

Menghitung Besar Sampel Kasus – Kontrol dan Eksperimen

Sri Poedji Hastoety Djaiman
Badan Litbangkes - Kemenkes RI

Sumber utama:

dr. Iwan Ariawan, MSc.PH
FKM-UI



■ **Populasi**

- **Populasi** adalah keseluruhan subyek penelitian yang memenuhi karakteristik yang ditentukan dimana penelitian akan di generalisir
- **Macam populasi:**
 - ❖ Populasi target
Kumpulan dari unit yang ingin kita buat generalisasinya/Lebih refresentatif untuk informasi yang dikehendaki.
 - ❖ Populasi studi
kumpulan dari unit dimana kita mengambil sampel
 - ❖ Populasi terjangkau
bagian dari populasi target yang dibatasi oleh tempat dan waktu → dari populasi terjangkau ini sampel yang akan diteliti diambil
- **Syarat populasi**
 - ❖ Meliputi seluruh unit sampel
 - ❖ Sampel tidak dihitung dua kali
 - ❖ Batas jelas
 - ❖ Informasi terbaru
 - ❖ Dapat dilacak dilapangan

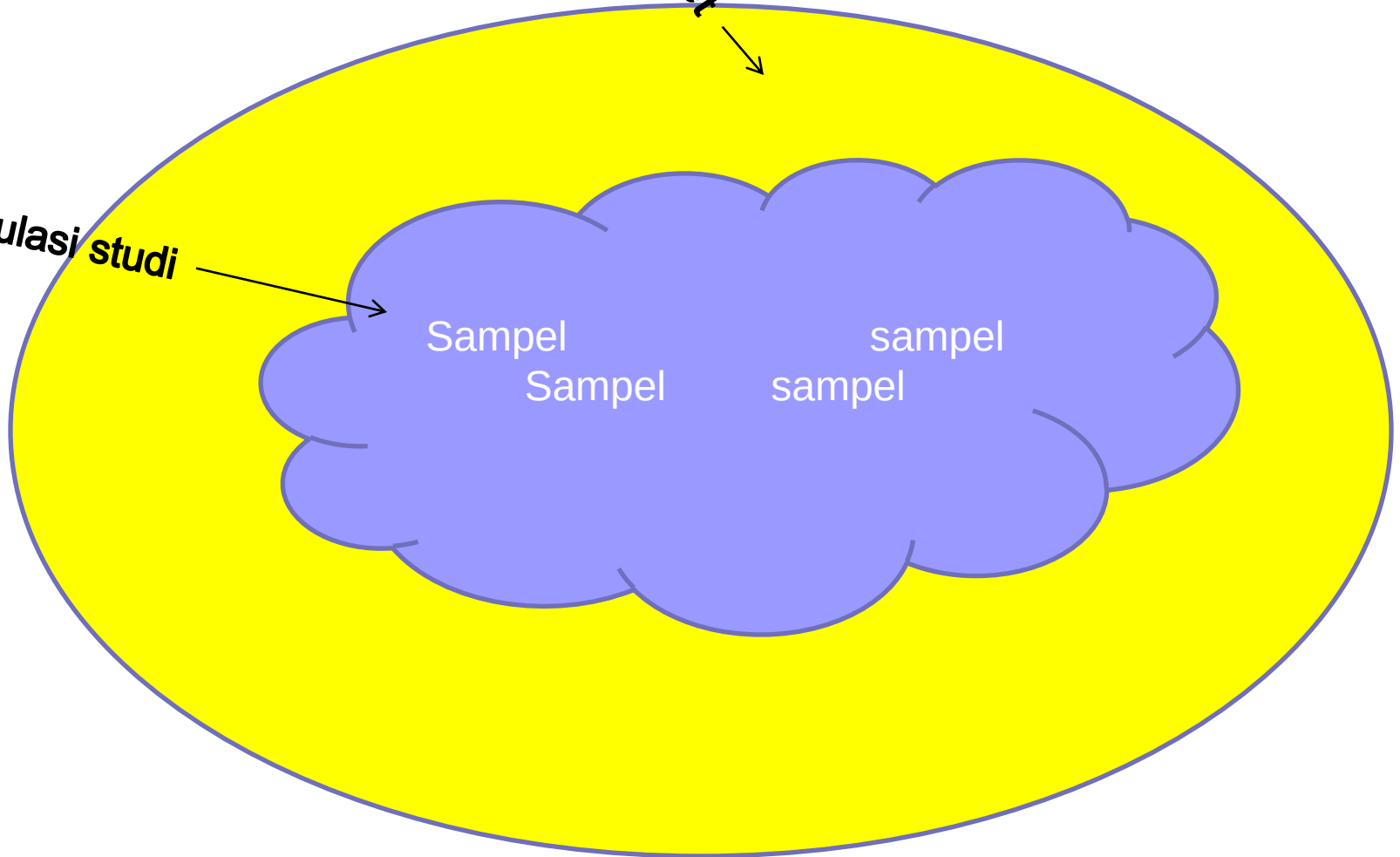
Keterkaitan populasi

Populasi target

Populasi studi

Sampel
Sampel

sampel
sampel



Sampel

- **Sampel** adalah kumpulan dari satuan/unit elemen yang kita ambil dari populasi studi dimana pengukuran dilakukan
- **Unit Elemen** adalah anggota individu dimana pengukuran dilakukan.
- **Sampling Frame** adalah suatu daftar sampling unit/unit elemen.
- **Mengapa perlu sampling**
 1. Ada Populasi yg anggota individunya jumlahnya sangat besar .
 2. Homogenitas, populasi yang homogen tidak perlu diperiksa semua.
 3. Menghemat biaya dan waktu
 4. Ketelitian/ketepatan pengukuran.
 5. Tidak cukup tenaga
- **Prosedur pengambilan sampel**
 1. Tujuan penelitian
 2. Populasi penelitian
 3. Besarnya sampel
 4. Memilih sampel

➤ Syarat sampel yang ideal:

1. Dapat menghasilkan gambaran yang tepat karakter populasi.
2. Dapat menentukan presisi (ketepatan) dari hasil penelitian.
3. Sederhana, mudah dilaksanakan
4. Dapat memberikan keterangan sebanyak mungkin dengan biaya serendah mungkin.

➤ Macam sampling

1. sampel yang dikehendaki
Merupakan merupakan bagian populasi target yang akan diteliti secara langsung
2. Sampel yang diteliti
subyek yang benar-benar diteliti, bagian dari sampel yang dikehendaki setelah dikurangi sampel “drop out”

MENGHITUNG BESAR SAMPEL

Beberapa hal yang perlu diperhatikan dalam menghitung besar sampel:

1. Biaya, waktu, dan tenaga yang tersedia
2. Variasi dalam variabel , semakin heterogen semakin besar sampel yang diperlukan
3. Banyaknya variabel yang diamati, semakin banyak variabel semakin besar sampel yang dibutuhkan
4. Presisi atau ketepatan yang dikehendaki.
5. Rencana analisis. Bila analisis hanya dengan manual tidak mungkin dapat menganalisis data yang banyak sekali

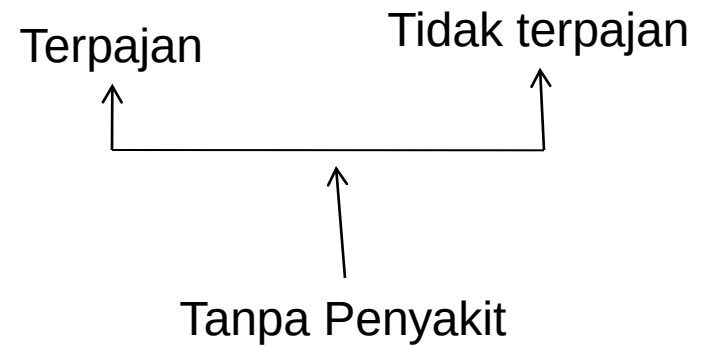
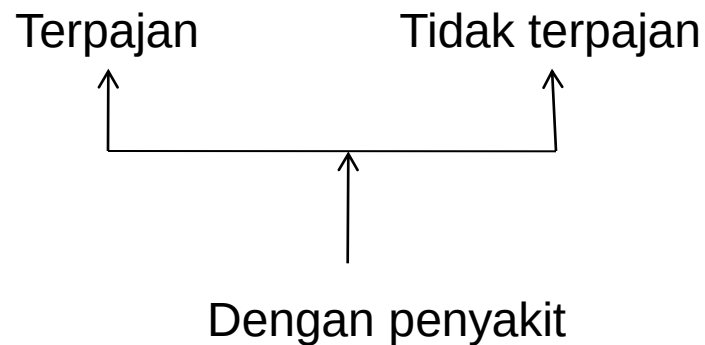


Pertimbangan dalam menentukan besar sampel

1. Desain penelitian
2. Tujuan penelitian
3. Jenis data dependen variable
4. Jenis analisis yang akan digunakan

Besar Sampel Pada Penelitian Kasus Control:

Prinsip dasar dalam penentuan sampel Kasus - Kontrol





Hal yang perlu diperhatikan pada kasus – kontrol:

- Pengambilan sampel dimulai dengan identifikasi
- Untuk memperoleh n kasus, perlu memeriksa n orang, yang jumlahnya tergantung prevalensi kasus di populasi
- Definisi kasus sangat penting
- Secara idealnya kontrol harus dari populasi yang sama
- Tidak dapat digunakan untuk menghitung prevalensi

Beberapa Cara Penghitungan Besar Sampel Pada Penelitian Casus Control:

1. Penghitungan sampel Estimasi Ratio Odds
2. Penghitungan sampel Uji Hipotesis Ratio Odds
3. Penghitungan sampel Kasus-Kontrol dengan K kontrol per kasus
4. Penghitungan sampel Kasus-Kontrol sepadan

1. Besar sampel Ratio Odds

- Ditentukan berdasarkan nilai ratio odds (OR)
- Selalu positif
- Jarang mencapai 100
- Distribusi sampel nilai odds cenderung tidak normal dengan kemencengan positif besar

Rumus penghitungan besar sampel estimasi odds ratio:

$$n = \frac{z^2_{1-\alpha/2} (1/P_1[1-P_1] + 1/P_2[1-P_2])}{[\ln\{1-\epsilon\}]^2}$$

$$\epsilon = \frac{|OR_u - OR_l|}{OR}$$

$$\ln\{1-\epsilon\} = -z^2_{1-\alpha/2} \sqrt{(1/n)[1/[1-P_1] + 1/[P_2(1-P_2)]]}$$

$$P_1 = \frac{(OR)P_2}{(OR)P_2 + (1-P_2)}$$

P_1 =proporsi subyek yang terpapar pada kelompok penyakit

P_2 =proporsi subyek yang terpapar pada kelompok tanpa penyakit

OR=nilai Ratio Odds

OR_u =nilai atas Ratio Odds dari 95% CI OR

OR_l = nilai bawah Ratio Odds dari 95% CI OR

ϵ =presisi relatif

Contoh:

- Suatu penelitian bermaksud untuk melihat efek terhadap kejadian penyakit jantung coroner. Penelitian dilakukan dengan desain kasus – kontrol, dari penelitian terdahulu besarnya nilai odds 2, sedangkan proporsi perokok yang tidak menderita jantung coroner 40%. Berapa besar sampel yang diperlukan bila nilai presisi relatif sebesar 20% dan derajat kepercayaan 95%

$$P_1 = \frac{2 \cdot 0,4}{(2 \cdot 0,4) + (1 - 0,4)} = 0,57$$

$$n = \frac{1,96^2 (1/0,57[1-0,57] + 1/0,4[1-0,4])}{[\ln(1-0,20)]^2} = 636,24$$

2. Penghitungan sampel Uji Hipotesis Ratio Odds

- Untuk menguji hipotesis dari ratio odds
- Hipotesis $H_0: OR=1$ dan $H_a: OR \neq 1$
- Hubungan antara OR, P_1 dan P_2

$$P_1 = \frac{(OR)P_2}{(OR)P_2 + (1-P_2)}$$

$$P = (P_1 + P_2) / 2$$

OR = besarnya rasio odds pada penelitian sebelumnya atau penelitian pendahuluan

P_2 = proporsi subyek terpajan pada kelompok bukan dengan penyakit

P_1 = proporsi subyek terpajan pada kelompok dengan penyakit

n = jumlah sampel

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

Contoh:

Seorang peneliti ingin menguji hipotesis pengaruh kegemukan terhadap kejadian diabetes. Hasil penelitian yang pernah dilakukan menemukan ratio odds sebesar 2,5. Prevalensi kegemukan diketahui yang tidak menderita diabetes sebesar 60%. Berapa besar sampel yang diperlukan jika peneliti menginginkan tingkat kepercayaan 5% dan kekuatan uji 80%

$$P_1 = \frac{2,5 * 0,6}{(2,5 * 0,6) + (1 - 0,6)} = 0,79$$

$$P = (0,6 + 0,79) / 2$$

$$n = \frac{|1,96 \sqrt{2 * 0,7((1 - 0,7) + 0,84 \sqrt{0,79(1 - 0,79 + 0,6(1 - 0,6))}}|^2}{(0,79 - 0,6)^2} = 93,17$$

3. Penghitungan sampel Kasus-Kontrol dengan K kontrol per kasus:

- Memanfaatkan nilai ratio odds
- Kasus sulit dicari, kontrol mudah dicari
- 1 kasus dapat digunakan k kontrol
- Pemanfaatan sampel jadi lebih efisien

$$P_1 = \frac{(OR)P_2}{(OR)P_2 + (1-P_2)}$$

$$P = (P_1 + k \cdot P_2) / (1 + k)$$

$$n = \frac{|z_{1-\alpha/2} \sqrt{(1+1/k)P(1-P)} + z_{1-\beta} \sqrt{P_1(1-P_1) + (P_2(1-P_2))/k}|^2}{(P_1 - P_2)^2}$$

Contoh

Seorang peneliti ingin mengetahui hubungan antara obese stunting terhadap kejadian hipertensi. Prevalensi obese stunting 40%, sedangkan rasio odds dari penelitian yang pernah dilakukan adalah 1,5. Peneliti ingin menggunakan desain kasus kontrol dengan 3 kontrol per kasus. Berapa besar sampel yang diperlukan jika peneliti menginginkan tingkat kemaknaan 5% dan kekuatan uji 80%.

$$P_1 = \frac{1,5 * 0,4}{1,5 * 0,4 + (1 - 0,4)} = 0,50$$

$$P = (0,5 + 3 * 0,4) / (1 + 3) = 0,425$$

$$n = \frac{|1,96 \sqrt{(1 + 1/3) 0,425(1 - 0,425) + 0,84 \sqrt{0,5(1 - 0,5) + (0,4(1 - 0,4)) / 3}}|^2}{(0,5 - 0,4)^2} = 255,98$$

Sampel yang dibutuhkan 256 penderita hipertensi dan
 $3 \times 256 = 768$ bukan penderita hipertensi

4. Penghitungan sampel Kasus-Kontrol sepadan:

- Desain kasus kontrol sepadan = matched case control
- Ada 4 kemungkinan pajanan:
 - Kasus terpajan – kontrol terpajan
 - Kasus terpajan – kontrol tidak terpajan
 - Kasus tidak terpajan – kontrol terpajan
 - Kasus tidak tepajan – kontrol tidak tepajan
- Untuk kasus terpajan – kontrol tidak tepajan atau kasus tidak terpajan dan kontrol terpajan disebut pasangan diskordan (tidak sama)
- Untuk kasus terpajan – kontrol terpajan atau kasus tidak terpajan dan kontrol tidak terpajan disebut pasangan konkordan (sama)

$$m = \frac{|(z_{1-\alpha/2})/2 + z_{1-\beta} \sqrt{P(1-P_1)}|^2}{(P-1/2)^2}$$

$$P = \frac{OR}{1+OR}$$

$$M = \frac{m}{P_e}$$

$$M = \frac{m}{P_2(1-P_1) + P_1(1-P_2)}$$

m=total sampel yang dibutuhkan
 M=pasangan kasus kontrol
 sepadan yang dibutuhkan
 P_e=pasangan diskordan
 OR=perbedaan resiko yang
 diinginkan
 P₂=proporsi subyek yang
 terpajan pada kelompok
 kontrol
 P₁=proporsi subyek terpajan
 pada kelompok kasus

$$P_e = P_2(1-P_1) + P_1(1-P_2)$$

Contoh:

Seorang peneliti bermaksud meneliti pengaruh pemberian asam salisilat sebagai obat antipiretik pada anak balita yang diduga meningkatkan risiko terjadinya sindroma reye. Untuk membuktikan hal tersebut maka akan dilakukan penelitian dengan desain kasus kontrol. Oleh karena sindroma reye diduga dipengaruhi oleh umur serta penyakit yang diderita anak, maka kasus dan kontrol akan disepadankan menurut umur dan jenis penyakit yang dialami anak. Dari informasi yang ada diketahui 30% anak balita dengan demam memperoleh obat asam salisilat. Berapa jumlah sampel yang diperlukan untuk mendeteksi perbedaan risiko sebesar 2 dengan tingkat kemaknaan 5% dan kekuatan uji 90%

$$P = \frac{2}{(1+2)} = 0,67$$

$$m = \frac{|1,96/2 + 1,28\sqrt{0,67(1-0,67)}|^2}{(0,67-0,5)^2} = 90$$

$$P_1 = \frac{2*0,30}{2*0,30 + (1-0,30)} = 0,46$$

$$P_e = 0,3(1-0,46) + 0,46(1-0,30) = 0,484$$

$$M = \frac{90}{0,486} = 185,95$$

***Dibutuhkan 186 pasangan kasus kontrol
Yang sepadan***

Penghitungan sampel eksperimen:

Besar Sampel untuk data numeric

1. Estimasi beda rata-rata pada 2 kelompok independen
2. Uji hipotesis beda rata-rata pada 2 kelompok independen
3. Uji hipotesis beda rata-rata pada 2 kelompok dependen
 - Variabel dependen bersifat kontinyu
 - Ada 2 atau lebih kelompok yang akan dibandingkan
 - Ada nilai presisi yang diperhitungkan
 - Semakin kecil nilai presisi semakin besar jumlah sampel
sebaliknya semakin besar nilai presisi semakin kecil jumlah sampel

Besar Sampel untuk data kategorik

4. Uji hipotesis beda proporsi
5. Analisis ketahanan hidup (survival)

1. Penghitungan sampel untuk estimasi beda rata-rata pada 2 kelompok independen:

- *Mengukur variabel bersifat kontinyu*
- *Kedua nilai rata-rata tidak bergantung satu sama lainnya*
- *Membandingkan antara nilai rata-rata penelitian terdahulu atau penelitian pendahuluan dengan penelitian yang dilakukan*

$$S_p^2 = \frac{|(n_1-1)s_1^2 + (n_2-1)s_2^2|}{(n_1-1) + (n_2-1)}$$

$$n = \frac{z_{1-\alpha/2}^2 2\sigma^2}{d^2}$$

S_p^2 = varians gabungan
 S_1 = varians kelompok 1
 S_2 = varians kelompok 2
 n_1 = jumlah sampel kelompok 1 pada penelitian terdahulu/pendahuluan
 n_2 = jumlah sampel kelompok 2 pada penelitian terdahulu/pendahuluan
 n = jumlah sampel yang dibutuhkan

Contoh:

Seorang peneliti ingin membandingkan efek penurunan gula darah antara obat anti diabetes A dan B. Pada penelitian pendahuluan, diketahui 3 minggu pengobatan A rata-rata penurunan kadar gula darah sebesar 40 mg/dl dengan standar deviasi 20 mg/dl. Sedangkan obat B rata-rata menurunkan kadar gula darah sebesar 30 mg/dl dengan standar deviasi 15 mg/dl. Pada penelitian awal peneliti hanya menggunakan 5 pasien pada masing-masing kelompok. Berapa besar sampel yang diperlukan jika peneliti ingin menunjukkan ada perbedaan rata-rata penurunan kadar gula darah antara pasien memperoleh obat A dan obat B dengan simpangan maksimum 5 mg/dl dengan derajat kepercayaan 95%

$$S_p^2 = \frac{|(5-1)20^2 + (5-1)15^2|}{(5-1)(5-1)} = 312,5$$

$$n = \frac{1,96^2 * 2 * 312,5}{5^2} = 96,04$$

Jumlah sampel yang dibutuhkan 97

2. Penghitungan sampel untuk uji hipotesis beda rata-rata pada 2 kelompok independen:

- Menguji perbedaan antara 2 rata-rata dalam kelompok atau populasi independen
- Hipotesis nol yang ingin dibuktikan $H_0: \mu_1 = \mu_2$
- Hipotesis alternatif yang ingin dibuktikan $H_a: \mu_1 \neq \mu_2$

$$S_p^2 = \frac{|(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2|}{(n_1 - 1) + (n_2 - 1)}$$

$$n = \frac{2\sigma^2 [Z_{1-\alpha/2}^2 + z_{1-\beta}^2]}{(\mu_1 - \mu_2)^2}$$

$S_p^2 = \sigma^2$ = varians gabungan

S_1 = standar deviasi kelompok 1

S_2 = standar deviasi kelompok 2

n_1 = jumlah sampel kelompok 1 penelitian sebelumnya

n_2 = jumlah sampel kelompok 2 penelitian sebelumnya

μ_1 = rata-rata kelompok 1 pada penelitian sebelumnya

μ_2 = rata-rata kelompok 2 pada penelitian sebelumnya

Contoh:

Seorang peneliti ingin mengetahui efek asupan natrium terhadap tekanan darah. Pada penelitian sebelumnya dengan jumlah sampel 20 masing-masing kelompok diketahui kelompok masyarakat konsumsi natrium rendah rata-rata tekanan diastolik 72mmHg, sedangkan masyarakat konsumsi natrium tinggi rata-rata tekanan darah diastoliknya 85mmHg dengan standar deviasi 12mmHg. Berapa sampel yang dibutuhkan untuk membuktikan uji hipotesis adanya perbedaan tekanan darah diastolik pada kedua kelompok dengan derajat kemaknaan 5%, kekuatan uji 80%

$$S_p^2 = \frac{|(20-1)10^2 + (20-1)12^2|}{(20-1) + (20-1)} = 122$$

$$n = \frac{2 * 122 [1,96 + 0,84]^2}{(82 - 75)^2} = 39,04$$

Jumlah sampel yang dibutuhkan 40 orang kelompok kasus dan 40 orang kelompok kontrol

3. Penghitungan sampel untuk uji hipotesis beda rata-rata pada 2 kelompok dependen:

- *Menguji rata-rata sebelum dan sesudah intervensi*
- *Pengukuran kedua bergantung pada pengukuran pertama*

$$n = \frac{\sigma^2 [Z_{1-\alpha/2} + z_{1-\beta}]^2}{(\mu_1 - \mu_2)^2}$$

σ^2 = varians gabungan
 $Z_{1-\alpha/2}$ = derajat kepercayaan
 $z_{1-\beta}$ = kekuatan uji

Contoh:

Seorang peneliti ingin menguji efek latihan aerobik terhadap penurunan kadar kolesterol LDL pada orang dewasa normal. Dari penelitian awal pada 5 orang dewasa, diketahui rata-rata kolesterol LDL sebelum latihan aerobik adalah 185mg/dl dan setelah 4 minggu berlatih aerobik adalah 165 mg/dl. Jadi ada penurunan kadar kolesterol LDL rata-rata 20 mg/dl dan standar deviasi 15 mg/dl. Peneliti ingin menguji hipotesis dengan perbedaaan rata-rata minimum yang ingin dideteksi sebesar 10 mg/dl, tingkat kemaknaan 5% dan kekuatan uji 90% berapa besar sampel yang diperlukan?

$$n = \frac{15^2 [1,96 + 1,28]^2}{(10)^2} = 23,62$$

Jadi dibutuhkan sampel minimal 24 orang dewasa

4. Besar sampel uji hipotesis beda proporsi

- Untuk beda proporsi 2 kelompok
- P_1 dan P_2 bergantung pada desain
- Jumlah untuk masing-masing kelompok
- $P_1 - P_2$ = beda minimal yang dianggap bermakna secara substansi

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

$$\bar{P} = (P_1 + P_2) / 2$$

Contoh:

Suatu obat A dikatakan dapat menghilangkan nyeri pada 80% pasien osteoarthritis. Sedangkan paracetamol analgetik lama dapat menghilangkan nyeri pada 50% pasien osteoarthritis. Seorang peneliti ingin menguji apakah obat A memang lebih efektif dari paracetamol. Berapa besar sampel yang dibutuhkan jika peneliti menginginkan derajat kemaknaan 1% dan kekuatan uji 80%

$$n = \frac{(1,96 \sqrt{2 \cdot 0,65(1-0,65)} + 0,84 \sqrt{0,80(1-0,80) + 0,50(1-0,50)})^2}{(0,80-0,50)_2}$$

=41,39 pasien perkelompok

Jadi untuk membuktikan obat A diperlukan 42 pasien osteoarthritis diobati obat A dan 42 pasien osteoarthritis diobati paracetamol

5. Besar sampel analisis ketahanan hidup

- Probabilitas untuk bertahan hidup selama waktu t adalah $S(t) = e^{-\lambda_t}$
- λ adalah rate kematian atau rate suatu kejadian
- Penghitungan sampel analisis ketahanan hidup digunakan apabila salah satu tujuan dari penelitian melihat probabilitas kematian atau suatu kejadian dalam kurun waktu tertentu
- Desain penelitian dapat berupa ekperiment, kohor atau kohor hipotetik
- Pencatatan waktu harus dalam interval tertentu
- Minimal titik pengamatan 3 kali

$$n = \frac{|z_{1-\alpha/2} \sqrt{2\lambda^2} + z_{1-\beta} \sqrt{\lambda_1^2 + \lambda_2^2}|^2}{[\lambda_1 - \lambda_2]^2}$$

λ =rate ketahanan hidup atau rate suatu kejadian

λ_1 = rate ketahanan hidup atau rate suatu kejadian pada kelompok 1

λ_2 =rate ketahanan hidup atau rate suatu kejadian pada kelompok 2

Contoh:

Seorang peneliti ingin menguji apakah terapi radiasi lebih baik dari terapi sitostatika untuk kasus kanker payudara stadium III. Dari rekam medik pasien diketahui 10% pasien yang diberi terapi radiasi meninggal setelah 1 tahun menjalani terapi.

Sedangkan pada pasien yang diberi terapi sitostatika 20% pasien meninggal setelah 1 tahun menjalani terapi. Dari data tersebut berapa besar sampel yang diperlukan jika peneliti menginginkan tingkat kemaknaan 5% dan kekuatan uji 80%

$$\bullet S(t) = e^{-\lambda t}$$

$$\bullet S_1(t) = e^{-\lambda_1 t} \rightarrow$$

$$\bullet S_2(t) = e^{-\lambda_2 t}$$

$$S1 = 100\% - 10\% = 90\%$$

$$S2 = 100\% - 20\% = 80\%$$

$$S_1 = 0,9 = e^{-\lambda_1 \cdot 1}$$

$$\ln 0,9 = \ln e^{-\lambda_1 \cdot 1}$$

$$\ln 0,9 = -\lambda_1$$

$$\lambda_1 = 0,10$$

$$S_2 = 0,8 = e^{-\lambda_2 \cdot 1}$$

$$\ln 0,8 = \ln e^{-\lambda_1 \cdot 1}$$

$$\ln 0,8 = -\lambda_1$$

$$\lambda_1 = 0,22$$

$$n = \frac{| Z_{1-\alpha/2} \sqrt{2\lambda^2} + z_{1-\beta} \sqrt{\lambda_1^2 + \lambda_2^2} |^2}{[\lambda_1 - \lambda_2]^2}$$

$$n = \frac{| 1,96 \sqrt{2 \cdot 0,16\lambda^2} + 0,84 \sqrt{0,10^2 + 0,22^2} |^2}{[0,10 - 0,22]^2} = 29,03$$

***Dibutuhkan 30
pasient untuk
setiap terapi***

Catatan penting dalam penghitungan sampel:

1. Besarnya nilai berdasarkan power penelitian

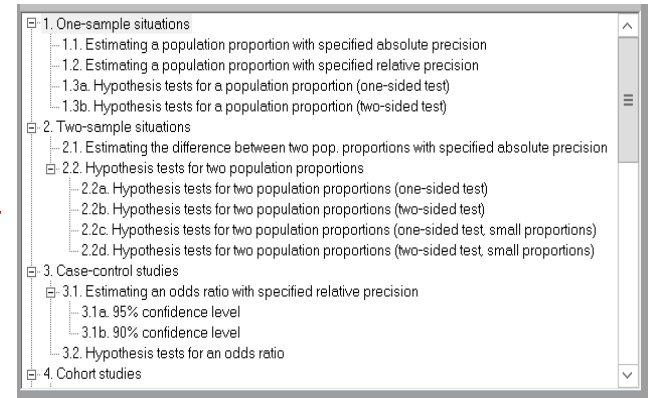
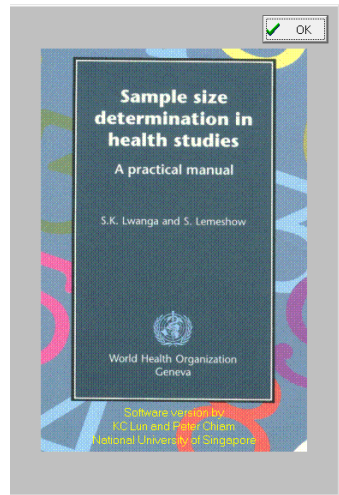
Power	nilai
80%	0,84
90%	1,28
95%	1,64
99%	2,33

2. Besarnya nilai berdasarkan nilai α

α (alpha)	nilai
1%	1,64
5%	1,96
10%	2,58

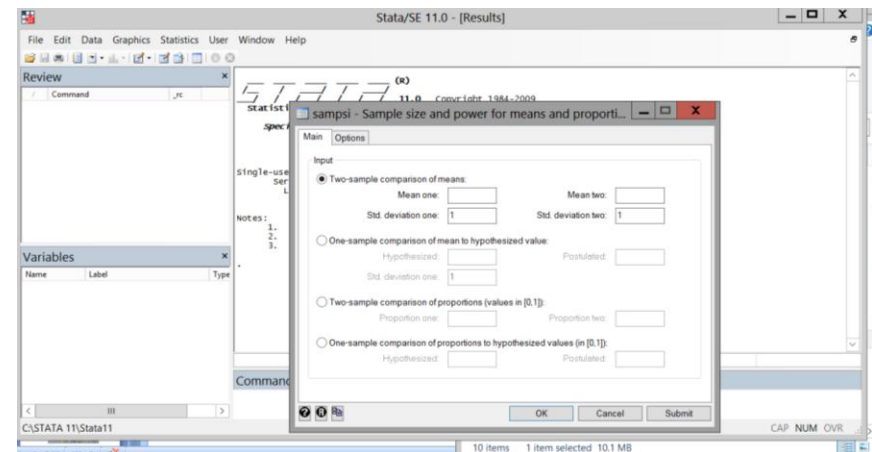
Beberapa software yang dapat digunakan untuk menghitung sampel:

Sample Size



Stata

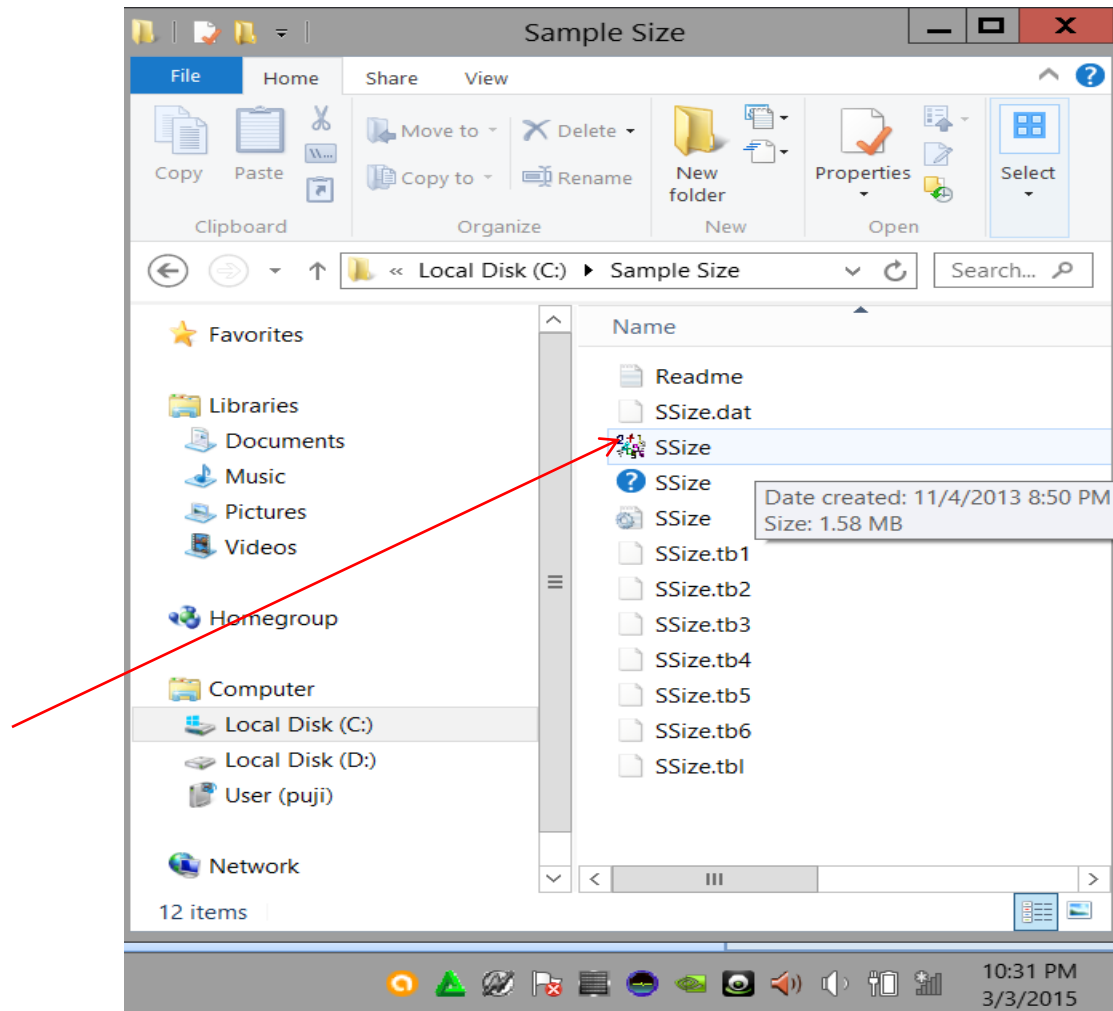
Tests of means and proportions
Tests of means with repeated measures
Cox proportional hazards model
Log-rank test
Exponential test

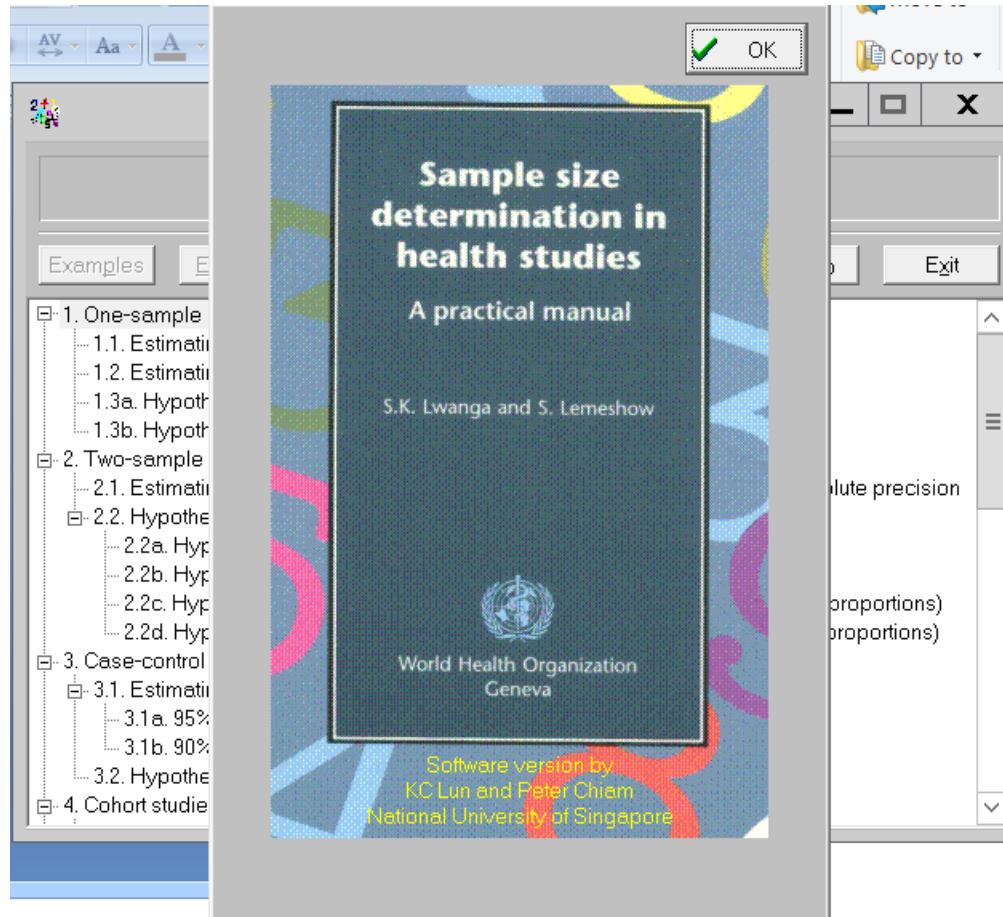


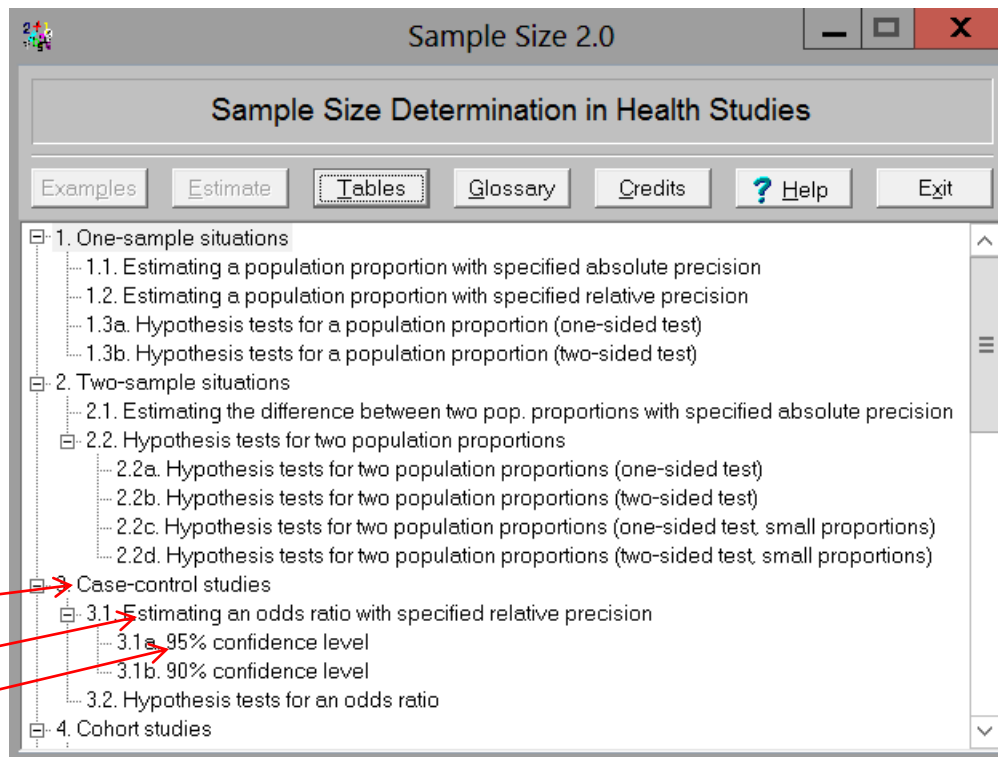
Contoh penghitungan sampel dengan menggunakan SSize

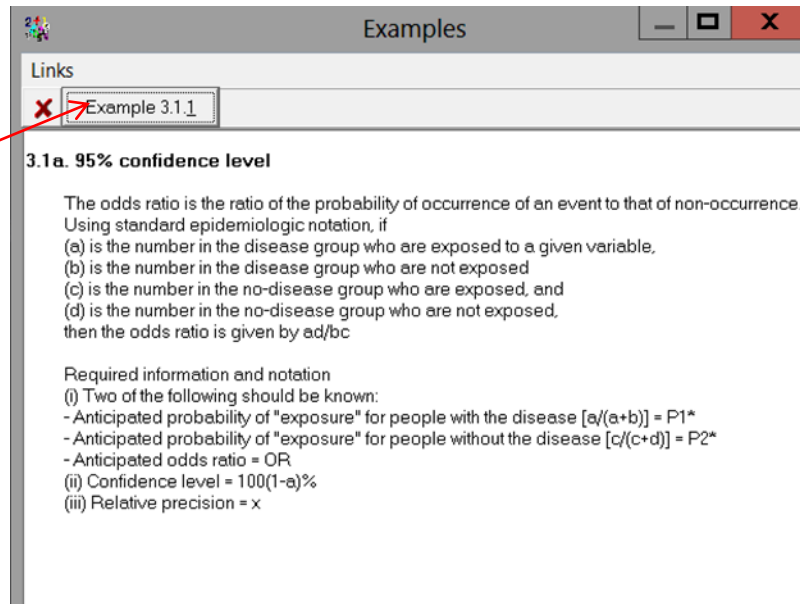
- Suatu penelitian bermaksud untuk melihat efek terhadap kejadian penyakit jantung coroner. Penelitian dilakukan dengan desain kasus – kontrol, dari penelitian terdahulu besarnya nilai odds 2, sedangkan proporsi perokok yang tidak menderita jantung coroner 40%. Berapa besar sampel yang diperlukan bila nilai presisi relatif sebesar 20% dan derajat kepercayaan 95%

Menghitung dengan software ssize:









Examples

Links

    Example 3.1.1

3.1a. 95% confidence level

The odds ratio is the ratio of the probability of occurrence of an event to that of non-occurrence. Using standard epidemiologic notation, if

- (a) is the number in the disease group who are exposed to a given variable,
- (b) is the number in the disease group who are not exposed
- (c) is the number in the no-disease group who are exposed, and
- (d) is the number in the no-disease group who are not exposed,

then the odds ratio is given by ad/bc

Required information and notation

- (i) Two of the following should be known:
 - Anticipated probability of "exposure" for people with the disease $[a/(a+b)] = P1^*$
 - Anticipated probability of "exposure" for people without the disease $[c/(c+d)] = P2^*$
 - Anticipated odds ratio = OR
- (ii) Confidence level = $100(1-\alpha)\%$
- (iii) Relative precision = x



Links



Introduction

3.1a. 95% confidence level

Example 3.1.1

In a defined area where cholera is posing a serious public health problem, about 30% of the population are believed to be using water from contaminated sources. A case-control study of the association between cholera and exposure to contaminated water is to be undertaken in the area to estimate the odds ratio to within 25% of the true value, which is believed to be approximately 2, with 95% confidence. What sample sizes would be needed in the cholera and control groups?

In this case, the 'Anticipated probability of "exposure" given disease' is not given
'Anticipated probability of "exposure" given "no disease"' is 30%, or 0.30
'Anticipated odds ratio' is 2
'Confidence level' is 95%
'Relative precision' is 25%

Evaluate



Worked Examples

Perform Estimation

3.1. Estimating an odds ratio with specified relative precision

Please select the desired unknown:

- ☐ Confidence level (%)
- ☐ Relative precision
- ☐ Anticipated probability of exposure given disease
- ☐ Anticipated probability of exposure given no disease
- ☐ Anticipated odds ratio
- ☒ Sample size

Please enter the remaining values:

$1 - \alpha$ 95

ε 0.20

P_1^* 0.46153846153

P_2^* 0.40

OR_a 1.28571428571

n 632

$$n = \frac{z_{1-\alpha/2}^2}{[\log_e(1 - \varepsilon)]^2} \left[\frac{1}{P_1^*(1 - P_1^*)} + \frac{1}{P_2^*(1 - P_2^*)} \right]$$

Print

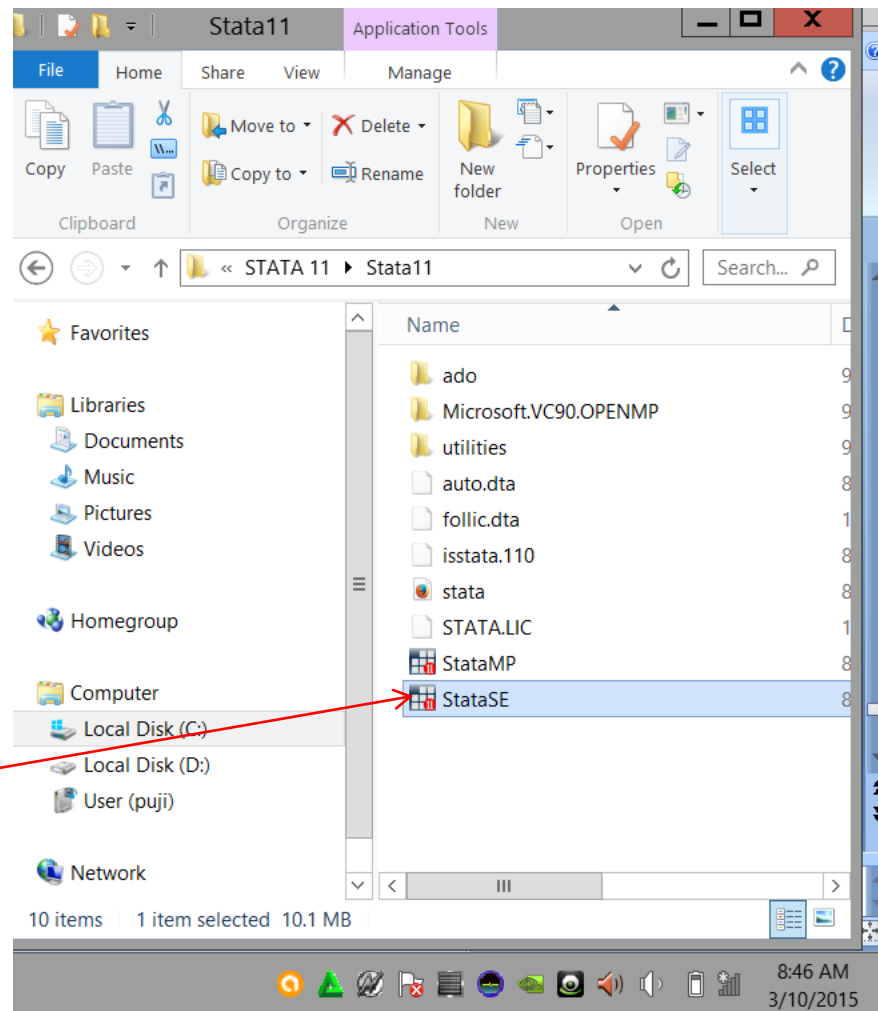
Help

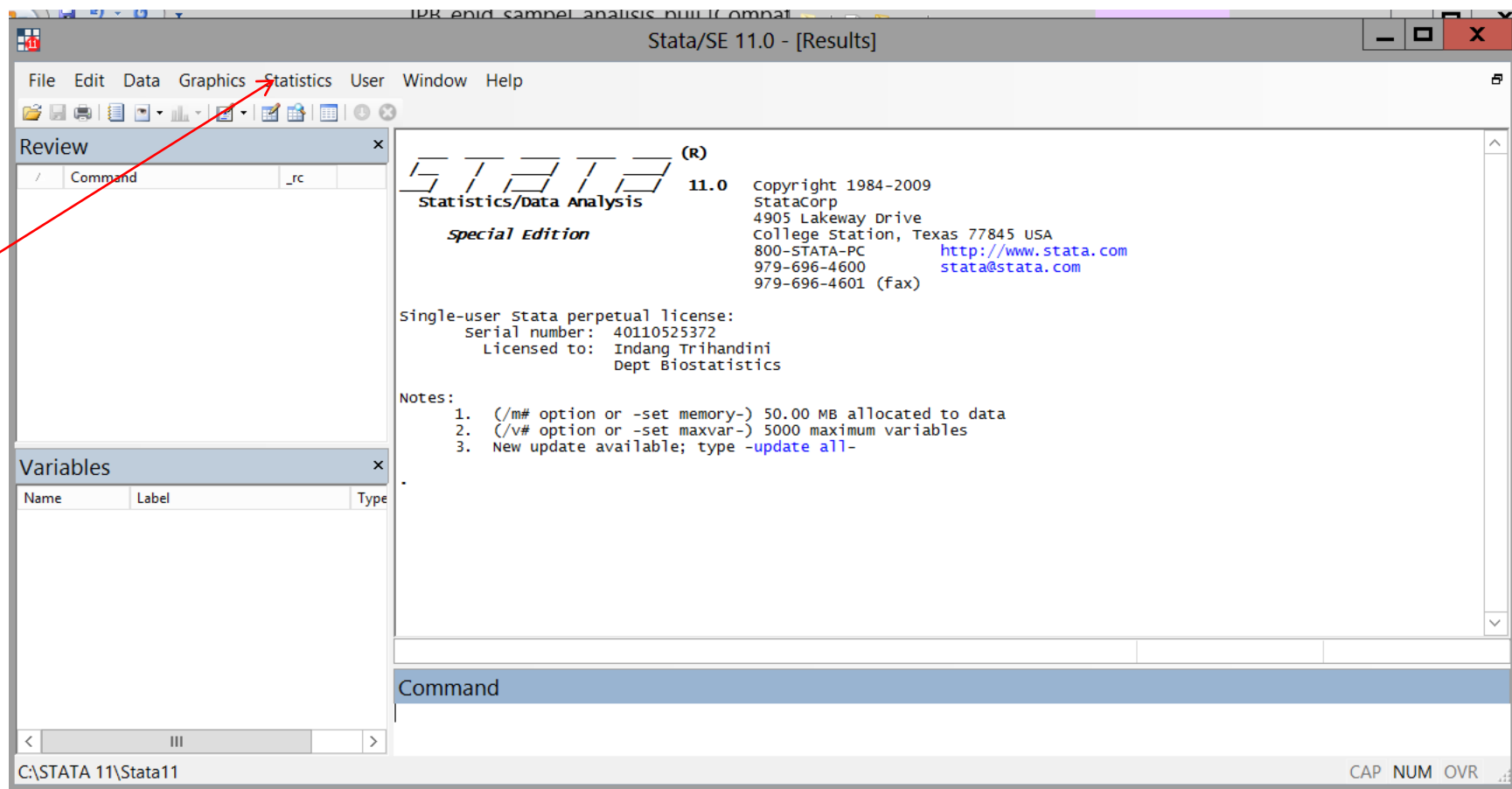
Close

OR=2
 $P_2=40\%$
 $\varepsilon = 20\%$
 CI =95%
 Maka jumlah
 sampel=632

Contoh penghitungan sampel dengan stata:

Seorang peneliti ingin menguji efek latihan aerobik terhadap penurunan kadar kolesterol LDL pada orang dewasa normal. Dari penelitian awal pada 5 orang dewasa, diketahui rata-rata kolesterol LDL sebelum latihan aerobik adalah 185mg/dl dan setelah 4 minggu berlatih aerobik adalah 165 mg/dl. Jadi ada penurunan kadar kolesterol LDL rata-rata 20 mg/dl dan standar deviasi 15 mg/dl. Peneliti ingin menguji hipotesis dengan perbedaaan rata-rata minimum yang ingin dideteksi sebesar 10 mg/dl, tingkat kemaknaan 5% dan kekuatan uji 90% berapa besar sampel yang diperlukan?





Stata/SE 11.0 - [Results]

File Edit Data Graphics **Statistics** User Window Help

Review

/ Command

Variables

Name	Label
------	-------

C:\STATA 11\Stata11

- Summaries, tables, and tests
- Linear models and related
 - Binary outcomes
 - Ordinal outcomes
 - Categorical outcomes
 - Count outcomes
- Exact statistics
- Endogenous covariates
- Sample-selection models
- Multilevel mixed-effects models
- Generalized linear models
- Nonparametric analysis
- Time series
 - Multivariate time series
- State-space models
- Longitudinal/panel data
- Survival analysis
- Epidemiology and related
- Survey data analysis
- Multiple imputation
- Multivariate analysis
- Power and sample size

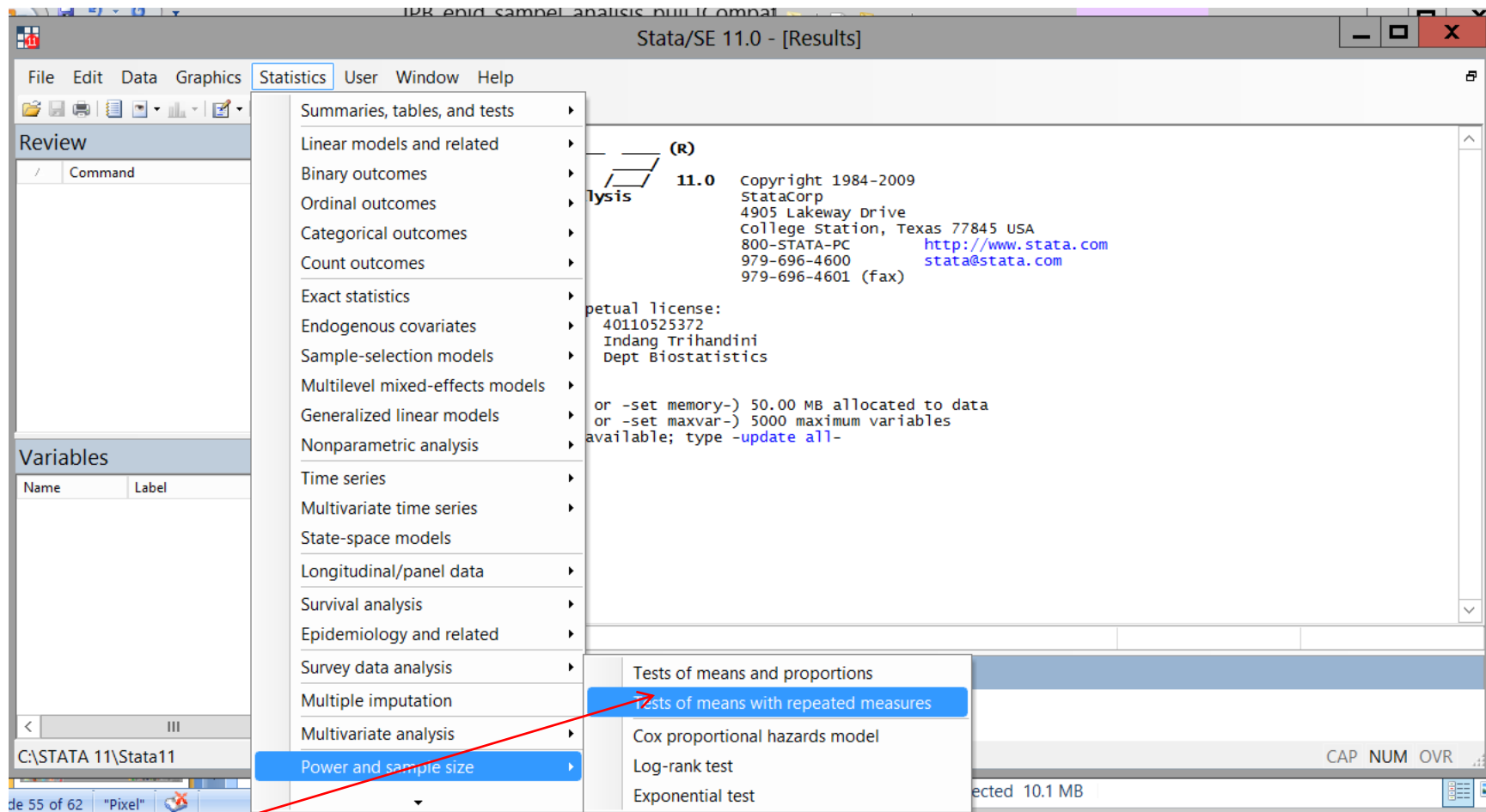
(R)

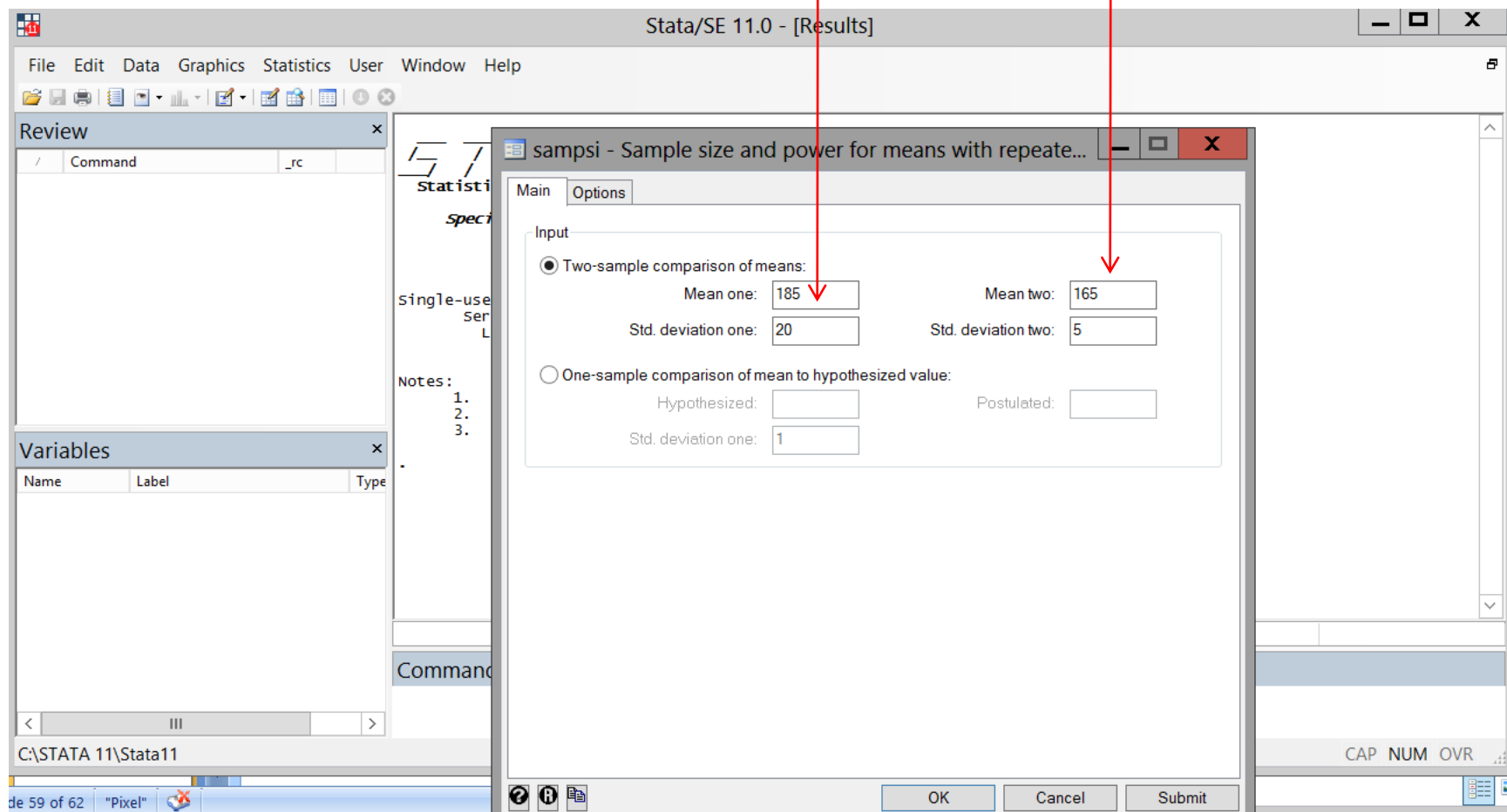
Stata 11.0 Copyright 1984-2009
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC <http://www.stata.com>
979-696-4600 stata@stata.com
979-696-4601 (fax)

perpetual license:
40110525372
Indang Trihandini
Dept Biostatistics

or -set memory-) 50.00 MB allocated to data
or -set maxvar-) 5000 maximum variables
available; type -update all-

CAP NUM OVR





sampsi - Sample size and power for means with repeate... — □ ✕

Main Options

Output

☒ Compute sample size

☐ Compute power

Sample-based calculations

Sample one size

☐ Sample two size

Ratio of sample sizes

Repeated measurements

Baseline measurements

Follow-up measurements

Correlations (between -1 and 1)

Baseline measurements

Follow-up measurements

Baseline/follow-up

Significance level (alpha)

Power of the test

Type of test

Analysis method

OK Cancel Submit

Review

	Command	_rc
1	sampsi 185 165, sd1(20) sd2(5) ...	

Variables

Name	Label	Type
------	-------	------

Notes:

1. (/m# option or -set memory-) 50.00 MB allocated to data
2. (/v# option or -set maxvar-) 5000 maximum variables
3. New update available; type `-update all-`

. **sampsi 185 165, sd1(20) sd2(5) alpha(0.05) pre(1) post(1) r01(0)**

Estimated sample size for two samples with repeated measures

Assumptions:

alpha =	0.0500	(two-sided)
power =	0.9000	
m1 =	185	
m2 =	165	
sd1 =	20	
sd2 =	5	
n2/n1 =	1.00	
number of follow-up measurements =	1	
number of baseline measurements =	1	
correlation between baseline & follow-up =	0.000	

Cek

Method: **POST**

relative efficiency =	1.000
adjustment to sd =	1.000
adjusted sd1 =	20.000
adjusted sd2 =	5.000

Estimated required sample sizes:

n1 =	12
n2 =	12

Command

Stata/SE 11.0 - [Results]

File Edit Data Graphics Statistics User Window Help

Review

	Command	_rc
1	sampsi 185 165, sd1(20) sd2(5) ...	

Variables

Name	Label	Type
------	-------	------

Method: **POST**
relative efficiency = 1.000
adjustment to sd = 1.000
adjusted sd1 = 20.000
adjusted sd2 = 5.000
Estimated required sample sizes:
n1 = 12
n2 = 12
Method: **CHANGE**
relative efficiency = 0.500
adjustment to sd = 1.414
adjusted sd1 = 28.284
adjusted sd2 = 7.071
Estimated required sample sizes:
n1 = 23
n2 = 23
Method: **ANCOVA**
relative efficiency = 1.000
adjustment to sd = 1.000
adjusted sd1 = 20.000
adjusted sd2 = 5.000
Estimated required sample sizes:
n1 = 12
n2 = 12
.

Command

C:\STATA 11\Stata11 CAP NUM OVR

*Malu bertanya sesat
di jalan loh!*

