Power and precision for physical intervention against microvascular complications

SDC

August 2017
Version 2

Compiled Friday 4th August, 2017, 15:15
from: `/home/bendix/sdc/coll/maej/microPower/mp.tex`

Bendix Carstensen Steno Diabetes Center Copenhagen, Gentofte, Denmark
& Department of Biostatistics, University of Copenhagen
bcar0029@regionh.dk b@bxc.dk
http://BendixCarstensen.com

Contents

1 Study set up

- We enroll N persons in total,
- allocate half to intervention half to conventional
- event rate is 7 / 100 PY
- 20% drop out over the first 2 years (10 / 100 PY)
- persons are followed over 5 years

The calculations rely on a simple assumption of a constant event rate of 7% resp. $\theta \times 7\%$ per year in each of the two groups.

2 Theory

If we assume a constant event rate of λ, then the survival function is $S(t) = \exp(-\lambda t)$, so we can simulate an event time by generating a uniform variate $u \in [0, 1]$ and solve:

$$u = \exp(-\lambda t) \quad \Leftrightarrow \quad t = -\log(u)/\lambda$$

Once we have generated an event time (or censoring if the generated event time is beyond the study period) for each person, we analyze the data by a constant rate Poisson model; which is just a standard `glm` with the log risk time as offset.

3 Implementation

3.1 Sample data

This is used to generate survival times in the following function that generates a dataset of size `N`, with event rate `lambda` and intervention effect `theta`, and dropout rate during the first two years of `delta`; the latter meaning that we assume that the drop-out rate only take place during the first two years and those who survive that remain in the study. We assume that `lambda` and `delta` are measured in %/year, and `Slength` in years.

```
> gm <-
+ function( N, lambda = 7, theta = 0.85, delta = 10, Slength=5 )
+ {
+ lambda <- lambda/100
+  delta <-  delta/100
+ allo <- factor( rep(c("int","ctr"), each=N/2) )
+   ue <- runif( N )
+   ud <- runif( N )
+  doM <- -log(ue)/(lambda*(theta^(allo=="int")))
+  doD <- -log(ud)/delta
+  doD <- ifelse( doD<2, doD, Slength )
+  doX <- pmin( Slength, doM, doD )
+  Xst <- doX==doM
+ data.frame( id = 1:N,
+           allo = allo,
+            doX = doX,
+            Xst = Xst )
+ }
> gm( 10 )
```

```
   id allo       doX   Xst
1   1  int 0.4357615 FALSE
2   2  int 0.5021269 FALSE
3   3  int 5.0000000 FALSE
4   4  int 1.0329377 FALSE
5   5  int 5.0000000 FALSE
6   6  ctr 5.0000000 FALSE
7   7  ctr 2.6259503  TRUE
8   8  ctr 3.9330196  TRUE
9   9  ctr 5.0000000 FALSE
10 10  ctr 5.0000000 FALSE
```

3.2 Analysis of generated data

Now for a dataset generated this way we can analyze it and derive the estimated effect and the precision and power. By the precision we mean the multiplicative factor (error factor, `erf`) with which we must multiply/divide the effect estimate in order to obtain the 95% confidence interval for the intervention effect. Where this is equal to the (inverse) effect estimate we have 50% power.

So here is an analysis of a generated dataset

```
> library( Epi )
> ci.lin( glm( Xst ~ allo,
+              offset=log(doX),
+              family=poisson,
+                data=gm(100) ) )
```

```
              Estimate    StdErr          z            P      2.5%     97.5%
(Intercept) -2.9303065 0.3333265 -8.7910996 1.481031e-18 -3.583614 -2.276999
alloint     -0.3537064 0.5039222 -0.7019067 4.827374e-01 -1.341376  0.633963
```

In this example we are interested in the `StdErr` and the `P` from the `alloint` line.

3.3 Deriving power and precision

We now devise a function that does this `nsim` times and returns precision and power:

```
> prp <-
+ function( nsim=1000, N, lambda = 7, theta = 0.85, delta = 10, Slength=5 )
+ {
+ res <- matrix( NA, nsim, 2 )
+ colnames( res ) <- c("erf","pval")
+ for( i in 1:nsim )
+ res[i,] <-
+ ci.lin( glm( Xst ~ allo,
+              offset = log(doX),
+              family = poisson,
+                data = gm(N,lambda,theta,delta,Slength) ) )[2,c(2,4)]
+ c( exp(1.96*mean(res[,1])), mean(res[,2]<0.05) )
+ }
> prp(nsim=100,N=200)
```

```
[1] 1.78314 0.08000
```

4 Results for power and precision

Now, in order to get a picture of how the power and precision of the study varies with the anticipated event rate and intervention effect we devise an array to hold the results from this small simulation model, and subsequently print them in a readable format.

```
> rArr <- NArray( list( "rate(%/y)" = 6:8,
+                      "effect(RR)" = seq(0.8,0.9,0.05),
+                                N = seq(1000,3500,500),
+                             what = c("erf","power") ) )
> system.time(
+ for( ir in dimnames(rArr)[[1]] )
+ for( ie in dimnames(rArr)[[2]] )
+ for( iN in dimnames(rArr)[[3]] )
+ rArr[ir,ie,iN,] <- prp( nsim = 500,
+                            N = as.numeric(iN),
+                       lambda = as.numeric(ir),
+                        theta = as.numeric(ie) ) )

   user  system elapsed
686.749   0.053 686.772

> round( ftable( rArr, col.vars=4:3), 2 )

                     what  erf                                 power
                     N     1000 1500 2000 2500 3000 3500  1000 1500 2000 2500 3000 3500
rate(%/y) effect(RR)
6         0.8              1.32 1.25 1.22 1.19 1.17 1.16  0.33 0.50 0.60 0.69 0.78 0.87
          0.85             1.31 1.25 1.21 1.19 1.17 1.16  0.19 0.31 0.37 0.50 0.49 0.59
          0.9              1.31 1.25 1.21 1.19 1.17 1.15  0.12 0.17 0.22 0.26 0.30 0.33
7         0.8              1.30 1.24 1.20 1.18 1.16 1.15  0.38 0.55 0.66 0.76 0.85 0.88
          0.85             1.29 1.23 1.20 1.18 1.16 1.15  0.23 0.35 0.41 0.48 0.55 0.69
          0.9              1.29 1.23 1.20 1.17 1.16 1.14  0.12 0.13 0.22 0.26 0.28 0.31
8         0.8              1.28 1.22 1.19 1.17 1.15 1.14  0.43 0.61 0.70 0.79 0.88 0.92
          0.85             1.27 1.22 1.19 1.17 1.15 1.14  0.26 0.35 0.53 0.58 0.65 0.69
          0.9              1.27 1.22 1.18 1.16 1.15 1.14  0.16 0.17 0.26 0.29 0.33 0.36
```

Note here that `N` is the *total* number of persons in the study, thus two groups of each `N/2` persons, and that the rate refer to the event rate in the control group.

The conclusion is that with a total of 1000 persons in the study, the relative uncertainty of the RR will be 30%, with 3000 or more it will be about 15%.