Survey design: power, precision and sample size

29 April 2017

This vignette covers the use of functions `incpower` and `incprecision`.

Function `incpower`

For this related set of calculations, we use the term "power" to mean the probability of obtaining a "statistically significant" result, *of the correct sign* in the estimation of an incidence difference, given some assumptions on effect size, recency test properties, and sample specification. See the vignette "introduction", especially the glossary, for some crucial details.

Function `incpower` primarily calculates samples sizes required to achieve desired power, or the power available at specified sample size(s). This requires study context defining parameter values–such as hypothetical incidence rates, prevalences, coverage rates, design effects, and the assay characteristics known as *mean duration of recent infection* (MDRI) and *false recent rate* (FRR)–and returns .

A number of supplementary outputs are also supplied, such as

- demonstrative relative standard errors and confidence limits, in the case that point estimates attain the true/expected values.
- expected survey counts, assuming a non structured sampling frame.

Examples

Consider calculating the power to infer the correct ordering of an incidence of 5% and one of 3%, at a shared prevalences of 20%, given a single set of recency test property estimates i.e. one value for each of MDRI, the RSE of MDRI, FRR, the RSE of FRR, and time cutoff T namely, in order: 200 days, 5%, 1%, 20%, and 730 days. Assume complete coverage of recency status ascertainment, and no survey design effects. Finally, envision a common study sample size of 5000 persons and set α to 5%. That power, as opposed to sample size (the default) is the desired output, is captured in the specification of the argument `Power = "out"` and `SS = NULL`.

```
incpower(I1 = 0.05, I2 = 0.03, PrevH1 = 0.20, PrevH2 = 0.20, n1 = 5000,
        n2 = 5000, alpha = 0.05,Power = "out", SS = NULL, CR = 1, DE_H = 1,
        DE_R = 1, BMest = "same.test", MDRI = 200, RSE_MDRI = 0.05, FRR = 0.01,
        RSE_FRR = 0.20, BigT = 730)
```

```
## $Inc.Difference.Statistics
##   deltaI_Est RSE_deltaI RSE_deltaI.infSS Power Power.infSS CI.low  CI.up
## 1       0.02       0.33            0.052 0.858       >0.99 0.0071 0.0329
##
## $Implied.Incidence.Statistics
##   Survey Given.I RSE_I CI.low CI.up
## 1      1    0.05 0.115  0.039 0.061
## 2      2    0.03 0.147  0.021 0.039
##
## $Implied.MDRI.Statistics
##   Given.MDRI CI.low CI.up
## 1        200  180.4 219.6
##
```

```
## $Implied.FRR.Statistics
##   Given.FRR CI.low CI.up
## 1      0.01  0.006 0.014
##
## $Implied.Subject.Counts
##                             Survey.1 Survey.2
## HIV.negative                    4000     4000
## HIV.positive                    1000     1000
## HIV.post.tested.for.recent      1000     1000
## Recency.test.pos                 116       73
```

Here the output returns that the power of this particular test is 0.858. In the limit of infinite sample size power approaches one.

For the benefit of survey planning (such as costing) the returned `Implied.Subject.Counts` object captures demonstrative survey counts in the case that expectation values are precisely attained.

To calculate the required (common) sample size for two surveys, to obtain a desired power:

- omit `n1` and `n2` or set both to `"both"`
- set `SS = "out"`
- set `Power` to the desired value.

```
incpower(I1 = 0.05, I2 = 0.03, PrevH1 = 0.20, PrevH2 = 0.15, alpha = 0.05,
        Power = 0.8, SS = "out", CR = 1, DE_H = 1, DE_R = 1,
        BMest = "FRR.indep", MDRI = 200, RSE_MDRI = 0.05, FRR = c(0.01,0.009),
        RSE_FRR = c(0.20,0.22), BigT = 730)
```

```
## $Minimum.Common.SS
## [1] 4122
##
## $Inc.Difference.Statistics
##   deltaI_Est RSE_deltaI RSE_deltaI.infSS Power Power.infSS CI.low CI.up
## 1       0.02      0.357            0.063   0.8       >0.99  0.006 0.034
##
## $Implied.Incidence.Statistics
##   Survey Given.I RSE_I CI.low CI.up
## 1      1    0.05 0.124  0.038 0.062
## 2      2    0.03 0.150  0.021 0.039
##
## $Implied.MDRI.Statistics
##   Given.MDRI CI.low CI.up
## 1        200  180.4 219.6
##
## $Implied.FRR.Statistics
##   Given.FRR CI.low  CI.up
## 1     0.010 0.0061 0.0139
## 2     0.009 0.0051 0.0129
##
## $Implied.Subject.Counts
##                             Survey.1 Survey.2
## HIV.negative                    3298     3504
## HIV.positive                     824      618
## HIV.post.tested.for.recent       824      618
## Recency.test.pos                  95       61
```

The function call outputs that the necessary common study sample size is 4122 persons per study to achieve

the desired 80% power given the specified population parameters and assay characteristics.

Function `incprecision`

This function summarizes performance of a recent infection test into a standard error of the incidence estimate, given the estimated test properties and hypothetical survey context or the sample size necessary for a given level of precision.

Examples

The function invocation below returns the necessary sample size to have RSE of the incidence estimator equal to 25%, given a hypothetical prevalence, coverage rate, and recency test parameter estimates. Note that `n = "out"`.

```
incprecision(I = 0.015, RSE_I = 0.25, PrevH = 0.2, CR = 1, MDRI = 200,
             RSE_MDRI = 0.05, FRR = 0.01, RSE_FRR = 0.2, BigT = 730,
             DE_H = 1.1, DE_R = 1.1, n = 'out')
```

```
## $sample.size
## [1] 3985
##
## $Prev.HIV.and.recent
## [1] 0.00833
##
## $Prev.HIV.and.nonrecent
## [1] 0.19167
##
## $RSE.I.inf.sample
## [1] 0.07606
##
## $RSE.PrevH
## [1] 0.03323
##
## $RSE.PrevR
## [1] 0.03564
```

Up to two arguments can be specified as ranges, with the input parameter `step` specifying the number of increments between the endpoints of the two ranges that are supplied under the argument name. Consider the calculation of sample size requirements for prevalence and incidence varied from 10 to 20% and 1.5 to 2% respectively:

```
incprecision(I = c(0.015,0.02), RSE_I = 0.25, PrevH = c(0.10,0.20), CR = 1,
             MDRI = 200, RSE_MDRI = 0.05, FRR = 0.01, RSE_FRR = 0.2, BigT = 700,
             DE_H = 1, DE_R = 1, n = 'out', step = 3)
```

```
## $sample.size
##             PrevH = 0.1 PrevH = 0.15 PrevH = 0.2
## I = 0.015          2660          2660         2660
## I = 0.0175         2547          2547         2547
## I = 0.02           2489          2489         2489
##
## $Prev.HIV.and.recent
##             PrevH = 0.1 PrevH = 0.15 PrevH = 0.2
## I = 0.015       0.00813       0.00813      0.00813
```

```
## I = 0.0175     0.00936      0.00936     0.00936
## I = 0.02       0.01045      0.01045     0.01045
##
## $Prev.HIV.and.nonrecent
##            PrevH = 0.1 PrevH = 0.15 PrevH = 0.2
## I = 0.015      0.09187      0.09187     0.09187
## I = 0.0175     0.14064      0.14064     0.14064
## I = 0.02       0.18955      0.18955     0.18955
##
## $RSE.I.inf.sample
##            PrevH = 0.1 PrevH = 0.15 PrevH = 0.2
## I = 0.015      0.05583      0.05583     0.05583
## I = 0.0175     0.06033      0.06033     0.06033
## I = 0.02       0.06549      0.06549     0.06549
##
## $RSE.PrevH
##            PrevH = 0.1 PrevH = 0.15 PrevH = 0.2
## I = 0.015      0.05817      0.05817     0.05817
## I = 0.0175     0.04717      0.04717     0.04717
## I = 0.02       0.04009      0.04009     0.04009
##
## $RSE.PrevR
##            PrevH = 0.1 PrevH = 0.15 PrevH = 0.2
## I = 0.015      0.02061      0.02061     0.02061
## I = 0.0175     0.02975      0.02975     0.02975
## I = 0.02       0.03817      0.03817     0.03817
```

To calculate the RSE of incidence over a range of 5 values of prevalence of positivity, at a sample size of 5000:

```
incprecision(I = 0.017, RSE_I = 'out', PrevH = c(0.10,0.20), CR = 1, MDRI = 211,
             RSE_MDRI = 0.05, FRR = 0.009, RSE_FRR = 0.2, BigT = 720, n = 5000,
             step = 5)
```

```
## $RSE_I
##           PrevH   RSE_I
## 1   PrevH = 0.1 0.16868
## 2 PrevH = 0.125 0.17352
## 3  PrevH = 0.15 0.17885
## 4 PrevH = 0.175 0.18470
## 5   PrevH = 0.2 0.19112
##
## $Prev.HIV.and.recent
##           PrevH Prev.HIV.and.recent
## 1   PrevH = 0.1             0.00947
## 2 PrevH = 0.125             0.00945
## 3  PrevH = 0.15             0.00944
## 4 PrevH = 0.175             0.00943
## 5   PrevH = 0.2             0.00942
##
## $Prev.HIV.and.nonrecent
##           PrevH Prev.HIV.and.nonrecent
## 1   PrevH = 0.1                0.09053
## 2 PrevH = 0.125                0.11555
## 3  PrevH = 0.15                0.14056
## 4 PrevH = 0.175                0.16557
```

```
## 5   PrevH = 0.2               0.19058
##
## $RSE.I.inf.sample
##           PrevH RSE.I.inf.sample
## 1   PrevH = 0.1          0.05363
## 2 PrevH = 0.125          0.05557
## 3  PrevH = 0.15          0.05824
## 4 PrevH = 0.175          0.06166
## 5   PrevH = 0.2          0.06585
##
## $RSE.PrevH
##           PrevH RSE.PrevH
## 1   PrevH = 0.1   0.04243
## 2 PrevH = 0.125   0.03742
## 3  PrevH = 0.15   0.03367
## 4 PrevH = 0.175   0.03071
## 5   PrevH = 0.2   0.02828
##
## $RSE.PrevR
##           PrevH RSE.PrevR
## 1   PrevH = 0.1   0.01383
## 2 PrevH = 0.125   0.01748
## 3  PrevH = 0.15   0.02113
## 4 PrevH = 0.175   0.02479
## 5   PrevH = 0.2   0.02845
```