Power Revisited

Matthew W Wheeler

Last time…

We looked at power in the case of simple tests, but we have been looking at more complicated models. What do we do with these models? How can we estimate power in these cases?

Also, how do we easily design an experiment. Does SAS have a proc for that?

In SAS we have: proc glmpower

This proc will handle our glm power needs. It is a little different than proc power, because experimental designs are more complicated it requires a dataset to define our assumptions. We use it much like we used proc glm.

The name of the game is pretty much the same: We make assumptions about what a critical effect is, the overall variability, and the differences between groups. The proc does the rest.

Example 1

Suppose I am investigating the MPG using three different engines. Here I randomly sample car chassis produced for one type of car put in the engine and look at MPG. The first is a traditional engine, the second two have modifications made for fuel efficiency.

```
/*Example 1. Car example */;
/*DOE-GLMPOWER Example 1*/

/*Unbalanced sample size*/
DATA car;
        INPUT engine
                 MPG;
        DATALINES;
1 30
2 31
3 33
;
RUN;
```

```
* ENGINE DATA LOOK AT THE CONTRAST -
*
;
PROC GLMPOWER DATA=CAR;
        CLASS engine;
        MODEL MPG = engine;
        CONTRAST 'engine 1 VS. 2 contrast' engine 1 -1 0;
        CONTRAST 'engine 1 VS. 3 contrast' engine 1  0 -1;
        CONTRAST 'engine 2 VS. 3 contrast' engine 0  1 -1;
        POWER
                ALPHA= 0.01667 /*FOR A BF ADJUSTMENT*/
                POWER=0.80
                NTOTAL=.
                STDDEV= 1 2; * ASSUME A SD OF 1 or 2 ;
RUN;
```

Computed N Total							
Index	Type	Source	Std Dev	Test DF	Error DF	Actual Power	N Total
1	Effect	engine	1	2	12	0.904	15
2	Effect	engine	2	2	36	0.834	39
3	Contrast	engine 1 VS. 2 contrast	1	1	63	0.801	66
4	Contrast	engine 1 VS. 2 contrast	2	1	252	0.802	255
5	Contrast	engine 1 VS. 3 contrast	1	1	9	0.874	12
6	Contrast	engine 1 VS. 3 contrast	2	1	30	0.830	33
7	Contrast	engine 2 VS. 3 contrast	1	1	18	0.852	21
8	Contrast	engine 2 VS. 3 contrast	2	1	63	0.801	66

Now we must specify a dataset for each of our options we consider. The nice thing is that with normal models we are really only interested in the differences between the means.

That is if I change the location, the power answer doesn’t change.

We can APPROXIMATE the power from dichotomous data models using GLM too.

Example 2

Your client is interested in a mail marketing campaign for a credit card and has two different factors of interest: the introductory rate (HIGH/LOW) and the final rate (HIGH/LOW). You are interested in designing a study to see which set of factors have the highest response rate. We also assume this design is not balanced, that is more sample units will go to certain treatments than others.

```
/*Example 2
 *GLMPOWER */

/*Unbalanced sample size
  with 2 factors*/
/*THIS DATASET ASSUMES THAT THERE IS NO INTERACTION
  BETWEEN FACTORS*/
DATA CREDIT;
        INPUT intro $1-4
                  goto $6-9
                  responserate
                  size; /*unbalanced sample sizes*/
        DATALINES;
LOW  LOW         0.0135   10
LOW  HIGH        0.0125   1
HIGH LOW         0.0110   1
HIGH HIGH        0.010    10
;
RUN;
```

```
/* ASSUME THAT THERE IS NO INTERACTION!*/
PROC GLMPOWER DATA=CREDIT;
        CLASS intro goto;
        MODEL responserate= intro goto;
        WEIGHT size;
        POWER
                POWER=0.80
                NTOTAL=.
                STDDEV= %SYSFUNC(SQRT(0.01*(1-0.01)))
%SYSFUNC(SQRT(0.0135*(1-0.0135)));

*standard deviation options
*based on binomial proportion
*approximating normality!;
RUN;
```

Our sample size weights LOW LOW and
HIGH HIGH get the largest proportion of weights

Assuming there is not an interaction

Computed N Total						
Index	Source	Std Dev	Test DF	Error DF	Actual Power	N Total
1	intro	0.0995	1	104475	0.800	104478
2	intro	0.1154	1	140533	0.800	140536
3	goto	0.0995	1	417887	0.800	417890
4	goto	0.1154	1	562141	0.800	562144

What if there was an interaction:

```
/*THIS DATASET ASSUMES THAT THERE IS AN
INTERACTION
  BETWEEN FACTORS*/
DATA CREDIT2;
       INPUT intro $1-4
                goto $6-9
                responserate
                size; /*unbalanced sample
sizes*/
       DATALINES;
LOW  LOW      0.0135 10
LOW  HIGH     0.0125 1
HIGH LOW      0.0110 1
HIGH HIGH     0.007  10
;
RUN;
```

There is now an interaction.

```
/* ASSUME THAT THERE IS NO INTERACTION!*/
PROC GLMPOWER DATA=CREDIT2;
       CLASS intro goto;
       MODEL responserate= intro goto
intro*goto;
       WEIGHT size;
       POWER
              POWER=0.80
              NTOTAL=.
              STDDEV=
%SYSFUNC(SQRT(0.007*(1-0.007)))
%SYSFUNC(SQRT(0.0135*(1-0.0135)));
         *standard deviation options
         *based on binomial proportion
         *approximating normality!;
RUN;
```

Computed N Total						
Index	Source	Std Dev	Test DF	Error DF	Actual Power	N Total
1	intro	0.0834	1	41268	0.800	41272
2	intro	0.1154	1	79064	0.800	79068
3	goto	0.0834	1	105640	0.800	105644
4	goto	0.1154	1	202374	0.800	202378
5	intro*goto	0.0834	1	293410	0.800	293414
6	intro*goto	0.1154	1	562140	0.800	562144

That is a pretty big sample size.

At this point you need to ask a question.

Is the experiment worth it?

It is reasonable to assume that each credit card offer will cost about \$0.50 is it reasonable to spend between 50 to 250 thousand dollars?

That depends up on a large number of factors specific to the problem, and each of these factors should be considered.

With a covariate.

You can also look at the power when you add a specific covariate. This methodology can be used for general regression problems too. What is do below is coerce a covariate into the mix. It is assumed linear.

Example 3

There are two drugs on the market that increase the lung function in elderly COPD sufferers. You are interested in determining if one drug is superior to the other, and looking at the efficacy over no intervention.

It is thought that lung function decreases linearly with age. What is the sample required to determine if the two drugs have different efficacies?

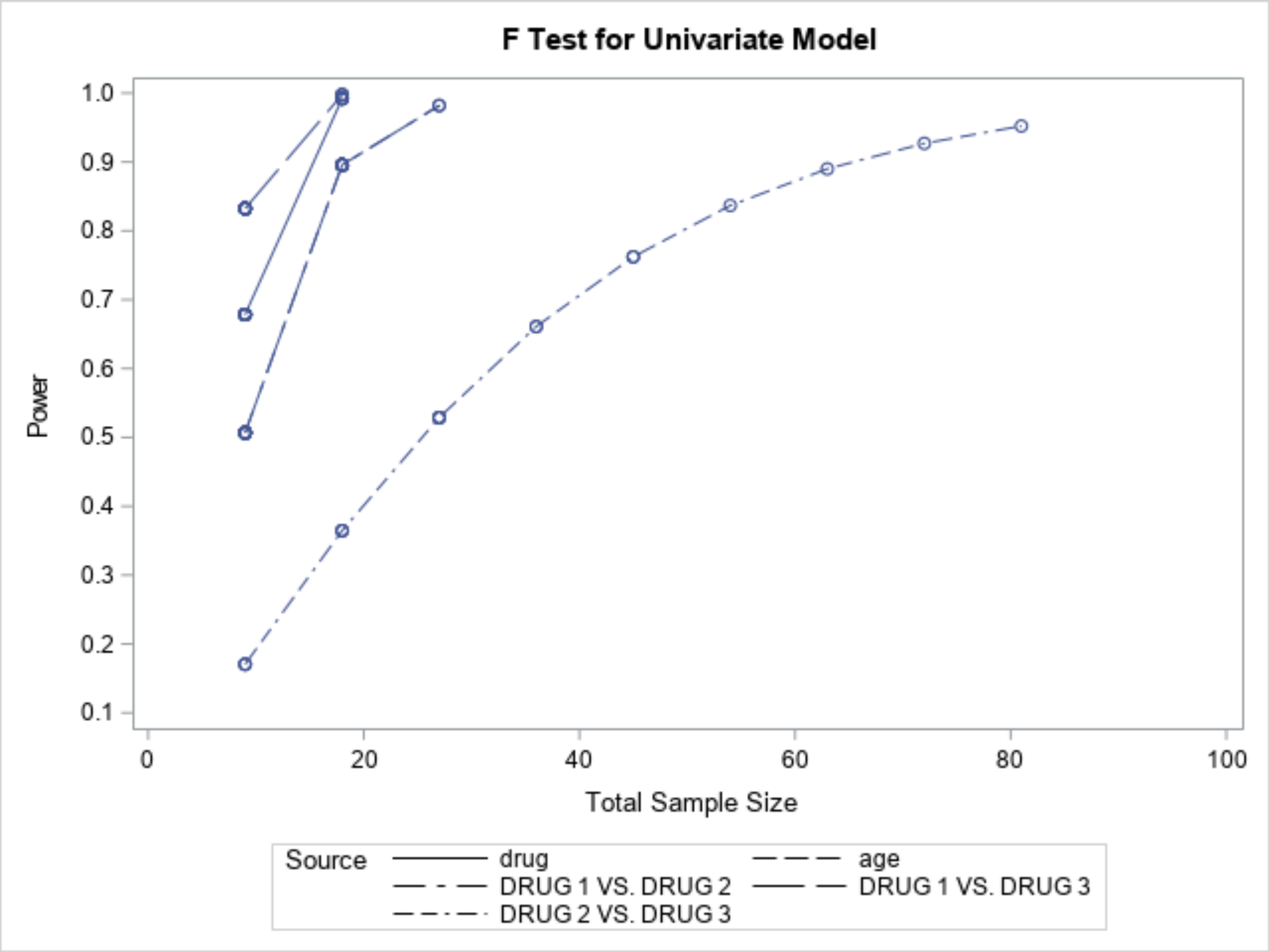
```
/*Example 3
*/

DATA LUNGCAPACITY;
        INPUT drug
                age
                cap;
DATALINES;
1    50   10
1    60   9
1    70   8
2    50   12
2    60   11
2    70   10
3    50   13
3    60   12
3    70   11
;
RUN;
```

```
PROC GLMPOWER DATA=LUNGCAPACITY;
        CLASS DRUG;
        MODEL CAP = DRUG AGE;
        CONTRAST 'DRUG 1 VS. DRUG 2' DRUG 1 -1 0;
        CONTRAST 'DRUG 1 VS. DRUG 3' DRUG 1 0 -1;
        CONTRAST 'DRUG 2 VS. DRUG 3' DRUG 0 1 -1;
        POWER
                POWER = 0.80
                NTOTAL= .
                STDDEV= 1;
                plot Y=POWER min=0 max=1;
RUN;
```

Computed N Total						
Index	Type	Source	Test DF	Error DF	Actual Power	N Total
1	Effect	drug	2	14	0.992	18
2	Effect	age	1	14	0.896	18
3	Contrast	DRUG 1 VS. DRUG 2	1	14	0.896	18
4	Contrast	DRUG 1 VS. DRUG 3	1	5	0.833	9
5	Contrast	DRUG 2 VS. DRUG 3	1	50	0.837	54

Our power for our tests over a range of N values.



More on GLMPOWER

The SAS procedure GLMPOWER as well as POWER have many options that we can’t cover fully in class.

If you have a question, please come to me and we can figure it out. Additionally the SAS manual gives detailed guidance on how to use the procedures with all of the possible options.

Other Topics

Experimental design is a vast topic, and there are many different types of designs you can consider. In fact, an entire course can be developed just on types of designs and when to use them. I am trying to focus on practical issues so I am not going to go into this realm.

Everything we have dealt with assumes you have two factors and enough data to estimate what you are interested in. For you, this is 99.9% of the time going to be the case.

When you have multiple variables (including BLOCKS) there are times when you can not estimate something.

Light switch problem:

Imagine a situation where every time you turn a switch to turn on a light. Another person flips a switch. The light turns on and off, but can you be sure the light switch actually controls the light? There are three options:

1. Your switch controls the light.
2. The other switch controls the light.
3. Both switches must be flipped to turn on the light.

In the form of a matrix:

You never observe these

Your Switch	Their Switch	Light
+	-	?
-	+	?
+	+	On

When an experimental design gets complicated and it is not complete it is possible you can't model between two factors because they are always "on" or "off" at the same time. One treatment is aliased with another. If we are interested in modeling it, we need to make sure we have the data before hand.

There is an entire field of study that builds designs that allow us to estimate the effects we are interested in, but do not waste resources doing it.

For example: Consider an experiment with three treatments the full design matrix is

I	A	B	C	AB	AC	BC	ABC
+	+	-	-	-	-	+	+
+	-	+	-	-	+	-	+
+	-	-	+	+	-	-	+
+	+	+	+	+	+	+	+
+	+	+	-	+	-	-	-
+	+	-	+	-	+	-	-
+	-	-	+	-	-	+	-
+	-	+	-	+	+	+	-

Where a + is treatment 1 – is treatment 2

This design allows us to estimate all main effects as well as 2 and 3 way interactions.

But suppose we can't afford the 2 or 3 way interaction. Can we get by with just estimating the main effects?

We can if we do the following design:

I	A	B	C	AB	AC	BC	ABC
+	+	-	-	-	-	+	+
+	-	+	-	-	+	-	+
+	-	-	+	+	-	-	+
+	+	+	+	+	+	+	+

But like the light switch I can't tell the difference between A and BC, B and AC or C and AB

This is not fun in general, but SAS can do it

The SAS procedure FACTEX can be used to automatically design our fractional factorial experiment. It designs factorial designs that can then be used directly when assigning experimental units.

If you need to you can go into this deeper, but I am only going to highlight some functionality

Definitions

Resolution III Design: No main effects are aliased with any other main effect, but main effects are aliased with two factor interactions. The example is an example of a Resolution III Design.

Resolution IV Design: No main effect is aliased with any other main effect or two factor interaction, but two-factor interactions are aliased with each other.

Resolution V Design: No main effect or two factor interaction is aliased with any other main effect or two factor interaction, but two-factor interactions are aliased with three-factor interactions.

Using FACTEX

Consider our credit card example. Here we have six factors and we want a Resolution V design and want 32 unique rows in our design matrix. What does the design look like?

```
/*Example 4 Using Fact ex*/
PROC FACTEX;
        FACTORS Intro Duration Goto Color Postage Rewards;
        MODEL resolution=5; *resolution 5 design;
        SIZE design = 32;   *2^(6-1) = 32 required to
                                      *create design;
        EXAMINE confounding design;
        OUTPUT OUT = designM /*output our info in a
                                          human readable form*/
                Intro        nvals = (0 2.99)
                Duration     nvals = (9 12)
                Goto         nvals = (4.99 7.99)
                Color        cvals = ("white" "blue")
                 Postage     cvals = ("first class" "third class")
                Rewards      cvals = ("yes" "no");
RUN;
QUIT;
```

Partial output

Design Points						
Experiment Number	Intro	Duration	Goto	Color	Postage	Rewards
1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	1	1
3	-1	-1	-1	1	-1	1
4	-1	-1	-1	1	1	-1
5	-1	-1	1	-1	-1	1
6	-1	-1	1	-1	1	-1
7	-1	-1	1	1	-1	-1
8	-1	-1	1	1	1	1
9	-1	1	-1	-1	-1	1
10	-1	1	-1	-1	1	-1

What about a resolution 4 design?

```
/* what about a resolution 4 design*/
PROC FACTEX;
      FACTORS Intro Duration Goto Color Postage Rewards;
      MODEL resolution=4; *resolution 4 design;
      SIZE design = 16;   *2^(6-2) = 32 required to
                                    *create design;
      EXAMINE confounding design;
      OUTPUT OUT = designM /*output our info in a
                                         human readable form*/
            Intro         nvals = (0 2.99)
            Duration      nvals = (9 12)
            Goto          nvals = (4.99 7.99)
            Color         cvals = ("white" "blue")
             Postage      cvals = ("first class" "third class")
            Rewards       cvals = ("yes" "no");
RUN;
QUIT;
```

Design Points						
Experiment Number	Intro	Duration	Goto	Color	Postage	Rewards
1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	1	1	1
3	-1	-1	1	-1	1	1
4	-1	-1	1	1	-1	-1
5	-1	1	-1	-1	1	-1
6	-1	1	-1	1	-1	1
7	-1	1	1	-1	-1	1
8	-1	1	1	1	1	-1
9	1	-1	-1	-1	-1	1
10	1	-1	-1	1	1	-1
11	1	-1	1	-1	1	-1
12	1	-1	1	1	-1	1
13	1	1	-1	-1	1	1
14	1	1	-1	1	-1	-1
15	1	1	1	-1	-1	-1
16	1	1	1	1	1	1

Full design….

FYI

These more complicated designs are probably going to be rare in your professional career (for example I have never done anything close to this). They are mainly used in industrial process improvement (think Six Sigma super ninja black belt). But if you go into this type of work I recommend a book on Response Surface Methodology.

The one I used for these slides is “Response Surface Methodology” Meyers and Montgomery.

If you do have a problem like this, you can build the design, and then feed it up into proc glmpower. From there you can go to town and build your experiment.