Models In Epidemiology And Biostatistics
Gordon Hilton Fick

Session 18 :
Size, Power, Confidence Intervals and Models

From a recent research protocol:

"Entry criteria and evaluation for response will be as above, limited to 2 cohorts of 11 patients each with XXX. With a sample size of 11, there is an 80% power to show that an observed response of 10% is significantly different from a proportion of 0.1% representing no effect. The sample size was calculated based on a one-stage, single-arm, Phase II design such that a meaningful increase in the overall response rate, if present, could be detected with high probability, while maintaining an upper limit on the type 1 error rate at conventional levels. "

It would appear that the investigator determined the sample size of 11 with the following:

```
. sampsi .001 .1, p(.8) onesample

Estimated sample size for one-sample comparison of proportion
  to hypothesized value

Test Ho: p = 0.0010, where p is the proportion in the population

Assumptions:
         alpha =   0.0500  (two-sided)
         power =   0.8000
 alternative p =   0.1000

Estimated required sample size:
             n =       11
```

But this calculation is based on the normal approximation to the binomial. Lets now consider some exact calculations:

```
. set obs 12
. gen y=_n-1
. gen py001=binomialp(11,y,0.001)
. gen py1=binomialp(11,y,0.1)
. gen tail001=binomialtail(11,y,0.001)
. gen tail1=binomialtail(11,y,0.1)
. gen tail14=binomialtail(11,y,0.14)
```

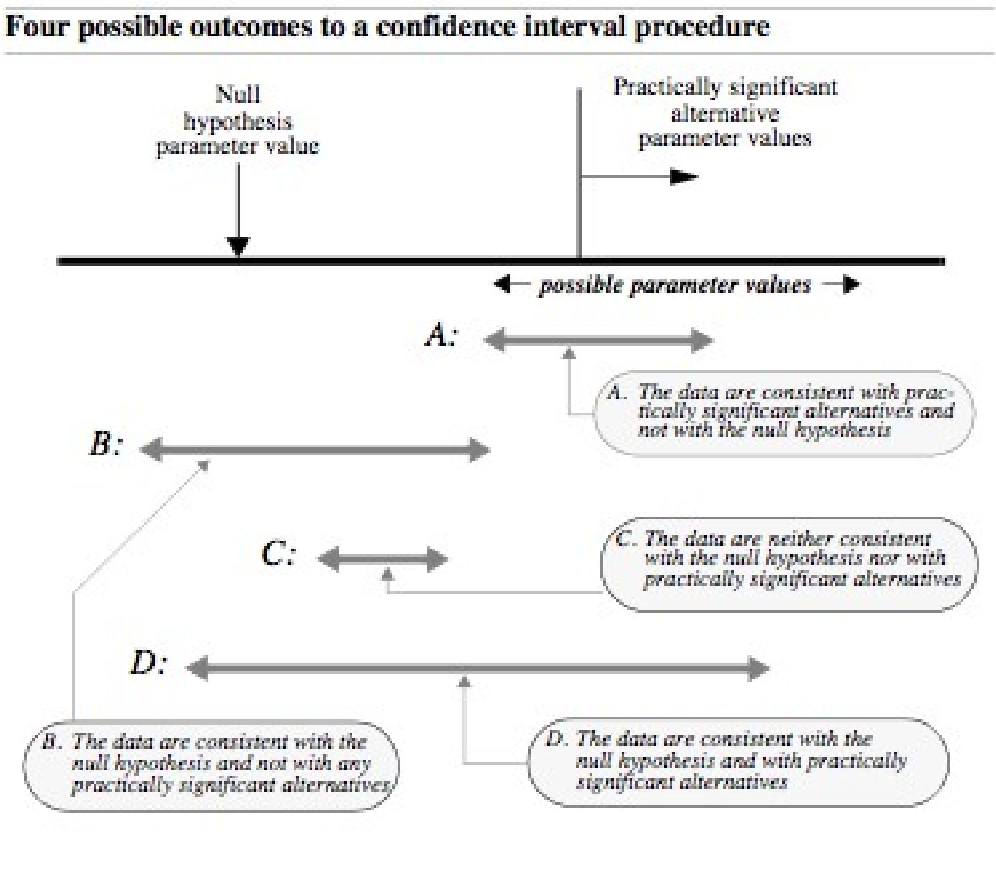
```
. list

     +---------------------------------------------------------------+
     |  y      py001        py1    tail001      tail1     tail14 |
     |---------------------------------------------------------------|
  1. |  0   .9890549   .3138106          1          1          1 |
  2. |  1   .0108905   .3835463   .0109452   .6861894   .8096806 |
  3. |  2   .0000545   .2130813   .0000547   .3026431   .4688762 |
  4. |  3   1.64e-07   .0710271   1.64e-07   .0895618   .1914773 |
  5. |  4   3.28e-10   .0157838   3.28e-10   .0185348   .0560033 |
     |---------------------------------------------------------------|
  6. |  5   4.59e-13   .0024553   4.60e-13    .002751   .0118956 |
  7. |  6   4.60e-16   .0002728   4.60e-16   .0002957   .0018431 |
  8. |  7   3.29e-19   .0000217   3.29e-19   .0000229   .0002066 |
  9. |  8   1.65e-22   1.20e-06   1.65e-22   1.25e-06   .0000164 |
 10. |  9   5.49e-26   4.45e-08   5.49e-26   4.55e-08   8.68e-07 |
     |---------------------------------------------------------------|
 11. | 10   1.10e-29   9.90e-10   1.10e-29   1.00e-09   2.78e-08 |
 12. | 11   1.00e-33   1.00e-11   1.00e-33   1.00e-11   4.05e-10 |
     +---------------------------------------------------------------+
```

The 'null' hypothesis considered is H_0: p=0.001. From py001 and tail001, we can see that a decision rule to "Reject H_0: if y > 0 " has size $\alpha = 0.0109452$ and, from py1 and tail1, has power $1-\beta = 0.6861894$ when p=0.1. To get at least 80% power would require p=0.14 or higher. [see tail14]

A consideration of possible confidence intervals seems prudent:

```
. cii 11 0

                                                         -- Binomial Exact --
    Variable |        Obs        Mean    Std. Err.       [95% Conf. Interval]
-------------+---------------------------------------------------------------
             |         11           0           0               0    .2849142*

(*) one-sided, 97.5% confidence interval

. cii 11 1

                                                         -- Binomial Exact --
    Variable |        Obs        Mean    Std. Err.       [95% Conf. Interval]
-------------+---------------------------------------------------------------
             |         11    .0909091    .0866784         .002299    .4127799

. cii 11 2

                                                         -- Binomial Exact --
    Variable |        Obs        Mean    Std. Err.       [95% Conf. Interval]
-------------+---------------------------------------------------------------
             |         11    .1818182    .1162913        .0228312    .5177559

. cii 11 3

                                                         -- Binomial Exact --
    Variable |        Obs        Mean    Std. Err.       [95% Conf. Interval]
-------------+---------------------------------------------------------------
             |         11    .2727273    .1342816        .0602177    .6097426
```

Notice that for all four 'possible' outcomes, the lower limit to the interval allows for very low values for p and the intervals are so wide so as to be next to useless.



: from Ramsay/Schafer (2002) - “Statistical Sleuth”

Sample Size when Modelling is Planned:

The literature on sample size matters relating to model building is very sketchy. Perhaps, the most widely quoted paper these days advocates for about 10 'events' per 'covariate'. (EPV)
See: 'Peduzzi et al'. Give this paper a read to get the gist of the matters considered there. You do not need to spend time with this paper's 'methods' section.

1) What are the main messages to emerge from this paper?
2) Have the authors made a strong case for EPV >10?
3) Review Table 1 in detail. Interpret the estimates carefully as we have discussed in this course. Consider the column of 'Wald p-values'. What questions might you ask that relate to the development of this model? With 6 indicator variables and one variable with 3 levels, there are 192 possible rates to study. If you had the actual dataset, how might you consider an analysis strategy?
4) Do you feel that the results and recommendations from this paper extend to your field of health research?