikut.

$$n = \frac{Z_{\alpha}^2 \sum_{h=1}^h \frac{N_h P_h (1 - P_h)}{w_h}}{N^2 d^2 + Z_{\alpha}^2 \sum_{h=1}^h N_h P_h (1 - P_h)}$$

Keterangan: n = besar sampel	N=jumlah semua populasi
N _h =jumlah populasi strata ke h	
P _h = kejadian pada strata ke h	
w _h = <i>weighting</i> strata ke h = N/N _h	
Z _α = nilai standar untuk kepercayaan sampel	
d= <i>margin of error</i>	

1) Penentuan besar sampel secara manual

Contoh:

Survei akses air bersih dan jamban pada rumah tangga (RT) di Kabupaten Badung. Jumlah populasi RT adalah 87.500 RT dengan rincian di kota sebanyak 37.000 RT dan di desa sebanyak 50.500 RT. Tingkat akses air bersih di kota sebesar 75% dan di desa sebesar 45%. Sampel akan dipilih dengan metode acak berlapis proporsional (*proportional stratified random sampling*). Besar sampel yang dibutuhkan bila sampel akan dipilih secara acak berlapis dengan *margin of error* $d = 5\%$ dan tingkat kepercayaan sampel 95%, dihitung dengan langkah seperti dibawah ini.

Cara penentuan besar sampel

Step 1: membuat karakteristik strata, sbb:

Parameter	Strata 1 Perkotaan	Strata 2 Pedesaan
N_h	37.000	50.500
P_h	0,75	0,45
w_h	2,36	1,73

Keterangan: N_h = jumlah populasi strata ke h
 P_h = prevalensi kejadian pada strata ke h
 w_h = *weighting* strata ke h = N/N_h

Step 2: menghitung besar sampel

$$n = \frac{1,96^2 [(37.000 \times 0,75 \times (1-0,75))/2,36 + \{(50.500 \times 0,45 \times (1-0,45))/1,73\}]}{87.500^2 \times 0,05^2 + 1,96^2 \times [\{37.000 \times 0,75 \times (1 - 0,75)\} + \{50.500 \times 0,45 \times (1-0,45)\}]}$$

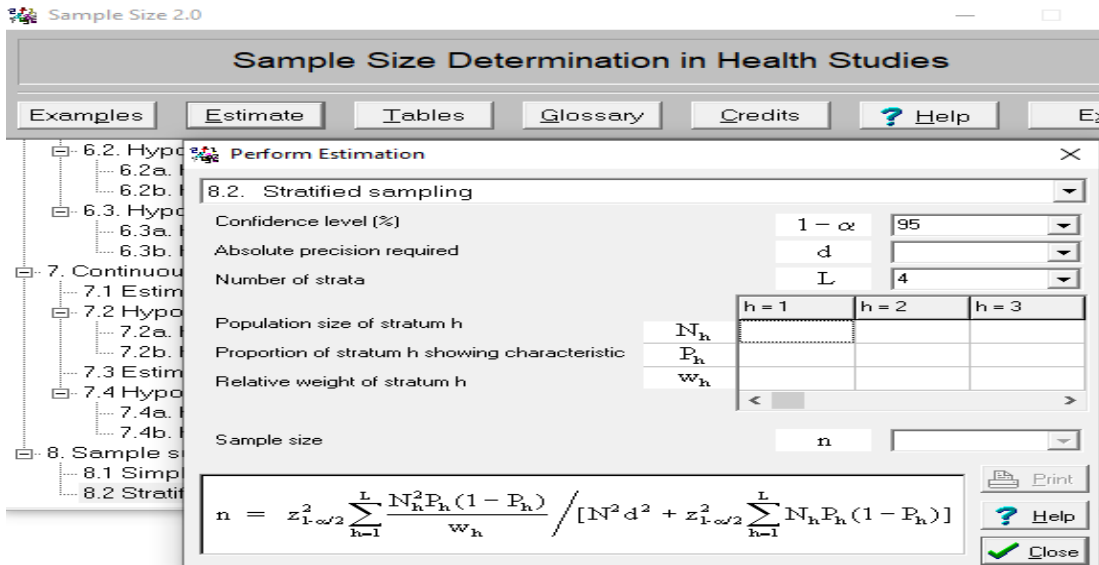
$$n = 386$$

Step 3: menentukan jumlah sampel setiap strata

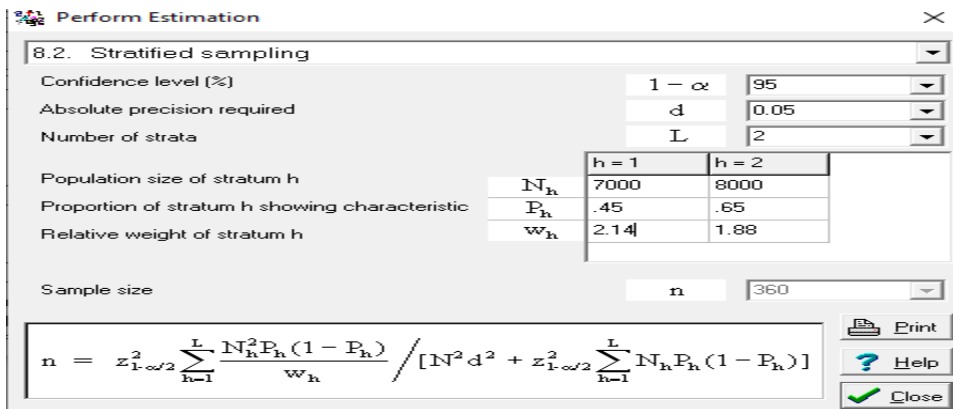
Jumlah sampel setiap strata ditentukan secara proporsional dimana:

$$\text{Jumlah sampel strata 1 (perkotaan)} = N_h/N \times n = 37.000/87.500 \times 386 = 163 \text{ RT}$$

$$\text{Jumlah sampel strata 2 (pedesaan)} = N_h/N \times n = 50.500/87.500 \times 386 = 223 \text{ RT}$$



Step 3: masukan data % confidence level (95), margin of error $d = 0.05$, jumlah strata $L = 2$, $N_1=7.000$, $N_2=8000$, $P_1 = 0.45$, $P_2 = 0.65$, dan $w_1 = N/Nh_1 = 15.000/7.000 = 2,14$, $w_2 = 15000/8000 = 1,88$ untuk $h=1$ dan $h=2$, sbb:



Total sampel yang dibutuhkan adalah $n = 360$.

Sampel untuk di kota $n_1 = N_h/N \times n = 7000/15000 \times 360 = 168$

Sampel untuk di desa $n_2 = N_h/N \times n = 8000/15000 \times 360 = 192$

2.5.6 Estimasi Parameter Populasi

Penghitungan statistik sampel seperti mean (\bar{x}) dan proporsi (p) dari sampel yang dipilih dengan *stratified random sampling* dapat dipakai sebagai estimator dari parameter populasi seperti rerata populasi (μ) dan prevalensi di populasi (π) serta standar deviasi sampel (SD) juga bisa dipakai sebagai estimator standar deviasi populasi (σ).

Cara penghitungan nilai statistik sampel seperti *mean* dan *sd* harus memperhatikan besarnya *weighting* dari masing-masing strata dengan rumus penghitungan sebagai berikut.

Penghitungan besar *weighting* setiap strata:

Strata	Juml populasi (N_h)	Jumlah sampel (n_h)	Weighting (w_h) N_h/n_h
Mahasiswa	30	10	3
Mahasiswi	60	10	6

Penghitungan statistik sampel:

Rerata	Standar Deviasi
$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$	$sd = \sqrt{\frac{\sum_{i=1}^n w_i (x_i - \bar{x})^2}{\sum_{i=1}^n w_i}}$
Keterangan: \bar{x} = rerata setelah weighting sd = standar deviasi w_i = <i>weighting</i> n = jumlah sampel	

Misalnya dari 30 mahasiswa dan 60 mahasiswi kedokteran diketahui IP kumulatifnya sebagai berikut:

Data mahasiswa:

2.5	2.8	3.0	3.0	3.2	3.5	3.6	3.7	3.9	4.0
2.7	3.0	3.0	3.0	3.1	3.1	3.3	3.4	3.5	3.5
2.8	3.0	3.0	3.1	3.2	3.2	3.5	3.6	3.7	4.0

Data mahasiswi:

2.7	3.8	3.0	3.0	3.2	3.5	3.6	3.7	3.9	4.0
2.9	3.0	3.0	3.0	3.1	3.1	3.3	3.4	3.5	3.5
3.8	3.0	3.0	3.1	3.2	3.2	3.5	3.6	3.7	4.0

3.7 3.8 3.0 3.0 3.2 3.5 3.6 3.7 3.9 4.0
 2.9 3.0 3.0 3.0 3.1 3.1 3.3 3.4 3.5 3.5
 3.8 3.0 3.0 3.1 3.2 3.2 3.5 3.6 3.7 3.0

Dari data populasi mahasiswa dan mahasiswi tersebut diketahui bahwa mean IPK (μ) = 3,316 dengan standar deviasi (σ) = 0,356

Dari masing-masing strata dipilih 10 sampel secara acak sederhana, dan didapatkan data sbb:

Strata 1: mahasiswa

No.	22	17	14	1	12	23	20	10	2	30
IPK	3.0	3.3	3.0	2.5	3.0	3.0	3.5	4.0	2.8	4.0

Strata 2: mahasiswi

No.	30	22	25	59	1	30	40	57	44	41
IPK	4.0	3.0	3.2	3.7	2.7	4.0	4.0	3.5	3.0	2.9

Dari data sampel didapatkan nilai rata-rata setelah di *weighting* (\bar{x}) = 3,3367 dengan standar deviasi (SD) = 0,233.

Estimasi IPK di populasi dengan kepercayaan 95% adalah 3,234 – 3,438. Jadi dengan kepercayaan 95%, estimasi rerata IPK populasi berada di dalam nilai CI rerata sampel yaitu antara 3,234 – 3,438 dimana rerata IPK populasi adalah 3,316. Hal ini membuktikan bahwa rerata sampel dengan *stratified random sampling* tidak bias dan pada umumnya memiliki CI yang lebih sempit dari acak sederhana.

2.5.7 Latihan

Kasus:

Ida adalah mahasiswa semester VII Program Studi Kedokteran sedang mengerjakan penelitian untuk tugas akhir pendidikan Sarjana Kedokteran (S.Ked). Dalam skripsinya, dia meneliti prevalensi merokok pada siswa SMU di Kota Denpasar. Terdapat 45.000 siswa SMU di Kota Denpasar dengan rincian 24.000 perempuan dan 21.000 laki-laki. SKRT (2018) melaporkan terdapat 5% siswi SMU merokok dan 25% pada siswa SMU.

Jelaskan:

- 1) Siapa populasi terjangkau dari survei tersebut?

- 2) Kenapa sampel dipilih dengan acak berlapis, berikan penjelasannya!
- 3) Apa kelebihan dan kekurangan dari metode acak berlapis?
- 4) Berapa sampel minimal yang dibutuhkan apabila diperkirakan proporsi merokok pada siswa perempuan 5% dan pada siswa laki-laki 25%, *sampling error* $D = 5\%$ dan tingkat kepercayaan sampel 95%?
- 5) Berapa jumlah sampel dari masing-masing strata secara proporsional!
- 6) Demonstrasikan cara pemilihan sampel pada kasus tersebut!

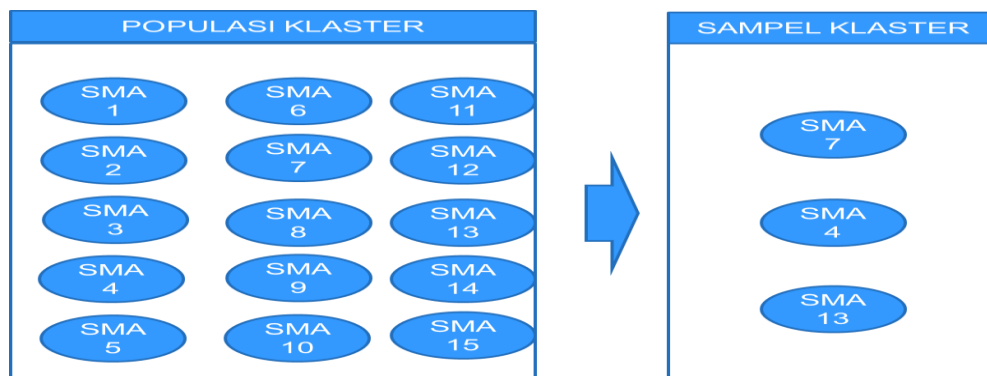
2.6 Cluster sampling (Acak Kelompok)

2.6.1 Pengertian

Klaster adalah sekelompok individu yang tinggal atau berkegiatan di satu wilayah atau institusi tertentu dengan ciri-ciri yang sama. *Cluster sampling* (acak kelompok) adalah metode sampling yang diterapkan terhadap “kelompok individu” atau klaster. Dalam *cluster sampling*, klaster ditetapkan sebagai unit sampel primer atau *primary sampling unit* (PSU). Pada langkah pertama akan dipilih sampel klasternya, setelah itu dilakukan pemilihan sampel individu dari klaster terpilih. Pemilihan sampel klaster (PSU) bisa dilakukan secara acak sederhana dan juga bisa menggunakan metode *Probability Proportional to Size* (PPS). Misalnya, pada survei rumah tangga tentang akses air minum dan jamban saniter, populasi surveinya adalah rumah tangga (RT). Populasi tersebut berkelompok menurut wilayah administrasi desa, maka kelompok RT di masing-masing desa menjadi klaster RT dan desa merupakan PSU dalam pemilihan sampel secara acak kelompok tersebut. Dalam kasus ini, pada tahap pertama akan dipilih sampel desa (klaster/PSU) dan pada tahap berikutnya baru dipilih sampel RT dari Desa yang terpilih. Contoh lain, pada survei siswa sekolah menengah atas (SMA) tentang perilaku *bullying*, siswa merupakan populasi dari survei tersebut. Populasi siswa tersebut berkelompok menurut sekolah masing-masing, maka pada survei ini sekolah ditetapkan sebagai klaster siswa SMA. Pada contoh ini, pemilihan sampel pada tahap pertama akan dipilih sampel sekolah (PSU) dan setelah itu dilakukan pemilihan sampel siswa dari sekolah yang terpilih (Gambar 2.4).

2.6.2 Indikasi

Metode klaster dipilih apabila area surveinya luas seperti misalnya satu kabupaten, provinsi atau satu negara dan populasi secara alami terdistribusi dalam kelompok-kelompok, misalnya siswa SMA, penduduk yang tinggal di asrama, penduduk yang tinggal di satu kawansan seperti desa atau blok sensus. Berikut adalah ilustrasi indikasi pemilihan sampel klaster.



Gambar 2.4. Bagan pemilihan sampel acak kelompok (*cluster sampling*)

2.6.3 Kelebihan dan kekurangan

Pemilihan sampel akan lebih mudah, cepat, dan murah bila dibandingkan dengan metode acak sederhana. Pada kluster sampling, kerangka sampel berisikan daftar klasternya saja, seperti daftar desa atau daftar posyandu sehingga dalam pembuatan kerangka sampelnya menjadi lebih mudah, cepat dan murah. Sebaliknya, pada pemilihan sampel secara acak sederhana, kerangka sampel berisi daftar nama semua populasi dan untuk membuat kerangka sampel tersebut memerlukan waktu dan biaya serta tenaga yang banyak. Di sisi lain, metode kluster memiliki kelemahan yaitu tidak setiap individu di populasi tersebut mempunyai kesempatan untuk terpilih. Hal ini akan berdampak terhadap representativitas dan varian sampel yang umumnya akan lebih besar dibandingkan bila sampel dipilih secara acak sederhana.

2.6.4 Cara pemilihan sampel

Metode pemilihan sampel kluster dapat dilakukan dengan acak sederhana atau dengan *probability proportional to size* (PPS). Perbedaan kedua metode tersebut terletak pada cara pemilihan sampel klasternya. Pada metode acak sederhana, semua kluster mempunyai kesempatan yang sama untuk terpilih sebagai sampel. Sebaliknya pada metode PPS, probabilitas kluster terpilih menjadi sampel berkaitan dengan ukuran (*size*) dari kluster atau jumlah populasi di kluster tersebut. Kluster dengan jumlah populasi yang besar akan memiliki kemungkinan terpilih lebih besar dari kluster dengan jumlah populasi yang kecil. Pemilihan sampel kluster pada umumnya dilakukan dengan dua tahap (*two stages cluster sampling*). Pada *stage 1* memilih sampel kluster (PSU) dan pada *stage 2* memilih sampel individu dari kluster terpilih. Cara pemilihan sampel kluster akan dijelaskan pada uraian di bawah ini.

1) Pemilihan sampel kluster secara acak sederhana

Step 1: memilih sampel kluster

Pada tahap ini dilakukan pemilihan kluster yang menjadi unit sampel pertama (*Primary Sampling Unit* atau PSU) dengan langkah sebagai berikut:

- Membuat kerangka sampel yang berisikan daftar nama kluster, misalnya nama semua posyandu
- Selanjutnya memilih sejumlah kluster sebagai sampel secara acak sederhana. Misalnya pada survei tersebut terdapat 450 kluster posyandu dan dari jumlah tersebut akan dipilih

sebanyak 30 sampel posyandu secara acak sederhana seperti yang sudah dibahas sebelumnya.

Step 2: memilih sampel individu

Setelah sampel kluster terpilih, maka tahap selanjutnya adalah memilih sampel individu dari kluster tersebut sebanyak yang ditetapkan dengan cara acak sederhana, dengan langkah sebagai berikut:

- Membuat kerangka sampel yang berisi daftar nama individu pada semua kluster terpilih, misalnya nama anak balita yang ada di masing-masing posyandu terpilih.
- Selanjutnya memilih sejumlah balita sebagai sampel dari masing-masing posyandu terpilih secara acak sederhana. Misalnya dari setiap posyandu ditetapkan akan dipilih 10 balita secara acak sederhana. Pada contoh di atas dipilih sebanyak 10 sampel balita per posyandu sehingga jumlah sampel balita menjadi sebanyak 300 sampel.

2) Pemilihan sampel kluster secara PPS

Step 1: Memilih sampel kluster

Pada tahap ini, yang menjadi unit sampel pertama (*Primary Sampling Unit* atau PSU) adalah kluster. Langkah pemilihan sampel kluster adalah sebagai berikut:

- Membuat kerangka sampel:

Kerangka sampel pertama (PSU) pada pemilihan sampel dengan PPS berbeda dengan pemilihan secara acak sederhana. Kerangka PSU pada pemilihan dengan PPS terdiri dari nama kluster dan ukuran (*size*) di setiap kluster. Ukuran kluster ditentukan dari jumlah populasi di setiap kluster. Misalnya sekolah dasar (SD) sebagai kluster (PSU) dan jumlah siswa di SD tersebut merupakan *size* dari kluster tersebut, maka kerangka sampel kluster sekolahnya seperti contoh berikut.

Kluster	Size (jumlah siswa)
SD1	100
SD1	125
SD3	120
SD4	90
SD5	120
SD6	100
SD7	160

SD8	100
SD9	125
Sd10	100
SD11	100
SD11	150
SD13	140
SD14	100
SD15	100
SD16	120
SD17	125
.....
.....
SD45	100
SD46	155
SD47	100
SD48	122
SD49	100
SD50	120

- Menentukan jumlah populasi kumulatif:

Langkah berikutnya adalah membuat jumlah populasi kumulatif dengan menambahkan jumlah populasi klaster-klaster sebelumnya. Berikut adalah contoh pembuatan jumlah populasi kumulatif dari klaster SD.

Klaster	Size (jumlah siswa)	Jumlah Kumulatif
SD1	100	100
SD2	125	225
SD3	120	345
SD4	90	435
SD5	120	555
SD6	100	655
SD7	160	815
SD8	100	915
SD9	125	1040
SD10	100	1140
SD11	100	1240

SD11	150	1390
SD13	140	1530
SD14	100	1630
SD15	100	1730
SD16	120	1850
SD17	125	1975
.....
.....
SD45	100	5350
SD46	155	5505
SD47	100	5605
SD48	122	5727
SD49	100	5827
SD50	120	5947

- Menentukan Interval sampel:

Pemilihan sampel dengan PPS mirip dengan pemilihan sampel dengan metode acak sistematis, tetapi bedanya pada PPS yang dijadikan kerangka sampel adalah jumlah populasi kumulatif. Setelah kerangka sampel dibuat, maka langkah selanjutnya adalah menetapkan interval sampel $I = \text{jumlah total populasi} / \text{jumlah sampel kluster}$. Pada contoh di atas diketahui jumlah total siswa adalah 5.947 orang dan jumlah sampel kluster ditentukan sebanyak 30 SD, maka intervalnya adalah:

$$I = 5.947/30 = 198,23$$

- Memilih sampel pertama:

Sama seperti pada pemilihan sampel acak sistematis, sampel pertama dipilih secara *random* dari I populasi pertama. Pada contoh di atas, $I = 198$, maka sampel pertama dipilih satu siswa dari 198 siswa pertama. Misalnya yang terpilih adalah siswa nomor 155. Berarti sampel kluster pertama adalah SD dimana siswa ke 155 itu bersekolah, yaitu SD 2, karena di SD 2 terdapat siswa yang ke 101 sampai 225.

- Penentuan sampel kluster berikutnya:

Penentuan sampel kluster berikutnya dilakukan dengan menambahkan nomor sampel siswa sebelumnya dengan interval I, maka sampel yang kedua adalah $155 + 198,23$ yaitu siswa ke 353,23, maka sampel SD yang kedua adalah SD 4, karena siswa yang ke 353,23 ada di SD 4. Penentuan sampel kluster ke-3 sampai yang ke-30 dapat dilakukan dengan cara seperti di atas.

Seperti contoh berikut.

Klaster	Size (jumlah siswa)	Jumlah Kumulatif	Sampel klaster
SD1	100	100	
SD2	125	225	155
SD3	120	345	
SD4	90	435	353,23
SD5	120	555	551,46
SD6	100	655	
SD7	160	815	749,69
SD8	100	915	
SD9	125	1040	947,92
Sd10	100	1140	
SD11	100	1240	1146,15
SD11	150	1390	1344,38
SD13	140	1530	
SD14	100	1630	1542,61
SD15	100	1730	
SD16	120	1850	1740,84
SD17	125	1975	1939,07
.....	
.....	
SD45	100	5350	5308,98
SD46	155	5505	
SD47	100	5605	5507,21
SD48	122	5727	5705,44
SD49	100	5827	
Sd50	120	5947	5903,67

Step 2: Memilih sampel individu

Setelah sampel klaster terpilih, maka tahap selanjutnya adalah memilih sampel individu dari klaster tersebut sebanyak yang ditetapkan dengan cara acak sederhana, dengan langkah sebagai berikut:

- Membuat kerangka sampel yang berisikan daftar nama individu dari semua kluster terpilih, misalnya nama anak SD yang ada di masing-masing sekolah terpilih.
- Selanjutnya memilih sejumlah siswa sebagai sampel dari masing-masing SD terpilih secara acak sederhana. Misalnya dari setiap SD ditetapkan akan dipilih 10 siswa secara acak sederhana, maka akan terpilih 300 sampel siswa pada survei ini.

2.6.5 Cara penentuan besar sampel:

Survei dengan metode kluster sampling memiliki kelemahan yaitu varians sampel akan lebih besar dari varians sampel yang dipilih secara acak sederhana. Upaya yang harus dilakukan untuk memperkecil varians sampel tersebut adalah dengan memperbesar jumlah sampel. Oleh karena itu, jumlah sampel survei dengan metode kluster sampling dihitung sama dengan metode acak sederhana tapi perlu dikalikan dengan efek desain (*Design Effect*) yang besarnya antara 1,75 sampai 2,0. Berdasarkan hal tersebut, maka besar sampel survei dengan metode kluster sampling dihitung dengan rumus berikut.

$$n = \frac{Z_{\alpha}^2 P(1-P)}{D^2} \times DE$$

Keterangan:

- n = jumlah sampel minimal
- Z_{α} = nilai Z score untuk kepercayaan (1- α) yang besarnya 1,96 untuk kepercayaan 95% dan 2,57 untuk kepercayaan 99%
- P = estimasi kejadian
- D = presisi atau *margin of error*
- DE = *design effect*

Contoh:

Survei prevalensi *bullying* pada siswa SD di Kota Denpasar dilakukan dengan metode kluster sampling dengan menggunakan sekolah sebagai kluster. Besar sampel siswa dan sampel SD yang dibutuhkan apabila jumlah sampel siswa untuk tiap sekolah (m) = 15 siswa, perkiraan prevalensi *bullying* 20%, *margin of error* 5% dan tingkat kepercayaan sampel 95%, serta DE = 1,75; dapat dihitung seperti dibawah ini.

Penghitungan jumlah sampel siswa SD:

$$n = \frac{Z_{\alpha}^2 P(1-P)}{D^2} x DE = \frac{1,96^2 (0,2 \times 0,8)}{0,05^2} \times 1,75 = 430,25 \text{ (dibulatkan menjadi 431 siswa)}$$

Penghitungan jumlah sampel kluster (sekolah):

Jumlah sampel kluster dihitung dengan cara membagi jumlah sampel siswa “n” dengan jumlah sampel siswa per SD “m”, dengan rumus berikut:

$$n_k = \frac{n}{m} = 431/15 = 22,7 \text{ (dibulatkan menjadi 23 SD)}$$

2.6.6 Estimasi Parameter Populasi

Penghitungan statistik sampel kluster seperti *mean* (\bar{x}) dan proporsi (p) dari sampel yang dipilih dengan PPS dapat dipakai sebagai estimator dari parameter populasi seperti rerata populasi (μ) dan prevalensi di populasi (π) serta standar deviasi sampel (SD) juga bisa dipakai sebagai estimator standar deviasi populasi (σ).

Cara penghitungan nilai statistik sampel seperti *mean* dan SD harus memperhatikan besarnya *weighting* dari masing-masing kluster dengan rumus penghitungan sebagai berikut.

Penghitungan besar *weighting* setiap kluster:

Contoh: sampel mahasiswa yang dipilih berdasarkan kluster angkatan

Kluster terpilih	Juml populasi (N _h)	Jumlah sampel (n _h)	Weighting (w _h) N _h /n _h
Angkatan 2015	50	10	5
Angkatan 2017	90	10	9
Angkatan 2019	155	10	15

Penghitungan statistik sampel

Rerata	Standar Deviasi
$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$	$sd = \sqrt{\frac{\sum_{i=1}^n w_i (x_i - \bar{x})^2}{\sum_{i=1}^n w_i}}$
<p>Keterangan:</p> <p>\bar{x} = rerata setelah <i>weighting</i> sd = standar deviasi w_i = <i>weighting</i></p>	

n = jumlah sampel

2.6.7 Latihan

Kasus

Dinas Kesehatan Provinsi Bali akan melakukan survei cakupan imunisasi dasar dan lanjutan di Provinsi Bali pada tahun 2020. Populasi survei adalah anak usia 12-24 bulan di seluruh Bali yang terdaftar di 4000 posyandu di Bali. Sampel akan dipilih secara *cluster* dua tingkat dengan posyandu sebagai *primary sampling unit* (PSU) atau sebagai klaster dan sebagai unit analisis adalah anak usia 12-24 bulan. Cakupan imunisasi dasar diperkirakan 90% dan imunisasi lanjutan sekitar 80%. Tingkat kepercayaan sampel 95%, *margin of error* $D = 5\%$ dan efek desain (DE) = 1,75. Jumlah sampel anak usia 12-24 bulan per posyandu sebanyak 10 anak.

TUGAS:

1. Tentukan siapa yang menjadi populasi target, terjangkau, PSU dan unit sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel anak usia 12-24 bulan yang dibutuhkan?
4. Berapa jumlah sampel klaster posyandu yang dibutuhkan?
5. Jelaskan cara pemilihan sampel secara acak sederhana dan secara PPS!

2.7 Multistage Random Sampling (Acak Bertingkat)

2.7.1 Pengertian

Multistage sampling adalah metode sampling dimana sampel dipilih secara berjenjang karena wilayah survei atau wilayah penelitiannya sangat luas seperti satu provinsi atau satu negara. Misalnya akan diteliti proporsi penggunaan metode kontrasepsi di Bali. Pemilihan sampel dilakukan secara berjenjang, pada jenjang 1 dipilih 2 kabupaten, pada jenjang 2 dipilih 2 kecamatan dari kabupaten terpilih, pada jenjang 3 dipilih 2 desa dari kecamatan terpilih, pada jenjang 4 dipilih 2 banjar pada desa terpilih, pada jenjang 5 dipilih 30 RT pada banjar terpilih. Penelitian akan dilakukan pada sampel RT dengan menanyakan jumlah anggota RT dan penggunaan metode kontrasepsi pada pasangan usia subur di RT tersebut.

2.7.2 Indikasi:

Metode multistage dipilih apabila wilayah survei atau penelitian sangat luas, misalnya satu provinsi atau satu negara dan di sisi lain terdapat keterbatasan sumber daya manusia, biaya, dan peralatan lainnya, serta keterbatasan waktu penelitian.

2.7.3 Kelebihan dan kekurangan:

Metode ini memiliki beberapa kelebihan seperti waktu penelitian bisa lebih singkat, biaya lebih rendah, dan sarana pendukung juga lebih sedikit. Walaupun ada kelebihan dibandingkan *simple random sampling*, metode ini juga memiliki kelemahan seperti tingkat representativitas sampel lebih rendah, bias lebih tinggi karena tidak setiap individu di populasi memiliki kesempatan yang sama menjadi sampel. Kelemahan dari desain sampling ini dapat diatasi dengan memperbanyak jumlah sampel dengan melakukan koreksi mempertimbangan faktor efek desain (*design effect*). *Design effect* yang umum dipakai antara 1,75 sampai 2, sehingga jumlah sampel hampir menjadi 2 kali lebih banyak dari *simple random sampling*.

2.7.4 Langkah pemilihan sampel

Pada setiap jenjang pemilihan sampel harus dilakukan:

- 1) Membuat kerangka sampel dari *list* populasi di jenjang tersebut. Misalnya pada jenjang pertama yang dipilih 2 kabupaten dari 9 kabupaten yang ada, maka harus dibuat kerangka sampel jenjang pertama yang berisi *list* semua kabupaten yang ada di wilayah penelitian. Pada jenjang ke dua akan dipilih kecamatan dari kabupaten terpilih, maka pada tahap ke dua, harus dibuat *list*

kecamatan yang ada di kabupaten terpilih. Hal yang sama dilakukan pada setiap jenjang pemilihan sampel selanjutnya.

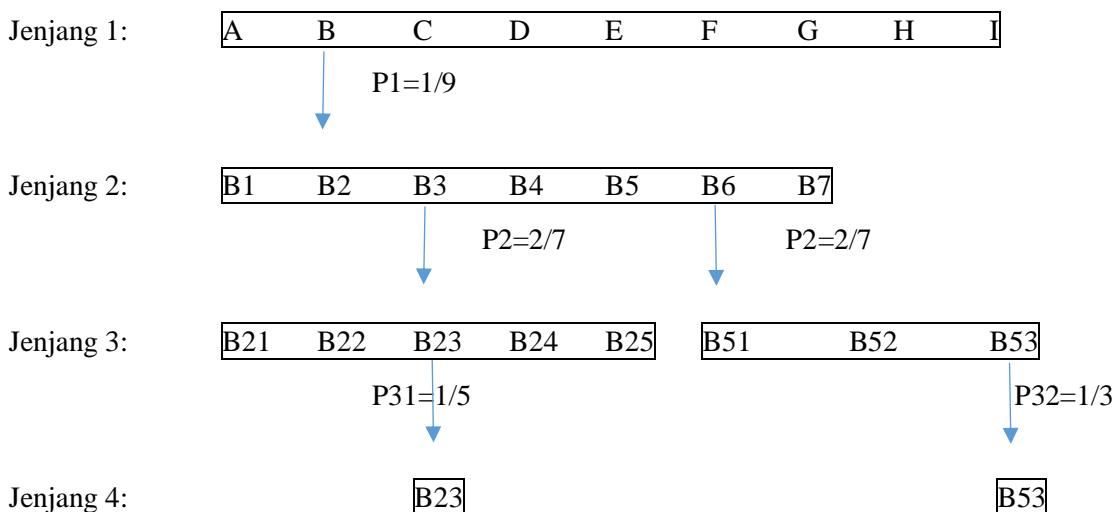
- 2) Menetapkan jumlah sampel dan cara pemilihan sampel di setiap jenjang pemilihan, contoh akan dipilih 2 sampel dan pemilihan sampel di setiap jenjang dilakukan dengan metode *simple random sampling* (undian/bilangan *random/computer assisted*).
- 3) Membuat daftar sampel terakhir berdasarkan hasil pemilihan dari jenjang pertama sampai terakhir.

2.7.5 Penentuan *sampling fraction*

Sampling fraction merupakan probabilitas unit sampel terakhir untuk terpilih menjadi sampel yang merupakan hasil perkalian probabilitas pemilihan sampel dari semua jenjang. Jadi probabilitas terpilihnya unit sampel dari keseluruhan populasi yang ada di tiap jenjangnya ditentukan terlebih dahulu

Berikut adalah ilustrasi penghitungan *sampling fraction* untuk tiap jenjang dari pemilihan sampel dengan *multistage sampling*.

Bila pada jenjang pertama dipilih 1 kabupaten (Kabupaten B) dari total 9 kabupaten yang ada, maka *sampling fraction* untuk jenjang pertama, $p_1 = 1/9$. Pada jenjang ke dua dipilih 2 kecamatan, B2 dan B5 dari 7 kecamatan yang ada, maka *sampling fraction* jenjang kedua, $p_2 = 2/7$. Selanjutnya di jenjang ketiga, misalkan dari kecamatan B2 dipilih satu desa B23 dari 5 desa yang ada dan dari B5 dipilih satu desa B53 dari 3 desa yang ada, maka *sampling fraction* untuk desa B23, $p_{31}=1/5$ dan untuk desa B53, $p_{32}=1/3$ (Gambar 2.5)



Gambar 2.5. Bagan pemilihan sampel acak bertingkat (*multi-stages sampling*)

Sampling fraction:

Sampling fraction untuk dua unit sampel terakhir di jenjang 4, sebagai berikut:

Sampling fraction sampel di B23 = $p_1 \times p_2 \times p_3 = 1/9 \times 2/7 \times 1/5 = 0,0063$

Sampling fraction sampel di B53 = $1/9 \times 2/7 \times 1/3 = 0,0106$

Weighting

Weighting merupakan nilai penimbang atau pembobotan yang besarnya sama dengan *inverse* dari *sampling fraction*, sehingga dihitung 1 dibagi *sampling fraction*.

Weighting (w) = $1/\text{sampling fraction}$

Weighting untuk sampel B23 (w_1) = $1/0,0063 = 158,73$

Weighting untuk sampel B54 (w_2) = $1/0,0106 = 94,34$

2.7.6 Estimasi parameter populasi

Dari contoh diatas sampel ibu hamil dipilih di desa B23 dan di desa B54 dengan *sampling fraction* yang berbeda dan faktor penimbang yang berbeda. Misal jumlah sampel n , maka akan terdapat data $x_1 \times x_2 \times x_3 \dots \dots X_4$, untuk perhitungan statistik pada desain *multistage* sampel, maka perlu dikalikan dengan *weight*/penimbangnya.

Statistik	Estimasi statistik sampel <i>simple random</i>	Estimasi statistik sample <i>multistage random</i>
Rerata (<i>Mean</i>)	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$
Simpang Baku (Standar Deviasi)	$sd = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$	$sd = \frac{\sum_{i=1}^n w_i (x_i - \bar{x})^2}{\sum_{i=1}^n w_i}$
Keterangan:		
X_i = nilai observasi ke I	\bar{x} = rerata	Sd = standar deviasi
w_i = <i>weighting</i>	n = jumlah sampel	

2.7.7 Penentuan besar sampel

Penentuan besar sampel survei dengan metode pemilihan sampel secara bertingkat (*multi stage*) sama dengan penentuan jumlah sampel survei dengan rancangan acak sederhana dengan koreksi efek desain, dengan rumus sebagai berikut.

$n = \frac{Z_{\alpha}^2 p(1-p)}{d^2} x DE$
Keterangan: n = jumlah sampel p = estimasi kejadian di populasi d = <i>margin of error</i> (presisi) Z_{α} = Z-skor untuk tingkat kepercayaan (1- α)

Catatan: Cara penghitungan besar sampel acak bertingkat sama dengan acak kelompok, seperti contoh penghitungan besar sampel penelitian acak kelompok di atas.

2.7.8 Latihan

Dinas Kesehatan Provinsi Bali ditugaskan melaksanakan Survei Rumah Tangga di wilayahnya dengan tujuan untuk mengetahui jumlah anggota keluarga dan pengeluaran kesehatan per RT. Jumlah RT 36.900 yang tersebar di 9 kabupaten/kota, 57 kecamatan, dan 420 desa (data pada tabel 2). Sampel akan dipilih secara *multi stage* dengan tingkat kepercayaan 95% dan *margin of error*/presisi (d) = 0,05. Tingkat pertama dipilih 2 kabupaten secara *random*, pada tingkat kedua dipilih 2 kecamatan dari setiap sampel kabupaten terpilih, tahap 3 dipilih 2 desa dari setiap sampel kecamatan terpilih, dan terakhir dipilih sampel RT dari setiap desa sampel.

Tugas:

1. Tentukan siapa yang menjadi populasi terjangkau, unit sampel dan unit analisisnya!
2. Apa informasi yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel yang diperlukan?
4. Bagaimana cara pemilihan sampel secara *multistage*?

2.8 Metode Non-Random

2.8.1 Pengertian *non-random sampling*

Non random sampling juga disebut *non-probability sampling* adalah metode sampling dimana sampel dipilih berdasarkan subyektivitas peneliti dan bukan dipilih secara acak. Teknik sampling ini sangat ditentukan oleh kepakaran peneliti dan banyak dipakai pada penelitian kualitatif. Pada metode ini, tidak semua individu dari populasi tersebut memiliki kesempatan (probabilitas) terpilih sebagai sampel seperti pada metode *random sampling*. Metode ini sangat baik dipakai pada penelitian eksplorasi seperti *pilot survey*. Metode ini juga bisa dipilih apabila peneliti tidak bisa membuat kerangka sampel (*sampling frame*) karena kesulitan untuk menjangkau kelompok populasi yang akan diteliti, masalah waktu dan biaya yang tersedia.

2.8.2 Jenis *non-random sampling*

Terdapat beberapa metode pemilihan sampel *non random*, seperti *convenience sampling*, *consecutive sampling*, *incidental sampling*, *quota sampling*, *purposive sampling*, dan *snowball sampling*. Berikut dijelaskan pengertian, indikasi, kelebihan dan kekurangan, serta cara pemilihan sampel dari masing-masing metode *non-random* tersebut.

- ***Convenience sampling:***

Pada metode ini sampel dipilih dari populasi yang mudah dijangkau dan tersedia di sekitar peneliti (*conveniently available to the researcher*). Peneliti memilih sampel yang mudah dipilih dan peneliti tidak mempertimbangkan representativitas terhadap semua populasi. Metode ini dipilih karena peneliti tidak mungkin mengakses semua populasi atau tidak bisa membuat kerangka sampel karena populasinya sangat besar dan area penelitian yang sangat luas. Pertimbangan untuk menggunakan metode ini adalah dari segi waktu yang lebih singkat, biayanya murah (*cost-effectiveness*) dan mudah dilakukan.

Contoh:

Penelitian dengan tujuan mengetahui sero-konversi yang terjadi pada orang setelah mendapat vaksin secara lengkap. Di lokasi penelitian terdapat 50 daftar nakes yang sudah mendapatkan vaksin secara lengkap. Sampel dipilih dari daftar nakes yang sudah mendapatkan vaksin secara lengkap tersebut.

- **Consecutive sampling:**

Pemilihan sampel secara konsekutif hampir sama dengan *convenience sampling* dengan sedikit variasi. Pada metode ini peneliti memilih individu yang memenuhi kriteria sampel dan diulang pada individu yang lain sampai jumlah sampel yang diperlukan terpenuhi. Metode ini banyak dilakukan pada penelitian klinik yang menggunakan pasien yang datang sebagai sampel. Sangatlah tidak memungkinkan penelitian membuat kerangka sampel karena tidak bisa diketahui siapa yang akan datang sebelum mereka secara langsung datang ke rumah sakit tersebut. Oleh karena itu, sampel dipilih dari pasien yang datang dan memenuhi kriteria sampel dan diulang sampai jumlah sampel yang dibutuhkan terpenuhi.

Contoh:

Penelitian klinik dilakukan untuk mengetahui kelainan genetik pada penderita DM yang menjalani rawat jalan di RSUD di Denpasar. Jumlah sampel yang dibutuhkan adalah 120 orang dengan kriteria usia antara 30-50 tahun, berdomisili di Kota Denpasar dan bersedia ikut dalam penelitian. Setiap pasien DM yang berobat ke RSUD Kota Denpasar, yang memenuhi kriteria dipakai sebagai sampel. Pemilihan sampel berikutnya dipilih dengan cara yang sama sampai jumlah sampel yang dibutuhkan terpenuhi.

- **Quota sampling:**

Quota sampling adalah satu metode sampling dimana populasinya tidak homogen sehingga populasinya perlu dipilah menjadi strata-strata dan masing-masing strata diberikan jatah sampel agar sampel mewakili semua strata. Besar kuota untuk masing-masing strata bisa ditentukan secara proporsional (berdasarkan jumlah masing-masing strata) atau dengan jumlah yang sama (*equal sample size*). Pemilihan sampel pada masing-masing strata dilakukan secara non *random*. Quota sampling ini merupakan versi non-*random* dari *stratified random sampling*.

Contoh:

Penelitian bertujuan mengetahui perilaku seksual berisiko pada supir. Populasi supir dikelompokkan menjadi supir truk antar kota, sopir travel, sopir taksi, dan sopir angkot. Jadi ada tiga kelompok atau strata berdasarkan jenis angkutan yang dikemudikan. Dari masing-masing kelompok dipilih 30 sampel secara *consecutive*.

- **Judgmental or Purposive sampling:**

Purposive sampling adalah metode sampling dimana peneliti memilih sampel didasarkan pengetahuan penelitian tentang sampel yang akan dipilih. Peneliti hanya memilih sampel yang

oleh peneliti dianggap tepat untuk penelitiannya dan bisa memberikan informasi sesuai dengan kebutuhan penelitian. Metode ini tidak termasuk ke dalam metode pemilihan sampel yang *scientific (non-scientific method)* sehingga dampaknya dapat berpengaruh terhadap hasil penelitian karena pada metode ini terdapat cukup banyak kerancuan (“*involves a high amount of ambiguity*”). Namun, metode ini banyak dilakukan pada penelitian eksplorasi dengan pendekatan kualitatif, karena pada rancangan penelitian ini pemahaman dan pertimbangan peneliti dalam pemilihan sampel yang sesuai dengan tujuan penelitian merupakan aspek yang penting

Contoh:

Penelitian dilakukan untuk mengetahui pendapat masyarakat tentang penerapan kawasan tanpa rokok di Kota Denpasar. Peneliti menetapkan sampelnya adalah masyarakat umum dewasa, pejabat publik, pengelola kawasan tanpa rokok, dengan pertimbangan variasi informasi dari responden dengan sudut pandang yang berbeda.

- ***Snowball sampling:***

Pemilihan sampel secara *snowball* (bola salju) merupakan cara pemilihan sampel yang dimulai dari satu atau beberapa sampel awal (*initial samples*) yang bisa ditemukan, kemudian sampel tersebut diminta bantuannya untuk menunjukkan teman atau kenalannya yang memiliki karakteristik yang sesuai dengan karakteristik sampel dan memperkenalkannya kepada peneliti. Cara penemuan sampel tersebut diulang-ulang sampai jumlah sampel yang dibutuhkan terpenuhi. Dari satu sampel bertambah menjadi 2, dari dua menjadi 4, dari 4 menjadi 8, dan akan terus bertambah seperti bola salju bila digelindingkan maka akan bertambah besar. Metode ini dipilih apabila peneliti tidak bisa dan tidak mengetahui dimana populasi tersebut berada, misalnya populasi penelitiannya adalah pemakai obat bius sementara keberadaan mereka selalau diawasi oleh kepolisian atau kita sebut populasinya adalah kelompok *underground*.

Contoh:

Suatu penelitian akan mempelajari perilaku seksual lelaki yang suka selingkuh. Peneliti tidak mungkin bisa membuat *list* dari mereka, maka peneliti datang ke tempat-tempat minum dan bar, bercakap-cakap dengan pengunjung sampai ketemu dengan orang yang selingkuh. Lalu peneliti melakukan pendekatan sampai sangat akrab dan melakukan wawancara secara mendalam. Selanjutnya peneliti minta bantuan sampel tadi memperkenalkan dirinya kepada teman lainnya yang memiliki perilaku yang sama. Peneliti melakukan pendekatan kepada mereka dan bila sudah

dekat dilakukan wawancara. Hal ini dilakukan terus-menerus sampai jumlah sampel yang dibutuhkan terpenuhi.

2.8.3 Indikasi *non-random sampling*

Pemilihan sampel secara *non-random* dipakai apabila:

- Penelitiannya merupakan studi kualitatif (*qualitative research*), *pilot studies*, atau *exploratory research*.
- Waktu dan biaya penelitian sangat terbatas.
- Penelitiannya memerlukan pengamatan secara mendalam.
- Peneliti tidak bisa membuat *list* populasi atau populasi sulit untuk diidentifikasi keberadaannya
- Peneliti tidak bermaksud untuk mengeneralisasi hasil penelitian ke seluruh populasi.

2.8.4 Kelebihan metode *non-random sampling*

Metode *non-random* mempunyai beberapa keuntungan seperti:

- Metode *non random* lebih kondusif dan lebih praktis bagi penelitian survei. Hasil penelitian yang menggunakan pemilihan sampel secara *non-random* mirip dengan hasil bila sampel dipilih secara *random* asalkan jumlah sampelnya adekuat dan cara penggumpulan datanya benar.
- Pemilihan sampel secara *non-random* lebih cepat dan lebih murah dibandingkan dengan metode *random*.
- Respon responden yang dipilih secara *non-random* cenderung lebih baik dibandingkan sampel yang dipilih secara *random*, karena motivasi mereka untuk berpartisipasi lebih tinggi.

2.8.5 Perbedaan metode *non-random* dengan metode *random*.

Tabel 2.2. Perbedaan metode *non random* dan metode *random*

<i>Non-probability sampling</i>	<i>Probability sampling</i>
1. Sampel dipilih berdasarkan pertimbangan subyektif dari peneliti.	1. Sampel dipilih secara <i>random</i>
2. Tidak setiap individu mempunyai kesempatan yang sama untuk terpilih.	2. Setiap individu mempunyai kesempatan yang sama terpilih atau kesempatannya bisa dihitung

<i>Non-probability sampling</i>	<i>Probability sampling</i>
3. Dipakai bila sampelnya berasal dari satu kelompok yang memiliki keterkaitan (<i>has similar traits</i>).	3. Dipakai bila populasinya bervariasi.
4. Sampel tidak akurat mewakili populasi.	4. Sampel akurat mewakili populasi.

2.8.6 Latihan

Kasus

Dinas Kesehatan Provinsi Bali akan melakukan survei perilaku seksual berisiko laki-laki seks laki-laki (LSL) di Kota Denpasar dan Kabupaten Badung. Berdasarkan penelitian di kota Jakarta dan Bandung dilaporkan sekitar 60% LSL memiliki perilaku seksual berisiko. Kelompok LSL di Denpasar dan Badung tidak diketahui keberadaannya.

Tugas:

1. Tentukan siapa yang menjadi populasi target, terjangkau, dan sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel yang dibutuhkan bila tingkat kepercayaan sampel yang diinginkan 95%, *margin of error* 10%, dan jumlah populasi LSL diperkirakan $N = 550$ orang?
4. Jelaskan metode sampling apa yang akan dipakai dan cara pemilihan sampelnya!

2.9 Survei Cepat (*Rapid survey*)

2.9.1 Pengertian

Metode Survei Cepat (*Rapid Survey Method*) saat ini memiliki peran yang sangat besar pada Ilmu Kesehatan Masyarakat, khususnya Epidemiologi. Penelitian ini merupakan upaya pengumpulan informasi dari sebagian populasi yang dianggap dapat mewakili populasi tertentu. Metode survei cepat pertama kali diperkenalkan pada tahun 1965. Pada saat itu, metode ini digunakan untuk membantu dinas kesehatan daerah mempelajari status imunisasi balita di wilayahnya (Serfling & Sherman, 1965). Sebagai metode pengumpulan informasi yang berasal dari masyarakat (*population-based information*), metode survei cepat memiliki beberapa ciri khas, yaitu:

- 1) Dipergunakan untuk mengukur kejadian yang sering terjadi di masyarakat.
- 2) Pengambilan sampel secara klaster dua tahap, dimana untuk tiap wilayah diambil sebanyak 30 klaster dan pada masing-masing klaster diambil sebanyak 7 responden (Metode WHO-EPI/*extended program for immunization*)
- 3) Jumlah pertanyaan hanya dibatasi 20 sampai dengan 30 item pertanyaan saja.
- 4) Rancangan sampel, pemilihan sampel, pemasukan, pengolahan dan analisis data dilakukan dengan bantuan komputer (program CSURVEY dan CSAMPLE yang menggunakan asumsi klaster dua tahap).
- 5) Waktu pelaksanaan sampai dengan penyusunan laporan hanya berkisar 2-3 minggu saja.
- 6) Hasil survei disajikan dengan menggunakan teknik statistik yang sederhana dengan tetap memperhatikan kaidah statistik yang berlaku.

2.9.2 Indikasi

Survei cepat banyak dipakai untuk mendapatkan informasi penting yang bisa diperoleh secara cepat untuk mengkonfirmasi informasi yang bersumber dari program atau sumber lain atau digunakan sebagai dasar dalam pengambilan keputusan. Metode survei cepat di bidang kesehatan sering dipakai untuk mengevaluasi keberhasilan atau capaian program kesehatan, misalnya program imunisasi dasar atau program lainnya.

2.9.3 Kelebihan dan kekurangan

Survei cepat merupakan satu metode survei yang dilakukan dalam waktu yang singkat, sederhana, dan murah. Hasil survei cepat dapat dipakai dasar bagi pengambil keputusan dalam waktu yang singkat.

Survei cepat juga memiliki kekurangan karena variabel yang disurvei jumlahnya terbatas dan superfisial sehingga tidak dapat menjelaskan secara mendalam mengapa peristiwa tersebut terjadi.

2.9.4 Cara penentuan besar sampel

1) Penghitungan dengan *design effect*

Penentuan besar sampel survei cepat sama dengan penentuan besar sampel survei dengan metode acak kelompok (*cluster sampling*), karena cara pemilihan sampel pada survei cepat menggunakan metode acak kelompok dua tingkat dengan pemilihan sampel klaster secara PPS. Besar sampel ditentukan dengan rumus seperti berikut.

$n = \frac{Z_{\alpha}^2 p(1-p)}{d^2} \times DE$
<p>Keterangan: n = jumlah sampel individu p = estimasi kejadian di populasi d = <i>margin of error</i> (presisi) Z_{α} = Z-skor untuk tingkat kepercayaan (1-α) DE = <i>Design Effect</i></p>

Rumus di atas tersebut dipakai menentukan jumlah sampel individu yang dibutuhkan. Misalnya dalam survei cepat cakupan imunisasi dasar di Kota Denpasar, anak balita menjadi populasi terjangkaunya yang sekaligus juga menjadi unit analisis. Dalam survei cepat pada contoh di atas, yang dijadikan unit sampel adalah Rumah Tangga (RT) bukan individu. Penentuan jumlah sampel RT dari survei cepat tersebut dihitung dari jumlah sampel individu yang dibutuhkan dibagi dengan rerata jumlah *eligible* individu (dalam contoh di atas adalah jumlah anak balita) di setiap RT. Misalnya rerata jumlah *eligible* individu adalah r, maka jumlah sampel RT menjadi seperti dibawah ini.

$n = \frac{Z_{\alpha}^2 p(1-p)}{d^2 r} \times DE$
<p>Keterangan: n = jumlah sampel RT p = estimasi kejadian di populasi d = <i>margin of error</i> (presisi)</p>

r = rerata jumlah *eligible* individu di setiap unit sampel
 Z_{α} = Z-skor untuk tingkat kepercayaan $(1-\alpha)$
 DE = *Design Effect*

Contoh:

Survei cepat dilakukan untuk mengetahui cakupan imunisasi dasar di Kota Denpasar dengan populasi terjangkau adalah anak Balita. Sampel akan dipilih dengan metode kluster dua tingkat, dimana pada tingkat pertama akan dipilih desa (sebagai kluster atau PSU) dan pada tahap ke dua akan dipilih sampel RT. Semua anak balita pada RT terpilih akan menjadi sampel Balita. Besar sampel minimal bila diasumsikan capaian imunisasi dasar “P” = 85%, estimasi rerata Balita di setiap RT “r” = 1,5, *margin of error* D = 5%, reliabilitas sampel 95%, dan *Design Effect* “DE” = 1,75 bisa ditentukan seperti di bawah ini.

Penentuan jumlah sampel RT:

Dari kasus di atas diketahui bahwa:

$$Z_{\alpha} = 1,96$$

$$P = 0,85$$

$$\text{Margin of error “D”} = 0,05$$

$$\text{Rerata jumlah balita per RT “r”} = 1,5$$

Berdasarkan parameter tersebut, maka jumlah sampel RT yang dibutuhkan menjadi:

$$n = 1,96^2 \times 0,85 \times 0,15 \times 1,75 / (0,05^2 \times 1,5) = 228,57 \text{ atau dibulatkan menjadi } 229 \text{ RT.}$$

Penentuan jumlah sampel Desa (kluster):

Apabila besar sampel RT di setiap kluster (desa) ditetapkan sebanyak 7 RT, maka jumlah sampel kluster (desa) yang dibutuhkan menjadi $229/7 = 32,7$ atau dibulatkan menjadi 33 kluster (desa).

2) Cara menentukan besar sampel survei cepat menurut metode WHO-EPI:

WHO menetapkan penentuan besar sampel survei cepat dengan metode 30×7 , yaitu jumlah sampel kluster sebanyak 30 kluster untuk setiap wilayah dan jumlah sampel RT per kluster sebanyak 7, sehingga total sampel menjadi 210 RT per wilayah. Misalnya, untuk contoh survei cepat cakupan imunisasi dasar di Kota Denpasar di atas, besar sampel minimalnya adalah 210 RT yang dipilih dari 30 sampel desa (kluster) dan 7 rumah tangga per kluster

Koreksi Rumus Sampel Survei Cepat metode WHO

Untuk menilai apakah besar sampel menurut metode WHO (30 x 7) itu sudah cukup bila dibandingkan dengan besar sampel yang dihitung dengan rumus besar sampel kluster, maka perlu dihitung Varians maximum (V_{\max}) berdasarkan metode WHO dan besar Varians maximum menurut perhitungan besar sampel kluster. Kemudian kedua varians tersebut dibandingkan, dikatakan besar sampel metode WHO tersebut adekuat bila nilai V_{\max} lebih kecil dari perhitungan besar sampel kluster

Cara penghitungan Varians Maximum metode sampling kluster:

$$V_{\max} = (p \times q/n) \times Deff$$

Pada contoh di atas: $p = 0,85$, $n = 229$, dan $DEF = 1,75$, sehingga:

$$V_{\max} = 0,85 \times 0,15 \times 1,75 / 229 = 0,00097$$

Cara penghitungan Varians Maximum metode WHO:

$$V_{\max} = (p \times q/n) \times Deff$$

Pada contoh di atas: $p = 0,85$, $n = 210$, dan $DEF = 1,75$, sehingga:

$$V_{\max} = 0,85 \times 0,15 \times 1,75 / 210 = 0,0011$$

Simpulan:

Pada contoh di atas didapatkan bahwa V_{\max} metode WHO = 0,0011 dan V_{\max} metode kluster = 0,00097. Berarti V_{\max} sampel menurut metode WHO > V_{\max} sampel menurut rumus besar sampel kluster, oleh karena itu disimpulkan bahwa jumlah sampel RT menurut metode WHO belum memadai, maka perlu ditambah jumlah sampel klasternya.

2.9.5 Cara pemilihan sampel

Pemilihan sampel pada survei cepat dilakukan dengan metode kluster sampling dua tingkat. Pada tingkat pertama dilakukan pemilihan sampel kluster (PSU) dengan metode *Probability Proportional to Size* (PPS) sebanyak sampel kluster yang dibutuhkan, seperti yang sudah dijelaskan sebelumnya. Pada tahap kedua akan dilakukan pemilihan sampel RT dengan cara berikut.

Step 1: petugas survei mendatangi lokasi sampel kluster, misalnya desa A dan menuju ke pusat desa atau persimpangan jalan di pusat desa. Kemudian memilih secara acak satu arah di persimpangan tersebut dengan memutar botol atau pensil dan memilih arah yang ditunjukkan

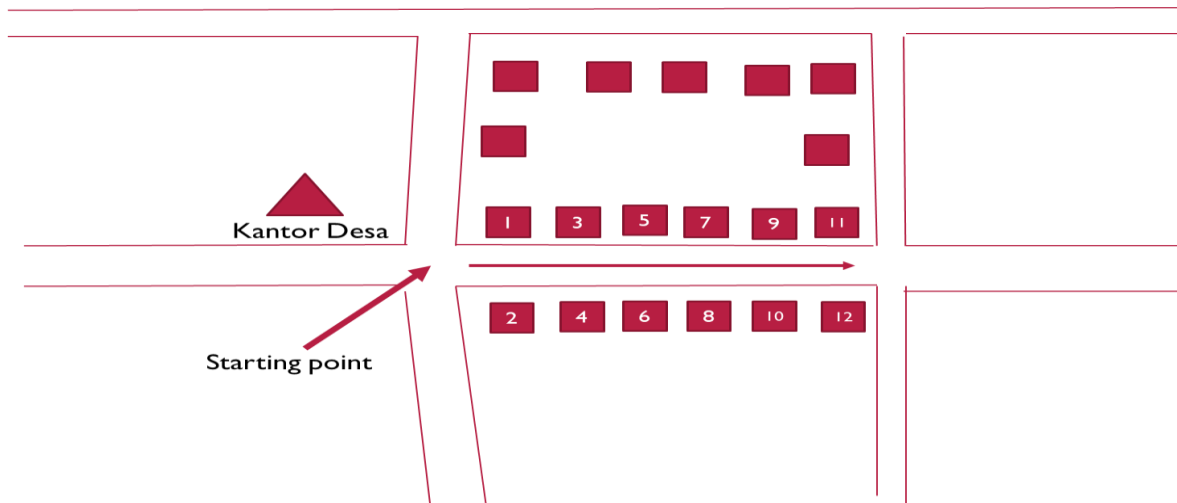
oleh ujung botol/pensil atau bisa juga menggunakan *random spin dial* pada software CSURVEY

Step 2: membuat *list* RT yang terdapat di satu arah dari simpang jalan tadi sampai persimpangan berikutnya atau sampai ujung.

Step 3: memilih satu RT dari *list* tadi secara *simple random* sebagai *starting point* pemilihan sampel.

Step 4: pengumpulan data survei dimulai dari RT yang menjadi titik awal sampel tadi. Apabila RT tersebut memenuhi persyaratan sampel (misalkan ada anak balita seperti contoh kasus diatas), maka data imunisasi dasar pada semua balita yang ada di RT tersebut didata.

Step 5: Sampel berikutnya dipilih dari tetangga sebelah (*next door*) sampel pertama, bila RT tersebut memenuhi persyaratan (ada balita), maka dilakukan pendataan kelengkapan imunisasi pada balita yang ada. Sampel berikutnya, dipilih dari tetangga sampel ke dua dan seterusnya sampai 7 sampel RT di kluster tersebut terpenuhi. Berikut adalah bagan cara pemilihan sampel di satu kluster.



Gambar 2.5. Bagan pemilihan sampel RT di satu Kluster (desa)

2.9.6 Latihan

Kasus:

BKKBN Provinsi Bali akan melakukan survei cepat untuk mengetahui akseptor *rate* dan jenis metode kontrasepsi yang sedang dipakai. Populasi survei adalah Pasangan Usia Subur (PUS) yang berdomisili di Provinsi Bali yang tersebar di 6500 desa. Sampel akan dipilih menggunakan kluster sampling dua tingkat dengan metode pemilihan PPS dimana Desa sebagai kluster (PSU). Estimasi

angka akseptor sebesar 64%, *margin of error* survei $d = 5\%$, tingkat kepercayaan sampel 95%, *Design Effect* (DE) = 1,75, dan rerata jumlah PUS per RT = 1,2.

TUGAS:

1. Jelaskan populasi terjangkau dari survei tersebut!
2. Jelaskan apa yang menjadi PSU dari survei tersebut!
3. Jelaskan siapa yang menjadi unit analisis dari survei tersebut!
4. Demonstrasikan cara penghitungan besar sampel minimal dari survei tersebut!
5. Analisis apakah penentuan besar sampel dengan metode 30 x 7 (WHO) memadai!
6. Jelaskan *Step-Step* pemilihan sampelnya!

BAGIAN 3. METODE SAMPLING PENELITIAN TRIAL

dr. I Ketut Tangking Widarsa, MPH.*

dr. Ayu Swandewi, MPH., PhD.**

*)Dosen Biostatistiks Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com

***) Dosen Biostatistiks Fakultas Kedokteran Universitas Udayana, Bali. E-mail: ayuswandewi@unud.ac.id*Dosen

Tujuan Pembelajaran:

- Memahami pengertian penelitian trial.
- Memahami cara penentuan besar sampel penelitian trial.
- Memahami cara pemilihan sampel dan alokasi sampel

Pokok Bahasan:

3.1 Pengantar

3.2 Trial Klinik Paralel

3.3 Trial Klinik Menyalang (*Cross-over Design*)

3.4 Cara penentuan besar sampel

3.5 Cara alokasi sampel secara *random*

Bahan Bacaan

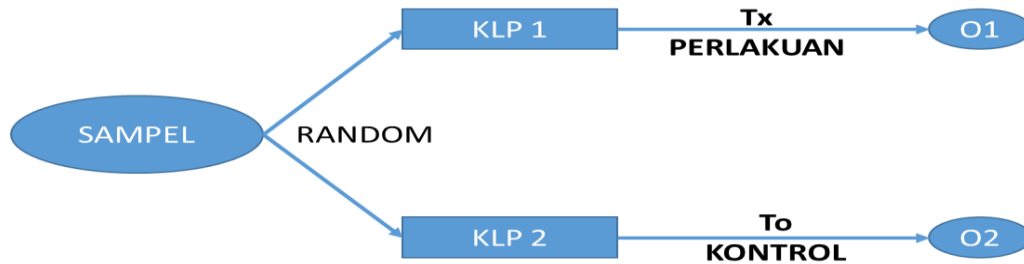
1. Alenjandro R. Jayad, Murray W. Enkin. 2007. Randomized Controlled Trial: Question, Answers and Musings, Oxford, Blackwell.
2. Cochran W.G. 1977. Sampling Techniques, Third edition, John Wiley&Sons, New York.
3. Levy P.S. and Lemeshow S. 1980. Sampling for Health Professionals, Lifetime Learning Publication, California.
4. Lwanga S.K. and Lemeshow S. 1991. Sample Size Determination in Health Studies, World Health Organization, Geneva.
5. Pocock SJ. 2008. Clinical Trial: A Practical Approach, John Wiley&Sons, New York,

3.1 Pengantar

Penelitian kedokteran dan kesehatan dibedakan menjadi penelitian eksperimental dan observasional berdasarkan ada atau tidaknya perlakuan. Penelitian eksperimen adalah penelitian yang ditandai oleh adanya perlakuan, randomisasi, replikasi, dan adanya kelompok kontrol. Trial klinik merupakan salah satu jenis penelitian eksperimental dengan subyek pasien (orang sakit). Trial klinik dibedakan berdasarkan tahap atau fasenya menjadi 4 tahap, yaitu: 1) trial klinik tahap 1 yang biasanya dilakukan pada sampel kecil untuk melihat keamanan (*safety*) dan menentukan dosis yang sesuai (*dosing*), 2) tahap 2 dilakukan pada lebih banyak pasien/individu untuk menilai efektivitas dari obat/intervensi dan juga menilai lebih lanjut keamanannya, 3) tahap 3 dilakukan pada sampel yang besar untuk konfirmasi efektivitas, menilai keamanan, efek samping, dan 4) tahap 4 yang juga disebut *post marketing clinical trial* yang bertujuan untuk melihat efektivitas obat, efek samping di setting sebenarnya setelah obat diijinkan beredar di populasi. Berdasarkan perlakuan yang diterima oleh subyek, rancangan trial klinik dibedakan menjadi trial klinik paralel dan menyilang atau *cross-over design*. Validitas trial klinik ditentukan oleh tingkat validitas eksternal dan validitas internal. Validitas interval berkaitan dengan kualitas data dan pengendalian perancu, sedangkan validitas eksternal ditentukan oleh representativitas sampel penelitian. Representativitas berhubungan dengan besar dan cara pemilihan serta alokasi perlakuan kepada sampel/subyek penelitian. Pada sesi ini akan dibahas tentang cara penentuan besar sampel serta alokasi sampel secara *random*.

3.2 Trial klinik paralel

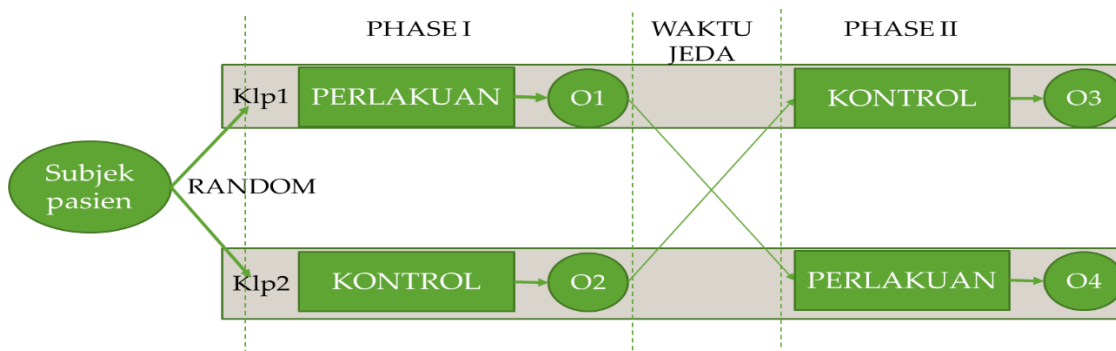
Trial klinik paralel adalah salah satu bentuk rancangan trial klinik dimana sampel penelitian dialokasikan secara *random* menjadi kelompok-kelompok trial dan perlakuan yang berbeda diberikan pada tiap-tiap kelompok. Misalnya pada trial efek obat *Candesartan* terhadap penurunan proteinuri (ditemukannya protein dalam urine) penderita hipertensi dengan kelainan fungsi ginjal. Sebanyak 120 penderita hipertensi dengan gangguan fungsi ginjal dipakai sebagai sampel, setengahnya yaitu 60 diberikan obat dan sisanya diberikan *placebo*. Pengalokasian sampel ke masing-masing kelompok trial dilakukan secara *random* (*random allocation*). Pada rancangan ini, sampel hanya mendapat satu perlakuan saja, bisa mendapat obat atau placebo saja, seperti bagan berikut.



Gambar 3.1. Bagan rancangan trial klinik paralel

3.3 Trial klinik menyilang (Cross-over design)

Trial klinik menyilang atau juga disebut trial klinik *cross-over design* adalah rancangan trial dimana sampel dibagi menjadi dua kelompok atau lebih dan tiap kelompok mendapatkan semua jenis perlakuan. Perlakuan diberikan sebanyak dua tahap dengan diselingi adanya waktu jeda yang disebut “*washing-out*”. Sampel dibagi menjadi dua kelompok dan pada tahap pertama kelompok pertama mendapat perlakuan obat (perlakuan) dan pada tahap kedua mendapat *placebo* (kontrol). Sebaliknya, kelompok ke dua pada tahap pertama mendapat *placebo* (kontrol) dan pada tahap kedua mendapatkan obat (perlakuan), seperti bagan (Gambar 3.2).



Gambar 3.2. Bagan rancangan trial klinik menyilang (*cross-over design*)

3.4 Cara penentuan besar sampel

3.4.1 Parameter penentuan besar sampel

Penentuan besar sampel penelitian trial klinik tergantung dari variasi kejadian yang akan diteliti di populasi, besarnya efek yang ingin diuji, level kepercayaan dan *power* penelitian yang diinginkan.

Variabilitas:

Variabilitas kejadian yang diteliti akan memengaruhi besar sampel penelitian. Makin lebar variabilitas kejadian yang diteliti, makin besar sampel yang dibutuhkan. Variabilitas ditentukan dari angka standar deviasi (SD) pada variabel numerikal. Misalnya akan diteliti efek probiotik terhadap waktu perawatan pasien diare akut pada anak balita, maka angka SD lama rawat dipakai sebagai ukuran variabilitas kejadian lama perawatan tersebut. Pada variabel kategorikal, variabilitas ditentukan dari $\sqrt{p(1-p)}$. Misalnya akan diteliti efek probiotik terhadap angka kesembuhan diare akut pada anak balita, maka angka kesembuhan (p) dipakai menentukan besarnya variabilitas kejadian kesembuhan di populasi dengan rumus $SD = \sqrt{p(1-p)}$.

Besar Efek:

Besar efek yang ingin dibuktikan juga memengaruhi besar sampel penelitian. Makin kecil besar efek yang harus dapat dibuktikan sebagai efek yang nyata, maka makin banyak jumlah sampel yang dibutuhkan. Besar efek yang harus dapat dibuktikan dapat ditentukan dari besarnya perbedaan rerata ($\mu_1 - \mu_2$) antara kelompok perlakuan dengan kelompok kontrol atau dengan perbedaan risk ($p_1 - p_2$) atau resiko relatif (RR) antara kelompok perlakuan dengan kelompok kontrol.

Level kepercayaan dan *Power* Penelitian:

Besar sampel penelitian trial klinik juga dipengaruhi oleh level kepercayaan dan *power* penelitian. Makin tinggi level kepercayaan dan *power* penelitian yang diinginkan, maka makin banyak jumlah sampel yang dibutuhkan dan sebaliknya. Pada pengujian hipotesis nihil trial klinik terdapat dua kemungkinan kesalahan dalam penetapan hasil uji, yaitu *type I error* yang diberi simbol “ α ” dan *type II error* yang diberi simbol “ β ”. *Type I error* adalah menolak H_0 yang seharusnya diterima dan *type II error* adalah menerima H_0 yang seharusnya ditolak, seperti tabel berikut.

SAMPEL	POPULASI	
	Ho benar (tidak ada efek)	Ho salah (ada efek)
Ho diterima	Benar	Type II error (β)
Ho ditolak	Type I error (α)	Benar

Untuk meningkatkan reliabilitas dan *power* penelitian, maka kesalahan type I dan type II harus dirancang sekecil mungkin. Pada umumnya ditetapkan besarnya kesalahan type I (α) adalah 0,01 atau

0,05 atau tingkat kepercayaan (1- α) sampel 99% atau 95%. Sedangkan kesalahan type II (β) ditetapkan 10% atau 20% atau *power* penelitian (1- β) sebesar 90% atau 80%.

3.4.2 Penghitungan besar sampel untuk hipotesis perbedaan *mean* (untuk data numerik).

Hipotesis statistik yang diuji:

Ho: $\mu_0 = \mu_1$ (tidak ada perbedaan *mean*) dan Ha: $\mu_0 \neq \mu_1$ (terdapat perbedaan *mean*)

Parameter penentuan besar sampel

- 1) SD hasil pada terapi standar atau kontrol
- 2) *Effect size* atau beda mean ($\mu_c - \mu_t$) terkecil yang harus dapat diidentifikasi nyata (*clinically significant*).
- 3) Alpha (kesalahan type 1), biasanya dipakai 0,01 atau 0,05.
- 4) Beta (kesalahan type 2), biasanya dipakai 0,10 sampai 0,20.

Rumus besar sampel trial untuk variabel outcome numerik	
$n = \frac{2\sigma^2}{(\mu_2 - \mu_1)^2} \cdot x f(\alpha, \beta)$	
Keterangan: n=jumlah sampel untuk satu kelompok σ = SD <i>outcome</i> variabel μ_1 = rerata <i>outcome</i> variabel kelompok trial μ_2 = rerata <i>outcome</i> variabel kelompok standar/pembanding α = tingkat kesalahan type I β = tingkat kesalahan type II	

Tabel 3.1. Nilai $f(\alpha, \beta)$ yang digunakan untuk menentukan besar sampel trial

		β (type II error)			
		0,05	0,10	0,20	0,50
α (type I error)	0,10	10,8	8,6	6,2	2,7
	0,05	13,0	10,5	7,9	3,8
	0,02	15,8	13,0	10,0	5,4
	0,01	17,8	14,9	11,7	6,6

Sumber: Pocock SJ. 2008. Clinical Trial: A Practical Approach, New York, John Wiley&Sons: 125.

Cara penggunaan tabel $f(\alpha, \beta)$ adalah sbb:

Misalnya kesalahan type I = 0,05 dan kesalahan type II = 0,20, maka nilai $f(\alpha, \beta) = 7,9$ (lihat arah panah pada tabel di atas).

Contoh:

Misalnya akan diteliti efektivitas obat baru (B) terhadap gula darah penyakit DM. Dari data sekunder diketahui kadar gula setelah mendapatkan obat standar (A) adalah 130 mg/dl dengan SD 25. Besar sampel yang diperlukan agar obat B dapat dinyatakan lebih efektif bila menurunkan kadar gula paling sedikit 10 mg/dl dari obat standar dengan tingkat kesalahan type 1 sebesar 0,01 (reliabilitas sampel 99%) dan kesalahan type 2 sebesar 0,20 (*power* penelitian 80%) serta kemungkinan *drop-out* 10%, dapat dihitung sebagai berikut.

Diketahui:

$$\mu_c = 130 \text{ mg/dl} \quad \mu_t = 130 - 10 = 120 \text{ mg/dl}$$

$$SD = 25 \text{ mg/dl}$$

$$f(\alpha, \beta) = 11,7$$

Jumlah sampel yang diperlukan dihitung dengan rumus berikut:

$$n = \frac{2\sigma^2}{(\mu_2 - \mu_1)^2} \times f(\alpha, \beta)$$
$$n = \frac{2 \times 25^2}{10^2} \times 11,7 = 146,25$$

Berdasarkan perhitungan tersebut di atas, maka jumlah sampel minimal yang diperlukan adalah $2 \times 147 = 294$.

3.4.3 Penghitungan besar sampel untuk hipotesis perbedaan insiden (variabel nominal)

Hipotesis:

Ho: $\pi_0 = \pi_1$ (tidak ada perbedaan insiden)

Ha: $\pi_0 \neq \pi_1$ (terdapat perbedaan insiden)

Parameter penentuan besar sampel:

- 1) Insiden pada terapi standar atau kontrol (p_2).
- 2) *Effect size* atau beda insidens rate ($p_1 - p_2$) terkecil yang harus dapat diidentifikasi sebagai efek yang nyata atau signifikan.
- 3) Alpha (kesalahan type 1), biasanya dipakai 0,01 atau 0,05.
- 4) Beta (kesalahan type 2), biasanya dipakai 0,10 sampai 0,20.

Rumus besar sampel trial untuk variabel *outcome* nominal

$$n = \frac{p_1(1-p_1) + p_2(1-p_2)}{(p_1 - p_2)^2} x f(\alpha, \beta)$$

Keterangan:

n=jumlah sampel untuk satu kelompok
 p1= kejadian *outcome* pada obat trial
 p2= kejadian *outcome* pada obat standar
 α = tingkat kesalahan tipe I
 β = tingkat kesalahan tipe II

Contoh (data nominal)

Misalnya akan diteliti efek samping muntah berat obat baru (B) pada pasien. Dari data sekunder diketahui insiden kejadian muntah berat dengan obat standar (A) adalah $p_2 = 20\%$. Besar sampel yang diperlukan agar obat B dapat dinyatakan memiliki efek samping muntah berat yang lebih rendah bila perbedaannya sebesar 10% dengan tingkat kesalahan type 1 sebesar 1% (reliabilitas sampel 99%) dan kesalahan type 2 sebesar 20% (*power* penelitian 80%).

Diketahui:

$$P_1 = 20\%$$

$$p_1 - p_2 = 10\%$$

$$p_2 = 20\% - 10\% = 10\%$$

$$f(\alpha, \beta) = 11,7$$

Jumlah sampel yang diperlukan dihitung dengan rumus berikut:

$$n = \frac{p_1(1-p_1) + p_2(1-p_2)}{(p_1 - p_2)^2} x f(\alpha, \beta)$$

$$n = ((0.20(0.8) + 0.1(0.9)) / (0.1)^2) (11.7)$$

$$n = ((0.16 + 0.09) / 0.01) x 11.7$$

$$n = 292.5$$

Berdasarkan perhitungan tersebut di atas, maka jumlah sampel minimal yang diperlukan adalah $2 \times 293 = 586$.

3.4.3 Penghitungan besar sampel Trial Negatif

Pengertian:

Trial negatif adalah trial yang tujuannya untuk menguji efektivitas satu obat baru dibandingkan dengan obat standard dimana efek obat yang baru diasumsikan lebih rendah dari obat standard. Penelitian trial negative dilakukan untuk mengatasi masalah efek samping yang berat dan atau harga obat standar yang tidak terjangkau, sehingga perlu dicari obat lain yang efeknya dihipotesiskan lebih rendah dari obat standar tetapi efek samping lebih ringan dan harga lebih terjangkau. Misalnya A adalah obat standard yang sangat efektif tetapi efek samping berat dan harga tidak terjangkau dengan efektivitas 85%. Obat B adalah obat lain yang sejenis dengan efektivitasnya 10% di bawah obat standard, tetapi efek samping lebih kecil dan harga lebih terjangkau. Obat B akan diuji apakah efeknya sama dengan obat standard A dengan rancangan trial klinik parallel tersamar ganda dengan hipotesis statistik:

Ho: $\pi_1 = \pi_2$ (tidak ada perbedaan)

Ha: $\pi_1 \neq \pi_2$ (ada perbedaan)

Cara penghitungan besar sampel:

Cara penentuan besar sampel pada trial negatif ditentukan dengan rumus sebagai berikut.

Rumus Penghitungan Besar Sampel Trial Negatif
$n = \frac{2p(1-p)}{d^2} f(\alpha, \beta)$
<p>Keterangan:</p> <p>n=jumlah sampel untuk satu kelompok</p> <p>p= kejadian <i>outcome</i> pada obat standar</p> <p>d = <i>effect size</i> (beda <i>mean</i> atau beda proporsi)</p> <p>α = tingkat kesalahan tipe I</p> <p>β = tingkat kesalahan tipe II</p>

CONTOH 3 (trial negatif)

Misalnya diketahui bahwa obat *amitriptyline* memberikan efek penyembuhan terhadap 70% pasien, tetapi obat ini memiliki banyak efek samping serta harganya tidak terjangkau. Oleh karena itu peneliti akan menguji efek *antidepressant* B. Suatu *random* trial klinik dilakukan terhadap obat *antidepressant* B. Obat ini baru dapat dipakai sebagai pengganti *amitriptyline* bila memberi efek samping 10% di bawah *amitriptyline*. Besar sampel yang diperlukan bila peneliti menginginkan sampelnya memiliki tingkat reliabilitas 95% dan *power* 80%, dihitung sebagai berikut

Diketahui:

$$p = 0,70$$

$$d = 0,10$$

$$\alpha = 0,05$$

$$\beta = 0,20$$

Jumlah sampel yang diperlukan dihitung dengan rumus berikut.

$$n = \frac{2p(1-p)}{d^2} f(\alpha, \beta)$$

$$n = \frac{2 \times 0,70 \times 0,30}{0,10^2} \times 7,9 = 331,8$$

Berdasarkan perhitungan tersebut di atas, maka jumlah sampel minimal yang diperlukan adalah $2 \times 332 = 664$.

3.5 Alokasi Sampel Secara *Random*

Alokasi subyek trial secara *random* dapat dilakukan dengan acak sederhana menggunakan tabel bilangan *random* dan juga dapat dilakukan dengan menggunakan blok *random* (*permuted random*).

3.5.1 Alokasi *Random* dengan Bilangan *Random*

Misalnya jumlah sampel sebuah trial dengan dua perlakuan A dan B adalah 10 orang. Alokasi sampel akan dilakukan secara *random* menggunakan tabel bilangan *random*. Cara alokasinya adalah sebagai berikut:

- 1) Buat *list* nomor urut subyek dari 1 s/d 10
- 2) Pilih sepuluh bilangan *random* pada tabel bilangan *random* dan tulis pada *list* nomor urut subyek
- 3) Subyek dengan bilangan *random* yang termasuk 50% terendah diberikan perlakuan A dan sisanya diberikan perlakuan B, seperti contoh berikut.

Nomor Urut Subyek	Bilangan <i>Random</i>	Perlakuan
1.	118	A
2.	728	B
3.	754	B
4.	346	A
5.	964	B
6.	178	A
7.	105	A
8.	433	B
9.	961	B
10.	101	A

3.5.2 Alokasi *Random* dengan Blok *Random*

Misalnya jumlah sampel sebuah trial dengan dua perlakuan A dan B adalah 10 orang. Alokasi sampel akan dilakukan secara *random* menggunakan blok *random*. Cara alokasinya adalah sebagai berikut:

- 1) Buat blok *random*, misalnya dalam satu blok terdapat 4 orang dengan 2 jenis perlakuan A dan B, maka terdapat 6 kemungkinan kombinasi A dan B. Kemudian buat juga pembagian bilangan *random* antara 1 s/d 100 menjadi 6 rentangan sesuai dengan jumlah kombinasi blok.:

Bilangan <i>Random</i>	1 - 17	18 - 34	35 -51	52 - 66	67 - 83	84-100
Blok <i>Random</i>	AABB	ABBA	BBAA	BAAB	ABAB	BABA

- 2) Tentukan jumlah blok *random* yang akan dipilih. Pada kasus di atas akan dipilih secara *random* sebanyak 3 blok (berarti akan terdapat 12 sampel individu)
- 3) Pilih 3 bilangan *random* (sesuai jumlah blok sampel yang diperlukan) secara *random* dengan nilai tiga digit dari belakang antara 1 s/d 100. Misalnya, bilangan *random* yang terpilih adalah: 047, 069, 057. Blok sampel yang terpilih adalah blok 3 (BBAA), blok 5 (ABAB), dan blok 4 (BAAB)
- 4) Buat alokasi subyek berdasarkan blok sampel yang terpilih, sbb:

Nomor Urut Subyek	Jenis Perlakuan	Block <i>Random</i> terpilih
1.	B (placebo)	BBAA
2.	B (placebo)	
3.	A (obat)	
4.	A (obat)	
5.	A (obat)	ABAB
6.	B (placebo)	
7.	A (obat)	
8.	B (placebo)	
9.	B (placebo)	BAAB
10.	A (obat)	
11.	A (obat)	
12.	B (placebo)	

Hasil alokasi sampel akan dipakai oleh petugas Farmasi (yang diminta membuat obat dengan bentuk dan ukurannya sama persis antara perlakuan dan *placebo*) untuk membuat alokasi perlakuan sesuai hasil alokasi di atas. Kantong obat no 1 isinya *placebo*, no 2 isinya *placebo*, no 3 isinya obat, no 4 isinya obat, no 5 isinya obat, no 6 isinya *placebo*, no 7 isinya obat, no 8 isinya *placebo*, no 9 isinya *placebo*, no 10 isinya obat, no 11 isinya obat, dan no 12 isinya *placebo*.

Pemberi perlakuan akan memberikan obat kepada pasien sesuai no urut kantong obat 1 sampai 12. Pasien dan petugas yang memberikan obat tidak mengetahui jenis obat apa yang diberikan (*blind*).

3.6 Latihan

Kasus

Dr Indah, M.Sc, Sp.S adalah dokter spesialis saraf di Rumah Sakit Sanjiwani Gianyar akan meneliti efek simvastatin terhadap kekambuhan stroke pada pasien stroke non haemorrhagik yang dirawat di RSUD Sanjiwani. Pasien dengan serangan stroke non haemorrhagik pertama, usia 30 – 60 tahun, tidak menderita penyakit DM dijadikan subyek penelitian. Perlakuan penelitian adalah obat simvastatin dengan dosis 20 mg yang diberikan 1 kali sehari selama 30 hari dan kontrolnya diberikan *placebo*. Sampel akan dialokasikan secara *random* menggunakan bilangan *random* menjadi dua kelompok trial, yaitu kelompok yang mendapat obat dan kelompok *placebo*.

Tugas:

1. Tentukan siapa yang menjadi populasi target, terjangkau, dan sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya!
3. Berapa jumlah sampel yang dibutuhkan bila tingkat kesalahan tipe I $\alpha = 0,05$ dan kesalahan tipe II $\beta = 0,20$, dan *effect size* $RR = 1,5$?
4. Jelaskan metode sampling apa yang akan dipakai dan cara pemilihan sampelnya!
5. Jelaskan cara alokasi sampel secara *random* dengan bilangan *random*!

BAGIAN 4. METODE SAMPLING PENELITIAN

UJI DIAGNOSTIK

dr. I Ketut Tangking Widarsa, MPH.*

dr. Ayu Swandewi, MPH., PhD.**

**) Dosen Biostatistiks Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com*

****) Dosen Biostatistiks Fakultas Kedokteran Universitas Udayana, Bali. Email: ayuswandewi@unud.ac.id*

Tujuan Pembelajaran:

1. Memahami pengertian penelitian uji diagnostik.
2. Memahami cara penentuan besar sampel penelitian uji diagnostik.
3. Memahami cara pemilihan sampel penelitian uji diagnostik

Pokok Bahasan:

4.1 Pengantar

4.2 Penentuan Besar Sampel

4.3 Cara Pemilihan Sampel

Bahan Bacaan

1. Cochran W.G. 1977. Sampling Techniques, Third edition, John Wiley&Sons, New York.
2. Levy P.S. and Lemeshow S. 1980. Sampling for Health Profesionals, Lifetime Learning Publication, California.
3. Lwanga S.K. and Lemeshow S. 1991. Sample Size Determination in Health Studies, World Health Organization, Geneva.
4. Sopiudin Dahlan M. 2009. Besar Sampel dan Cara Pengambilan Sampel Dalam Penelitian Kedokteran dan Kesehatan, Edisi 2, Salemba Medika, Jakarta

4.1 Pengantar

Penelitian uji diagnostik adalah salah satu penelitian kedokteran yang digunakan untuk menguji akurasi satu alat *test* baru yang akan dipakai sebagai alat diagnostik atau alat skrining suatu penyakit tertentu. Untuk menentukan akurasi alat *test* baru tersebut, maka hasil pemeriksaan alat *test* baru akan dicocokkan (*cross-check*) dengan hasil pemeriksaan alat standar baku (*gold standard*). Akurasi alat *test* baru akan dinilai dari angka sensitivitas, spesifisitas, nilai *Likelihood Ratio* positif (LR+), nilai *Likelihood Ratio* negatif (LR-), nilai duga positif, dan nilai duga negatif.

Validitas internal hasil uji diagnostik ditentukan oleh beberapa faktor antara lain adalah akurasi *gold standard* yang dipakai, rancangan penelitian, penerapan *blindness* pada pengukuran dan analisis data. Sedangkan validitas eksternal penelitian ditentukan oleh beberapa faktor antara lain variabilitas subyek penelitian, jumlah sampel, dan cara pemilihan sampel. Pada Bagian ini akan dibahas tentang cara penentuan besar sampel dan cara pemilihan sampel untuk penelitian uji diagnostik.

4.2 Parameter penentuan besar sampel

Parameter yang menentukan besar sampel pada penelitian uji diagnostik adalah angka sensitivitas atau spesifisitas (p) yang ingin dimiliki *test* baru yang akan diuji, tingkat presisi sampel (D), level kepercayaan ($1-\alpha$), dan asumsi prevalensi penyakit di populasi (p_x).

Angka sensitivitas atau spesifisitas:

Besarnya angka sensitivitas atau spesifisitas terkecil yang harus dapat diidentifikasi sebagai angka yang nyata akan memengaruhi jumlah sampel yang dibutuhkan. Makin kecil angka sensitivitas atau spesifisitas yang harus dapat diidentifikasi sebagai tingkat akurasi yang nyata, makin banyak sampel yang dibutuhkan dan sebaliknya.

Level kepercayaan:

Besar sampel juga ditentukan oleh level kepercayaan sampel yang diinginkan. Makin tinggi level kepercayaan yang diinginkan, maka makin banyak sampel yang dibutuhkan. Level kepercayaan ditentukan dari nilai standar Zscore dari level kepercayaan yang dipakai. Apabila level kepercayaan yang dipakai adalah 95%, maka nilai $Z_{1-\alpha} = 1,96$ dan apabila level kepercayaan yang dipakai 99%, maka nilai $Z_{1-\alpha} = 2,575$.

Tingkat presisi:

Tingkat presisi hasil *test* menggambarkan seberapa mirip atau seberapa besar perbedaan hasil *test* pada sampel dengan hasil test di populasi. Tingkat presisi diberi simbol “d” dan makin tinggi tingkat presisi yang diinginkan atau semakin kecil perbedaan hasil *test* pada sampel dengan populasi, makin banyak sampel yang dibutuhkan. Umumnya tingkat presisi yang dipakai adalah 5%.

Prevalensi penyakit:

Kejadian atau prevaensi penyakit di komunitas diberi simbol p_x . Kejadian penyakit yang akan diuji menentukan berapa besar sampel yang dibutuhkan. Apabila kejadian penyakit yang dijadikan sasaran *test* adalah jarang, maka sampel yang dibutuhkan akan makin banyak dan sebaliknya.

4.3 Penentuan besar sampel

Berdasarkan uraian di atas, maka parameter yang dibutuhkan dalam penentuan besar sampel penelitian uji diagnostik adalah: level kepercayaan / reliabilitas sampel ($1-\alpha$), tingkat presisi atau deviasi (d), prevalensi penyakit yang akan ditest (p_x), tingkat sensitivitas atau spesifisitas yang harus dapat diidentifikasi sebagai tingkat akurasi yang signifikan (p). Rumus penghitungan besar sampel penelitian uji diagnostik disajikan di bawah ini.

Rumus Besar Sampel Uji Diagnostik
$n = \frac{Z_{\alpha}^2 P(1 - p)}{d^2 p_x^2}$
Keterangan: n = jumlah sampel; Z_{α} = Zskor untuk kepercayaan ($1-\alpha$); d = <i>margin of error</i> P = sensitivitas atau spesifisitas yang diinginkan; p_x = prevalensi penyakit “X” yang di test

Contoh kasus

Satu uji diagnostik akan menguji akurasi *pap smear* sebagai alat skrining kanker serviks dengan hasil pemeriksaan PA sebagai *gold standard*. Diharapkan, sensitivitas papsmear adalah 75%. Penelitian dilakukan di rumah sakit dimana diketahui prevalensi kanker serviks sebesar 20%. Jika tingkat kepercayaan ditetapkan sebesar 95% dan *margin of error* atau tingkat akurasi alat sebesar (d) = 10%, berapa jumlah sampel yang diperlukan?

Cara penentuan jumlah sampel adalah sebagai berikut.

Parameter penentuan besar sampel yang ditetapkan adalah:

- 1) $p = 0,75$ (sensitivitas alat yang diinginkan)
- 2) $d = 0,10$ (akurasi alat)
- 3) $Z_{\alpha} = 1,96$ (koefisien reliabilitas untuk tingkat kepercayaan 95%)
- 4) $p_x = 0,20$ (prevalensi penyakit yang akan diuji)

Sampel minimal yang diperlukan menjadi:

$$n = \frac{(Z_{\alpha})^2 p(1-p)}{d^2 p_x^2}$$
$$n = \frac{(1,96)^2 \times 0,75 \times 0,25}{0,10^2 \times 0,20^2} = 350$$

Jumlah sampel minimal yang dibutuhkan sebanyak 350 orang.

4.4 Cara Pemilihan Sampel

Berikut adalah tahapan dalam pemilihan sampel pada uji diagnostik.

- 1) Penetapan unit analisis:

Pada tahap pertama peneliti harus menetapkan siapa yang akan menjadi unit analisis dari penelitian uji diagnostik tersebut dengan batasan yang jelas. Misalnya akan diuji akurasi *pap smear* sebagai alat skrining lesi prakanker servik. Pada penelitian ini ditetapkan unit analisisnya adalah wanita usia di atas 15 tahun yang sudah pernah menikah. Kriteria tersebut mengacu kepada kelompok yang rentan terhadap lesi prakanker servik.

- 2) Penetapan populasi target:

Populasi target penelitian adalah kumpulan dari unit analisis dari penelitian tersebut. Untuk contoh penelitian uji diagnostik di atas, yang menjadi populasi target penelitian adalah wanita usia di atas 15 tahun yang sudah menikah. Di dalam penetapan populasi target penelitian uji diagnostik agar diperhatikan variabilitas populasinya disesuaikan dengan khalayak sasaran yang akan menggunakan alat *test* tersebut baik dari segi umur, jenis kelamin, suku dan ras. Apabila alat *test* tersebut ditujukan untuk mendiagnosis suatu penyakit, maka harus diperhatikan agar populasi target *test* mencakup suspek dengan tingkat keparahan yang paling ringan sampai yang paling berat.

3) Penetapan populasi terjangkau:

Penelitian uji diagnostik pada umumnya memiliki keterbatasan dalam hal waktu, sumber daya manusia, biaya, dan sarana penelitian. Oleh karena itu penelitian akan dilakukan pada individu, waktu dan lokasi terbatas. Misalnya penelitian akan dilakukan pada wanita usia di atas 15 tahun yang berkunjung ke Rumah Sakit X dalam periode waktu tertentu. Populasi yang akan dijangkau dalam penelitian tersebut adalah bagian dari populasi target yang didapatkan dari lokasi tertentu pada waktu tertentu. Bagian dari populasi target yang akan dijangkau tersebut dinamakan populasi terjangkau, maka dari itu karakteristik dari populasi terjangkau harus sama dengan populasi target dan ditambahkan dengan karakteristik tempat dan waktu, seperti contoh di atas tadi.

4) Penetapan unit sampel:

Unit sampel pada penelitian uji diagnostik umumnya sama dengan unit analisisnya. Yaitu penduduk sasaran penggunaan alat *test* tersebut. Kriteria eligibilitas sampel harus dijabarkan di dalam kriteria inklusi, eksklusi, dan *drop out*. Kriteria inklusi harus sama dengan kriteria populasi target, misalnya untuk uji diagnostik *pap smear* di atas, yang boleh dipilih sebagai sampel adalah wanita usia di atas 15 tahun yang sudah pernah menikah dan bersedia ikut serta di dalam penelitian. Individu yang telah memenuhi kriteria inklusi harus dikeluarkan dari subyek penelitian apabila memenuhi kriteria eksklusi. Misalnya ibu yang uterusnya sudah diangkat secara total harus dieksklusi dari penelitian karena tidak ada potensi menderita lesi prakanker servik. Ibu yang mengalami *fluor albus* juga dieksklusi karena bisa memberikan hasil yang palsu. Sebagai contoh kriteria eligibilitas sampel *pap smear* sebagai alat skrining lesi prakanker servik adalah:

Kriteria inklusi:

1. Wanita usia di atas 15 tahun
2. Sudah pernah menikah
3. Berkunjung ke Rumah Saki X pada periode waktu Januari – Desember 2021.
4. Bersedia secara sukarela sebagai subyek penelitian dengan menandatangani formulir *Informed consent*.

Kriteria eksklusi:

1. Telah melakukan histerektomi total
2. Menderita *fluor albus*

5) Pemilihan sampel:

Penelitian uji diagnostik pada umumnya menggunakan pendekatan insiden dalam penentuan subyek penelitian, yaitu dengan menggunakan kasus baru sebagai subyek. Maka dari itu, sangat tidak mungkin peneliti bisa membuat kerangka sampelnya. Oleh karena itu, teknik pemilihan sampel secara acak tidak dapat diterapkan dan teknik yang paling tepat dipakai adalah metode *consecutive sampling*, yaitu subyek yang dipakai adalah sesuai dengan urutan kedatangan. Sebaliknya, apabila penelitian akan menggunakan data pasien yang sudah ada seperti data yang terdapat di rekam medis, maka pemilihan sampel dapat dilakukan secara *random* dengan menggunakan data pasien sebagai kerangka sampelnya.

Teknik pemilihan sampel secara *consecutive*:

Setiap wanita yang berobat ke RS X akan dilakukan penapisan dengan kriteria eligibilitas sampel. Apabila semua kriteria telah terpenuhi, maka pasien tersebut ditetapkan sebagai sampel dan selanjutnya dilakukan pemeriksaan dengan alat *test* yang diuji dan dengan alat *gold standard*. Pemilihan sampel dengan cara tersebut dilanjutkan kepada pasien berikutnya sampai jumlah sampel minimal yang dibutuhkan terpenuhi.

Teknik pemilihan sampel secara acak:

Apabila daftar populasi terjangkau tersedia atau dapat dibuat, maka daftar tersebut bisa dijadikan kerangka sampel. Sampel selanjutnya dipilih secara acak dari kerangka sampel tersebut sebanyak yang dibutuhkan.

4.4 Latihan

Kasus:

Dr Bayu, Sp.OG. meneliti akurasi IVA sebagai alat skrining lesi pra-kanker leher rahim dengan hasil pemeriksaan PA dari swab serviks sebagai *gold standard*. Populasi terjangkau penelitian adalah wanita usia 30-45 tahun yang melakukan pemeriksaan obstetri dan ginekologi di RSUP. Perkiraan prevalensi lesi prakanker leher rahim (p) = 45%.

Tugas:

1. Tentukan siapa yang menjadi populasi target, terjangkau, dan sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel yang dibutuhkan bila tingkat kesalahan tipe I $\alpha = 0,05$ dan *margin of error* $d = 10\%$, sensitivitas (p_s) = 85%?
4. Jelaskan metode sampling apa yang akan dipakai dan cara pemilihan sampelnya!

BAGIAN 5. METODE SAMPLING PENELITIAN KOHORT

dr. I Ketut Tangking Widarsa, MPH.*

* Ni Made Dian Kurnia, S.KM, MPH.**

**) Dosen Biostatistiks Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com*

****) Dosen Biostatistiks Fakultas Kedokteran Universitas Udayana, Bali.*

Tujuan Pembelajaran:

- Memahami pengertian penelitian kohort.
- Memahami cara penentuan besar sampel penelitian kohort.
- Memahami cara pemilihan sampel penelitian kohort.

Pokok Bahasan:

5.1 Pengantar

5.2 Pengertian

5.3 Cara penentuan besar sampel

5.4 Metode pemilihan sampel

Bahan bacaan:

1. Cochran W.G. 1977. Sampling Techniques, Third edition, John Wiley&Sons, New York.
2. Levy P.S. and Lemeshow S. 1980. Sampling for Health Profesionals, Lifetime Learning Publication, California.
3. Lwanga S.K. and Lemeshow S. 1991. Sample Size Determination in Health Studies, World Health Organization, Geneva.
4. Sopiudin Dahlan M. 2009. Besar Sampel dan Cara Pengambilan Sampel Dalam Penelitian Kedokteran dan Kesehatan, Edisi 2, Salemba Medika, Jakarta

5.1 Pengantar

Penelitian kohort adalah salah satu metode penelitian observasional di bidang kedokteran dan kesehatan. Rancangan ini dipakai untuk mempelajari faktor risiko penyakit atau mempelajari prognosis suatu penyakit. Validitas internal penelitian kohort ditentukan oleh adanya *bias* seperti *selection bias*, *hystorical bias*, *instrumental bias*, dan *regression bias*. Sedangkan validitas eksternal penelitian kohort dipengaruhi oleh representativitas, jumlah dan cara pemilihan sampel. Pada Bagian ini akan dibahas tentang cara penentuan besar sampel dan cara pemilihan sampel untuk penelitian kohort.

5.2 Penelitian Kohort

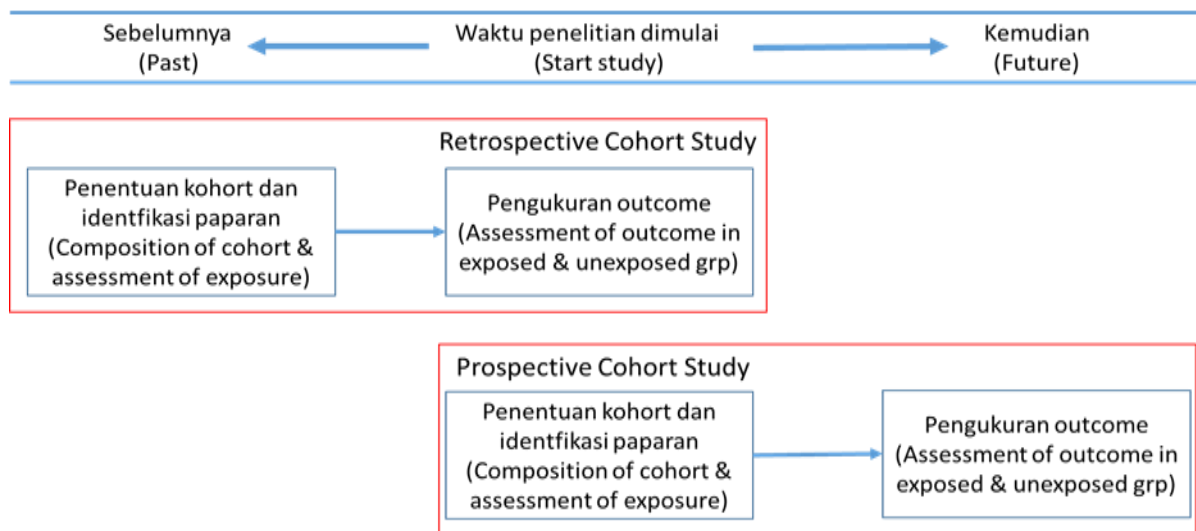
Kohort adalah sekelompok individu dengan karakteristik yang sama atau mirip seperti misalnya pekerja di pabrik rokok, pasien stroke iskemik, lansia, dan lainnya yang bebas dari penyakit yang akan diteliti. Penelitian kohort adalah penelitian yang meneliti hubungan satu paparan dengan satu atau beberapa kejadian suatu penyakit atau fenomena lainnya. Misalnya penelitian efek paparan nikotin terhadap fungsi paru pada pekerja di bagian produksi di pabrik rokok. Pada penelitian ini kohortnya adalah pekerja di pabrik rokok tersebut yang belum mengalami gangguan fungsi paru. Secara prinsip, pada penelitian kohort akan dilakukan pengamatan terjadinya kejadian penyakit yang diteliti secara berkelanjutan terhadap semua anggota kohort mulai ditetapkan sebagai anggota kohort sampai akhir penelitian. Berdasarkan kepada waktu penetapan kohort dari waktu penelitian dimulai, Penelitian kohort dibedakan menjadi penelitian kohort prospektif dan penelitian kohort retrospektif.

Penelitian kohort prospektif:

Penetapan anggota kohort pada penelitian kohort prospektif dilakukan ke depan dari saat penelitian dimulai. Misalnya pada penelitian efek simvastatin terhadap kekambuhan serangan stroke pada pasien stroke iskemik yang dirawat di RSUP. Penelitian dimulai pada 1 Januari 2021, maka pemilihan anggota kohort dimulai pada 1 Januari 2021 ke depan sampai jumlah sampel minimal terpenuhi. Kasus yang mendapatkan obat simvastatin menjadi kohort terpapar dan yang tidak menjadi kohort kontrol. Setelah seorang pasien ditetapkan sebagai kohort, selanjutnya akan diamati terjadinya kekambuhan serangan stroke mulai dari ditetapkan sebagai anggota kohort sampai penelitian berakhir.

Penelitian kohort retrospektif:

Penetapan anggota kohort pada penelitian kohort retrospektif berbeda dengan penelitian kohort prospektif. Pada penelitian kohort retrospektif, pemilihan anggota kohort dilakukan mundur dari waktu penelitian dimulai. Misalnya yang diteliti adalah efek simvastatin terhadap pencegahan serangan ulang pada pasien stroke iskemik yang dirawat di RSUP Sanglah. Penelitian dimulai pada bulan Januari 2021 dan pemilihan anggota kohort mundur yaitu mulai 1 Januari 2020. Anggota kohort pada penelitian kohort retrospektif dipilih dari rekam medis pasien stroke iskemik yang dirawat di RSUP Sanglah sejak tahun 2020. Pasien stroke iskemik baru yang dirawat mulai bulan Januari 2020 sampai Januari 2021 dijadikan sampel kohort. Pasien yang mendapatkan obat simvastatin dijadikan kohort terpapar dan yang tidak mendapat simvastatin dijadikan kohort kontrol. Selanjutnya, semua rekam medis sampel diteliti terhadap kejadian serangan stroke ulang mulai ditetapkan sebagai sampel kohort sampai penelitian berakhir. Kelebihan dari kohort retrospektif adalah dapat mengatasi permasalahan waktu dan biaya penelitian dan kelemahannya adalah sering tidak tersedia rekam medis yang lengkap sesuai dengan data yang dibutuhkan. Berikut adalah bagan penelitian kohort *retrospective* dan *prospective*.



Gambar 6.1. Perbedaan *cohort retrospective* dengan *cohort prospective*

5.3 Parameter penentuan besar sampel

Besar sampel penelitian kohort ditentukan oleh beberapa parameter seperti kejadian penyakit yang diteliti di populasi, besar efek paparan terhadap risiko kejadian sakit yang ditentukan dari angka *Relative Risk* (RR), level kepercayaan dan *power* penelitian yang diinginkan.

Efek paparan:

Penelitian kohort umumnya dipakai meneliti faktor risiko suatu penyakit atau efek satu paparan terhadap kejadian suatu penyakit. Besarnya efek paparan pada penelitian kohort ditentukan dari angka *Relative Risk* (RR). Besar sampel penelitian kohort dipengaruhi oleh berapa besar efek (*effect size*) terkecil yang harus dapat diidentifikasi sebagai efek yang signifikan. Makin kecil RR yang ditetapkan sebagai efek yang dapat diidentifikasi sebagai efek yang nyata, maka makin besar jumlah sampel yang dibutuhkan dan sebaliknya. Pada umumnya nilai RR yang dipakai adalah 1,5 atau 1,75 sebagai nilai efek terkecil yang harus dapat diidentifikasi sebagai efek yang nyata.

Angka kejadian di populasi:

Variabilitas kejadian di populasi akan memengaruhi besar sampel penelitian. Makin lebar variasi kejadian di populasi akan makin banyak sampel yang dibutuhkan dan sebaliknya. Ukuran variasi kejadian di populasi ditentukan dari kejadian penyakit yang diteliti "P". Dalam penelitian kohort, kejadian penyakit yang diteliti di populasi $P = (p_1 + p_2)/2$. Pada umumnya angka kejadian penyakit pada kelompok terpapar (p_1) belum diketahui, sebaliknya angka kejadian pada kelompok kontrol (p_2) dapat diasumsikan dari kejadian di populasi. Sedangkan asumsi kejadian pada kelompok terpapar (p_1) akan dihitung dari besarnya efek (RR) yang diinginkan dengan rumus penghitungan p_1 berikut:

$$p_1 = RR \times p_2$$

Contoh:

Apabila diasumsikan kejadian sakit pada kelompok kontrol $p_2 = 30\%$ dan $RR = 1,75$ sebagai efek paparan terkecil yang harus dapat diidentifikasi signifikan, maka asumsi kejadian efek pada kohort terpapar p_1 adalah:

$$p_1 = 1,75 \times 0,3 = 0,53$$

Level kepercayaan:

Level kepercayaan ditetapkan untuk mengantisipasi adanya *type I error* atau disingkat dengan α pada pengujian hipotesis penelitian. Supaya hasil penelitian dapat dipertanggungjawabkan secara ilmiah, maka kesalahan tersebut harus dikendalikan. Umumnya level kepercayaan yang dipakai adalah 95% untuk penelitian sosial dan 99% untuk penelitian klinik. Apabila level kepercayaan ditetapkan 95%, maka terdapat kemungkinan akan terjadi kesalahan dalam pengambilan keputusan hasil uji hipotesis sebesar 5% dan apabila menggunakan level kepercayaan 99%, maka akan terdapat kemungkinan

kesalahan 1%. Dalam penentuan besar sampel, level kepercayaan ditentukan dari nilai Zskor. Apabila level kepercayaan yang ditetapkan 95%, maka nilai $Z_{1-\alpha} = 1,96$ dan bila level kepercayaan 99%, maka nilai $Z_{1-\alpha} = 2,575$. Makin tinggi level kepercayaan yang diinginkan, maka makin banyak jumlah sampel yang dibutuhkan dan sebaliknya.

Power penelitian:

Penetapan besar *power* penelitian ditujukan untuk meminimalisir adanya kesalahan *type II* atau juga disebut β . Pada umumnya, pada penelitian kedokteran dan kesehatan, tingkat kesalahan *type II* yang masih diterima adalah 10% dengan *power* penelitian $(1-\beta)$ sebesar 90% atau tingkat kesalahan *type II* sebesar 20% dengan *power* penelitian $(1-\beta)$ sebesar 80%. Dalam penentuan jumlah sampel, besarnya *power* penelitian ditentukan dengan nilai $Z_{1-\beta}$. Untuk *power* penelitian $(1-\beta)$ 80%, nilai $Z_{1-\beta} = 0,84$ dan untuk *power* 90% nilai $Z_{1-\beta} = 1,24$. Besar *power* penelitian yang direncanakan akan memengaruhi besar sampel penelitian. Makin besar *power* yang direncanakan, makin besar sampel yang dibutuhkan.

5.4 Cara penentuan besar sampel

Penentuan besar sampel penelitian kohort ditentukan dari beberapa parameter berikut:

1. *Effect size* yang ditentukan dari $RR = Relative Risk$ terkecil yang dapat diidentifikasi bermakna (presisi sampel)
2. Kejadian penyakit pada kelompok kohort kontrol atau masyarakat umum P_2
3. Kejadian penyakit pada kelompok terpapar (P_1)
4. Kejadian penyakit pada populasi $(P) = \{(p_1) + (p_2)\}/2$
5. Tingkat kesalahan *type I* (α) atau level kepercayaan $(1-\alpha)$ yang diinginkan
6. Tingkat kesalahan *type II* (β) atau *power* penelitian $(1-\beta)$ yang diinginkan

Besar sampel ditentukan dengan menggunakan rumus seperti di bawah ini.

Rumus Besar Sampel Penelitian Kohort
$n = \frac{\left(Z_{\alpha} \sqrt{2P(1-P)} + Z_{\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right)^2}{(P_1 - P_2)^2}$
Keterangan:
P2=kejadian kasus pada kelompok kontrol
P1=kejadian kasus pada kelompok terpapar

$P_1 = P_2 \times RR$ $P = (P_1 + P_2) / 2$ $Z_\alpha = \text{koefisien reliabilitas } (1-\alpha)$ $Z_\beta = \text{koefisien power } (1-\beta)$
--

1) Contoh penghitungan sampel secara manual:

Akan diteliti pengaruh merokok terhadap risiko kejadian Penyakit Jantung Koroner (PJK) dengan rancangan kohort. Besar sampel yang dibutuhkan apabila dari data sekunder diketahui kejadian PJK $P_2 = 35\%$, RR penelitian yang dianggap bermakna 1,75, dengan tingkat reliabilitas 95% dan *power* penelitian 80%, dihitung seperti dibawah ini

Asumsi yang dipakai menentukan besar sampel adalah:

1. $P_2 = 0,35$
2. $RR = 1,75$
3. $P_1 = P_2 \times RR = 0,35 \times 1,75 = 0,61$
4. $P = (P_1 + P_2) / 2 = (0,35 + 0,61) / 2 = 0,48$
5. $Z_\alpha = 1,96$
6. $Z_\beta = 0,84$

$$n = \frac{\left(Z_\alpha \sqrt{2P(1-P)} + Z_\beta \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right)^2}{(P_1 - P_2)^2}$$

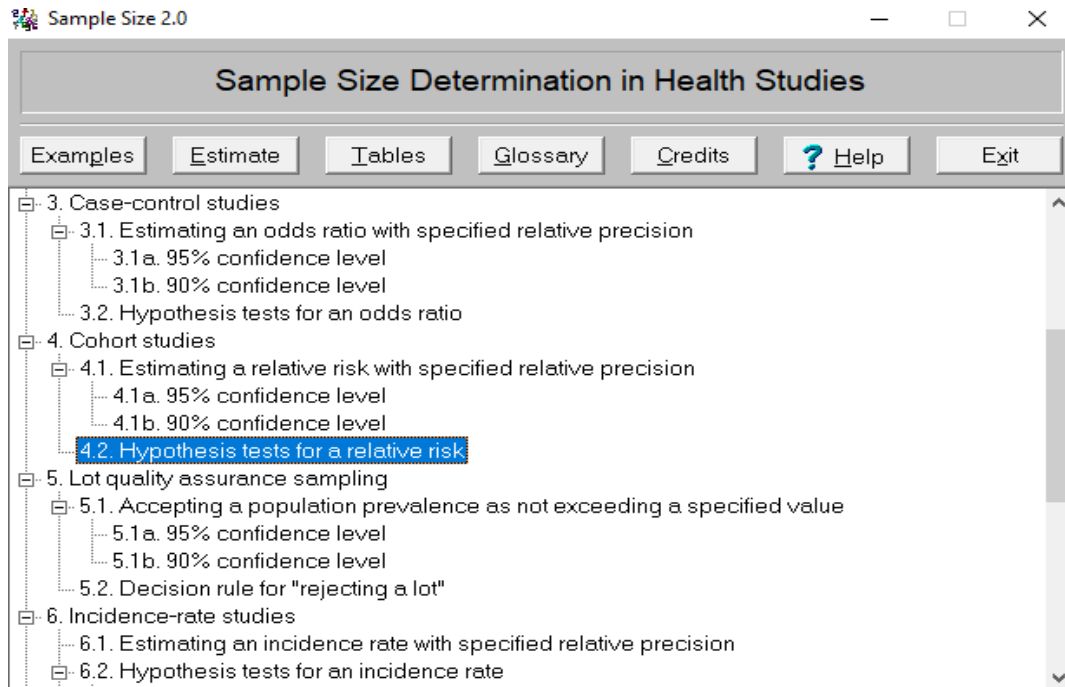
$$n = \frac{\left(1,96 \sqrt{2 \times 0,48 \times 0,52} + 0,84 \sqrt{0,61 \times 0,39 + 0,35 \times 0,65} \right)^2}{(0,61 - 0,35)^2} = 56,7$$

Dari perhitungan di atas diketahui sampel minimal dari setiap kelompok kohort $n_1 = n_2 = 57$.

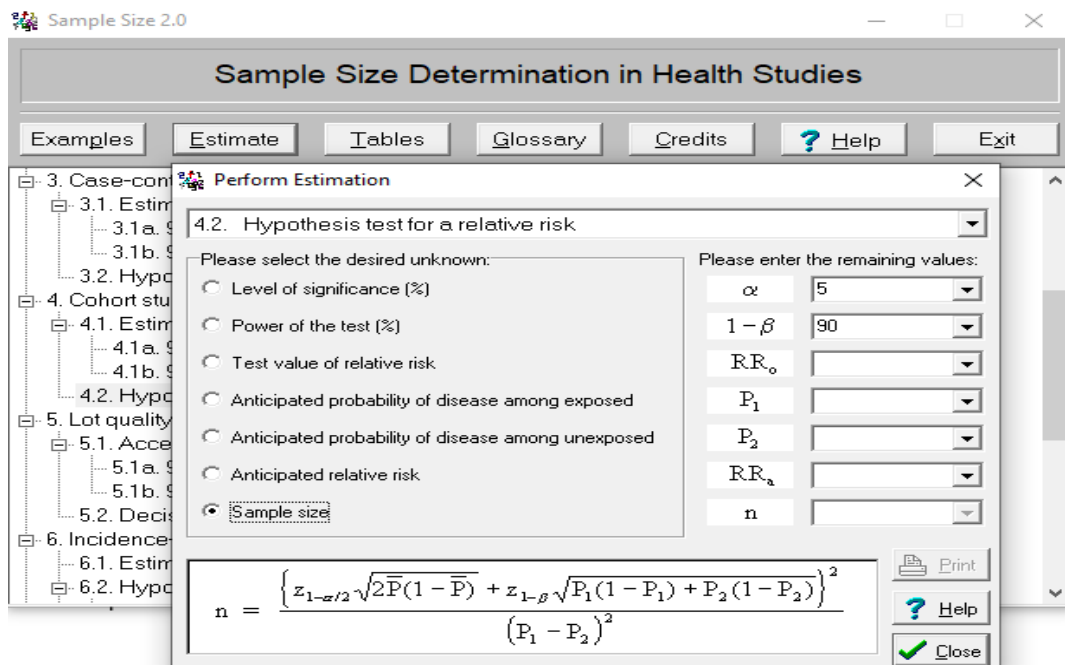
2). Contoh Penghitungan besar sampel menggunakan program Sample Size Determination in Health Studies Versi 2.0 (WHO)

Penghitungan besar sampel menggunakan bantuan program *sample size* (WHO) dengan langkah-langkah seperti berikut.

Step 1: aktifkan program *Sample Size*, sehingga tampak sbb:



Step 2: klik “4.2 Hypothesis tests for a relative risk “ pada Cohort studies dan klik “Estimate”, maka akan tampak sbb:



Step 3: masukan parameter penentuan sampel sesuai contoh di atas:

Level of significance (%) yaitu (α) = 5

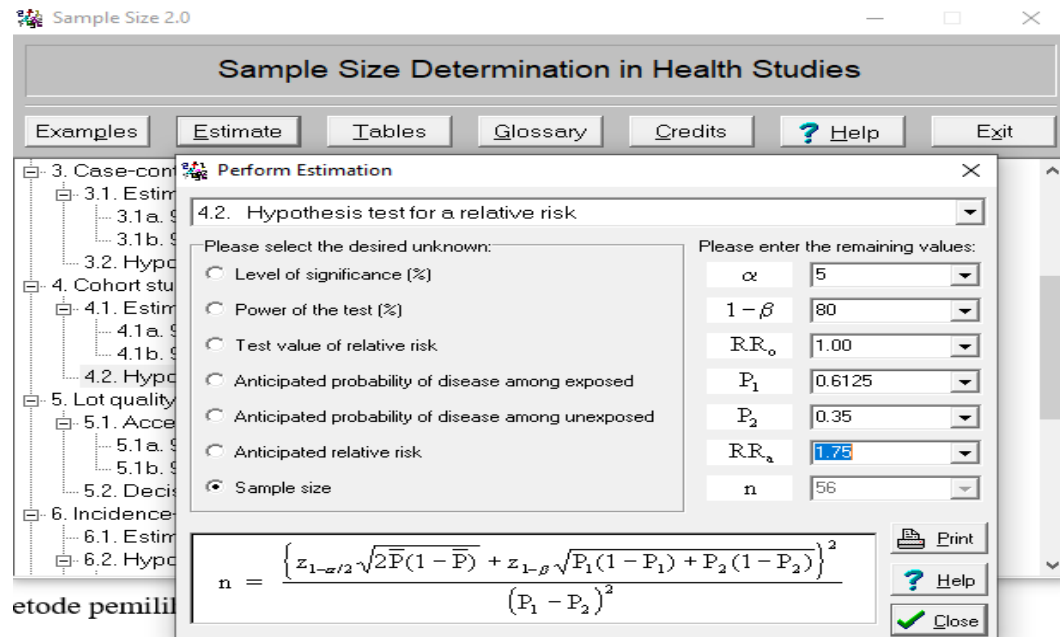
Power of the test (%) yaitu ($1-\beta$) = 80

Test value of relative risk yaitu (RR_0) = 1

Anticipated cases probability of disease among unexposed yaitu (P_2) = 0.35

Anticipated relative risk yaitu (RR_a) = 1.75

Pada kotak “Please enter the remaining values”, sehingga tampak sbb:



Dari perhitungan tersebut didapatkan besar sampel minimal $n_1 = n_2 = 56$.

5.4 Metode pemilihan sampel

Pemilihan sampel pada penelitian kohort dilakukan dalam beberapa tahapan sebagai berikut.

1) Penetapan unit analisis

Langkah pertama dalam pemilihan sampel adalah menetapkan unit analisis. Unit analisis pada penelitian kohort adalah individu yang memiliki potensi terpapar dari paparan yang diteliti. Misalnya yang diteliti adalah paparan debu tepung terigu pada pekerja pabrik roti terhadap kejadian dermatitis atopi, maka yang menjadi unit analisisnya adalah pekerja di pabrik roti tersebut. Apabila yang diteliti adalah pemberian obat simvastatin terhadap serangan ulang stroke pada penderita stroke ischemic, maka unit analisisnya adalah penderita stroke ischemic.

2) Penetapan populasi target

Populasi target adalah khalayak yang akan diteliti yang merupakan kumpulan dari unit analisis suatu penelitian dan kemana hasil penelitian akan digeneralisasikan. Apabila yang akan diteliti adalah efek paparan debu tepung terigu terhadap risiko kejadian dermatitis atopi, maka populasi

targetnya adalah orang yang memiliki potensi terpapar debu tepung terigu seperti pekerja pabrik roti. Pada penelitian pengaruh pemberian obat simvastatin terhadap risiko serangan ulang pada penderita stroke iskemik, penderita stroke ischemi menjadi populasi target.

3) Penetapan populasi terjangkau

Populasi terjangkau adalah bagian dari populasi target yang akan dijadikan sebagai sumber sampel. Pada contoh di atas adalah pekerja di pabrik Roti X, populasi terjangkau akan dikelompokkan menjadi kelompok terpapar atau kohort terpapar dan kelompok tidak terpapar atau kelompok kontrol. Pekerja di bagian produksi ditetapkan sebagai kohort terpapar dan pekerja di bidang administrasi dan pemasaran ditetapkan sebagai kohort kontrol. Untuk penelitian pengaruh pemberian obat simvastatin terhadap risiko serangan ulang pada pasien stroke iskemik di Rumah Sakit Pusat Sanglah, pasien stroke baru yang dirawat di RSUP pada periode penelitian menjadi populasi terjangkau. Mereka dikelompokkan menjadi kohort terpapar apabila diberikan obat simvastatin dan yang tidak mendapat obat simvastatin dijadikan kohort kontrol.

4) Penetapan unit sampel

Unit sampel adalah satuan yang dipakai memilih sampel penelitian. Pada penelitian kohort, umumnya unit sampel sama dengan unit analisis. Dari contoh di atas, unit analisisnya adalah pekerja pabrik roti X dan unit samplingnya juga pekerja di pabrik roti X tersebut. Pada penelitian pengaruh obat simvastatin, penderita stroke menjadi unit sampel yang juga merupakan unit analisisnya.

5) Pemilihan sampel

Pemilihan sampel pada penelitian kohort dapat dilakukan dengan metode *random* dan juga non *random*.

Pemilihan sampel secara *random*

Indikasi:

Indikasi penggunaan metode pemilihan sampel secara *random* adalah apabila daftar anggota kohort atau kerangka sampel bisa dibuat baik dari data rekam medis yang ada atau dari daftar pekerja atau peserta asuransi yang ada. Misalnya pada penelitian kohort pengaruh paparan debu tepung terhadap risiko dermatitis atopik pada pekerja pabrik Roti X. Daftar pekerja di bagian produksi dijadikan kerangka sampel dari kohort terpapar dan daftar pekerja di bagian administrasi

dan pemasaran dijadikan kerangka sampel dari kohort kontrol. Contoh lain, pada penelitian paparan obat simvastatin terhadap kekambuhan serangan stroke pada penderita stroke ischemic. Pada penelitian ini, daftar pasien stroke ischemic yang terdapat di register atau rekam medis di rumah sakit dapat dijadikan sebagai kerangka sampel.

Cara pemilihan sampel:

Membuat kerangka sampel:

Kerangka sampel dari masing-masing kelompok kohort dibuat berdasarkan data populasi terjangkau yang sudah ada, seperti daftar pekerja pabrik, daftar pasien yang terdapat di register atau rekam medis, daftar peserta asuransi, dan lainnya. Individu yang terpapar dimasukkan sebagai kelompok kohort terpapar dan yang tidak dimasukkan ke dalam kelompok kohort kontrol. Langkah selanjutnya adalah menetapkan kriteria eligibilitas sampel. Kriteria eligibilitas dijabarkan menjadi kriteria inklusi dan eksklusi. Dalam kriteria inklusi harus memuat karakteristik populasi target dan persyaratan sebagai anggota kohort yaitu terbebas dari penyakit atau fenomena yang diteliti. Misalnya pada penelitian pengaruh paparan debu tepung terhadap kejadian dermatitis atopi pada pekerja pabrik roti X, maka yang menjadi kriteria inklusi adalah: 1) sebagai pekerja pabrik Roti X paling sedikit 6 bulan, 2) tidak pernah dan sedang menderita dermatitis atopi pada saat penetapan sebagai anggota kohort, dan 3) bersedia ikut di dalam penelitian dengan menandatangani formulir *informed consent*. Sedangkan sebagai kriteria eksklusi adalah: 1) memiliki riwayat alergi berat terhadap debu, 2) menderita penyakit asma (sebagai prasyarat diterima sebagai karyawan). Semua yang memenuhi kriteria akan dijadikan kerangka sampel dari kelompok kohort terpapar dan kelompok kohort kontrol. Selanjutnya sampel dipilih secara *random* dari masing-masing kelompok kohort sesuai dengan jumlah yang dibutuhkan. Pemilihan sampel secara *random* dapat dilakukan dengan menggunakan bantuan Tabel Bilangan *Random* atau menggunakan aplikasi. Berikut adalah bagan alir pemilihan sampel secara *random*.



Gambar 5. Bagan alir pemilihan sampel secara *random* pada penelitian kohort

Contoh 1:

Penelitian efek paparan debu tepung terhadap kejadian dermatitis atopi pada pekerja pabrik roti di Denpasar, 2021 dilakukan dengan rancangan kohort prospektif. Penelitian dimulai pada bulan Januari 2021 dan berakhir pada bulan Desember 2021. Populasi terjangkau adalah semua pekerja di pabrik roti A yang berjumlah 1200 orang. Jumlah sampel yang dibutuhkan adalah 60 orang untuk masing-masing kelompok kohort. Kriteria sampel adalah: 1) paling sedikit sudah menjadi karyawan selama 6 bulan, 2) tidak pernah atau sedang menderita dermatitis atopi, 3) tidak memiliki alergi berat terhadap debu, dan tidak memiliki penyakit asma. Langkah pemilihan sampel.

1) Penetapan populasi terjangkau

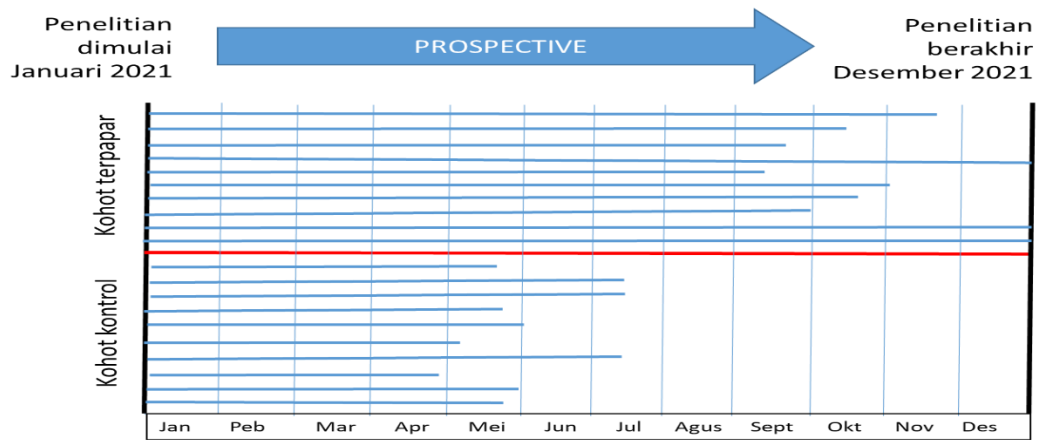
Sebelum penelitian dimulai, peneliti melakukan penapisan terhadap semua pekerja untuk menetapkan populasi terjangkau penelitian dengan menerapkan kriteria eligibilitas penelitian. Pertama didaftar pekerja yang masa kerjanya minimal 6 bulan. Dari daftar tersebut lalu diwawancara apakah dia pernah alergi debu atau pernah menderita asma. Hasil wawancara tersebut dikuatkan dengan melihat rekam medis pekerja bersangkutan. Mereka yang lolos kriteria tersebut selanjutnya dilakukan pemeriksaan terhadap penyakit dermatitis atopi atau riwayat dermatitis atopi. Mereka yang sedang atau pernah menderita dermatitis atopi dikeluarkan karena sebagai anggota kohort mereka harus terbebas dari penyakit yang diteliti. Selanjutnya, kepada semua pekerja yang lolos seleksi tersebut dibagi menjadi dua kelompok, yaitu kelompok kohort terpapar dan kontrol. Mereka yang bekerja di bagian produksi dijadikan anggota kohort terpapar dan yang bekerja di bagian produksi dan pemasaran dijadikan anggota kohort kontrol.

2) Pemilihan sampel:

Semua anggota kohort baik untuk kohort terpapar atau kontrol diberikan nomor urut yang unik. Selanjutnya dipilih sebanyak n sampel dari masing-masing kelompok secara *random* dengan bantuan Tabel Bilangan *Random* atau dengan bantuan aplikasi komputer.

3) Kepada semua sampel terpilih dari kedua kelompok, selanjutnya diikuti mulai dari bulan Januari 2021 sampai bulan Desember 2021 dan dievaluasi apakah ada dari mereka yang mengalami dermatitis atopi.

Berikut bagan pemilihan sampel kohort pada saat yang sama dan selanjutnya pengamatan dilakukan secara prospective:



Keterangan: — lama pengamatan sampai terjadi outcome

Gambar 3. Bagan penetapan anggota kohort dan pengamatan efek yang diteliti secara prospektif
Penjelasan gambar:

Pada gambar di atas, semua sampel dari kelompok kohort terpapar dan kontrol dipilih dan ditetapkan pada awal penelitian, yaitu pada bulan Januari 2021. Selanjutnya semua sampel diamati secara berlanjut dengan melakukan pemeriksaan terjadinya penyakit dermatitis atopi setiap bulannya sampai ditemukan terjadinya dermatitis atopi. Apabila sampai akhir penelitian belum ditemukan terjadinya *outcome*, maka pengamatan akan dihentikan juga.

Contoh 2:

Penelitian efek obat simvastatin terhadap pencegahan serangan ulang pada penderita stroke iskemik yang dirawat di RSUP. Penelitian dirancang sebagai penelitian kohort restrospektif dan penelitian dilakukan pada bulan Januari 2021. Kriteria eligibilitas penelitian adalah: 1) penderita stroke ischemic baru yang dirawat pada periode Januari – Desember 2020, 2) usia antara 30-55 tahun, dan 3) memiliki rekam medis yang lengkap. Populasi terjangkau adalah semua penderita stroke ischemic baru yang dirawat di RSUP sejak bulan Januari sampai Desember 2020. Cara pemilihan sampelnya adalah sebagai berikut.

1) Penetapan populasi terjangkau

Peneliti membuat daftar semua pasien stroke iskemik baru yang dirawat di RSUP sejak bulan Januari sampai bulan Desember 2020. Kemudian satu persatu rekam medis dari pasien tersebut diteliti kriteria eligibilitas sampel. Mereka yang memenuhi kriteria selanjutnya didaftar sebagai populasi terjangkau. Selanjutnya, mereka dikelompokkan menjadi dua kelompok, yaitu kelompok kohort terpapar dan kelompok kontrol. Mereka yang pada rekam medisnya tercatat mendapatkan obat simvastatin dijadikan kelompok terpapar dan mereka yang tidak mendapat

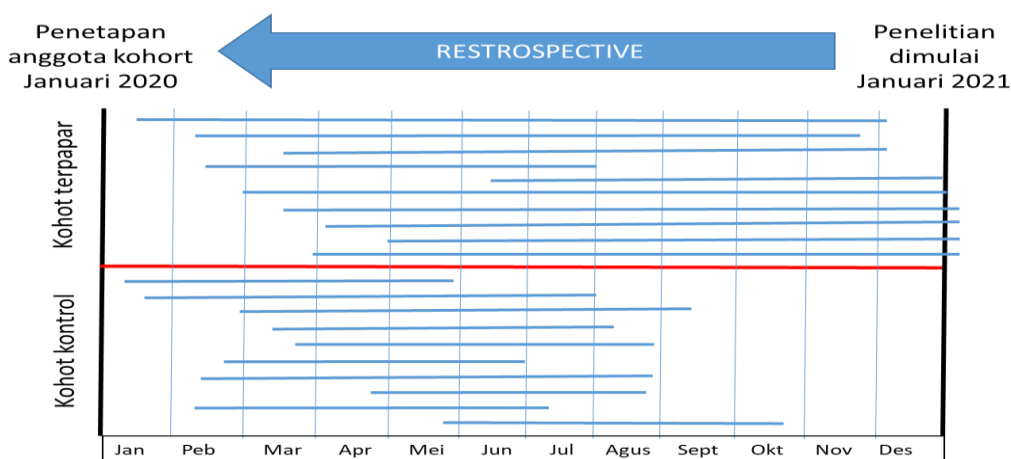
obat simvastatin dijadikan kelompok kontrol. Daftar dari kedua kelompok tersebut selanjutnya dijadikan sebagai kerangka sampel kohort terpapar dan kerangka sampel kohort kontrol.

2) Pemilihan sampel:

Semua anggota kohort baik untuk kohort terpapar atau kontrol diberikan nomor urut yang unik. Selanjutnya dipilih sebanyak n sampel dari masing-masing kelompok secara *random* dengan bantuan Tabel Bilangan *Random* atau dengan bantuan aplikasi komputer.

3) Kepada semua sampel terpilih dari kedua kelompok, rekam medisnya diteliti apakah selama periode mulai dirawat sebagai pasien baru sampai bulan Desember 2020 pernah terjadi serangan stroke ulang.

Berikut adalah bagan pemilihan sampel pada penelitian kohort retrospektif dan pengamatan terjadinya *outcome* yang diteliti.



Keterangan: — waktu mulai penetapan sebagai kohort sampai terjadinya outcome

Gambar 3. Bagan penetapan anggota kohort secara retrospektif dan pengamatan efek yang diteliti secara prospektif dari waktu penetapan sebagai anggota kohort

Keterangan gambar:

Pada gambar di atas, penetapan anggota kohort mundur satu tahun dari waktu penelitian. Penelitian dimulai pada tahun 2021 dan penetapan anggota kohort mulai dari bulan Januari sampai Desember 2020. Semua penderita stroke iskemik baru yang dirawat pada periode Januari sampai Desember 2020 dipakai sebagai kerangka sampel. Pasien stroke yang mendapat simvastatin menjadi kohort terpapar dan yang tidak mendapat simvastatin menjadi kohort kontrol. Sampel dipilih secara *random* dari kedua kerangka sampel tersebut. Pada umumnya, pasien yang terpilih sebagai sampel terdistribusi di sepanjang periode penelitian. Selanjutnya rekam medis dari setiap

pasien yang menjadi sampel diteliti secara prospektif dari waktu ditetapkan sebagai sampel sampai akhir penelitian untuk mengetahui apakah terjadi serangan stroke ulang.

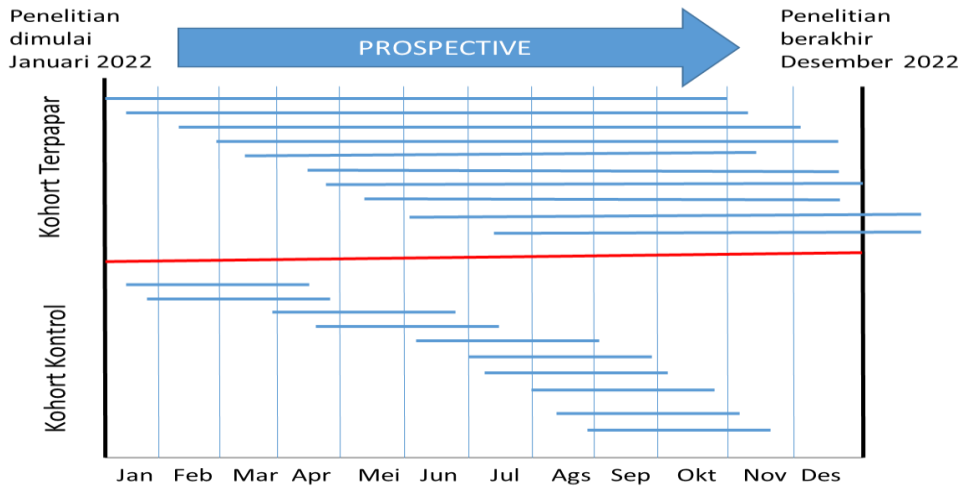
Pemilihan sampel secara *non-random*

Indikasi:

Pada penelitian kohort prospektif yang menggunakan pasien baru dalam periode penelitian sebagai anggota kohort, peneliti tidak bisa membuat kerangka sampel penelitiannya karena tidak diketahui siapa yang akan datang. Pada kondisi tersebut, pemilihan sampel secara *random* tidak dimungkinkan untuk dilakukan, maka metode sampling yang bisa dilakukan adalah metode sampling *non random*. Metode *non random* yang sering dipakai adalah metode *consecutive sampling* yaitu pemilihan sampel yang dilakukan berdasarkan urutan kedatangan. Misalnya pada penelitian pengaruh pemberian simvastatin terhadap serangan ulang pada penderita stroke iskemik dimana anggota kohort dipilih dari penderita baru yang didapat dari pengunjung RSUP sejak penelitian dimulai yaitu mulai bulan Januari 2022 sampai Desember 2022.

Cara pemilihan sampel:

Pemilihan sampel dengan metode *consecutive* dilakukan berdasarkan urutan kedatangan pasien. Setiap pasien stroke baru yang datang pada periode penelitian akan ditapis berdasarkan kriteria sampel baik untuk kohort terpapar atau kohort kontrol. Pasien yang memenuhi kriteria ditetapkan sebagai sampel, kemudian berdasarkan pemberian obat simvastatin, pasien dikelompokkan menjadi kohort terpapar dan kontrol. Setelah itu dilakukan pengamatan secara prospektif terjadinya serangan stroke ulang sampai waktu penelitian berakhir. Berikut adalah bagan pemilihan sampel dan pengamatan kejadian *outcome* pada penelitian kohort prospective yang menggunakan pasien baru yang datang selama periode penelitian.



Keterangan: — waktu mulai penetapan sebagai kohort sampai terjadinya *outcome*

Gambar 4. Bagan penetapan anggota kohort dan pengamatan efek yang diteliti secara prospektif dari waktu penetapan sebagai anggota kohort

5.5 Latihan

Kasus:

Dr. Bayu, Sp. OG, meneliti harapan hidup pasien ca-cerviks stadium II setelah tindakan operasi saja dan kombinasi operasi dan kemoterapi. Penelitian ini menggunakan rancangan kohort retrospektif. Semua pasien ca cerviks stadium II yang dirawat di rumah sakit di Bali dijadikan populasi terjangkau. Penelitian dilakukan pada tahun 2021-2022 dan data populasi terjangkau diambil dari catatan medis sejak tahun 2019. Variabel bebas adalah tidak operasi + terapi dan kontrol adalah yang hanya diberi terapi saja, sedangkan variabel terganggunanya adalah kejadian meninggal.

TUGAS:

1. Tentukan siapa yang menjadi populasi target, terjangkau, dan sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel yang dibutuhkan bila tingkat kesalahan tipe I $\alpha = 0,05$; tingkat kesalahan tipe II 20%, *effect size* RR = 2,5; dan kemungkinan meninggal setelah 1 tahun $P_2 = 20\%$?
4. Jelaskan metode sampling yang akan dipakai dan cara pemilihan sampelnya?

BAGIAN 6. METODE SAMPLING PENELITIAN CASE-CONTROL

dr. I Ketut Tangking Widarsa, MPH

Dosen Biostatistik Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com

Tujuan Pembelajaran:

- Memahami pengertian penelitian *case-control*.
- Memahami cara penentuan besar sampel penelitian *case-control*.
- Memahami cara pemilihan sampel penelitian *case-control*

Pokok Bahasan:

6.1 Pengantar

6.2 Rancangan penelitian *case-control*

6.3 Cara penentuan besar sampel

6.4 Cara pemilihan sampel

Bahan bacaan

1. Cochran W.G. 1977. Sampling Techniques, Third edition, John Wiley&Sons, New York
2. James J. Schlesseman. 1982. Case-Control Studies: Design, Conduct, Analysis. Oxford University Press, New York.
3. Lwanga S.K. and Lemehow S. 1991. Sample Size Determination in Health Studies, World Health Organization, Geneva.
4. Sopiyyudin Dahlan M. 2009. Besar Sampel dan Cara Pengambilan Sampel Dalam Penelitian Kedokteran dan Kesehatan, Edisi 2, Salemba Medika, Jakarta.

6.1 Pengantar

Penelitian faktor risiko penyakit merupakan salah satu penelitian kedokteran dan kesehatan yang banyak dilaporkan di jurnal internasional atau nasional. *Case-Control* (Kasus Kontrol) adalah salah

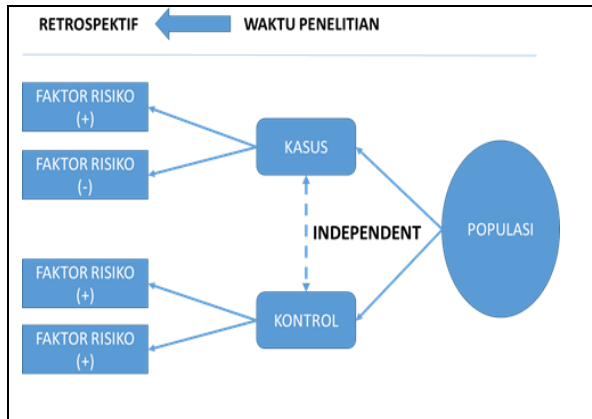
satu metode penelitian yang sering digunakan untuk tujuan meneliti faktor risiko suatu penyakit. Kelebihan rancangan *case-control* dibandingkan dengan penelitian kohort antara lain: jumlah sampel yang dibutuhkan secara umum akan lebih sedikit, waktu penelitian relatif lebih pendek, dan biaya lebih murah, serta bisa mempelajari banyak faktor risiko secara bersamaan. Disisi lain, *case-control* juga mempunyai kekurangan bila dibandingkan dengan penelitian kohort antara lain dalam menjelaskan hubungan sebab akibat. Pada penelitian *case-control* tidak bisa dipastikan bahwa faktor risiko terjadi sebelum terjadinya penyakit, karena pada studi *case-control* penelitian dimulai dari penentuan kasus dan kontrol baru setelah itu ditelusuri adanya faktor risiko. Validitas hasil penelitian *case-control* ditentukan oleh adanya bias dan *sampling error*. Bias yang sering terjadi adalah *selection bias* seperti pemilihan kasus dan kontrol tidak berasal dari populasi yang sama dan *recall bias* karena data paparan dikumpulkan secara retrospektif. *Sampling error* berhubungan dengan jumlah dan cara pemilihan sampel. Pada bagian ini akan dibahas cara penentuan besar dan metode pemilihan sampel pada penelitian *case-control*.

6.2 Rancangan Penelitian Case-Control

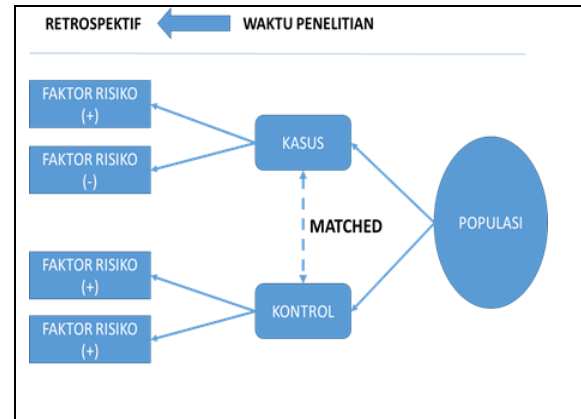
Penelitian *case-control* merupakan salah satu metode penelitian observasional yang banyak dilakukan pada penelitian di bidang kedokteran dan kesehatan. Penelitian ini digunakan untuk meneliti faktor risiko suatu penyakit. Pada prinsipnya, penelitian ini dimulai dari penentuan kasus dan kontrol dan keberadaan paparan ditelusuri kemudian. Orang yang menderita penyakit yang diteliti ditetapkan sebagai kasus dan orang yang tidak menderita penyakit yang diteliti dijadikan kontrol, misalnya: penderita PJK ditetapkan sebagai kasus dan penderita penyakit lain dijadikan kontrol pada penelitian faktor risiko PJK.

Berdasarkan cara pemilihan kasus dan kontrol, penelitian *case-control* dibedakan menjadi *independent case-control* dan *matched case-control*. *Independent case-control* adalah penelitian *case-control* dimana kasus dan kontrol dipilih secara bebas. Sebaliknya, pada *matched case-control*, pemilihan sampel kontrol dikaitkan dengan ciri-ciri kasus. Misal kasusnya anak laki-laki usia 3 tahun, maka kontrol dari kasus itu juga anak laki-laki usia 3 tahun. Berikut adalah gambar rancangan *independent* dan *dependent case-control study*.

Independent Case-Control



Matched Case-Control



Gambar 7.1. Bagan rancangan case-control

Berdasarkan sumber kasus dan kontrol, *case-control* dibedakan menjadi *community-based case-control*, *hospital-based case-control*, dan *nested case-control*. Pada *community-based case-control*, kasus dan kontrol bersumber dari register penyakit di masyarakat atau peserta asuransi kesehatan. Jenis penelitian ini bisa dilakukan apabila terdapat register penyakit di komunitas. Di negara berkembang termasuk di Indonesia, tidak tersedia data register penyakit komunitas sehingga *community-based case-control* jarang dilakukan. Pada penelitian *hospital-based case-control*, kasus dan kontrol bersumber dari pasien yang dirawat di pusat pelayanan kesehatan seperti rumah sakit, puskesmas, klinik dan lainnya. Sedangkan, pada penelitian *nested case-control*, kasus dan kontrol dipilih dari sampel penelitian dimana penelitian *case-control* ini ditautkan.

6.3 Parameter penentuan besar sampel

Besar sampel penelitian *case-control* ditentukan oleh beberapa parameter seperti proporsi paparan yang diteliti di populasi, besar efek paparan terhadap risiko kejadian sakit yang ditentukan dari angka *Odds Ratio* (OR), level kepercayaan dan *power* penelitian yang diinginkan.

Efek Paparan:

Besarnya efek paparan pada penelitian *case-control* ditentukan dari angka *Odds Ratio* (OR) yang juga disebut *effect size*. Besarnya efek terkecil yang harus mampu dideteksi sebagai efek yang nyata pada sampel akan menentukan jumlah sampel pada penelitian *case-control*. Makin kecil OR yang ditetapkan sebagai efek yang dapat diidentifikasi sebagai efek yang nyata, maka makin besar jumlah sampel yang

dibutuhkan dan sebaliknya. Pada umumnya nilai OR yang dipakai adalah 1,5 atau 2,5 sebagai nilai efek terkecil yang harus dapat diidentifikasi sebagai efek yang nyata.

Proporsi paparan di populasi:

Variabilitas kejadian paparan di populasi akan memengaruhi besar sampel penelitian. Makin lebar variasi kejadian paparan di populasi akan makin banyak sampel yang dibutuhkan dan sebaliknya. Ukuran variasi kejadian paparan di populasi ditentukan dari kejadian paparan yang diteliti "P". Dalam penelitian *case-control*, kejadian paparan yang diteliti di populasi $P = (p_1 + p_2)/2$. Pada umumnya angka kejadian paparan pada kelompok kasus (p_1) belum diketahui, sebaliknya angka kejadian paparan pada kelompok kontrol (p_2) dapat diasumsikan dari kejadian paparan di populasi. Asumsi paparan pada kelompok terpapar (p_1) akan dihitung dari besarnya efek (OR) yang diinginkan dengan rumus penghitungan p_1 sebagai berikut:

$$P_1 = (P_2 \times OR) / \{1 + p_2(OR - 1)\}$$

Contoh:

Apabila diasumsikan kejadian paparan pada kelompok kontrol $p_2 = 30\%$ dan $OR = 1,75$ sebagai efek paparan terkecil yang harus dapat diidentifikasi signifikan, maka asumsi proporsi paparan pada kasus p_1 adalah:

$$p_1 = 1,75 \times 0,3 = 0,53$$

Berdasarkan hasil penghitungan perkiraan proporsi paparan pada kasus P_1 seperti di atas, maka dapat diasumsikan paparan pada semua populasi $P = (p_1 + p_2)/2 = (0,53 + 0,30)/2 = 0,415$.

Level kepercayaan:

Seperti pada penelitian analitik lainnya, level kepercayaan ditetapkan untuk mengantisipasi adanya *type I error*. Umumnya level kepercayaan yang dipakai adalah 95% untuk penelitian sosial dan 99% untuk penelitian klinik. Level kepercayaan akan memengaruhi besar sampel penelitian *case-control*, dimana makin tinggi level kepercayaan yang dipakai maka makin banyak sampel yang dibutuhkan. Dalam penentuan besar sampel, level kepercayaan ditentukan dari nilai Zskor. Apabila level kepercayaan yang ditetapkan 95%, maka nilai $Z_{1-\alpha} = 1,96$ dan apabila level kepercayaan 99%, maka nilai $Z_{1-\alpha} = 2,575$.

Power penelitian:

Penetapan besar power penelitian ditujukan untuk meminimalisir adanya kesalahan *type II*. Pada umumnya, power penelitian ($1-\beta$) ditetapkan sebesar 80% atau 90%. Dalam penentuan jumlah sampel,

besarnya power penelitian ditentukan dengan nilai $Z_{1-\beta}$. Untuk power penelitian $(1-\beta)$ 80%, nilai $Z_{1-\beta} = 0,84$ dan untuk power 90% nilai $Z_{1-\beta} = 1,24$.

6.4 Cara penentuan besar sampel

Penentuan besar sampel penelitian *case-control* tergantung dari jenis rancangan yang dipakai, *independent* atau *matched case-control*. Berikut akan dijelaskan cara penentuan besar sampel minimal untuk kedua jenis rancangan tersebut.

6.4.1 Penentuan Besar Sampel *Independent Case-Control*

Parameter yang dibutuhkan dalam penentuan besar sampel pada *case-control* adalah:

1. Efek paparan terhadap kejadian sakit yang diinginkan (OR)
2. Perkiraan proporsi paparan pada populasi (P) yang dihitung dengan rumus: $P = (P_1 + P_2) / 2$
3. Perkiraan proporsi paparan pada kelompok kontrol (P_2)
4. Perkiraan proporsi paparan pada kelompok kasus yang dihitung dengan rumus:

$$P_1 = (P_2 \times OR) / \{1 + P_2(OR - 1)\}$$
5. Level kepercayaan sampel ($1-\alpha$)
6. *Power* penelitian ($1-\beta$)

Rumus Besar Sampel Minimal <i>Independent Case-Control</i>
$n = \frac{\left(Z_{\alpha} \sqrt{2P(1-P)} + Z_{\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right)^2}{(P_1 - P_2)^2}$
<p>Keterangan:</p> <p>P_2 = proporsi paparan pada kontrol</p> <p>P_1 = proporsi paparan pada kelompok kasus = $(P_2 \times OR) / \{1 + P_2(OR - 1)\}$</p> <p>$P = (P_1 + P_2) / 2$</p> <p>$Z_{\alpha}$ = koefisien reliabilitas ($1-\alpha$)</p> <p>Z_{β} = koefisien <i>power</i> ($1-\beta$)</p>

Penghitungan secara manual

Contoh:

Akan diteliti pengaruh merokok terhadap Penyakit Jantung Koroner (PJK) dengan rancangan *independent case-control*. Berapa besar sampel minimal yang dibutuhkan apabila dari data sekunder diketahui bahwa proporsi populasi dewasa merokok $P_2=35\%$, OR penelitian yang dianggap bermakna 2,5 dengan tingkat reliabilitas 95% dan *power* penelitian 80%. Penentuan besar sampel akan dihitung sebagai berikut.

Penetapan asumsi yang dipakai menentukan besar sampel yaitu:

1. $P_2 = 0,35$
2. $OR = 2,5$
3. $P_1 = 0,35 \times 2,5 / (1 + 0,35 \times (2,5 - 1)) = 0,57$
4. $P = (0,35 + 0,57) / 2 = 0,46$
5. $Z_\alpha = 1,96$
6. $Z_\beta = 0,84$

Menghitung jumlah sampel yang diperlukan sebagai berikut:

$$n = \frac{\left(Z_\alpha \sqrt{2P(1-P)} + Z_\beta \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right)^2}{(P_1 - P_2)^2}$$

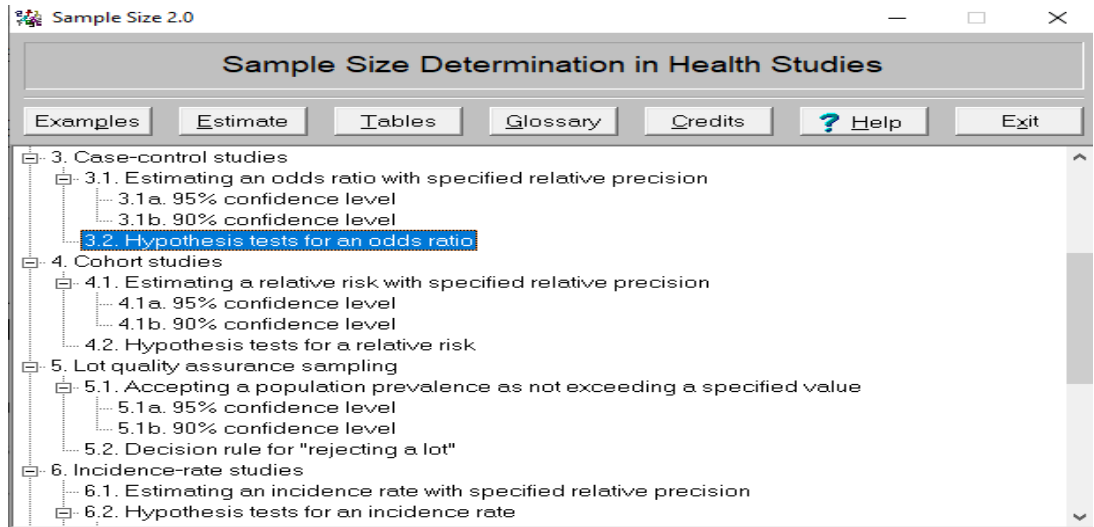
$$n = \frac{\left(1,96 \sqrt{2 \times 0,46 \times 0,54} + 0,84 \sqrt{0,57 \times 0,43 + 0,35 \times 0,65} \right)^2}{(0,57 - 0,35)^2} = 79,28$$

Jumlah sampel yang diperlukan menjadi $2 \times 80 = 160$

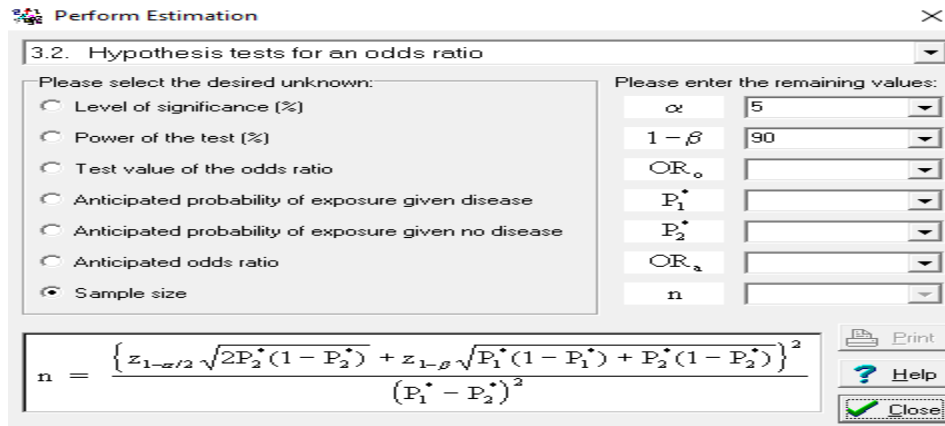
Penghitungan besar sampel menggunakan program Sample Size Determination in Health Studies Versi 2.0 (WHO)

Penghitungan besar sampel *case-control* menggunakan bantuan program *sample size* (WHO) dengan langkah-langkah seperti berikut.

Step 1: Aktifkan program Sample Size, sehingga tampak sbb:



Step 2: Klik “3.2 Hypothesis tests for a relative risk “ pada *Case-control studies* dan klik “Estimate”, maka akan tampak sbb:



Step 3. Masukkan parameter penentuan besar sampel penelitian di atas:

Level of significance (%) yaitu: $\alpha = 5$

Power of the test (%) yaitu: $1 - \beta = 80$

Test value of the odds ratio yaitu: $OR_0 = 1$

Anticipate probability of exposed given no disease, yaitu: $P_2 = 0,35$

Anticipated odds ratio, yaitu: $OR_a = 2,5$

Pada kotak: please enter the remaining values, sehingga tampak sebagai berikut.

Perform Estimation X

3.2. Hypothesis tests for an odds ratio

Please select the desired unknown:

- Level of significance (%)
- Power of the test (%)
- Test value of the odds ratio
- Anticipated probability of exposure given disease
- Anticipated probability of exposure given no disease
- Anticipated odds ratio
- Sample size

Please enter the remaining values:

α 5

$1 - \beta$ 80

OR_o 1.00

P_1^* 0.57377049180

P_2^* 0.35

OR_a 2.50

n 73

$$n = \frac{\left\{ z_{1-\alpha/2} \sqrt{2P_2^*(1-P_2^*)} + z_{1-\beta} \sqrt{P_1^*(1-P_1^*) + P_2^*(1-P_2^*)} \right\}^2}{(P_1^* - P_2^*)^2}$$

Print Help Close

Besar sampel minimal yang dibutuhkan adalah $n_1 = n_2 = 73$.

6.4.2 Penentuan besar sampel *Case-Control* dengan C kontrol per kasus

Pada kondisi dimana jumlah kasus yang diteliti sangat jarang, sehingga untuk mendapatkan jumlah sampel kasus sesuai dengan yang dibutuhkan tidak mungkin dapat dipenuhi, sebaliknya jumlah kontrol sangat banyak. Untuk mengatasi masalah tersebut, maka jumlah kontrol per kasus digandakan dengan “C” kali, dimana $C > 1$. Misalnya jumlah kontrol per kasus adalah 1:3 yang artinya untuk setiap 1 sampel kasus terdapat 3 sampel kontrol. Perhitungan besar sampel case-control dengan C kontrol per kasus sebagai berikut.

Rumus Besar Sampel Case-Control dengan C kontrol per kasus ($C > 1$)
$n'_c = \{(C+1) \times n\} / 2C$
<p>keterangan:</p> <p>n'_c = jumlah sampel kasus</p> <p>n = jumlah sampel yang dihitung dengan rumus sampel <i>independent case-control</i></p> <p>C = jumlah kontrol per kasus</p>

Contoh:

Misalnya pada penelitian *case-control* di atas akan dipakai 3 kontrol per kasus dan diketahui jumlah sampel kasus menurut rumus besar sampel *independent case-control* $n = 73$ serta $C = 3$, maka jumlah sampel kasus menjadi:

$$n'_c = \{(3 + 1) \times 73\} / (2 \times 3) = 292/6 = 48,6 \text{ atau dibulatkan menjadi } 49.$$

Jadi jumlah sampel kasus $n'_c = 49$ dan jumlah sampel kontrol $= n'_c \times 3 = 147$, sehingga jumlah sampel keseluruhan menjadi $= 49$ (kasus) + 147 (kontrol) = 196 .

6.4.3 Penentuan besar sampel *Matched Case-Control*

Parameter yang diperlukan pada penentuan besar sampel pada *matched case-control* sama seperti pada rancangan *independent case-control*, yaitu:

1. *Effect size* (OR) terkecil yang dianggap bermakna
2. Estimasi proporsi paparan pada kontrol (P_2).
3. Estimasi proporsi paparan pada kelompok kasus (P_1) yang dihitung dengan rumus

$$P_1 = \frac{P_2 \times OR}{1 + P_2(OR - 1)}$$
4. Proporsi pasangan diskordan (π)
5. Level kepercayaan sampel ($1 - \alpha$)
6. *Power* penelitian ($1 - \beta$)

Rumus Besar Sampel Penelitian Matched Case-Control
$n = \frac{(Z_\alpha + Z_\beta)^2 \pi}{(P_1 - P_2)^2}$
<p>Keterangan:</p> <p>n = jumlah sampel pasangan</p> <p>Z_α = koefisien reliabilitas untuk α tertentu</p> <p>Z_β = koefisien <i>power</i> penelitian untuk β tertentu</p> <p>π = proporsi pasangan diskordan</p> <p>P_2 = proporsi ekspose pada kontrol</p> <p>P_1 = proporsi ekspose pada kasus (dihitung berdasarkan nilai P_2 dan OR)</p>

Contoh:

Tingkat efikasi vaksin BCG dalam mencegah terjadinya infeksi TBC pada anak akan dilakukan dengan rancangan *matched case-control* yang dipasangkan menurut umur dan jenis kelamin. Kasusnya adalah anak yang menderita TBC dan kontrolnya adalah anak yang tidak menderita TBC dan sebagai paparan adalah tidak mendapat vaksinasi BCG. Informasi yang tersedia adalah perkiraan kontrol yang mendapat vaksinasi BCG sekitar 20%. Peneliti menginginkan *power* dari penelitian adalah 90% dan menggunakan tingkat kemaknaan 5%, serta proporsi diskordan (pasangan kasus kontrol yang memiliki status paparan berbeda) 40%. Besar sampel minimal yang dibutuhkan untuk

penelitian tersebut bila OR terkecil yang dianggap bermakna adalah 2,5 dapat dihitung seperti berikut ini.

Dari rencana penelitian tersebut ditetapkan:

1. $P_2 = 0,20$ (proporsi kontrol yang mendapat BCG)
2. $OR = 2,5$ (akurasi sampel)
3. $P_1 = (P_2 \times OR) / \{1 + P_2(OR - 1)\} = 0,384$
4. Proporsi pasangan diskordan $\pi = 0,40$
5. $Z_\alpha = 1,96$
6. $Z_\beta = 1,28$

Jumlah sampel adalah:

$$n = \frac{(Z_\alpha + Z_\beta)^2 \pi}{(P_1 - P_2)^2} = \frac{(1,96 + 1,28)^2 \times 0,40}{(0,38 - 0,20)^2} = 129,5987$$

Jumlah pasangan sampel menjadi 130 pasangan (dibulatkan ke atas).

6.5 Cara Pemilihan Sampel

Pendekatan yang umumnya dipakai di dalam pemilihan sampel pada penelitian *case-control* adalah pendekatan insiden (kasus baru) atau pendekatan prevalens (kasus lama). Berikut adalah penjelasan pemilihan sampel dengan pendekatan prevalens dan insiden.

Pemilihan sampel dengan pendekatan prevalens:

Pemilihan sampel dengan pendekatan prevalen adalah sampel kasus dan kontrol dipilih dari pasien yang sudah ada. Misalnya dipilih dari catatan medik atau hasil penelitian cross-sectional atau kohort yang sudah dilakukan sebelumnya. Pada pendekatan ini, kerangka sampel untuk kasus dan kontrol dibuat dari catatan medik atau daftar sampel yang sudah ada pada penelitian dimana *case-control* tersebut ditautkan. Pada *independent case-control study*, pemilihan sampel kasus dan kontrol dilakukan secara *random* dari kerangka sampel masing-masing kelompok. Pada *matched case control study* pemilihan sampel kasus dilakukan dengan cara *random* (bilangan *random*) dari kerangka sampel kasus tersebut dan kontrol dipilih dari kerangka sampel kontrol yang mempunyai kriteria yang sama dengan kasus. Misalnya kriteria yang dipasangkan adalah jenis kelamin dan umur, apabila sampel kasus yang terpilih adalah wanita usia 25 tahun, maka kontrolnya harus mempunyai kriteria yang sama yaitu jenis kelaminnya wanita dan usianya juga 25 tahun. Pada *nested case-control*, sampel kasus

dan kontrol dipilih secara *random* dari sampel penelitian dimana penelitian *case-control* tersebut ditautkan.

Langkah pemilihan sampel:

1) Penetapan unit analisis

Unit analisis merupakan satuan analisis yang akan dipakai pada penelitian *case-control*. Misalnya akan diteliti efek paparan rokok terhadap kejadian PJK, maka yang menjadi unit analisis adalah orang yang potensial merokok dan berisiko menderita PJK.

Penetapan populasi terjangkau

Populasi terjangkau adalah kumpulan dari unit analisis dari mana sampel akan dipilih. Misalnya Penelitian efek paparan rokok terhadap kejadian PJK akan dilakukan pada pasien yang dirawat di RSUP Sanglah tahun 2020-2021, maka yang menjadi populasi terjangkau adalah pasien umur 30 tahun ke atas yang dirawat di RSUP Sanglah pada periode tahun 2020-2021. Populasi terjangkau dikelompokkan menjadi populasi terjangkau kelompok kasus dan kelompok kontrol. Pasien yang dimasukkan ke populasi kelompok kasus adalah pasien yang menderita penyakit yang diteliti. Sangat penting untuk menetapkan kriteria diagnosis yang dipakai menentukan kasus atau kontrol. Apabila kriteria yang dipakai tidak akurat, maka ada kemungkinan akan terjadi *selection bias*, dimana pasien yang seharusnya menjadi kasus justru dimasukkan sebagai kontrol dan sebaliknya. Pada contoh di atas, semua penderita PJK (sesuai kriteria kasus) yang dirawat pada tahun 2020-2021 di RSUP Sanglah menjadi kelompok populasi kasus. Sebaliknya, pasien yang bukan menderita PJK dijadikan populasi kontrol.

2) Penetapan unit sampel

Unit sampel adalah satuan yang dipakai dasar untuk memilih sampel. Pada contoh di atas unit sampel penelitiannya sama dengan unit analisisnya, yaitu pasien yang dirawat di RSUP Sanglah pada tahun 2020-2021. Unit sampel kasus adalah penderita PJK dan unit sampel kontrol adalah pasien yang bukan penderita PJK. Supaya sampel kasus dan kontrol berasal dari populasi yang sama, maka harus ditetapkan kriteria eligibilitas sampel yang dijabarkan ke dalam kriteria inklusi dan kriteria eksklusi. Kriteria inklusi dibuat sesuai dengan karakteristik populasi dan untuk contoh di atas adalah: (1) pasien yang dirawat di RSUP Sanglah pada periode tahun 2020-2021, (2) usia \geq 30 tahun, dan (3) bersedia ikut sebagai subyek penelitian. Sebagai kriteria eksklusi antara lain: (1) rekam medis tidak lengkap (data paparan, data perancu, data pendukung diagnosis tidak ada), (2) pasien sudah meninggal, (3) alamat pasien tidak jelas dan tidak lengkap. Semua pasien yang memenuhi kriteria eligibilitas selanjutnya didaftar dan dijadikan kerangka sampel.

3) Pemilihan sampel:

Pemilihan sampel bisa dilakukan secara *random* dan *non-random*.

Pemilihan sampel secara *random*:

Sampel kasus dan kontrol dipilih sebanyak sampel yang dibutuhkan dari masing-masing kelompok secara *random* dengan bantuan Tabel Bilangan *Random* atau dengan bantuan aplikasi komputer.

Pemilihan sampel secara *non-random*:

Pada pemilihan sampel secara *non-random*, rekam medis dari pasien kasus dan kontrol diteliti apakah pasien tersebut memenuhi kriteria eligibilitas sampel atau tidak. Apabila kriteria inklusi dan eksklusi dipenuhi, maka pasien tersebut dijadikan sampel kasus atau kontrol. Pemilihan rekam medis yang ditelaah dimulai dari rekam medis terbaru dan berlanjut ke rekam medis sebelumnya. Cara yang sama dilakukan untuk rekam medis pasien berikutnya sampai jumlah sampel yang dibutuhkan terpenuhi.

- 4) Setelah sampel rekam medis kasus dan kontrol terpilih, selanjutnya data paparan yang diteliti ditelusuri pada rekam medis tersebut dan dicatat pada formulir data sekunder penelitian. Bagan pemilihan sampel penelitian *case-control* dengan pendekatan prevalensi.

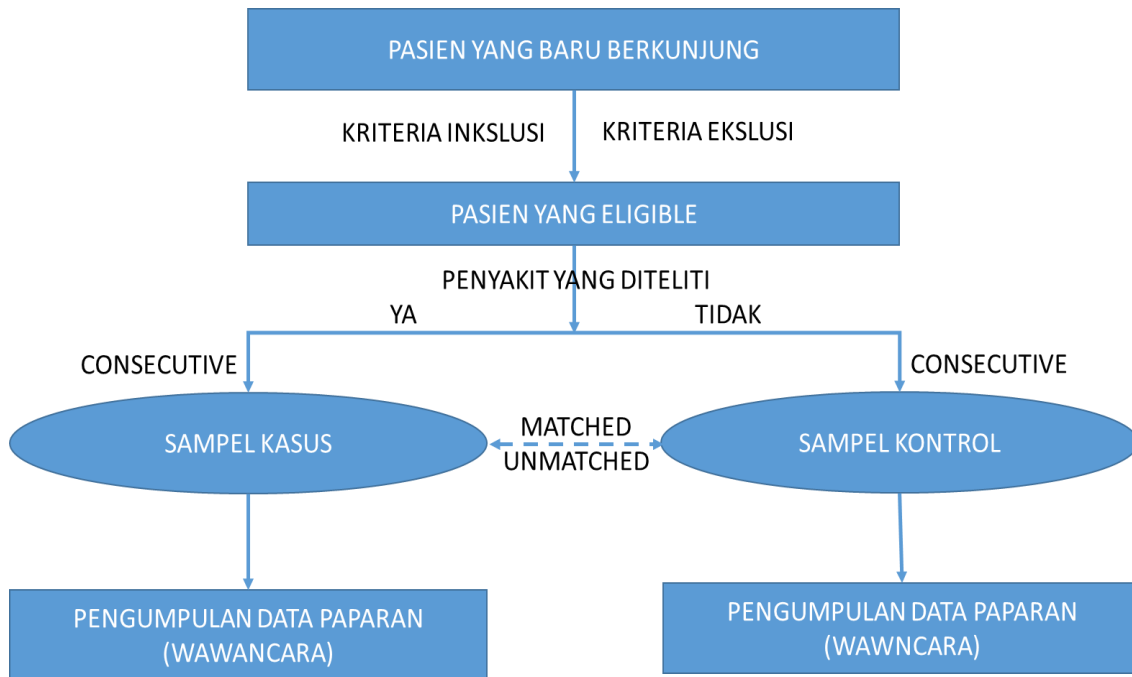


Gambar 6.1 Gambar bagan pemilihan sampel dengan pendekatan prevalensi

Pemilihan sampel dengan pendekatan insiden:

Pemilihan sampel dengan pendekatan insiden adalah pemilihan sampel kasus dan kontrol dari pasien yang baru. Misalnya penelitian dimulai pada bulan Januari 2021, sampel kasus dan kontrol dipilih dari pasien yang datang mulai bulan Januari 2021 ke depan sampai jumlah sampel yang dibutuhkan terpenuhi. Pada pendekatan ini peneliti tidak mungkin membuat kerangka sampel, oleh karena itu, cara

pemilihan sampel yang mungkin dilakukan adalah dengan metode *non-random* seperti *consecutive* atau *convenience sampling*. Setiap pasien baru yang menderita penyakit yang diteliti dan memenuhi kriteria inklusi dan eksklusi dipilih sebagai sampel kasus. Kontrol untuk kasus tersebut dipilih dari pasien baru yang bukan menderita penyakit yang diteliti dan datang pada hari yang sama dengan kasus. Untuk studi *matched case control*, tiap kasus yang datang dicarikan pasangannya yang sesuai dengan ciri kasus. Pemilihan sampel kasus dan kontrol seperti di atas diteruskan untuk pasien selanjutnya sampai jumlah sampel terpenuhi. Berikut adalah bagan pemilihan sampel dengan pendekatan insidens.



Gambar 6.2 Gambar bagan pemilihan sampel dengan pendekatan insidens

6.6 Latihan

Kasus:

Dr. Ida, M.Kes akan meneliti pengaruh merokok dan minum alkohol pada ibu hamil terhadap kejadian berat badan lahir rendah (BBLR). Populasi terjangkau adalah ibu yang melahirkan ditolong oleh tenaga kesehatan di Kabupaten Gianyar pada tahun 2021. Penelitian menggunakan rancangan *independent case-control*. Sebagai kasus adalah ibu yang melahirkan anak BBLR dan kontrol adalah ibu yang melahirkan bayi dengan berat lahir normal. Variabel yang diukur adalah kebiasaan minum alkohol dan merokok sebagai variabel bebas dan BBLR sebagai variabel tergantung.

TUGAS:

1. Tentukan siapa yang menjadi populasi target, terjangkau, dan sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel yang dibutuhkan bila tingkat kesalahan tipe I $\alpha = 0,05$; tingkat kesalahan tipe II 20%; *effect size* OR = 2,5, dan proporsi merokok pada ibu hamil (P2) = 15%?
4. Jelaskan metode sampling yang akan dipakai dan cara pemilihan sampelnya!

BAGIAN 7. METODE SAMPLING PENELITIAN KORELASI

dr. I Ketut Tangking Widarsa, MPH.*

* *Dosen Biostatistik Fakultas Kedokteran Universitas Udayana dan Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. E-mail: twidarsa@gmail.com dan twidarsa@yahoo.com*

Tujuan Pembelajaran

- Memahami pengertian penelitian korelasi.
- Memahami cara penentuan besar sampel penelitian korelasi.
- Memahami cara pemilihan sampel penelitian korelasi

Pokok Bahasan:

- 7.1 Pengantar
- 7.2 Pengertian
- 7.3 Cara penentuan besar sampel
- 7.4 Cara pemilihan sampel

Bahan Bacaan

1. Cochran W.G. 1977. *Sampling Techniques*, Third edition, John Wiley&Sons, New York.
2. Levy P.S. and Lemeshow S. 1980. *Sampling for Health Professionals*, Lifetime Learning Publication, California.
3. Lwanga S.K. and Lemeshow S. 1991. *Sample Size Determination in Health Studies*, World Health Organization, Geneva.
4. Sopiudin Dahlan M. 2009. *Besar Sampel dan Cara Pengambilan Sampel Dalam Penelitian Kedokteran dan Kesehatan*, Edisi 2, Salemba Medika, Jakarta

7.1 Pengantar

Penelitian korelasi atau asosiasi merupakan salah satu jenis penelitian kedokteran dan kesehatan. Pada penelitian ini akan diteliti hubungan kecenderungan kejadian satu fenomena dengan kecenderungan kejadian fenomena lainnya. Misalnya penelitian yang ingin meneliti hubungan antara tingkatan

aktifitas fisik dengan kapasitas paru penderita fibrosis paru. Contoh lain, seorang guru olah raga di sekolah menengah atas ingin meneliti hubungan antara panjang tungkai dengan jauhnya lompatan siswa saat melakukan olahraga lompat jauh.

Metode analisis pada penelitian korelasi ditentukan dari skala pengukuran dan normalitas distribusi data dari variabel yang akan dihubungkan. Ada beberapa metode analisis korelasi yang dapat diterapkan pada penelitian asosiasi yaitu metode korelasi Pearson, Spearman Rank, dan Kendal. Metode korelasi Pearson dipakai apabila variabel yang dihubungkan berskala interval/ratio dan berdistribusi normal. Metode korelasi Spearman Rank dipakai apabila variabel yang dihubungkan berskala ordinal atau interval yang tidak berdistribusi normal. Metode korelasi Kendal dipakai apabila variabel yang dihubungkan berskala ordinal dengan interval yang sempit misalnya hanya terdapat 4 kategori seperti sangat baik, baik, kurang, dan sangat kurang. Penentuan besar sampel pada penelitian korelasi ditentukan oleh beberapa parameter seperti besar *effect size*, tingkat kepercayaan dan *power* penelitian yang direncanakan. Pada bagian ini akan dibahas tentang pengertian, cara penentuan dan cara pemilihan sampel pada penelitian korelasi.

7.2 Penelitian korelasi

Penelitian korelasi adalah penelitian yang bertujuan mempelajari hubungan antara satu kejadian X dengan kejadian Y pada satu subyek pengamatan. Misalnya akan mempelajari korelasi antara aktivitas fisik dengan kapasitas paru penderita fibrosis paru. Pada penelitian ini, angka aktivitas fisik pada satu pasien dihubungkan dengan kapasitas paru dari pasien itu sendiri dan tidak boleh tingkat aktivitas fisik pasien A dihubungkan dengan kapasitas paru pasien B. Kejadian yang akan dihubungkan bisa dikumpulkan pada saat yang bersamaan seperti pada penelitian *cross-sectional* atau penelitian survei, bisa juga kedua kejadian yang dihubungkan dikumpulkan pada waktu yang berbeda seperti pada penelitian dengan rancangan *pretest-posttest* dimana kejadian pertama diukur pada awal studi dan kejadian kedua diukur pada akhir studi.

Metode analisis korelasi dipakai untuk menilai hubungan antara variabel yang akan diteliti. Metode korelasi dibedakan berdasarkan skala pengukuran dari variabel yang akan dihubungkan menjadi metode korelasi parametrik dan non-parametrik. Metode korelasi parametrik dipakai menganalisis hubungan antara variabel numerikal. Metode analisis korelasi yang termasuk metode korelasi parametrik adalah metode korelasi *product-moment* Pearson. Metode analisis korelasi *product moment* Pearson atau disebut Korelasi Pearson memiliki asumsi yaitu: data berskala numerikal dan

berdistribusi normal. Metode analisis korelasi non-parametrik dipakai apabila data yang akan dihubungkan berskala ordinal atau nominal dan tidak perlu asumsi data berdistribusi normal. Metode analisis korelasi yang termasuk korelasi non-parametrik adalah korelasi Spearman Rank atau disebut korelasi Spearman dan korelasi Kendal Tau atau disebut korelasi Kendal.

Koefisien korelasi Pearson yang diberi simbol “r” atau koefisien korelasi Spearman yang diberi simbol “rs” dipakai untuk menilai kuat dan arah hubungan. Arah korelasi dilihat dari tanda (+/-) dari koefisien. Apabila angka koefisien positif, maka hubungannya positif atau searah yang artinya semakin besar nilai variabel yang satu akan diikuti oleh semakin besar nilai variabel yang kedua. Sebaliknya, apabila koefisiennya negatif berarti terjadi hubungan yang berlawanan dimana semakin tinggi nilai variabel yang satu akan diikuti oleh penurunan nilai dari variabel kedua. Nilai absolut dari koefisien korelasi menggambarkan kuatnya hubungan. Hubungan dinyatakan kuat apabila nilai koefisien korelasi $\geq 0,7$ dan dinyatakan sempurna apabila nilai koefisien korelasi sama dengan satu ($r = 1$).

7.3 Parameter penentuan besar sampel

Parameter yang menentukan besar sampel pada penelitian korelasi adalah kuat hubungan (*effect size*) terkecil yang harus mampu diidentifikasi sebagai hubungan yang nyata (signifikan), level kepercayaan dan *power* penelitian yang diinginkan.

Effect size:

Kuat hubungan (*effect size*) terkecil yang harus mampu dideteksi sebagai hubungan yang bermakna ditentukan dengan nilai koefisien korelasi “r”. Makin kecil nilai r yang ditetapkan sebagai parameter sampel, maka makin besar sampel yang dibutuhkan. Umumnya dipakai nilai $r = 0,3$ sebagai acuan.

Level kepercayaan:

Parameter level kepercayaan menentukan tingkat reliabilitas sampel. Level kepercayaan yang dipakai dalam penentuan besar sampel ditunjukkan untuk meminimalisir kesalahan type I. Umumnya level kepercayaan yang dipakai adalah 95% atau 99%. Apabila level kepercayaan yang dipakai 95% maka ada kemungkinan terjadi kesalahan type I sebesar 5%. Kesalahan type I berkaitan dengan pengambilan keputusan uji hipotesis. Dengan type I error 5% berarti ada kemungkinan kesimpulan menyatakan ada hubungan dimana sebenarnya tidak ada hubungan sebesar 5%.

Power penelitian:

Power penelitian ditetapkan dengan tujuan untuk memperkecil adanya kesalahan type II. Besar nilai *Power* penelitian yang umum dipakai adalah 90% atau 80%. Apabila *power* penelitian ditetapkan sebesar 90%, maka ada kemungkinan 10% salah dalam pengambilan kesimpulan dimana yang seharusnya ada hubungan (H_0 ditolak) dinyatakan tidak ada hubungan (H_0 diterima).

7.4 Cara Penentuan Besar Sampel

Besar sampel penelitian korelasi ditentukan dengan parameter sebagai berikut.

- 1) Kuat hubungan (*Effect size*) yang diinginkan dan besarnya ditentukan dari nilai koefisien korelasi (r) yang diinginkan.
- 2) Tingkat kesalahan type I (α) atau tingkat reliabilitas sampel ($1-\alpha$)
- 3) Tingkat kesalahan type II (β) atau *power* penelitian ($1-\beta$)

Rumus Penentuan Besar Sampel Penelitian Korelasi
$n = \left[\frac{Z_{\alpha} + Z_{\beta}}{0,5 \ln \{ (1+r)/(1-r) \}} \right]^2 + 3$
<p>Keterangan:</p> <p>n = jumlah sampel pasangan Z_{α} = koefisien reliabilitas untuk α tertentu Z_{β} = koefisien <i>power</i> penelitian untuk β tertentu r = koefien korelasi yang ingin dideteksi</p>

Contoh kasus

Seorang peneliti ingin meneliti korelasi antara asupan buah merah dengan kadar vit A pada balita. Dari studi pendahuluan didapatkan koefisien korelasi antara asupan buah merah dengan kadar vit A sebesar 0,30. Besar sampel yang diperlukan pada penelitian ini, bila peneliti menginginkan tingkat reliabilitas sampel 95% dengan *power* penelitian 90% dengan hipotesis satu sisi, dihitung sebagai berikut.

Asumsi yang dipakai adalah:

- 1) $r = 0,30$ (perkiraan koefisien korelasi di populasi)
- 2) $Z_{\alpha} = 1,96$ (untuk tingkat reliabilitas 95%)
- 3) $Z_{\beta} = 1,28$ (untuk *power* penelitian sebesar 80%)

Jumlah sampel yang dibutuhkan menjadi:

$$n = \left[\frac{Z_{\alpha} + Z_{\beta}}{0,5 \ln\{(1+r)/(1-r)\}} \right]^2 + 3$$

$$n = \left[\frac{1,96 + 1,28}{0,5 \ln\{(1+0,3)/(1-0,3)\}} \right]^2 + 3 = 112,2$$

Dengan demikian, maka jumlah sampel minimal yang diperlukan adalah sebanyak 113 orang.

7.8 Cara pemilihan sampel

Sampel pada penelitian korelasi dapat dipilih secara *random* dan juga bisa secara *non-random*. Pemilihan secara *random* bisa dilakukan apabila kerangka sampel bisa dibuat. Misalnya akan diteliti hubungan pengetahuan, sikap, dan praktek pemberian ASI eksklusif pada ibu balita di Kecamatan Gianyar. Data ibu balita terdapat pada register posyandu di setiap banjar yang ada di Kecamatan Gianyar. Pada kasus penelitian ini, sampel ibu balita bisa dipilih secara *random*. Sebaliknya, bila kerangka sampel tidak bisa dibuat, misalnya akan diteliti korelasi antara pengetahuan dan praktek penerapan protokol kesehatan COVID-19 pada pengguna obat bius di Kota Denpasar. Daftar pengguna obat bius di Kota Denpasar tidak mungkin atau sangat sulit didapat, maka pemilihan sampel untuk penelitian ini dapat dilakukan secara *non-random* seperti *snowballing sampling*.

Tahapan pemilihan sampel

1) Penetapan unit analisis

Pada tahap pertama harus ditetapkan unit analisis dari penelitian. Misalnya akan diteliti hubungan aktivitas fisik dengan kapasitas paru pasien fibrosis paru, maka yang menjadi unit analisis adalah pasien fibrosis paru.

2) Penetapan populasi terjangkau

Populasi terjangkau adalah kumpulan dari unit analisis dari mana sampel akan dipilih. Misalnya Penelitian hubungan aktivitas fisik dengan kapasitas paru pasien fibrosis paru yang dirawat inap di RSUP Sanglah periode Januari sampai Desember 2020, maka yang menjadi populasi terjangkau adalah pasien fibrosis paru yang rawat inap di RSUP Sanglah pada periode Januari sampai Desember 2020

3) Penetapan unit sampel

Unit sampel pada penelitian korelasi umumnya sama dengan unit analisisnya. Pada contoh di atas unit sampel penelitiannya sama dengan unit analisisnya, yaitu pasien fibrosis paru yang dirawat di

RSUP Sanglah pada periode Januari sampai Desember 2020. Selanjutnya ditetapkan kriteria eligibilitas sampel yang dijabarkan ke dalam kriteria inklusi dan kriteria eksklusi agar sampel yang terpilih merupakan representasi dari populasi penelitian. Kerangka sampel dibuat dari daftar semua pasien yang memenuhi kriteria eligibilitas.

4) Pemilihan sampel:

Pemilihan sampel bisa dilakukan secara *random* dan *non-random*.

Pemilihan sampel secara *random*:

Sampel dipilih dari kerangka sampel penelitian sebanyak yang dibutuhkan secara *random* dengan bantuan Tabel Bilangan *Random* atau dengan bantuan aplikasi komputer.

Pemilihan sampel secara *non-random*:

Pada pemilihan sampel secara *non-random*, rekam medis dari pasien fibrosis paru yang dirawat di RSUP Sanglah pada periode Januari sampai Desember 2020 diperiksa datanya. Apabila kriteria inklusi dan eksklusi dipenuhi, maka pasien tersebut dijadikan sampel. Pemilihan sampel berikutnya dilakukan dengan cara yang sama sampai jumlah sampel yang dibutuhkan terpenuhi. Pemilihan rekam medis yang diperiksa biasanya dimulai dari yang terbaru.

7.9 LATIHAN

Kasus:

Dr Mirah, M.Kes akan meneliti korelasi antara kemampuan aktivitas fisik dengan fungsi paru pada penderita fibrosis paru yang dirawat di RSUP. Aktivitas fisik diukur dengan skor aktivitas fisik dan fungsi paru diukur dari VO_2Max .

TUGAS:

1. Tentukan siapa yang menjadi populasi target, terjangkau, dan sampelnya!
2. Apa parameter yang dibutuhkan dalam penentuan jumlah sampelnya?
3. Berapa jumlah sampel yang dibutuhkan bila tingkat kesalahan tipe I $\alpha = 0,05$; tingkat kesalahan tipe II 20%; *effect size* $r = 0,3$?
4. Jelaskan metode sampling apa yang akan dipakai dan cara pemilihan sampelnya?

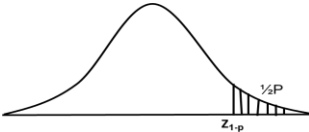
LAMPIRAN

LAMPIRAN 1. TABEL BILANGAN RANDOM

34	81	46	80	20	28	99	85	16	87	40	8	83	54	58	24	70	80	18	15	53	77	61	41
99	10	65	8	99	7	39	49	10	71	22	41	16	16	43	64	43	56	25	52	64	31	69	89
47	18	53	17	82	48	89	46	87	90	51	85	23	69	18	5	73	79	6	53	61	12	34	14
81	35	45	72	96	39	32	95	22	63	43	19	45	16	24	42	42	96	13	67	41	20	79	81
83	46	78	56	22	95	77	80	53	47	0	44	55	13	34	29	44	59	91	76	30	9	11	30
27	92	43	41	67	57	6	5	75	35	32	27	81	69	49	4	96	0	41	16	91	46	74	91
58	31	98	81	64	94	13	0	77	43	16	91	71	56	81	93	57	29	97	53	49	11	72	6
49	73	26	67	2	72	42	59	91	17	49	29	88	72	65	14	19	58	33	91	19	79	42	38
69	59	42	67	49	68	74	33	91	39	44	49	51	19	44	12	97	5	65	23	62	41	12	79
30	7	49	75	17	97	45	5	42	51	51	77	72	10	73	3	90	92	65	19	39	57	83	43
81	14	75	75	8	93	47	98	78	26	28	96	57	44	74	97	46	88	1	49	17	83	44	5
74	68	92	89	33	59	81	99	30	52	61	32	58	31	45	44	68	47	29	58	91	82	42	88
14	80	22	59	82	48	2	41	54	61	37	57	4	46	85	47	38	60	95	4	77	83	19	10
68	50	13	41	94	85	35	53	84	11	46	55	94	16	94	99	67	88	82	68	86	67	42	4
48	73	49	26	71	85	25	28	59	85	34	22	89	12	89	56	33	11	46	83	36	49	39	48
84	10	28	16	99	97	35	25	45	9	93	19	65	12	4	43	35	11	18	18	41	17	91	19
28	43	23	28	73	83	83	40	98	62	12	72	6	45	69	42	21	82	67	26	80	86	66	44
91	45	88	25	24	75	52	30	12	76	19	59	14	74	73	90	25	19	91	3	72	94	74	21
45	43	53	3	89	69	73	28	19	62	30	24	39	61	99	33	54	63	96	72	83	76	2	95
23	55	77	84	37	44	95	79	87	28	65	67	43	47	1	83	39	85	41	2	65	84	58	11
30	39	59	18	50	52	00	40	48	44	28	84	81	97	28	96	57	11	33	17	70	85	97	06
69	99	11	27	55	97	91	65	76	39	43	44	59	4	63	85	55	63	54	69	19	74	95	95
32	98	04	36	08	20	59	27	25	27	63	58	34	64	43	53	92	98	72	19	55	19	89	04
59	77	63	70	35	53	76	55	51	96	68	65	97	14	29	25	22	59	19	42	51	13	26	67
73	71	47	98	68	23	88	9	22	54	72	98	40	77	92	81	30	62	42	77	82	36	63	51
61	54	71	65	75	68	52	05	98	69	67	29	97	35	65	77	31	69	66	82	18	03	17	03
87	73	70	10	58	76	01	27	62	47	75	61	65	13	35	70	36	47	89	42	40	56	44	93
98	66	90	83	34	40	52	07	83	89	56	49	87	43	36	02	43	44	85	99	67	57	90	95
81	53	54	66	90	92	81	44	44	56	29	22	74	81	22	30	52	21	38	52	48	43	64	08
05	97	54	71	10	32	73	34	69	29	98	26	15	21	93	83	21	55	31	92	83	10	96	89
03	47	73	86	36	96	47	36	61	46	98	63	71	62	33	26	16	80	45	60	11	14	10	11
97	74	24	67	62	42	81	14	57	20	42	53	32	37	32	27	07	36	07	51	24	51	74	95
16	76	62	27	66	56	50	26	71	07	32	90	79	78	53	13	55	38	58	59	88	97	58	44
12	56	85	99	26	96	96	68	27	31	05	03	72	93	15	57	12	10	14	21	88	26	43	13

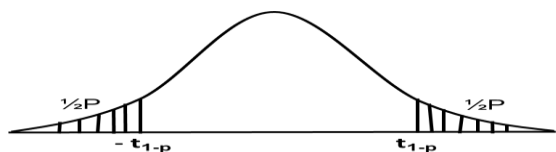
55	59	56	35	64	38	54	82	46	22	31	62	43	09	90	06	18	44	32	53	23	83	04	17
16	22	77	94	39	49	54	43	54	82	17	37	93	23	78	87	35	20	96	43	84	26	34	91
84	42	17	53	31	57	24	55	6	88	77	4	74	47	67	21	76	33	50	25	83	92	12	6
63	01	63	78	59	16	95	55	67	19	98	10	50	71	75	12	86	73	58	07	44	39	52	38
33	21	12	34	29	78	64	56	07	82	52	42	07	44	38	15	51	00	13	42	99	66	02	79
57	60	86	32	44	09	47	27	96	54	49	17	46	09	62	90	52	84	77	27	08	02	73	43
05	26	93	70	60	22	35	85	15	13	92	03	51	59	77	59	56	78	06	83	52	91	05	70
07	97	10	88	23	09	98	42	99	64	61	71	62	99	15	06	51	29	16	93	58	5	77	9
68	71	86	85	85	54	87	66	47	54	73	32	08	11	12	44	95	92	63	16	29	56	24	29
26	99	61	65	53	58	37	78	80	70	42	10	50	67	42	32	17	55	85	74	94	44	67	16
14	65	52	68	75	87	59	36	22	41	26	78	63	06	55	13	08	27	01	50	15	29	39	39
17	53	77	58	71	71	41	61	50	72	12	41	94	96	26	44	95	27	36	99	02	96	74	30
90	26	59	21	19	23	52	23	33	12	96	93	02	18	39	07	02	18	36	07	25	99	32	70
41	23	52	55	99	31	04	49	69	96	10	47	48	45	88	13	41	43	89	20	97	17	14	49
60	20	50	81	69	31	99	73	68	68	35	81	33	03	76	24	30	12	48	60	18	99	10	72
91	25	38	05	90	94	58	28	41	36	45	37	59	03	09	90	35	57	29	12	82	62	54	65
34	50	57	74	37	98	80	33	00	91	09	77	93	19	82	74	94	80	04	04	45	07	31	66
85	22	04	39	43	73	81	53	94	79	33	63	46	86	28	08	31	54	46	31	53	94	13	38
9	79	13	77	48	73	82	97	22	21	05	03	27	24	83	72	89	44	05	60	35	80	39	94
88	75	80	18	14	22	95	75	42	49	39	32	82	22	49	02	48	07	70	37	16	04	61	67
90	96	23	70	00	39	00	03	06	90	55	85	78	38	36	94	37	30	69	32	90	89	00	76
27	92	43	41	67	57	06	05	75	35	32	27	81	69	49	04	96	00	41	16	91	46	74	91
58	31	98	81	64	94	13	00	77	43	16	91	71	56	81	93	57	29	97	53	49	11	72	06
49	73	26	67	02	72	42	59	91	17	49	29	88	72	65	14	19	58	33	91	19	79	42	38
69	59	42	67	49	68	74	33	91	39	44	49	51	19	44	12	97	05	65	23	62	41	12	79
30	07	49	75	17	97	45	05	42	51	51	77	72	10	73	03	90	92	65	19	39	57	83	43
81	14	75	75	08	93	47	98	78	26	28	96	57	44	74	97	46	88	01	49	17	83	44	05
74	68	92	89	33	59	81	99	30	52	61	32	58	31	45	44	68	47	29	58	91	82	42	88
14	80	22	59	82	48	02	41	54	61	37	57	04	46	85	47	38	60	95	04	77	83	19	10
68	50	13	41	94	85	35	53	84	11	46	55	94	16	94	99	67	88	82	68	86	67	42	04
48	73	49	26	71	85	25	28	59	85	34	22	89	12	89	56	33	11	46	83	36	49	39	48
84	10	28	16	99	97	35	25	45	09	93	19	65	12	04	43	35	11	18	18	41	17	91	19
28	43	23	28	73	83	83	40	98	62	12	72	6	45	69	42	21	82	67	26	80	86	66	44
91	45	88	25	24	75	52	30	12	76	19	59	14	74	73	90	25	19	91	03	72	94	74	21
45	43	53	03	89	69	73	28	19	62	30	24	39	61	99	33	54	63	96	72	83	76	02	95
23	55	77	84	37	44	95	79	87	28	65	67	43	47	01	83	39	85	41	02	65	84	58	11

LAMPIRAN 2. TABEL DISTRIBUSI NORMAL



Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641
0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1819	0,1867
0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
2,0	0,02275	0,02222	0,02169	0,02118	0,02068	0,02018	0,01970	0,01923	0,01876	0,01831
2,1	0,01786	0,01743	0,01700	0,01659	0,01618	0,01578	0,01539	0,01500	0,01463	0,01426
2,2	0,01390	0,01355	0,01321	0,01287	0,01255	0,01222	0,01191	0,01160	0,01130	0,01101
2,3	0,01072	0,01044	0,01017	0,00990	0,00964	0,00939	0,00914	0,00889	0,00866	0,00842
2,4	0,00820	0,00798	0,00776	0,00755	0,00734	0,00714	0,00695	0,00676	0,00657	0,00639
2,5	0,00621	0,00604	0,00587	0,00570	0,00554	0,00539	0,00523	0,00508	0,00494	0,00480
2,6	0,00466	0,00453	0,00440	0,00427	0,00415	0,00402	0,00391	0,00379	0,00368	0,00357
2,7	0,00347	0,00336	0,00326	0,00317	0,00307	0,00298	0,00289	0,00280	0,00272	0,00264
2,8	0,00256	0,00248	0,00240	0,00233	0,00226	0,00219	0,00212	0,00205	0,00199	0,00193
2,9	0,00187	0,00181	0,00175	0,00169	0,00164	0,00159	0,00154	0,00149	0,00144	0,00139
3,0	0,00135									
3,5	0,00023									
4,0	0,00003									

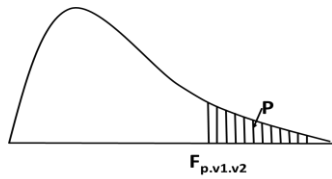
LAMPIRAN 3. TABEL DISTRIBUSI “t”



DF	$t_{0,90}$	$t_{0,80}$	$t_{0,50}$	$t_{0,40}$	$t_{0,30}$	$t_{0,20}$	$t_{0,10}$	$t_{0,05}$	$t_{0,01}$	$t_{0,001}$
1	0,158	0,325	1,000	1,376	1,963	3,078	6,314	12,706	63,657	636,619
2	0,142	0,289	0,816	1,061	1,386	1,886	2,920	3,303	9,925	31,598
3	0,137	0,277	0,765	0,978	1,250	1,638	2,353	3,182	5,841	12,924
4	0,134	0,271	0,741	0,941	1,190	1,533	2,132	2,776	4,604	8,610
5	0,132	0,267	0,727	0,920	1,156	1,476	2,015	2,571	4,032	6,869
6	0,131	0,265	0,718	0,906	1,134	1,440	1,943	2,447	3,707	5,959
7	0,130	0,263	0,711	0,896	1,119	1,415	1,895	2,365	3,499	5,408
8	0,130	0,262	0,706	0,889	1,108	1,397	1,860	2,306	3,355	5,041
9	0,129	0,261	0,703	0,883	1,100	1,383	1,833	2,262	3,250	4,781
10	0,129	0,260	0,700	0,879	1,093	1,372	1,812	2,228	3,169	4,587
11	0,129	0,260	0,697	0,876	1,088	1,363	1,796	2,201	3,106	4,437
12	0,128	0,259	0,695	0,873	1,083	1,356	1,782	2,179	3,055	4,318
13	0,128	0,259	0,694	0,870	1,079	1,350	1,771	2,160	3,012	4,221
14	0,128	0,258	0,692	0,868	1,076	1,345	1,761	2,145	2,977	4,140
15	0,128	0,258	0,691	0,866	1,074	1,341	1,753	2,131	2,947	4,073
16	0,128	0,258	0,690	0,865	1,071	1,337	1,746	2,120	2,921	4,015
17	0,128	0,257	0,689	0,863	1,069	1,333	1,740	2,110	2,898	3,965
18	0,127	0,257	0,688	0,862	1,067	1,330	1,734	2,101	2,878	3,922
19	0,127	0,257	0,688	0,861	1,066	1,328	1,729	2,093	2,861	3,883
20	0,127	0,257	0,687	0,860	1,064	1,325	1,725	2,086	2,845	3,850
21	0,127	0,257	0,686	0,859	1,063	1,323	1,721	2,080	2,831	3,819
22	0,127	0,256	0,686	0,858	1,061	1,321	1,717	2,074	2,819	3,792
23	0,127	0,256	0,685	0,858	1,060	1,319	1,714	2,069	2,807	3,767
24	0,127	0,256	0,685	0,857	1,059	1,318	1,711	2,064	2,797	3,745
25	0,127	0,256	0,684	0,856	1,058	1,316	1,708	2,060	2,787	3,725
26	0,127	0,256	0,684	0,856	1,058	1,315	1,706	2,056	2,779	3,707
27	0,127	0,256	0,684	0,855	1,057	1,314	1,703	2,052	2,771	3,690
28	0,127	0,256	0,683	0,855	1,056	1,313	1,701	2,048	2,736	3,674
29	0,127	0,256	0,683	0,854	1,055	1,311	1,699	2,045	2,756	3,659
30	0,127	0,256	0,683	0,854	1,055	1,310	1,697	2,042	2,750	3,646
40	0,126	0,255	0,681	0,851	1,050	1,303	1,684	2,021	2,704	3,551

60	0,126	0,254	0,679	0,848	1,046	1,296	1,671	2,000	2,660	3,460
120	0,126	0,254	0,677	0,845	1,041	1,289	1,658	1,980	2,617	3,373
∞	0,126	0,253	0,674	0,842	1,036	1,282	1,645	1,960	2,576	3,291

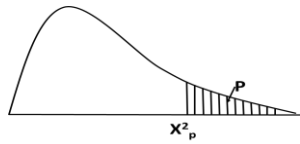
LAMPIRAN 4. TABEL DISTRIBUSI F



DF denominator		DF numerator, V1										
V2	P	1	2	3	4	5	6	7	8	12	24	∞
1	0,05	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	243,9	249,1	254,3
	0,01	4052	5000	5403	5625	5764	5859	5928	5981	6106	6235	6366
2	0,05	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,41	19,45	19,50
	0,01	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,42	99,46	99,50
3	0,05	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,74	8,64	8,53
	0,01	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,05	26,60	26,13
4	0,05	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	5,91	5,77	5,63
	0,01	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,37	13,93	13,46
5	0,05	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,68	4,53	4,36
	0,01	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	9,89	9,47	9,02
6	0,05	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,00	3,84	3,67
	0,01	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,72	7,31	6,88
7	0,05	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,57	3,41	3,23
	0,01	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,47	6,07	5,65
8	0,05	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,28	3,12	2,93
	0,01	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,67	5,28	4,86
9	0,05	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,07	2,90	2,71
	0,01	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,11	4,73	4,31
10	0,05	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	2,91	2,74	2,54
	0,01	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,71	4,33	3,91
12	0,05	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,69	2,51	2,30
	0,01	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,16	3,78	3,36
14	0,05	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,53	2,35	2,13
	0,01	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	3,80	3,43	3,00
16	0,05	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,42	2,24	2,01
	0,01	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,55	3,18	2,75

18	0,05	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,34	2,15	1,92
	0,01	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,37	3,00	2,57
20	0,05	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,28	2,08	1,84
	0,01	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,23	2,86	2,42
30	0,05	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,09	1,89	1,62
	0,01	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	2,84	2,47	2,01
40	0,05	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,00	1,79	1,51
	0,01	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,66	2,29	1,80
60	0,05	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	1,92	1,70	1,39
	0,01	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,50	2,12	1,60
120	0,05	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,83	1,61	1,25
	0,01	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,34	1,95	1,38
∞	0,05	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,75	1,52	1,00
	0,01	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,18	1,79	1,00

LAMPIRAN 5. TABEL DISTRIBUSI CHI-SQUARE



DF	$X^2_{0,975}$	$X^2_{0,900}$	$X^2_{0,750}$	$X^2_{0,500}$	$X^2_{0,250}$	$X^2_{0,100}$	$X^2_{0,050}$	$X^2_{0,025}$	$X^2_{0,010}$	$X^2_{0,001}$
1		0,02	0,10	0,45	1,32	2,71	3,84	5,02	6,63	10,83
2	0,05	0,21	0,58	1,39	2,77	4,61	5,99	7,38	9,21	13,82
3	0,22	0,58	1,21	2,37	4,11	6,25	7,81	9,35	11,34	16,27
4	0,48	1,06	1,92	3,36	5,39	7,78	9,49	11,14	13,28	18,47
5	0,83	1,61	2,67	4,35	6,63	9,24	11,07	12,83	15,09	20,52
6	1,24	2,20	3,45	5,35	7,84	10,64	12,59	14,45	16,81	22,46
7	1,69	2,83	4,25	6,35	9,04	12,02	14,07	16,01	18,48	24,32
8	2,18	3,49	5,07	7,34	10,22	13,36	15,51	17,53	20,09	26,12
9	2,70	4,17	5,90	8,34	11,39	14,68	16,92	19,02	21,67	27,88
10	3,25	4,87	6,74	9,34	12,55	15,99	18,31	20,48	23,21	29,59
11	3,82	5,58	7,58	10,34	13,70	17,28	19,68	21,92	24,72	31,26
12	4,40	6,30	8,44	11,34	14,85	18,55	21,03	23,34	26,22	32,91
13	5,01	7,04	9,30	12,34	15,98	19,81	22,36	24,74	27,69	34,53
14	5,63	7,79	10,17	13,34	17,12	21,06	23,68	26,12	29,14	36,12
15	6,27	8,55	11,04	14,34	18,25	22,31	25,00	27,49	30,58	37,70
16	6,91	9,31	11,91	15,34	19,37	23,54	26,30	28,85	32,00	39,25
17	7,56	10,09	12,79	16,34	20,49	24,77	27,59	30,19	33,41	40,79
18	8,23	10,86	13,68	17,34	21,60	25,99	28,87	31,53	34,81	42,31
19	8,91	11,65	14,56	18,34	22,72	27,20	30,14	32,85	36,19	43,82
20	9,59	12,44	15,45	19,34	23,83	28,41	31,41	34,17	37,57	45,32
21	10,28	13,24	16,34	20,34	24,93	29,62	32,67	35,48	38,93	46,80
22	10,98	14,04	17,24	21,34	26,04	30,81	33,92	36,78	40,29	48,27
23	11,69	14,85	18,14	22,34	27,14	32,01	35,17	38,08	41,64	49,73
24	12,40	15,66	19,04	23,34	28,24	33,20	36,42	39,36	42,98	51,18
25	13,12	16,47	19,94	24,34	29,34	34,38	37,65	40,65	44,31	52,62
26	13,84	17,29	20,84	25,34	30,43	35,56	38,89	41,92	45,64	54,05
27	14,57	18,11	21,75	26,34	31,53	36,74	40,11	43,19	46,96	55,48
28	15,31	18,94	22,66	27,34	32,62	37,92	41,34	44,46	48,28	56,89

29	16,05	19,77	23,57	28,34	33,71	39,09	42,56	45,72	49,59	58,30
30	16,79	20,60	24,48	29,34	34,80	40,26	43,77	46,98	50,89	59,70
40	24,43	29,05	33,66	39,34	45,62	51,80	55,76	59,34	63,69	73,40
50	32,36	37,69	42,94	49,33	56,33	63,17	67,50	71,42	76,15	86,66
60	40,48	46,46	52,29	59,33	66,98	74,40	79,08	83,30	88,38	99,61
70	48,76	55,33	61,70	69,33	77,58	85,53	90,53	95,02	100,42	112,32
80	57,15	64,28	71,14	79,33	88,13	96,58	101,88	106,63	112,33	124,84
90	65,65	73,29	80,62	89,33	98,64	107,56	113,14	118,14	124,12	137,21
100	74,22	82,36	90,13	99,33	109,14	118,50	124,34	129,56	135,81	149,45

Informasi Tentang Penulis

I Ketut Tangking Widarsa adalah alumnus Fakultas Kedokteran Universitas Udayana dan juga alumnus dari School of Public Health, Hawaii University dalam bidang Ilmu Biostatistika. Penulis merupakan pengajar ilmu biostatistika di Fakultas Kedokteran dan di Program Studi Ilmu Kesehatan Masyarakat Universitas Udayana serta sebagai dosen Ilmu Kedokteran Komunitas dan Ilmu Kedokteran Pencegahan di Fakultas Kedokteran dan Ilmu Kesehatan Universitas Warmadewa, Bali. Buku lain yang telah ditulis berupa Buku Ajar Biostatistik Dasar (2017), Buku Ajar Analisis Regresi Dengan SPSS: Konsep, Prosedur, dan Interpretasi (2010), Modul Praktikum Stata Biostatistik II (2017), dan Modul Analisis Data Untuk Variabel Outcome Berskala Nominal Dua Kategori (Binary Outcome) (2016).

Ni Made Dian Kurniasari merupakan lulusan Sarjana Kesehatan Masyarakat peminatan Biostatistik dari Program Studi Sarjana Kesehatan Masyarakat, Kedokteran Universitas Udayana tahun 2011 dan S2 Master of Public Health konsentrasi Biostatistik dari The University of Adelaide Australia tahun 2015. Penulis merupakan dosen dan ketua Divisi Biostatistik, Demografi dan KIA-KESPRO, Departemen Kesehatan Masyarakat dan Kedokteran Pencegahan Fakultas Kedokteran Universitas Udayana. Saat ini aktif mengajar di Program Studi Sarjana Kesehatan Masyarakat dan prodi lain dilingkungan FK Unud untuk mata kuliah biostatistik dasar maupun lanjutan.

Putu Ayu Swandewi Astuti adalah dokter lulusan Universitas Udayana yang kemudian menyelesaikan pendidikan S2 Kesehatan Masyarakat di bidang epidemiologi dan biostatistik dari The University of Melbourne dan pendidikan S3 di The University of Sydney. Penulis merupakan dosen pengajar untuk mata kuliah Biostatistik dasar dan peminatan di Fakultas Kedokteran Universitas Udayana.



BASWARA
PRESS

ISBN 978-623-99689-0-8 (PDF)



9 786239 968908