

SAMPLE SIZE CALCULATION TUTORIAL

DR. NURAH ALAMRO
CMED 304
NOV. 30TH 2011

QUESTION 1

We want to estimate the mean systolic blood pressure of Saudi females. The standard deviation is around 20 mmHg and we wish to estimate the true mean to within 10 mmHg with 95% confidence. What is the required sample size ?



STEPS!

- Type of Study
- Type of Estimate: Mean(s)/ Proportion(s)/OR/RR
- Highlight the following from the question:
 - Z : Standard Normal Deviate (Z table)
 - α : Significance Level (e.g. 0.05, 0.1,..etc)
 - $1-\beta$: Power (e,g, 90% Power = $1-\beta = 1-0.1 = 0.9$)
 - σ : Standard Deviation
 - p : proportion
 - d : Difference from the estimate (e,g, lower/higher than mean)

Calculation Methods

```
graph TD; A[Calculation Methods] --> B[Formulas]; A --> C[Tables]; A --> D[Softwares];
```

Formulas

Tables

Softwares

QUESTION 1

We want to estimate the mean systolic blood pressure of Saudi females. The standard deviation is around 20 mmHg and we wish to estimate the true mean to within 10 mmHg with 95% confidence. What is the required sample size ?

Remember the steps!

- Type of Study
- Type of Estimate: Mean(s)/ Proportion (s)/OR/RR
- Highlight the following from the question:
 - Z : Standard Normal Deviate (Z table)
 - α : Significance Level (e.g. 0.05, 0.1,...etc)
 - $1-\beta$: Power (e,g, 90% Power = $1-\beta = 1-0.1 = 0.9$)
 - σ : Standard Deviation
 - p : proportion
 - d : Difference from the estimate (e,g, lower/higher than mean)

QUESTION 1

We want to estimate the mean systolic blood pressure of Saudi females. The standard deviation is around 20 mmHg and we wish to estimate the true mean to within 10 mmHg with 95% confidence. What is the required sample size ?

Remember the steps!

- Type of Study: Survey
- Type of Estimate: Mean
- Highlight the following from the question:
 - Z : Standard Normal Deviate (Z table)
 - $\alpha = 0.05$
 - ~~$1 - \beta$: Power~~
 - $\sigma = 20$
 - ~~p : proportion~~
 - $d = 10$

CALCULATION METHOD: 1) FORMULAS

CHOOSE THE APPROPRIATE FORMULA:

☐ $n = Z_{(1-\alpha)}^2 \sigma^2 / d^2$

☐ $n = Z_{(1-\alpha)}^2 p(1-p) / d^2$

☐ $n = 2 \sigma^2 (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 / d^2$

☐ $n = (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 [p_1(1-p_1) + p_2(1-p_2)] / (p_1 - p_2)^2$

$$n = Z_{(1-\alpha)}^2 \sigma^2 / d^2$$

$$\alpha = 0.05 \rightarrow (1-\alpha) = 0.95 \rightarrow Z_{(1-\alpha)} = 1.96$$

$$\sigma = 20 \quad d = 10$$

$$n = (1.96)^2 \times (20)^2 / (10)^2 = 15.37$$

Since we cannot take 0.37 of a person, we round up to 16 women as our sample size

CALCULATION METHOD: 2) TABLES



QUESTION 2

We wish to estimate the proportion of Saudi males who smoke. What sample size do we require to achieve a 95% confidence interval of width $\pm 5\%$ (that is to be within 5% of the true value) ?
A study some years ago found approximately 30% were smokers ?

Remember the steps!

- Type of Study: Survey
- Type of Estimate: Proportion
- Highlight the following from the question:
 - Z : Standard Normal Deviate (Z table)
 - $\alpha = 0.05$
 - ~~$1 - \beta$: Power~~
 - σ : Standard Deviation
 - $p = 0.3$
 - $d = 0.05$

CALCULATION METHOD: 1) FORMULAS

CHOOSE THE APPROPRIATE FORMULA:

☐ $n = Z_{(1-\alpha)}^2 \sigma^2 / d^2$

☒ $n = Z_{(1-\alpha)}^2 p(1-p) / d^2$

☐ $n = 2 \sigma^2 (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 / d^2$

☐ $n = (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 [p_1(1-p_1) + p_2(1-p_2)] / (p_1 - p_2)^2$

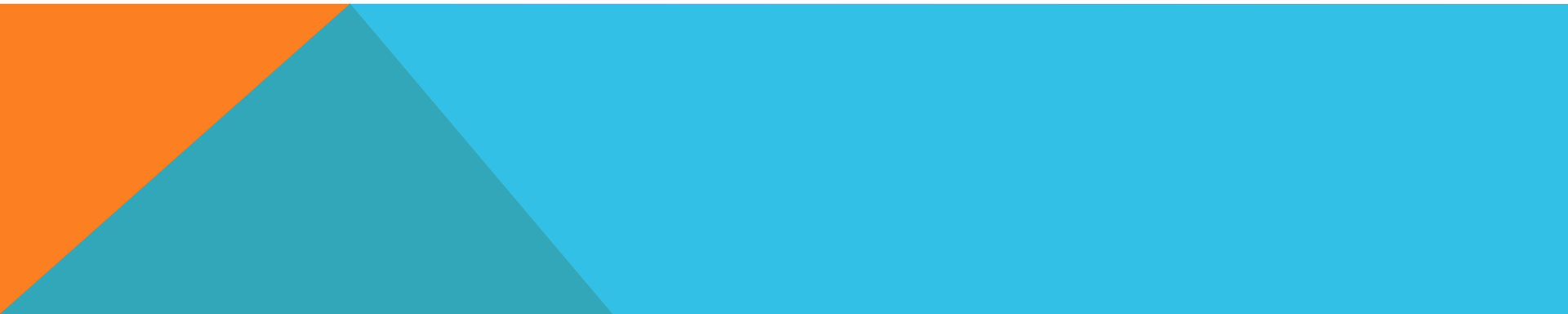
$$n = Z_{(1-\alpha)}^2 p(1-p)/d^2$$

$$\alpha = 0.05 \rightarrow (1-\alpha) = 0.95 \rightarrow Z_{(1-\alpha)} = 1.96$$
$$p=0.3 \quad d=0.05$$

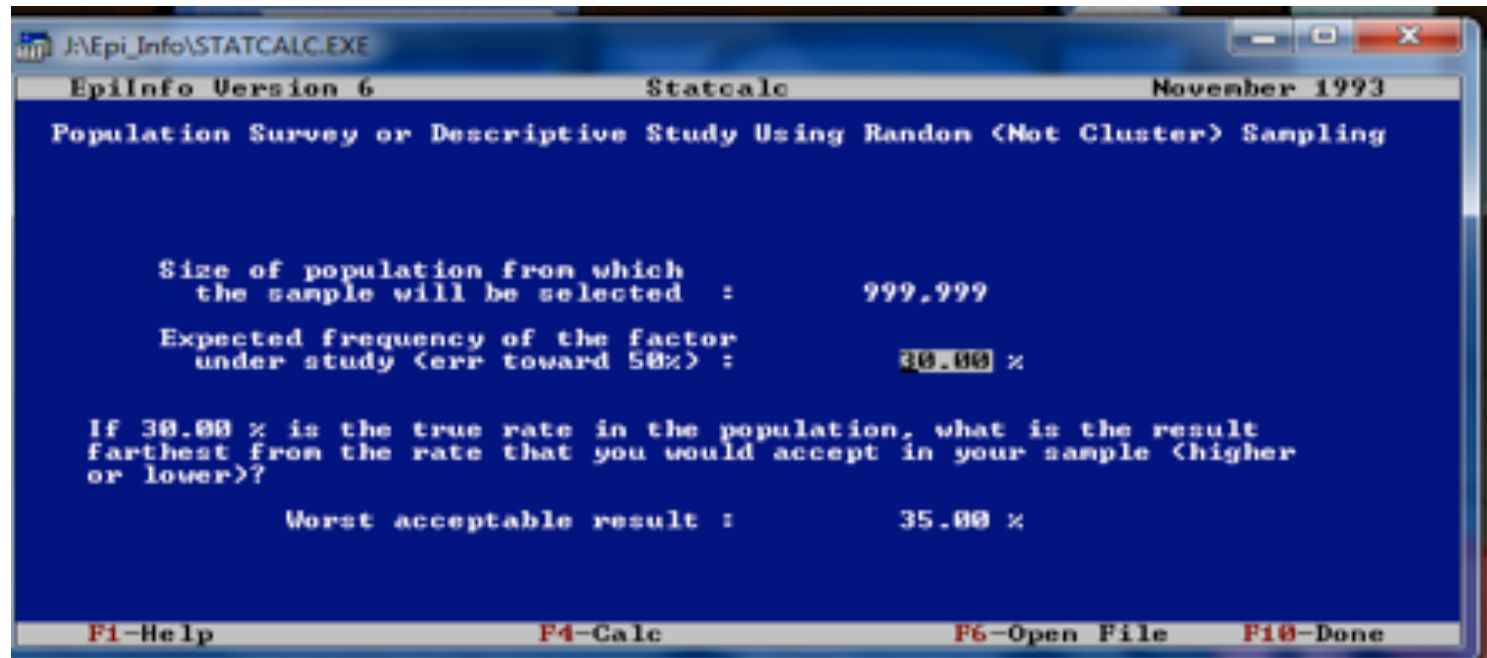
$$n = (1.96)^2 (0.3)(0.7)/(0.05)^2 = 322.7 \text{ rounded to } 323$$

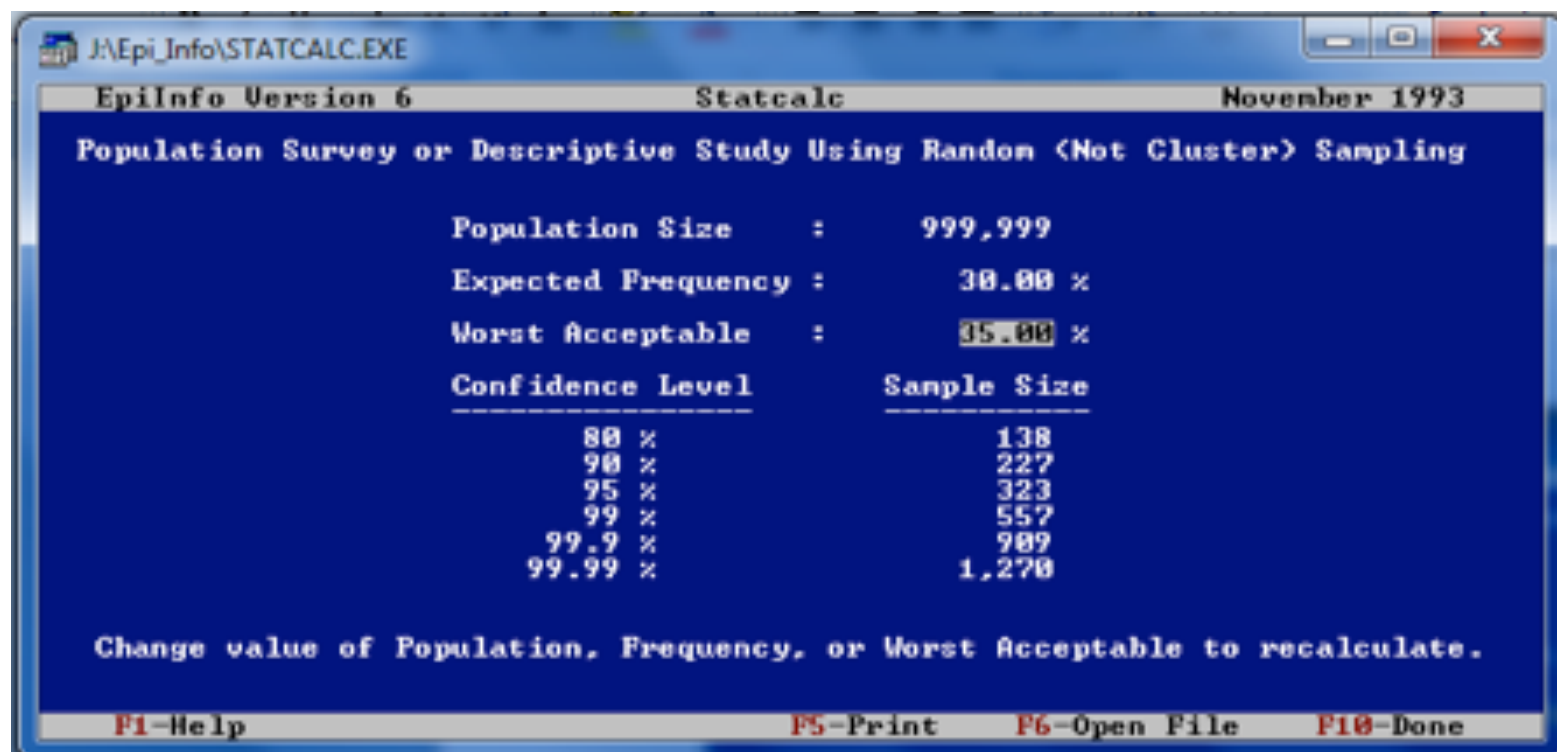
We need 323 Saudi males as our sample size to estimate the proportion of Saudi males who smoke

CALCULATION METHOD: 2) TABLES



CALCULATION METHOD: 3) EPI INFO





QUESTION 3

A new antihypertensive drug is to be tested against current treatment practice in people with systolic blood pressure > 160 mmHg and/or diastolic blood pressure >95 mmHg. It is felt that if the new drug can achieve blood pressure levels that are on the average 10 mmHg lower than those achieved using current treatment then it would be accepted by the medical community. The investigators would like at least 90% power and have chosen $\alpha = 0.01$ (two-sided) as the current therapy is quite acceptable and they want to be sure that the new therapy is superior before switching over. Blood pressure measurements have a standard deviation of 20 mmHg.

Remember the steps!

- Type of Study: RCT
- Type of Estimate: two means
- Highlight the following from the question:
 - Z: Standard Normal Deviate (Z table)
 - $\alpha = 0.01$
 - $1-\beta = 0.9 \rightarrow \beta = 0.1$
 - $\sigma = 20$
 - ~~p = proportion~~
 - $d = 10$

CALCULATION METHOD: 1) FORMULAS

CHOOSE THE APPROPRIATE FORMULA:

☐ $n = Z_{(1-\alpha)}^2 \sigma^2 / d^2$

☐ $n = Z_{(1-\alpha)}^2 p(1-p) / d^2$

☒ $n = 2 \sigma^2 (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 / d^2$

☐ $n = (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 [p_1(1-p_1) + p_2(1-p_2)] / (p_1 - p_2)^2$

$$n = 2 \sigma^2 (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 / d^2$$

$$\alpha = 0.01 \rightarrow (1-\alpha) = 0.99 \rightarrow Z_{(1-\alpha)} = 2.58$$

$$\beta = 0.10 \rightarrow (1-\beta) = 0.90 \rightarrow Z_{(1-\beta)} = 1.28$$

$$d=10 \quad \sigma=20$$

$$n = 2 \times 20^2 (2.58 + 1.28)^2 / (10)^2 = 119.2$$

We require at least 120 per group (240 hypertensive's in all)

CALCULATION METHOD: 2) TABLES



QUESTION 4

A standard regimen has an efficacy of 80% and a new regimen has been claimed to be 90% effective. What is the sample size required to test whether the new treatment is really effective at 5% level with 90% power ?

Remember the steps!

- Type of Study: Survey
- Type of Estimate: Two Proportions
- Highlight the following from the question:
 - Z : Standard Normal Deviate (Z table)
 - $\alpha = 0.05$
 - $1-\beta = 0.90$
 - ~~σ : Standard Deviation~~
 - $p_1 = 0.8$ $p_2 = 0.9$
 - ~~d = difference in estimate~~

CALCULATION METHOD: 1) FORMULAS

CHOOSE THE APPROPRIATE FORMULA:

☐ $n = Z_{(1-\alpha)}^2 \sigma^2 / d^2$

☐ $n = Z_{(1-\alpha)}^2 p(1-p) / d^2$

☐ $n = 2 \sigma^2 (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 / d^2$

☐ $n = (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 [p_1(1-p_1) + p_2(1-p_2)] / (p_1 - p_2)^2$

$$n = (Z_{(1-\alpha)} + Z_{(1-\beta)})^2 [p_1(1-p_1) + p_2(1-p_2)] / (p_1 - p_2)^2$$

$$\alpha = 0.05 \rightarrow (1-\alpha) = 0.95 \rightarrow Z_{(1-\alpha)} = 1.96$$

$$\beta = 0.10 \rightarrow (1-\beta) = 0.90 \rightarrow Z_{(1-\beta)} = 1.28$$

$$p_1 = 0.8 \quad p_2 = 0.9$$

$$n = (1.96 + 1.28)^2 ((0.8 \times 0.2) + (0.9 \times 0.1)) / (0.8 - 0.9)^2 = 263 \text{ patients for each treatment}$$

Total sample size=526

CALCULATION METHOD: 2) TABLE

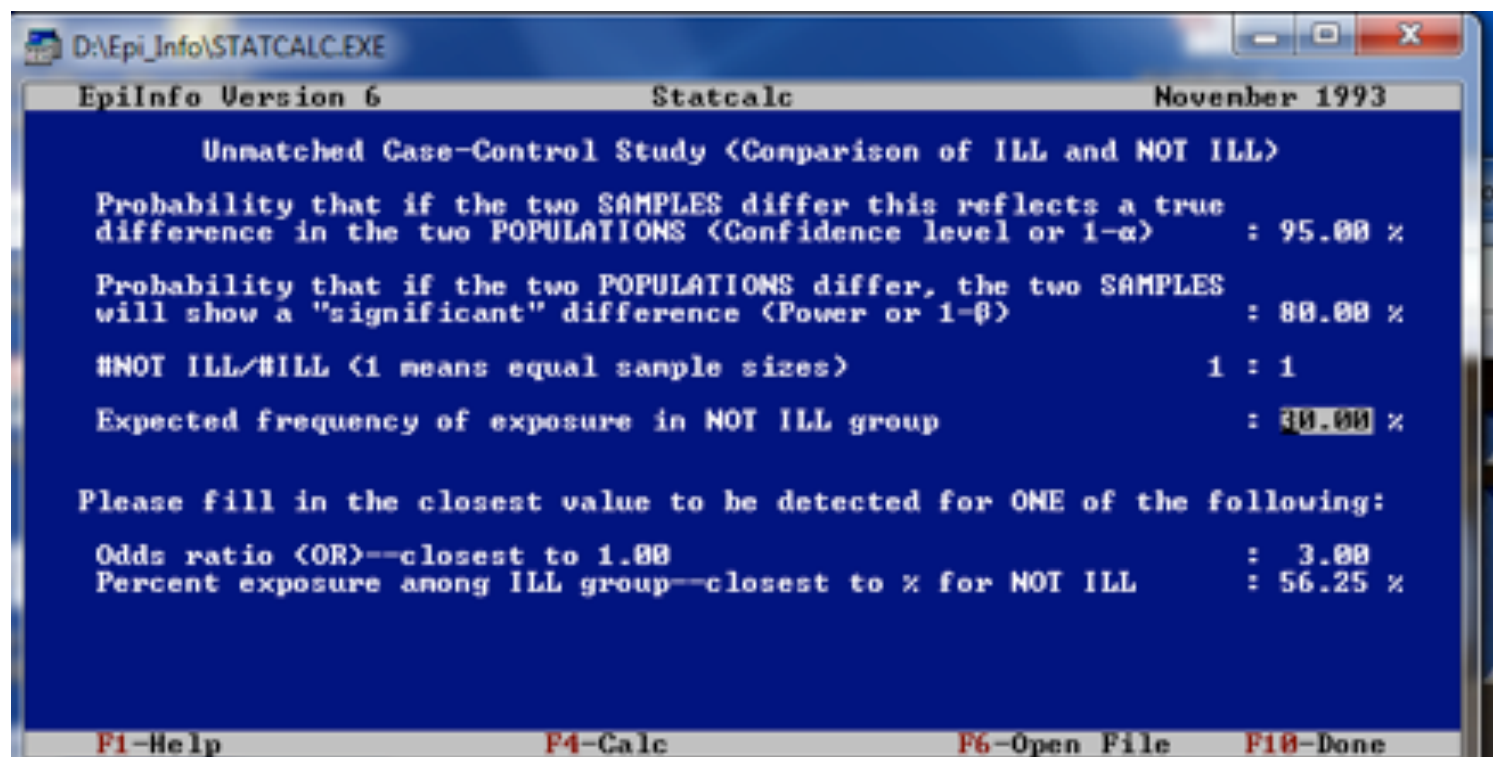


QUESTION 5

To find whether contaminated food had caused diarrhea in an area, a study is planned in that area where 30% individuals ate contaminated food but had no diarrhea and we want to detect an odds ratio of 3 in individuals who ate contaminated food and had diarrhea as compared to controls. How large sample size in each group is required to have an 80% chance of getting this odds ratio at the 5% level.

Let's Use Epi Info





D:\Epi_Info\STATCALC.EXE

Epilinfo Version 6 Statcalc November 1993

Unmatched Case-Control Study (Comparison of ILL and NOT ILL)
Sample Sizes for 30.00 % Exposure in NOT ILL Group

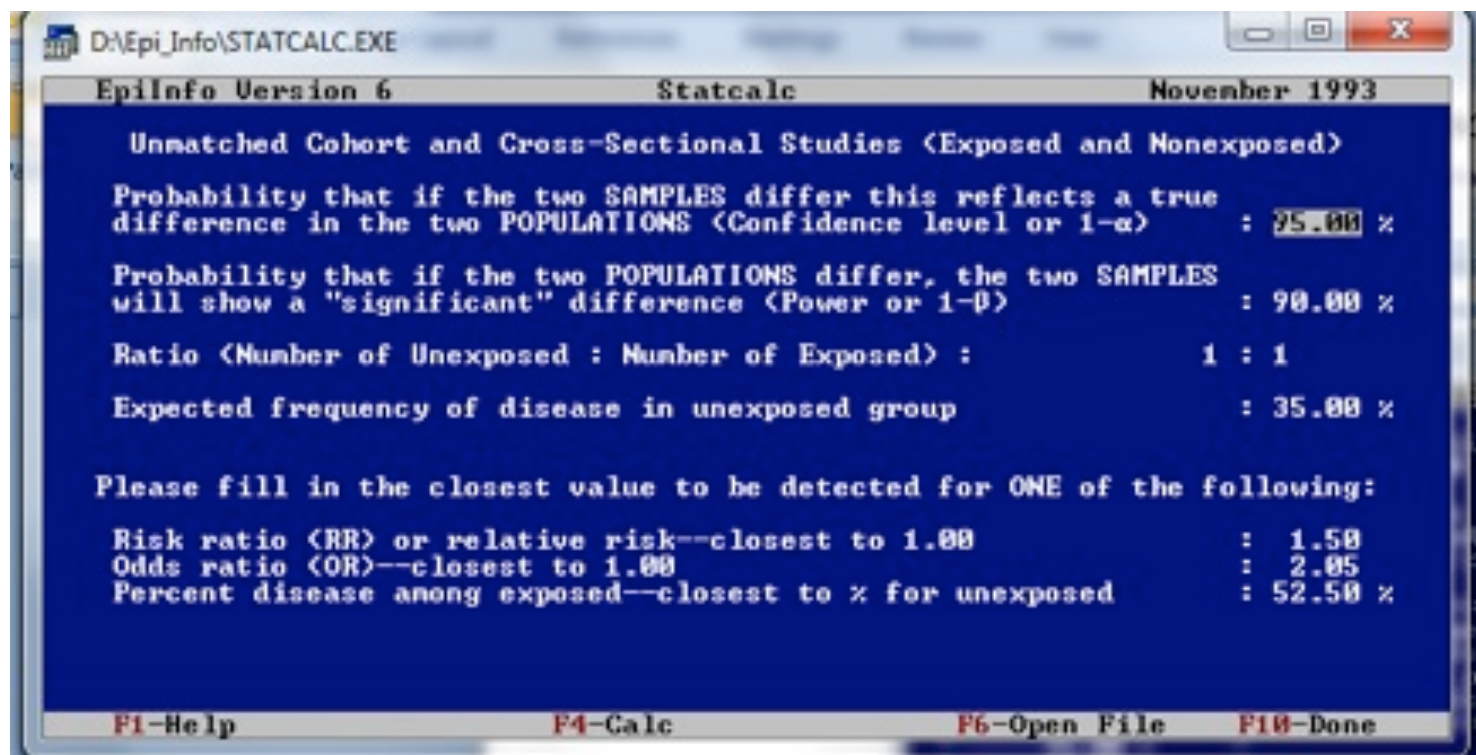
Conf.	Power	NOT ILL	ILL : ILL	Exposure in ILL	Odds Ratio	NOT ILL	ILL	Total
95.00 %	80.00 %	1:1		56.25 %	3.00	62	62	124
98.00 %	"	"				50	50	100
95.00 %	"	"				62	62	124
99.00 %	"	"				89	89	178
99.90 %	"	"				127	127	254
95.00 %	80.00 %	"				62	62	124
"	98.00 %	"				80	80	160
"	95.00 %	"				97	97	194
"	99.00 %	"				133	133	266
"	80.00 %	4:1				152	38	190
"	"	3:1				120	40	160
"	"	2:1				92	46	138
"	"	1:2				46	93	139
"	"	1:3				41	124	165
"	"	1:4				39	155	194

Change values for inputs as desired, then press F4 to recalculate.

F1-Help F5-Print F6-Open File F10-Done

QUESTION 6

Two competing therapies for a particular cancer are to be evaluated by the cohort study strategy in a multi center clinical trial. Patients are randomized to either treatment A or B and are followed for recurrence of disease for 5 years following treatment. How many patients should be studied in each of two arms of the trial in order to be 90% of rejecting $H_0: RR=1$ in favor of the alternative $H_a: RR \neq 1$, if the test is to be performed at the $\alpha=0.05$ level and if it is assumed that $p_2=0.35$ and $RR=1.5$.



D:\Epi_Info\STATCALC.EXE																
EpiInfo Version 6			Statcalc				November 1993									
Unmatched Cohort and Cross-Sectional Studies (Exposed and Nonexposed)																
Sample Sizes for 35.00 % Disease in Unexposed Group																
Conf.	Power	Unex:Exp	Disease in Exposed	Risk Ratio	Odds Ratio	Sample Size Unexp.	Sample Size Exposed	Total								
95.00 %	80.00 %	1:1	52.50 %	1.50	2.05	178	178	356								
90.00 %	"	"				147	147	294								
95.00 %	"	"				178	178	356								
99.00 %	"	"				248	248	496								
99.90 %	"	"				344	344	688								
95.00 %	80.00 %	"				136	136	272								
"	90.00 %	"				178	178	356								
"	95.00 %	"				217	217	434								
"	99.00 %	"				302	302	604								
"	90.00 %	4:1				440	110	550								
"	"	3:1				354	118	472								
"	"	2:1				266	133	399								
"	"	1:2				134	267	401								
"	"	1:3				119	356	475								
"	"	1:4				111	444	555								
Change values for inputs as desired, then press F4 to recalculate.																
F1-Help			F5-Print			F6-Open File		F10-Done								

THANK YOU

ANY QUESTIONS: NMALAMRO@KSU.EDU.SA