UNIFYPOW

A SAS MODULE/MACRO FOR SAMPLE-SIZE ANALYSIS

Ralph G. O'Brien
Department of Biostatistics and Epidemiology
The Cleveland Clinic Foundation
robrien@bio.ri.ccf.org

29 August 1998

UnifyPow Examples

UNIFYPOW Mission

To develop comprehensive SAS-based tools for research planning, in particular, for choosing sample sizes and assessing statistical power.

"The calculation and justification of sample size is at the crux of the design ...

... Ideally, clinical trials should have adequate **power, ≈90%**, to detect a clinically relevant difference between the experimental and control therapies. Unfortunately, the power of clinical trials is frequently influenced by budgetary concerns as well as pure biostatistical principles. Yet an underpowered trial is, by definition, unlikely to demonstrate a difference between the interventions assessed and may ultimately be considered of little or no clinical value. From an ethical standpoint, an underpowered trial may put patients needlessly at risk of a new therapy without being able to come to a clear conclusion."

Topol EJ, Califf RM, Van de Werf F, et. al.
Circulation, 1997, 95: 1072-1082

1998 SASware Ballot

Rank (out of 588)

1. Calculate confidence intervals for all appropriate procedures.
2. Provide ability to establish a secondary SASWORK area to allow the program to continue processing when the current work location runs out of space.
3. **Provide power analysis and sample-size determination in all applicable procedures.**

Main Things

A shameless boast:

UNIFYPOW is freeware that rivals freestanding commercial applications costing hundreds of dollars.

- You do not need to be a SAS expert. Easy to Learn
- Runs in base SAS system, thus under numerous platforms and configurations. Outstanding portability
- Builds a SAS data set of results. Thus, SAS users can develop customized reports, even merging results from two or more UNIFYPOW runs.

  Outstanding reporting flexibility

More Main Things

- Handles
  - Unbalanced sample sizes in G independent groups for all relevant problems.
  - One- and two-tailed tests when appropriate.
  - Any alpha level.
  - General contrasts (including $df_H > 1$) on cell means, logits, and Fisher's Z-transformed correlations.

  Outstanding depth
- Uses exact or virtually exact computations whenever feasible. Outstanding accuracy

- See SUGI 23 Proceedings paper for description of methods covered.
- Website contains Acrobat files of
  - current notes for full-day workshop; including examples covering more than 20 methods (suitably bookmarked).
  - SUGI 23 Proceedings paper, updated as UNIFYPOW evolves
- Website contains text files of
  - Test set of all examples from full-day workshop
  - UnifyPow code. Algorithms are described within code

A Simple Experiment

- Suppose Mystic Michelle is either "ordinary" or "gifted" in her ability to predict the toss of a fair coin.

  Ordinary: $\pi = .50$ (50% correct, on average)

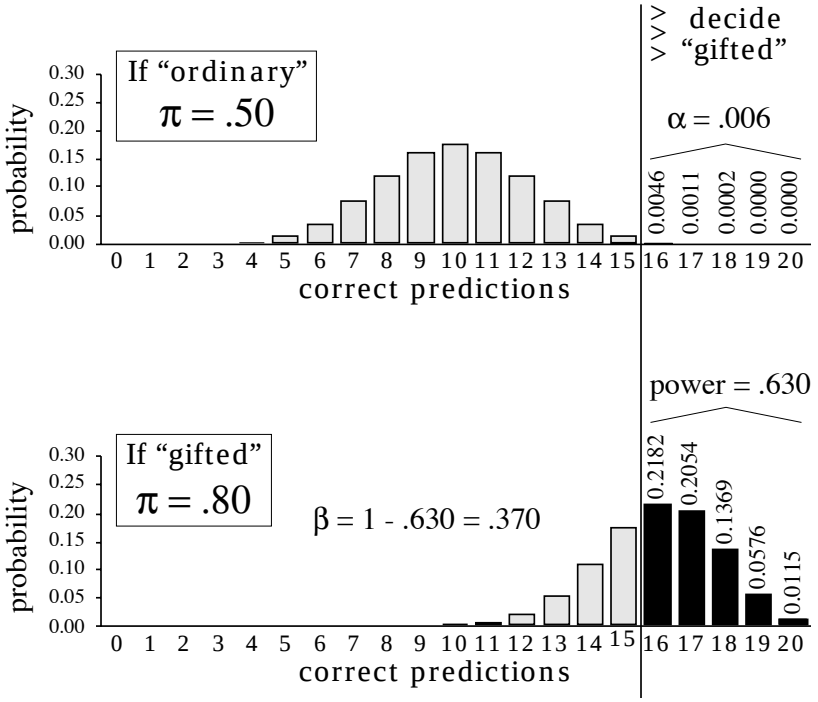
  Gifted: $\pi = .80$ (80% correct, on average)

- Perform study to decide whether she is gifted.
  - Predict the outcomes of 20 tosses.
  - Decide "gifted" if Michelle is correct on 16 (80%) or more tosses.
  - Do not decide "gifted" if correct on 15 or fewer.

Example is trite, but covers key concepts simply.

How good is this simple experiment?

Binomial Distributions, N = 20



Key Probabilities: N = 20

- α = p-value for 16 successes = $\Pr[X \geq 16; \pi = .50]$ =
  $.0046 + .0011 + .0002 \approx .006$
- power = $1 - \beta = \Pr[X \geq 16; \pi = .80]$ =
  $.2182 + .2054 + .1369 + .0576 + .0115 \approx .630$

Decision (sample data)	Truth ("infinite data") "ordinary"	Truth ("infinite data") "gifted"
"ordinary"	correct	Type II error (β)
"gifted"	Type I error (α)	correct (power)

Solving with UnifyPow

```
%let UnifyPow = file-specification-here;

*-------------------------------------------;
* Testing single proportion (sign test)    ;
*-------------------------------------------;
%include "&UnifyPow";
title2 "Mystic Michelle: Ordinary (pi = .50) or
Gifted (pi = .80)?";
title3 "Testing a single proportion";
datalines;
pi .80 . Mystic Michelle claims 80% accuracy.
null .50 . (This is default--not really needed.)
alpha .01 .05
Ntotal 20 40
%tables;
```

Alternative: `power .990 .995`

Scenario: pi .80 . Mystic Michelle claims 80% accuracy.

		ALPHA			
		0.01		0.05	
		Total N		Total N	
		20	40	20	40
		Power	Power	Power	Power
Method	Type				
Exact Binomial	2-tail bnml	.630	.912	.804	.981
	1-tail bnml	.630	.957	.804	.992

Critical values and actual alpha levels using binomial distribution.

			ALPHA					
			0.01			0.05		
Method	Total N	Type	Actual Alpha	Lower Crit Value	Upper Crit Value	Actual Alpha	Lower Crit Value	Upper Crit Value
Exact Binomial	20	2-tail bnml	0.007	3	16	0.041	5	15
		1-tail bnml	0.006	.	16	0.021	.	15
	40	2-tail bnml	0.006	11	29	0.038	13	27
		1-tail bnml	0.008	.	28	0.040	.	26

- By increasing N and adopting a decision rule that keeps α the same, we increase power.
- Coin tosses are cheap. Use many of them.
- Subjects in real studies are usually expensive to run. We are constantly fighting against low power.

Non-Directional vs. Directional Hypotheses

("Two-tailed" vs. "One-tailed" Tests)

- "Two-tailers always" say all hypotheses are really non-directional because researchers will usually rationalize and report any result that goes in the "wrong" direction. Some clinical trialists say that all trials should be non-directional, because one direction tests efficacy and other tests harm.
- "Flexible scientists" (Note my bias!) say that studies have unique backgrounds and goals. If the research hypothesis is directional, then so should be the statistical hypothesis. This will gain power—if you pick the right direction. If the research hypothesis is non-directional, then so should be the statistical hypothesis. This broadens the question, making any statistically significant result appear more "honest."

"Power" Statement and Comment Blocks

Michelle knows UnifyPow and decides to plan a study that has a very small chance of making either type of error.

```
pi .80
/#
Same problem, but now find minimum N to
achieve specified power at given alphas.
#/
power .99 .995
alpha .005
tails 1
%tables;
```

	ALPHA	
	0.005	
	Minimum Power	
	.990	.995
	Total N	Total N
Method: Exact Binomial	Type: 1-tail bnml	61	69

Graphical output?

GraphPow is a SAS freeware macro that plots UnifyPow results.

Actively being programmed by Christine Skibinski, Cleveland Clinic Foundation.

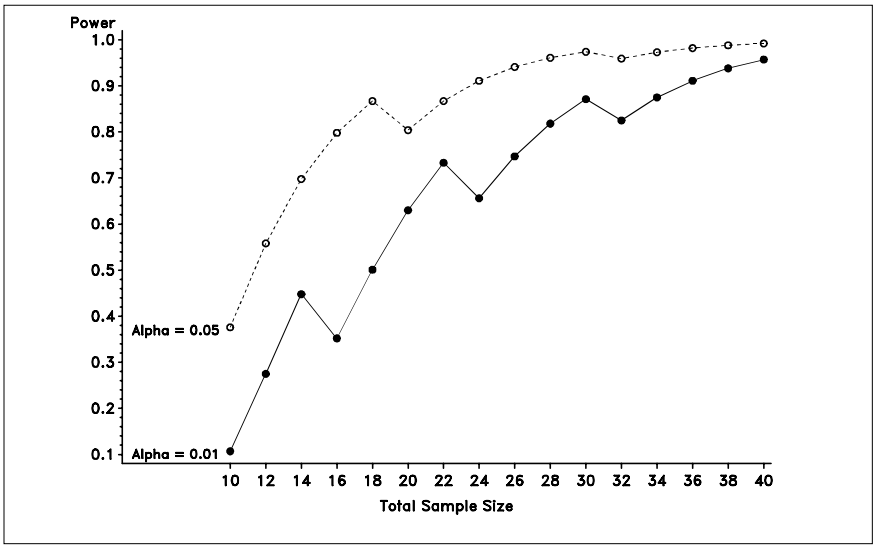
Beta release: Fall 1998 from the UnifyPow website. Registered UnifyPow users will be notified via email.

You must have SAS/GRAPH installed.

Example of GraphPow:

```
%include = "GraphPow file-specification";
%let UnifyPow = file-specification-here;
*-----------------------------------------------;
* A Preview of GraphPow                         ;
*-----------------------------------------------;
%include "&UnifyPow";
title2 "Mystic Michelle: Ordinary (pi = .50) or
Gifted (pi = .80)?";
datalines;
pi .80 . Mystic Michelle claims 80% accuracy.
alpha .01 .05
Ntotal 10 to 40 by 2 . <== new syntax!
tails 1
Graph output EPS
%graphs;
```


Mystic Michelle: H_0: $\pi = 0.50$ vs. H_a: $\pi > 0.50$
Power if $\pi = 0.80$



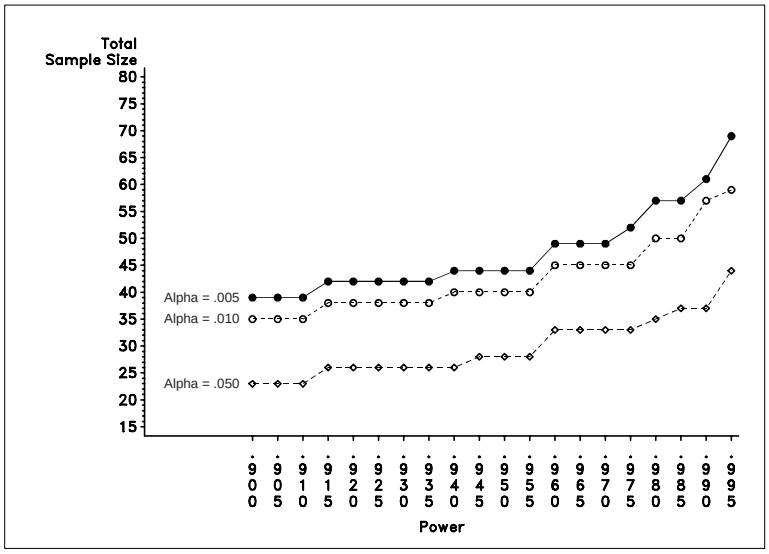
Yes, the sawtooth lines are accurate for this *exact* solution.

Another Example of GraphPow:

```
%include = "GraphPow file-specification";
%let UnifyPow = file-specification-here;
*-------------------------------------------;
* A Preview of GraphPow                    ;
*-------------------------------------------;
%include "&UnifyPow";
datalines;
pi .80 . Mystic Michelle claims 80% accuracy.
alpha .005 .01 .05
power .90 to .995 by .005 . <== new syntax!
tails 1
Graph output EPS
%graphs;
```


Mystic Michelle: H_0: $\pi = 0.50$ vs. H_a: $\pi > 0.50$

Sufficient Number of Coin Tosses if $\pi = 0.80$



```
*-------------------------------------------;
* Testing single proportion (custom null)  ;
*-------------------------------------------;
title2 "DCA for Lactic Acidosis in Children
with Severe Malaria";
title3 "Uncontrolled pilot study, comparing
mortality to 28%";
%include "&UnifyPow";
datalines;
pi .14 . DCA cuts mortality in half.
null .28 . est mortality rate from epi studies
alpha .05 .20 . high alpha OK for pilot study
Ntotal 40 60 100
%tables;
```



```
Testing Ho: pi = 0.28

NOTE: SETTING 2-TAILED CRITICAL REGIONS FOR THE BINOMIAL DISTRIBUTION.
      Denote the critical regions for a 2-tailed test as "major" and
      "minor" depending on which one is consistent with the true pi.
      Thus for Ho: pi = .35 with a conjecture of true pi = .20, the
      major critical region would be in the lower tail of the
      binomial(Ntotal, .35) distribution.  Let alpha_major and
      alpha_minor be the Type I error rates in these tails.  UnifyPow
      first finds the largest minor critical region such that
                  alpha_minor LE alpha/2,
      where "LE" stands for "less than or equal." Then it finds the
      largest major region such that
                  alpha_major LE alpha - alpha_minor.
      This ensures that
                  alpha_minor + alpha_major LE alpha,
      and yet favors the major tail, thus increasing power.

      The critical values tabled below are in the rejection region.
      For example, the alpha = .05, two-tailed test of Ho: pi = .35 with
      NTotal = 40 gives lower and upper critical values of 8 and 21
      if the conjectured true pi is less than .35.  Thus, the major
      critical region is r = 0, 1, ..., 8 and the minor one is
      r = 21, 22, ..., 40.  If the conjectured true pi exceeds .35,
      then the major region is r = 20, 21, ..., 40 and the minor one is
      r = 0, 1, ..., 7.
```

DCA for Lactic Acidosis in Children with Severe Malaria
Uncontrolled trial, comparing mortality to 28%

Scenario: pi .14 .

		ALPHA					
		0.05			0.2		
		Total N			Total N		
		40	60	100	40	60	100
		Power	Power	Power	Power	Power	Power
Method	Type						
Exact Binomial	2-tail bnml	.504	.788	.939	.811	.931	.990
	1-tail bnml	.676	.788	.964	.902	.965	.995

Critical values and actual alpha levels using binomial distribution.

			ALPHA					
			0.05			0.2		
Method	Total N	Type	Actual Alpha	Lower Crit Value	Upper Crit Value	Actual Alpha	Lower Crit Value	Upper Crit Value
Exact Binomial	40	2-tail bnml	0.033	5	18	0.161	7	16
		1-tail bnml	0.043	6	.	0.171	8	.
	60	2-tail bnml	0.046	10	25	0.196	12	22
		1-tail bnml	0.030	10	.	0.172	13	.
	100	2-tail bnml	0.045	19	38	0.184	22	35
		1-tail bnml	0.044	20	.	0.158	23	.

These critical values are part of the rejection region.
The note above describes how they are set.

```
*-------------------------------------------------------;
* Comparing 2 independent proportions                   ;
*-------------------------------------------------------;
%include "&UnifyPow";
title2
"Placebo vs. DCA for Lactic Acidosis in Children with
Malaria";
title3 "28% die untreated. What if DCA cuts this by 25%?";
datalines;
/#
o  Actual study of children with _severe_ malaria.
o  28% base mortality rate supported by meta-analysis of
   published surveys.
o  25% reduction in mortality supported by animal-model
   study.
o  Small pilot study had 2/10 deaths in each group.
o  2:1 randomization to DCA pleases local health officials.
o  One interim look using alpha = .01 & final analysis at
  .045.
o  Able to randomize 1500. Should interim look come at 750
   or 999?
#/
```


Design and Scenario

	Lived	Died	
Placebo	72%of n_1	28%of n_1	$n_1 = N/3$
DCA	79%of n_2	21%of n_2	$n_2 = 2N/3$

```
pi .28 .21
weight 1 2
alpha .01 .045
Ntotal 750 999 1500
%tables;
```

Testing Ho: pi1 - pi2 = 0

Scenario: pi .28 .21

		ALPHA					
		0.01			0.045		
		Total N			Total N		
		750	999	1500	750	999	1500
		Power	Power	Power	Power	Power	Power
Method	Type						
Approximate Uncondit'l "chi^2*"	2-tld t apr	.331	.458	.674	.554	.680	.847
	1-tld t apr	.426	.557	.759	.672	.781	.909
Exact Uncondit'l**	2-tld t apr	.307	.427	.639	.527	.651	.823
	1-tld t apr	.399	.526	.728	.647	.757	.892
Fisher's exact conditional	2-tld aprx	.288	.412	.633	.505	.636	.819
	1-tld aprx	.378	.511	.722	.626	.745	.889
Likhd Ratio for Log Odds Ratio	2-tail Z	.323	.447	.661	.544	.669	.838
	1-tail Z	.417	.546	.747	.663	.772	.903

Note: We will redo this one later to introduce how the "exemplary data" method can handle complex linear models.

Footnote to table

*The Approximate Unconditional corresponds to the Ordinary Pearson chi-square test for a 2 x 2 table. Technically, the method here uses a regular t test with Y = 0 (no) or 1 (yes), which is known to offer more accurate p-levels and can be done with any standard t-test routine. See D'Agostino, Chase, and Belanger (1988), American Statistician, 1988, 42:198-202.

**The Exact Unconditional corresponds to the test proposed by Suissa and Shuster (1985), J Royal Stat Soc A, 148:317-327).

```
*-------------------------------------------------------;
* Testing goodness of fit to a multinomial distribution ;
*-------------------------------------------------------;
%include "&UnifyPow";
title2 "Do the Biostat Boys use fair dice?";
datalines;
/#
A trio of shady characters called the Biostat Boys run a craps
game at lunch near the FDA.  One of the regular patrons, Lucky Luke,
has been losing lately and wonders if the dice are fair. The
known distribution for a fair pair of dice is:

    X:   2    3    4    5    6    7    8    9   10   11   12
 p(X): .028 .056 .083 .111 .139 .166 .139 .111 .083 .056 .028

Luke believes that one dice is weighted so that the "1" comes up 1/4
of the time, not 1/6. So he computes that the distribution may be:

    X:   2    3    4    5    6    7    8    9   10   11   12
 p(X): .042 .067 .092 .116 .142 .166 .125 .100 .075 .050 .025

How many tosses must he observe in order to have 90% power
(alpha = .05) to find such a discrepancy, based on the
full distribution?

#/
GoodnessOfFit .042 .067 .092 .116 .142 .166 .125 .100 .075 .050 .025
Null .028 .056 .083 .111 .139 .166 .139 .111 .083 .056 .028
alpha .05 .10
power .80 .90
%tables;
```


Scenario: {0.042 0.067 0.092 0.116 0.142 0.166 0.125 0.1 0.075 0.05 0.025}

		ALPHA			
		0.05		0.1	
		Minimum Power		Minimum Power	
		.800	.900	.800	.900
		Total N	Total N	Total N	Total N
Method	Statistic				
Ordinary Pearson	Chi-square	1109	1402	913	1187
Likhd Ratio	Chi-square	1186	1500	976	1270

```
%include "&UnifyPow";
datalines;
/#
What if Lucky Luke only counts how many "1"s
appear on a throw?
The distributions are
            X:   0    1    2
    fair p(X): .694 .278 .028
  biased p(X): .625 .333 .042
#/
GoodnessOfFit .625 .333 .042
Null .694 .278 .028
alpha .05 .10
power .80 .90
%tables;
```


Scenario: {0.625 0.333 0.042}

		ALPHA			
		0.05		0.1	
		Minimum Power		Minimum Power	
		.800	.900	.800	.900
		Total N	Total N	Total N	Total N
Method	Statistic				
Ordinary Pearson	Chi-square	390	512	312	424
Likhd Ratio	Chi-square	413	542	330	448

Lesson: It pays to test tighter hypotheses...if you guess right!

```
*-----------------------------------------------------;
* Testing association in an R x C contingency table  ;
*-----------------------------------------------------;

%include "&UnifyPow";
title2 "Variation in sarcoma type by region";
/*
Stimulated by example found in Section 3.4 of Freeman DH
(1987). Applied Categorical Data Analysis, New York: Marcel
Dekker.
*/
datalines;
/#
There are 3 different types of soft-tissue sarcomas of the
arms and legs:
      o  Fibroid
      o  Lipoid
      o  Mixed (or other)

Incidence data can be obtained from good cancer registries.
The Question: Do the relative proportions of these 3 types
differ among 4 geographic regions?


#/
```


	SARCOMA SOFT TISSUE TYPE			% OF TOTAL SAMPLE
REGION	Fibroid	Lipoid	Mixed	
A	.50	.20	.30	30%
B	.60	.25	.15	20%
C	.35	.35	.30	25%
D	.45	.20	.35	25%

How large must Ntotal be to have 90% or 95% power?

```
2WayContTable .50 .20 .30
 >            .60 .25 .15
 >            .35 .35 .30
 >            .45 .20 .35
weight .30 .20 .25 .25
alpha .01 .05
power .90 .95
%tables;
```

Scenario: {0.5 0.2 0.3} v. {0.6 0.25 0.15} v. {0.35 0.35 0.3} v. {0.45 0.2 0.35}

		ALPHA			
		0.01		0.05	
		Minimum Power		Minimum Power	
		.900	.950	.900	.950
		Total N	Total N	Total N	Total N
Method	Statistic				
Ordinary Pearson	Chi-square	500	580	380	440
Likhd Ratio	Chi-square	480	560	360	420

Common t-test comparing 2 independent means

Can biofeedback therapy reduce the frequency and severity of chronic vascular headaches?

- Double-blind, randomized trial
- Enhanced Thermal (ET) biofeedback therapy vs. Sham Placebo (SP)
- Each patient will be studied from a Monday to a Friday.
  - On Monday begin standard medical therapy.
  - On Monday and Tuesday evenings, determine Pretreatment Vascular Headache Index (**PreVHI**).
  - On Wednesday morning, patients will be randomized to ET or SP groups
  - Wednesday afternoon—main behavioral therapy session, with a followup session early Thursday morning
  - Thursday and Friday determine **PostVHI**.

2/3 randomized to ET

- $w_{ET} = 2/3$, $w_{SP} = 1/3$; 2:1 randomizaton
- will help with recruiting
  - *recruiting* efficiency outweighs the *statistical* efficiency of a balanced design

This sets the design.

What is the primary outcome measure?

Started with: Relative Change in VHI = $\frac{\text{PostVHI} - \text{PreVHI}}{\text{PreVHI}}$

Statistician suggests:

$$\begin{aligned}\text{VHchange} &= \log_2[\text{PostVHI}/\text{PreVHI}] \\ &= \log_2\text{PostVHI} - \log_2\text{PreVHI}.\end{aligned}$$

The "Elicitation Process"

What is the *infinite dataset?* (the population distributions)

- Alalgia has some pilot data, but not nearly enough to do anything formal. His conjectures:
  - ET: 45% reduction → $\log_2[0.55] = -0.86 \rightarrow \mu_{ET} = -0.86$
  - SP: 25% reduction → $\log_2[0.75] = -0.42 \rightarrow \mu_{SP} = -0.42$

Determining σ is difficult.

- Statistician must help, probe! Where is middle 95% of data?
- For ET group: –75% to +21% relative change → $-2.00 < \text{VHchange} < +0.28$
- Assuming VHchange ~ Normal, this defines a 4σ range
- Thus, $\sigma = 2.28/4 = 0.57$.
- For SP group, $\sigma = 0.57$ → mid-95% range of -1.56 < VHchange < +0.72
  → mid-95% range of relative change: –66% to +65%

Look reasonable?

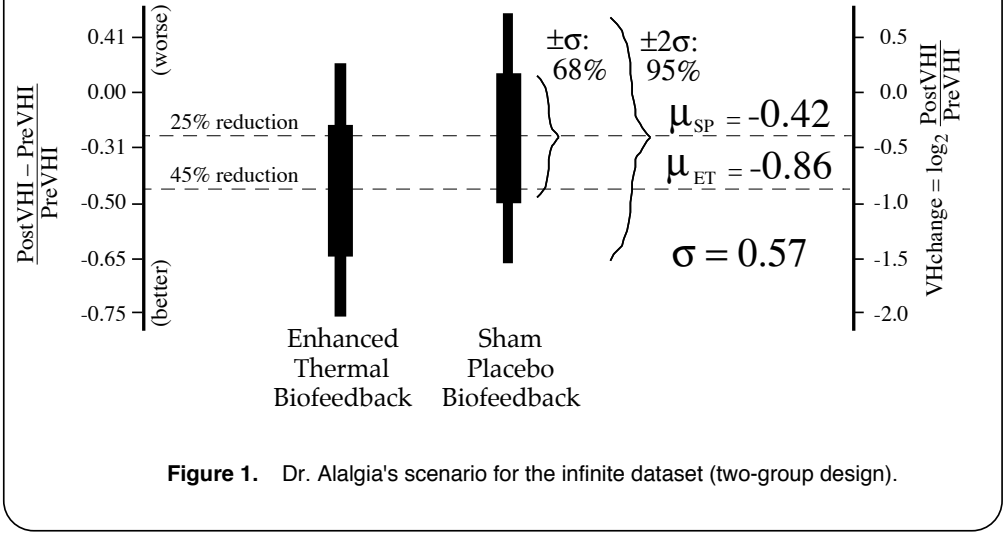


Figure 1. Dr. Alalgia's scenario for the infinite dataset (two-group design).

Prudent to bracket values for σ (sensitivity analysis)

- So, $\sigma = \{0.45, 0.57, 0.65\}$

Dr. Alalgia has enough resources to study N = 21 patients.

- 105 hospital bed-days (21 patients $\times$ 5 days/patient)
- Look at N = {15, 21, 33}

Biofeedback has a history of conflicting findings.

- Directional ("one-tailed") test unwarranted, but interesting to view one-tailed powers anyway.

Use $\alpha = .05$

- but Alalgia curious about $\alpha = .01$.

Under these conditions...

Will N = 21 provide acceptable power?

Would fewer suffice?

```
*-------------------------------------------;
* Ordinary 2-group t test                   ;
*-------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Alalgia's 2-Group Design";
title3 "Traditional t-Tests";
datalines;
/#
Scenario represents reductions of 45% vs. 25%.
#/
mu -.86 -.42
SD .45 .57 .65
alpha .05 .01
weight 2 1
NTotal 15 21 33
%tables;
```

Alternative: `power .80 .90`

Output Using **NTOTAL** Statement

Scenario: mu -.86 -.42 .
Ordinary t test

		Standard Deviation								
		0.45			0.57			0.65		
		Total N			Total N			Total N		
		15	21	33	15	21	33	15	21	33
		Power	Power	Power	Power	Power	Power	Power	Power	Power
Alpha	Type									
0.05	2-tail t	.380	.518	.727	.257	.353	.526	.209	.284	.427
	1-tail t	.519	.652	.828	.379	.485	.655	.318	.407	.559
0.01	2-tail t	.156	.260	.472	.091	.147	.276	.068	.108	.201
	1-tail t	.235	.358	.581	.145	.219	.372	.112	.167	.283

Note: We will redo this one later to introduce how the “exemplary data” method can handle complex linear models.

Output Using **POWER** Statement

Scenario: mu -.86 -.42
Ordinary t test

		Standard Deviation					
		0.45		0.57		0.65	
		Minimum Power		Minimum Power		Minimum Power	
		.800	.900	.800	.900	.800	.900
		Total N	Total N	Total N	Total N	Total N	Total N
Alpha	Type						
0.05	2-tail t	39	54	63	84	81	108
	1-tail t	33	42	51	69	63	87
0.01	2-tail t	60	75	93	117	120	150
	1-tail t	51	66	81	102	102	132

Even researchers with limited experience in statistical planning readily understand the rationale and implications of a power analysis like that given here.

Easy to see how much power increased by

- increasing α
- decreasing σ
- increasing N
- directional test.

Dr. Alalgia is discouraged about these results

- 105 bed-days for a project with weak power
- Even for one-tailed test and assuming $\sigma = 0.45$, 80% power requires N = 33.
- decides to restructure the design...

Matched-pairs t-test (2 correlated means)

Run in *pairs* after being matched on several factors

New protocol

- Monday and Tuesday, outpatients, begin standardized medical treatment
  - complete **PreVHI** measure,
- Wednesday morning, admit to GCRC, already randomized.
  - ET subject will be treated first, getting legitimate biofeedback.
  - SP subject will be treated as a "yoked control"
- Thursday
  - re-treat
- Thursday afternoon and Friday morning
  - **PostVHI**
- Friday morning
  - discharge

Outcome measure:

$$D = \text{VHchange}_{ET} - \text{VHchange}_{SP}$$

Hoped:

- matching and yoking pairs of patients → lower error variance

Note:

- 2 GCRC bed-days/subject (versus 5 before)

Dr. Alalgia keeps

$$[\mu_{ET}, \mu_{SP}] = [-0.86, -0.42],$$

so that

$$\mu_D = -0.44.$$

Believes that σ will be greater under this design, because the outpatient PreVHI measurements are not as well controlled and PostVHI measurement is taken over fewer hours. Further, he feels that the SP (random biofeedback) patients will have greater variability than the ET patients.

$$[\sigma_{ET}, \sigma_{SP}] = [0.60, 0.80].$$

The correlation between VHchange_{ET} and VHchange_{SP} is suspected to be at least

$$\rho = 0.50$$

It is convenient to display these values in an "SD-Corr" matrix,

$$\tilde{\boldsymbol{\Sigma}} = \begin{bmatrix} 0.60 & 0.50 \\ 0.50 & 0.80 \end{bmatrix}$$

$$\sigma_D = [\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2]^{1/2}.$$

$$\sigma_D = [0.60^2 + 0.80^2 - 2(0.50)(0.60)(0.80)]^{1/2} = 0.72.$$

Need one-group t test to assess

$$H_0: \mu_D = 0$$

assuming D is Normal with

$$\mu_D = -0.44 \text{ and } \sigma_D = 0.72.$$

σ_D can be varied according to general equation,

$$\sigma_D(\rho, m) = m[\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2]^{1/2}$$

m: "SD multiplier" (SDmult)

Dr. Alalgia is prepared to study up to 25 pairs of patients

- $25 \times 2 \times 2 = 100$ bed-days.
- What if the correlation is greater, say $\rho = 0.60$?
- What if the standard deviations are, say, 20% larger (m = 1.20)?

UnifyPow makes it easy!

```
*-------------------------------------------;
* Matched-pairs t test                      ;
*-------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Alalgia's Matched-Pairs Design";
title3 "Traditional t-Based Tests";
datalines;
PairedMu -.86 -.42
SD .60 .80 . give exactly 2 SDs
corr .5 .6
SDMult 1 1.2 . may give several, default = 1.0
alpha .05 .01
TotalPairs 17 25 . Any word with 'total' is OK
%tables;
```


Dr. Alalgia's Matched-Pairs Design
Traditional t-Based Tests

Scenario: PairedMu -.86 -.42 & SD 0.6 0.8
Matched-pairs t test

		x SD (SD Multiplier)							
		1				1.2			
		Corr(Y1, Y2)				Corr(Y1, Y2)			
		0.5		0.6		0.5		0.6	
		Total Pairs		Total Pairs		Total Pairs		Total Pairs	
		17	25	17	25	17	25	17	25
		Power	Power	Power	Power	Power	Power	Power	Power
Alpha	Type								
0.05	2-tail t	.657	.833	.744	.900	.504	.684	.588	.771
	1-tail t	.777	.906	.846	.949	.641	.795	.718	.862
0.01	2-tail t	.375	.604	.469	.715	.245	.418	.312	.518
	1-tail t	.491	.709	.588	.805	.343	.529	.421	.629

```
*------------------------------------------- ;
* 2-group, Wilcoxon-Mann-Whitney (using mu, SD) ;
*------------------------------------------- ;
%include "&UnifyPow";
title2 "Dr. Alalgia's 2-Group Design";
title3 "Wilcoxon Tests Based on mu & SD, Assuming
Logistic Parent";
datalines;
mu -.86 -.42
SD .45 .57 .65
alpha .05 .01
weight 2 1
NTotal 15 21 33
Wilcoxon
parent logistic
methods all
%tables;
```



```
Testing location difference between 2 groups:
   <Parametric>    Ho: mu1 - mu2 = 0
  <NonParametric>  Ho: p1 = .50
      where p1 = Pr[Y{i,1} - Y{i',2} > 0 ] + .50*Pr[Y{i,1} - Y{i',2} = 0 ].
            Y{i,g} = Y for case i in group g.

Nonparametric Moments (if no ties possible)
-------------------------------------------
Let Y{i,g} be the outcome score for case i in group g.  (For the
PairedMu problem, Y{i,g} is a difference score.)  Then,
  p1 = Pr[Y{i,1} - Y{i',2} > Null]
  p2 = Pr[(Y{i,1} - Y{k,2} > Null) and (Y{i',1} - Y{k,2} > Null)]
  p3 = Pr[(Y{i,1} - Y{k,2} > Null) and (Y{i,1} - Y{k',2} > Null)]


UnifyPow will reverse ordering relations in order to force p1 > .5
Ties are handled by partitioning probabilities appropriately, e.g.,
  p1 = Pr[Y{i,1} - Y{i',2} > Null] + .50*Pr[Y{i,1} - Y{i',2} = Null]

The (default) Lehmann-Hettsmansperger method uses p1, p2, and p3,
whereas 'METHOD NOETHER' uses only p1.

'METHOD ARE' does not use the p-type moments, but rather uses
the asymptotic relative efficiences of the Wilcoxon versus the t-test.
This includes using ARE = .864, the theoretical minimum.
```

```
Parent Distributions
--------------------
Powers for the Wilcoxon will be approximated assuming Normal,
Logistic, and Laplace parent distributions, thus giving a range of
tail thicknesses (kurtoses) and asymptotic relative efficiencies (ARE):

   Parent     Kurtosis     ARE
  --------    --------    -----
  Normal        0.0       0.955
  Logistic      1.2       1.097
  Laplace       3.0       1.500
  <<Lower limit of ARE>>  0.864

-------------------------------------------------------------------------------
```



```
Dr. Alalgia's Two-Group Design
Wilcoxon Tests Based on mu & SD, Assuming Logistic Parent

Scenario: mu -.86 -.42
```

	Nonparametric Moments								
	p1			p2			p3		
	Std Dev			Std Dev			Std Dev		
	0.45	0.57	0.65	0.45	0.57	0.65	0.45	0.57	0.65
Parent									
Normal	.755	.707	.684	.627	.566	.537	.627	.566	.537
Logistic	.768	.719	.695	.645	.582	.551	.645	.582	.551
Laplace	.788	.741	.716	.675	.612	.580	.675	.612	.580
min ARE	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

```
Nonparametric Moments (if no ties possible)
-------------------------------------------
Let Y{i,g} be the outcome score for case i in group g.  (For the
PairedMu problem, Y{i,g} is a difference score.)  Then,
  p1 = Pr[Y{i,1} - Y{i',2} > Null]
  p2 = Pr[(Y{i,1} - Y{k,2} > Null) and (Y{i',1} - Y{k,2} > Null)]
  p3 = Pr[(Y{i,1} - Y{k,2} > Null) and (Y{i,1} - Y{k',2} > Null)]
```

Scenario: mu -.86 -.42
AND Alpha: 0.05

			Standard Deviation							
			0.45			0.57			0.65	
			Total N			Total N			Total N	
			15	21	33	15	21	33	15	21	33
			Power	Power	Power	Power	Power	Power	Power	Power	Power
Method	Type	Parent									
Wilcoxon Mann-Whitney [Lehmann (p1, p2, p3) aprx]	2-tail W	Normal	.291	.438	.677	.200	.295	.472	.164	.238	.380
		Logistic	.321	.481	.727	.221	.327	.521	.181	.264	.422
		Laplace	.375	.553	.800	.265	.392	.610	.219	.321	.508
	1-tail W	Normal	.432	.588	.800	.312	.426	.612	.263	.356	.516
		Logistic	.468	.631	.839	.340	.464	.659	.286	.388	.560
		Laplace	.528	.699	.891	.395	.535	.740	.335	.455	.645

This method uses all three moments and seems to give the best results most of the time. UnifyPow provides for two other methods ...

output continued ...

			Standard Deviation							
			0.45			0.57			0.65	
			Total N			Total N			Total N	
			15	21	33	15	21	33	15	21	33
			Power	Power	Power	Power	Power	Power	Power	Power	Power
Method	Type	Parent									
Wilcoxon Mann-Whitney [Noether (p1) aprx]	2-tail W	Normal	.346	.463	.655	.245	.329	.483	.202	.269	.397
		Logistic	.374	.499	.697	.268	.360	.526	.222	.296	.437
		Laplace	.422	.557	.758	.313	.420	.604	.262	.352	.515
	1-tail W	Normal	.468	.588	.763	.354	.449	.607	.302	.382	.522
		Logistic	.498	.623	.797	.381	.483	.648	.326	.413	.562
		Laplace	.547	.677	.845	.432	.545	.719	.374	.474	.638

The Noether method uses only the p_1 parameter.

output continued ...

Wilcoxon Mann-Whitney [aprx via ARE W vs. t]	2-tail W	Normal	.363	.498	.706	.246	.339	.506	.200	.273	.410
		Logistic	.415	.560	.769	.281	.385	.566	.227	.309	.462
		Laplace	.549	.706	.888	.377	.506	.708	.303	.410	.595
		min ARE	.328	.455	.660	.224	.308	.465	.183	.249	.376
	1-tail W	Normal	.501	.633	.811	.366	.469	.637	.307	.394	.542
		Logistic	.554	.690	.858	.406	.518	.692	.340	.435	.594
		Laplace	.680	.811	.940	.510	.637	.810	.428	.542	.715
		min ARE	.465	.593	.774	.339	.436	.598	.285	.366	.506
Ordinary t test	2-tail t	Normal	.380	.518	.727	.257	.353	.526	.209	.284	.427
	1-tail t	Normal	.519	.652	.828	.379	.485	.655	.318	.407	.559

The “min ARE” results give lower limits for power, assuming that the parent distributions are continuous (not necessarily symmetric) and have bounded densities that differ only in their means.


```
*-------------------------------------------;
* 2-group, Wilcoxon-Mann-Whitney (using p1)     ;
*-------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Alalgia's Two-Group Design  ";
title3 "Wilcoxon Test, Specifying p1 Directly,
Assuming Logistic Parent";
datalines;
2Wilcoxon .75        ← specify p1 = 0.75
SD 1.00 1.25 1.40    ← 25% & 40% increase in σ
weight 2 1
NTotal 15 21 33
methods Noether   ← Do Noether's approximation, too.
parent logistic   ← The p1 = 0.75 is relative to the logistic parent.
%tables;
```

Parent Distributions

Powers for the Wilcoxon will be approximated assuming Normal, Logistic, and Laplace parent distributions, thus giving a range of tail thicknesses (kurtoses) and asymptotic relative efficiencies (ARE):

Parent	Kurtosis	ARE
Normal	0.0	0.955
Logistic	1.2	1.097
Laplace	3.0	1.500
<<Lower limit of ARE>>		0.864

You specified the LOGISTIC distribution and p1 = 0.750.
This equates to psi = (mu2 - mu1 - NullValue)/SD = 0.8999.

Scenario: 2Wilcoxon .75

	Nonparametric Moments								
	p1			p2			p3		
	Relative Std Dev			Relative Std Dev			Relative Std Dev		
	1	1.25	1.4	1	1.25	1.4	1	1.25	1.4
Parent									
Normal	.738	.695	.675	.604	.550	.527	.604	.550	.527
Logistic	.750	.706	.686	.621	.565	.540	.621	.565	.540

Scenario: 2Wilcoxon .75
AND Alpha: 0.05

			Relative Std Dev								
			1			1.25			1.4		
			Total N			Total N			Total N		
			15	21	33	15	21	33	15	21	33
			Power	Power	Power	Power	Power	Power	Power	Power	Power
Method	Type	Parent									
Wilcoxon Mann-Whitney [Lehmann (p1, p2, p3) aprx]	2-tail W	Normal	.254	.381	.601	.180	.263	.421	.153	.219	.348
		Logistic	.281	.421	.653	.199	.292	.466	.168	.243	.388
		Laplace	.332	.492	.735	.239	.353	.555	.203	.297	.471
	1-tail W	Normal	.385	.526	.734	.285	.387	.559	.247	.332	.481
		Logistic	.418	.568	.778	.310	.422	.605	.268	.362	.524
		Laplace	.477	.640	.843	.362	.491	.689	.314	.426	.608
Wilcoxon Mann-Whitney [Noether	2-tail W	Normal	.307	.412	.594	.221	.296	.436	.188	.249	.367
		Logistic	.334	.447	.637	.242	.325	.478	.206	.274	.405

continued ...

Wilcoxon Mann-Whitney [Noether (p1) aprx]	2-tail W	Normal	.307	.412	.594	.221	.296	.436	.188	.249	.367
		Logistic	.334	.447	.637	.242	.325	.478	.206	.274	.405
		Laplace	.382	.508	.707	.285	.383	.556	.244	.328	.481
	1-tail W	Normal	.425	.537	.710	.325	.412	.561	.284	.358	.490
		Logistic	.455	.573	.747	.351	.445	.602	.306	.388	.529
		Laplace	.505	.631	.805	.400	.507	.676	.353	.448	.606
Ordinary t test	2-tail t	Normal	.331	.455	.656	.230	.315	.472	.193	.261	.393
	1-tail t	Normal	.465	.591	.770	.345	.442	.603	.297	.380	.523


```
*-------------------------------------------;
* Matched-pairs design, Wilcoxon signed-rank   ;
*                                             ;
*-------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Alalgia's Matched-Pairs Design";
title3 "One-Group Wilcoxon Tests, Assuming Normal
Parent";
datalines;
PairedMu -.86 -.42
SD .60 .80
corr .5 .6
SDMult 1 1.2
TotalPairs 17 25
Wilcoxon
%tables;
```

```
Testing difference of single pair of correlated measures:
   <Parametric>    Ho: mu(Y1 - Y2) = 0

Testing location of a single group:
  <Nonparametric>  Ho: p1 = .50
      where p1 = Pr[D{i} > 0 ] + .50*Pr[D{i} = 0 ].
            D{i} = Y1{i} - Y2{i}.

Nonparametric Moments (if no ties possible)
-------------------------------------------
Let Y{i} be the outcome score for case i.  (For the PairedMu
problem, Y{i} is a difference score.)  Then,
  p1 = Pr[Y{i} > Null]
  p2 = Pr[Y{i} + Y{i'} > 2*Null]
  p3 = (p2 + p1**2)/2
  p4 = Pr[(Y{i} + Y{i'} > 2*Null) and (Y{i} + Y{i''} > 2*Null)]


UnifyPow will reverse ordering relations in order to force p1 > .5
Ties are handled by partitioning probabilities appropriately, e.g.,
  p1 = Pr[Y{i} > Null] + .50*Pr[Y{i} = Null]
```


Scenario: PairedMu -.86 -.42 & SD 0.6 0.8

	x SD Multiplier											
	1						1.2					
	Corr(Y1, Y2)						Corr(Y1, Y2)					
	0.5			0.6			0.5			0.6		
	Nonparametric Moments			Nonparametric Moments			Nonparametric Moments			Nonparametric Moments		
	p1	p2	p4	p1	p2	p4	p1	p2	p4	p1	p2	p4
Parent												
Normal	.729	.806	.695	.750	.830	.730	.694	.764	.639	.713	.787	.670
Logistic	.752	.818	.713	.773	.841	.747	.716	.776	.657	.735	.799	.688
Laplace	.789	.834	.740	.808	.855	.771	.756	.796	.686	.775	.817	.716

Scenario: PairedMu -.86 -.42 & SD 0.6 0.8
AND Alpha 0.05

			x SD (SD Multiplier)					
			1			1.2		
			Corr(Y1, Y2)			Corr(Y1, Y2)		
			0.5	0.6		0.5	0.6	
			Total Pairs	Total Pairs		Total Pairs	Total Pairs	
			17	25	17	25	17	25	17	25
			Power	Power	Power	Power	Power	Power	Power	Power
Method	Type	Parent								
Wilcoxon Signed Rank [Lehmann (p1, p2, p3, p4) aprx]	2-tail W	Normal	.589	.806	.686	.889	.438	.637	.518	.733
		Logistic	.639	.847	.732	.917	.485	.690	.568	.781
		Laplace	.708	.896	.788	.946	.564	.770	.643	.845
	1-tail W	Normal	.739	.900	.823	.952	.590	.769	.672	.847
		Logistic	.782	.927	.858	.967	.637	.812	.718	.882
		Laplace	.836	.954	.896	.980	.711	.871	.782	.923


```
*-------------------------------------------------------;
* Wilcoxon-Mann-Whitney                                 ;
* 2 groups, ordered categorical outcome                 ;
*-------------------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Alalgia's 2-Group Design";
title3 "7-pt Likert Scale for Improvement in Headache";
/*
Much   Somewhat  A Little   No    A Little  Somewhat   Much
Worse   Worse     Worse   Change   Better    Better   Better
 -3       -2       -1       0        1         2        3
*/
datalines;
2Wilcoxon
.03 .04 .08 .10 .27 .28 .20
.10 .10 .15 .25 .20 .15 .05
weight 2 1
power .90 .95
alpha .05 .01
%tables;
```

```
Testing location difference between 2 groups:
  <NonParametric>  Ho: p1 = .50
      where p1 = Pr[Y{i,1} - Y{i',2} > 0 ] + .50*Pr[Y{i,1} - Y{i',2} = 0 ].
            Y{i,g} = Y for case i in group g.

Nonparametric Moments (if no ties possible)
-------------------------------------------
Let Y{i,g} be the outcome score for case i in group g.  (For the
PairedMu problem, Y{i,g} is a difference score.)  Then,
  p1 = Pr[Y{i,1} - Y{i',2} > Null]
  p2 = Pr[(Y{i,1} - Y{k,2} > Null) and (Y{i',1} - Y{k,2} > Null)]
  p3 = Pr[(Y{i,1} - Y{k,2} > Null) and (Y{i,1} - Y{k',2} > Null)]


--------------------------------------------------------------------------------

Scenario: 2Wilcoxon
--------------------------
|          | Nonparametric |
|          |    Moments    |
|          |---------------|
|          | p1 | p2 | p3  |
|----------+----+----+-----|
|Parent    |    |    |     |
|----------|    |    |     |
|Custom    |.706|.560|.566 |
--------------------------
```



```
Dr. Alalgia's 2-Group Design
7-pt Likert Scale for Improvement in Headache

Scenario: 2Wilcoxon,
{.03 .04 .08 .10 .27 .28 .20},
{.10 .10 .15 .25 .20 .15 .05}
-------------------------------------------------------------
|                            |            Alpha             |
|                            |------------------------------|
|                            |    0.05      |    0.01       |
|                            |--------------+---------------|
|                            |Minimum Power |Minimum Power  |
|                            |--------------+---------------|
|                            | .900 | .950  | .900 | .950   |
|                            |------+-------+------+--------|
|                            |Total |Total  |Total |Total   |
|                            |  N   |  N    |  N   |  N     |
|----------------------------+------+-------+------+--------|
|Method       |Type          |      |       |      |        |
|-------------+--------------|      |       |      |        |
|Wilcoxon-    |2-tail W      |   87 |  105  |  126 |  147   |
|Mann-Whitney |--------------+------+-------+------+--------|
|[Lehmann     |1-tail W      |      |       |      |        |
|(p1, p2, p3) |              |      |       |      |        |
|aprx]        |              |   72 |   87  |  108 |  129   |
-------------------------------------------------------------
```

```
*-----------------------------------------------;
* Wilcoxon signed-rank                          ;
* 1 group, interval-level categorical outcome  ;
*-----------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Alalgia's Matched-Pairs Design/ 7-pt Likert
scale";
title3 "Ordered Categorical Outcome, 1-Group Wilcoxon
Test";
/*
D = Difference of pair of 7-pt Likert scores, scored
    Y = {-3, -2, -1, 0, 1, 2, 3}.  Thus, D can range from
    -3 - 3 = -6 (worst outcome favoring ET therapy) to
    3 - (-3) = +6 (best outcome favoring ET therapy).
*/
datalines;
1Wilcoxon
.02 .03 .05 .05 .05 .06 .09 .14 .14 .14 .09 .08 .06
Limits -6 6            ← Default: 1, 2, 3, ..., NumCat
Null 0                 ← Null must be specified.
Ntotal 50 75 100
alpha .05
method Noether
%tables;
```



```
Testing location of single group:
  <Nonparametric>  Ho: p1 = .50
      where p1 = Pr[Y{i} > 0 ] + .50*Pr[Y{i} = 0 ].

Category values (interval scale):
  -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6


Nonparametric Moments (if no ties possible)
-------------------------------------------
Let Y{i} be the outcome score for case i.  (For the PairedMu
problem, Y{i} is a difference score.)  Then,
  p1 = Pr[Y{i} > Null]
  p2 = Pr[Y{i} + Y{i'} > 2*Null]
  p3 = (p2 + p1**2)/2
  p4 = Pr[(Y{i} + Y{i'} > 2*Null) and (Y{i} + Y{i''} > 2*Null)]

UnifyPow will reverse ordering relations in order to force p1 > .5
Ties are handled by partitioning probabilities appropriately, e.g.,
  p1 = Pr[Y{i,1} - Y{i',2} > Null] + .50*Pr[Y{i,1} - Y{i',2} = Null]


-------------------------
|          | Nonparametic |
|          |   Moments    |
|          |--------------|
|          | p1 | p2 | p4 |
|---------+----+----+----|
|Parent    |    |    |    |
|---------|    |    |    |
|Custom    |.695|.713|.574|
-------------------------
```

```
Dr. Alalgia's Matched-Pairs Design/ 7-pt Likert scale
Ordered Categorical Outcome, 1-Group Wilcoxon Test

Scenario: 1Wilcoxon,
{.02 .03 .05 .05 .05 .06 .09 .14 .14 .14
.09 .08 .06},
AND NULL: 0
```

		Alpha		
		0.05		
		Total N		
		50	75	100
		Power	Power	Power
Method	Type			
Wilcoxon Signed Rank [Lehmann (p1, p2, p3, p4) aprx]	2-tail W	.765	.916	.974
	1-tail W	.859	.959	.989
Wilcoxon Signed Rank [Noether (p1) aprx]	2-tail W	.743	.892	.958
	1-tail W	.833	.940	.980


```
*-------------------------------------------------------------;
* One-way ANOVA with complex contrasts                       ;
*-------------------------------------------------------------;
%include "&UnifyPow";
title2 "Dr. Mindy Bowdy: 4-Group Design";
title3 "Example 3 in O'Brien and Muller (1993, Section
 8.2.3)";
title4 "(1) Dominators, (2) Ordinaries, (3) Lon
 (4) Friendlies";
datalines;
mu .35 .50 .52 .60 . Higher = better immune function.
weight .20 .50 .10 .20
SD .16 .19
alpha .05
Ntotal 60 80 100
contrasts
"Ord vs Loners" 0  1  -1  0
"Almost Overall (2 DF)" 1  -.833  -.167  0
>                       0  -.833  -.167  1
;
```

```
/* Temporarily store .05 results */
data PowData1; set PowData;

%include "&UnifyPow";
datalines;
mu .35 .50 .52 .60
weight .20 .50 .10 .20
SD .16 .19
alpha .0167  .05
Ntotal 60 80 100
NoOverall
contrasts
"{Ord & Loners} vs Fr" 0  -.833  -.167  1
"Dom vs {Ord & Loners}" 1  -.833  -.167  0
"Fr vs Dom" -1  0  0  1
;

/* Concatenate the two sets of results and table */
data PowData; set PowData1 PowData;
%tables;
```


Dr. Mindy Bowdy: 4-Group Design
(1) Dominators, (2) Ordinaries, (3) Loners, (4) Friendlies

Scenario: mu .35 .50 .52 .60

			Standard Deviation					
			0.16			0.19		
			Total N			Total N		
			60	80	100	60	80	100
			Power	Power	Power	Power	Power	Power
Test	Alpha	Type						
Overall test	0.05	Regular F	.899	.970	.992	.763	.887	.951
Ord vs Loners	0.05	2-tail t	.059	.062	.065	.056	.058	.060
		1-tail t	.086	.093	.099	.079	.084	.090
Almost Overall (2 DF)	0.05	Regular F	.933	.982	.996	.821	.923	.969

{Ord & Loners} vs Fr	0.05	2-tail t	.429	.542	.639	.323	.413	.496
		1-tail t	.558	.666	.751	.445	.541	.622
	0.0167	2-tail t	.265	.367	.465	.183	.253	.326
		1-tail t	.363	.474	.573	.264	.348	.429
Dom vs {Ord & Loners}	0.05	2-tail t	.807	.906	.957	.663	.788	.872
		1-tail t	.884	.950	.979	.772	.870	.928
	0.0167	2-tail t	.658	.806	.897	.486	.637	.754
		1-tail t	.755	.873	.938	.597	.735	.832
Fr vs Dom	0.05	2-tail t	.964	.992	.998	.886	.957	.985
		1-tail t	.984	.997	.999	.938	.979	.994
	0.0167	2-tail t	.909	.974	.993	.772	.896	.956
		1-tail t	.948	.987	.997	.849	.938	.976


```
*-------------------------------------------------------------;
* 2-group (AB/BA) cross-over design                          ;
* via t tests on differences                                 ;
*-------------------------------------------------------------;
%include "&UnifyPow";
title2
"Asthma (Active) Treaments A vs. B, Outcome is FEV1";
datalines;
/#
This follows Patel (1983) example (2.1) in Jones B,
Kenward MG (1989, Design and Analysis of Cross-Over
Trials, Chapman and Hall). Two active drugs, patients have
acute bronchial asthma, design is ordinary AB/BA crossover,
outcome measure is FEV1.

Scenario:                                   Mean FEV1
                                          Drug A        Drug B
                                      |   ------------------------
                              AB      |   mu11 = 1.6     mu12 = 1.7
                    Order             |
                              BA      |   mu21 = 1.9     mu22 = 2.3

    SD(A) = .60, SD(B) = .75, SDMult = {1.0, 1.2},
    corr(A,B) = {.60, .70, 80}
#/
```

```
PairedMu 1.6 1.7
>        1.9 2.3
SD .60 .75
SDMult 1.0 1.2
Corr .6 .7 .8
sides 2
Ntotal 50
NoOverall
contrast
"Drug x Order" 1 -1
"Drug A vs. Drug B" 1 1
%tables;
```


Scenario: PairedMu 1.6 1.7; 1.9 2.3 & SD 0.6 0.75
Drug x Order

		x SD (SD Multiplier)					
		1			1.2		
		Corr(Y1, Y2)			Corr(Y1, Y2)		
		0.6	0.7	0.8	0.6	0.7	0.8
		Total Pairs	Total Pairs	Total Pairs	Total Pairs	Total Pairs	Total Pairs
		50	50	50	50	50	50
		Power	Power	Power	Power	Power	Power
Alpha	Type						
0.05	2-tail t	.390	.485	.637	.288	.360	.486

Drug A vs. Drug B

Alpha	Type						
0.05	2-tail t	.800	.893	.971	.646	.761	.894

```
*-------------------------------------------------------;
* Blackwelder's one-sided "equivalency" testing strategy ;
*-------------------------------------------------------;
%include "&UnifyPow";
datalines4;
/#
Abciximab plus a low-dose of heparin reduces complications
and increases 30-day survival rates in patients undergoing
high-risk coronary angioplasty or other kinds of
revascularization (EPILOG Invetigators, 1997 [N Engl J Med,
336:1689]).  But abciximab is expensive. Giving bivalirudin
(Bittl et. al., 1995 [N Engl J Med, 333:764-9]) albeit with
provisional ("bail-out") use of abciximab may give
equivalent efficacy and safety at lower costs.

Design. Bivalirudin plus provisional abciximab (B+provA)
vs. abciximab plus a low-dose of heparin (A+lowH).
Randomization ratio is 2:1, that is, 2/3 get B+provA.

Primary end-point. As per EPILOG trial: "death by any
cause, myocardial infarction, or urgent revascularization
within 30 days."

Statistical considerations.  This equivalency trial will be
handled using Blackwelder's method (1982 [Controlled
Clinical Trials, 3:345-53]), which tests
```



```
  Ho: pi(B+provA) - pi(A+lowH) GE delta  [B+provA
                                         appreciably worse]

  Ha: pi(B+provA) - pi(A+lowH) < delta   [B+provA not
                                         appreciably worse;
                                          could be better]

Delta > 0 sets the region of equivalency, which we take to
be within 25% of the 30-day event rate of 0.052 found in the
EPILOG trial.  Thus, delta = 0.012.  Our sample-size
analysis will assume that pi(B+provA) = pi(A+lowH) = 0.052.

We seek a power of 0.90 for an alpha = .05 test of this one-
tailed hypothesis.  We shall use either the traditional
Pearson ("approximate") unconditional test or its exact
counterpart (Suissa & Shuster, 1985 [J Royal Stat Soc A,
48:317-27]), whichever is more powerful.  [Even though
their true alpha-levels are practically identical in large
studies, the two methods will have different powers when
the group pi conjectures are unequal and the sample-sizes
are not balanced. See O'Brien and Muller (1993).]
#/
```

```
pi .052 .052
weight 2 1
null .012
Ntotal 8001 9999 12000 14001
tails 1
;;;;

/* store results to later build custom table */
data StorePow; set PowData;
%include "&UnifyPow";
datalines4;
/#
Sponsor actually believes that bivaliridin may
cut primary events by at least 5%. That is, they
think that
   pi(B+provA) = .95*.052 = .0494.
#/
```



```
pi .0494 .052
weight 2 1
null .012
Ntotal 8001 9999 12000 14001
tails 1
;;;;

/* merge results with first set */
data PowData; set StorePow PowData;

/* build custom table */
proc tabulate format = 6.3 order=data;
 class alpha effctitl NullValu NTotal ProbStmt;
 var power;
 table
  ProbStmt="True State" * NullValu="Delta" * effctitl =
"Method",
  alpha = "Alpha" * NTotal ="Total N" *
power="Power"*mean=" "
           /rtspace=45;
```

Blackwelder's one-sided equivalency testing strategy

			Alpha			
			0.05			
			Total N			
			8001	9999	12000	14001
			Power	Power	Power	Power
True State	Delta	Method				
pi .052 .052	0.012	Approximate Uncondit'l "chi^2*"	0.737	0.817	0.874	0.915
		Exact Uncondit'l**	0.737	0.817	0.874	0.915
pi .0494 .052	0.012	Approximate Uncondit'l "chi^2*"	0.880	0.934	0.965	0.981
		Exact Uncondit'l**	0.875	0.930	0.962	0.980


```
*-------------------------------------------------------------;
* McNemar's test of 2 correlated proportions                  ;
*-------------------------------------------------------------;
%include "&UnifyPow";
title2
"Thromboembolism (case/control) & The Pill ('outcome')";
title3 "Matched Pairs, Binary Outcome";

/*
Example follows study by Sartwell PE, et. al. (1969,
Thromoembolism and oral contraceptives: an epidemiologic
case-control study. Am J Epi, 90: 365-380.)

Cases: Women of childbearing age who have some kind of
thromoembolism.

Matched Controls: One per case--treated in the same
hospital, matched on age, number of prior pregnancies,
income level.
```

```
Scenario:                                   Matched Control
                                            Using The Pill?
                                          No              Yes
                                    |   ----------------------
 Thromboembolism Case      No       |   p00 = .54       p01 = .08
   Using The Pill?
                           Yes      |   p10 = .32       p11 = .06
*/

datalines;
McNemar .32 .08
Ntotal 50 to 150 by 25
%tables;
```



```
Testing Ho: pi12 - pi21 = 0 using McNemar's test.


Scenario: McNemar .32 .08
-----------------------------------------------------
|                            |         ALPHA          |
|                            |------------------------|
|                            |          0.05          |
|                            |------------------------|
|                            |         Pairs          |
|                            |------------------------|
|                            | 50 | 75 |100 |125 |150 |
|                            |----+----+----+----+----|
|                            |Pow-|Pow-|Pow-|Pow-|Pow-|
|                            | er | er | er | er | er |
|----------------------------+----+----+----+----+----|
|Method       |Type          |    |    |    |    |    |
|-------------+--------------|    |    |    |    |    |
|McNemar      |2-tailed      |.798|.931|.980|.995|.999|
|(virtually   |--------------+----+----+----+----+----|
|exact)       |1-tailed      |.839|.954|.988|.997|.999|
-----------------------------------------------------
```

```
*-------------------------------------------------------;
* Testing a single Pearson correlation                  ;
*-------------------------------------------------------;

%include "&UnifyPow";
title2 "Jean Netticks: ARX Enzyme Levels in Father and
Son";
title3 "Long been assumed that rho[ARX(dad), ARX(son)] <
.55.";
title4 "New genetic theory predicts a correlation of rho >
.70.";
/*
    Testing Ho: rho le 0.55 versus Ha: rho > 0.55
*/
datalines;
rho .70
null .55
Ntotal 50 75 100
tails 1
%tables;
```



```
Testing single correlation coefficient using
Fisher's r-to-Z:

   <Parametric>    Ho: Z(rho) =  Z(0.55)


Scenario: rho .70
-------------------------------------------
|                            |    ALPHA     |
|                            |--------------|
|                            |     0.05     |
|                            |--------------|
|                            |   Total N    |
|                            |--------------|
|                            | 50 | 75 |100 |
|                            |----+----+----|
|                            |Pow-|Pow-|Pow-|
|                            | er | er | er |
|----------------------------+----+----+----|
|Method      |Type           |    |    |    |
|------------+---------------|    |    |    |
|Fisher's    |1-tail Z       |    |    |    |
|r-to-Z test |               |    |    |    |
|of one rho  |               |.525|.680|.790|
-------------------------------------------
```

```
*-------------------------------------------------------;
* Comparing two Pearson correlations                    ;
*-------------------------------------------------------;

%include "&UnifyPow";
title2 "rho[ARX(dad), ARX(son)]";
title3 "Is there a difference between obese and normal
fathers?";
title4 "Obese: rho = .55    Normal: rho = .70";
datalines;
rho .55 .70
weight 1 3
Ntotal 400 to 1600 by 400
tails 2
%tables;
```


Testing correlations using Fisher's r-to-Z:

<Parametric> Ho: Z(rho1) - Z(rho2) = 0

Obese: rho = .55 Normal: rho = .70

Scenario: rho .55 .70

		ALPHA			
		0.05			
		Total N			
		400	800	1200	1600
		Pow-er	Pow-er	Pow-er	Pow-er
Method	Type				
Comparing two correlations (r-to-Z)	2-tail Z	.567	.858	.961	.990

```
*-------------------------------------------------------;
* Comparing nested (full vs. reduced) OLS regressions    ;
*-------------------------------------------------------;

%include "&UnifyPow";
title2 "Corey Latour: OLS Modeling to Predict Job
Satisfaction";
title3 "Adding a 4-level nominal predictor (3 dummy
variables).";
title4;
datalines;
R**2 .45 .50
NumParms 5 8
Ntotal 100 125 150
alpha .05 .01
%tables;
```



```
Testing Ho: Beta_6 = Beta_7 = Beta_8 = 0.


Scenario: R**2 .45 .50
```

		ALPHA					
		0.05			0.01		
		Total N			Total N		
		100	125	150	100	125	150
		Power	Power	Power	Power	Power	Power
Method	Type						
Comparing nested R**2 values	Regular F	.741	.843	.909	.506	.648	.762

```
*-------------------------------------------------------------------------;
* Testing a single beta in a k-variable OLS regression model             ;
*-------------------------------------------------------------------------;
title2 "Corey Latour: Predicting Salaries:";
title3 "Will the TeamValu scale signicantly improve the model?";
%include "%UnifyPow";
/*
Predicting Y = annual salary (example: 32.4 = $32,400 US).
X3 = TeamValu is a 0-4 visual analog (continuous) scale measuring the
employee's value to his/her teammates as determined by an external
evaluator. TeamValu's (unstandardized) beta coefficient, beta3, is the only
one under examination here. It corresponds to the increase in annual salary
(in $1000 US) associated with a 1-point increase in TeamValu given 2 other
predictors, X1 and X2, whatever they are. The R**2 from predicting TeamValu
from X1 and X2 is 0.30-0.40, e.g. TeamValu's tolerance is 0.60-0.70.
*/
datalines;
1betaOLS 1.5 2.0 . Some conjectures for beta3
SDX .6 .8  . Conjectures for SD(TeamValu)
Tolerance .6 .7 . Conjectures for Tol(TeamValu)
SD 3 3.5 . Conjectures for SD(error)
NumParms 4 .  (= 3 Xs + 1 for intercept)
Ntotal 50 75
tails 1
%tables;
;
```

Model:

$$\text{Salary} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3(\text{TeamValu}) + \varepsilon$$


```
/*
I didn't like the default table for this problem,
so I wrote a custom one.  This kind of thing works
for SAS/GRAPH nicely, too.
*/

proc tabulate format=4.3 order=data;
 class
  alpha BetaWt tolernce SDx testtype SD Ntotal;
 var power;
 table
  alpha='Alpha:',BetaWt='Beta for TeamValu'*
  tolernce='Tol(X)' *SDx='SD(X)' *
testtype='Type',
  SD='SD(Resid)' * Ntotal='Total N' *
       Power='Power'*mean=' '
/rtspace=35;
```

Alpha: 0.05

				SD(Resid)			
				3		3.5	
				Total N		Total N	
				50	75	50	75
				Power	Power	Power	Power
Beta Coefficient:	Tol(X)	SD(X)	Type				
1.5	0.6	0.6	1-tail t	.490	.636	.399	.525
		0.8	1-tail t	.696	.844	.581	.737
	0.7	0.6	1-tail t	.541	.694	.442	.579
		0.8	1-tail t	.754	.890	.638	.793
2	0.6	0.6	1-tail t	.696	.844	.581	.737
		0.8	1-tail t	.891	.971	.794	.918
	0.7	0.6	1-tail t	.754	.890	.638	.793
		0.8	1-tail t	.928	.985	.846	.949


```
*-------------------------------------------------------------------------;
* General linear models (intro.): traditional t-test the hard way        ;
*-------------------------------------------------------------------------;
title2 "Dr. Alalgia's Two-Group Design (revisited)";
title3 "Simple Example to Introduce EXEMPLARY SSH problem";
/*
   Set the groups, their conjectured means, and
   give sample sizes that reflect the cell weights.
   This creates an exemplary data set. For this problem,
   an exemplary data set is one having:
      sample means = "true" means, no error variance.

   The theory justifying this computational trick is described
   in O'Brien and Muller (1993, Section 8.3).

   The EXEMPLARY SSH problem type in UnifyPow replaces the
   PowSetUp module described in O'Brien and Muller.
*/

data;
input feedback $ VHchange  n_exemp  ;  datalines;
         ET       -.86        20
         SP       -.42        10
proc glm order=data;
  class feedback;  freq n_exemp;
   model VHchange = feedback;  ← regular data analysis on exemplary data

/*
```

Key part of output

Dependent Variable: VHCHANGE
Frequency: N_EXEMP

Source	DF	Type III SS	Mean Square	Value	Pr > F
FEEDBACK	**1**	**1.29066667**	1.29066667	9999.99	0.0001

```
/* Use UnifyPow to compute and table powers or sample sizes */

%include "&UnifyPow";  datalines;
/#
Nobody should use UnifyPow this way to do a traditional t test!
It is only done here to introduce the use of the "exemplary SSH"
method to compute and table power or sample sizes for complex general
linear models.
#/
exemplary SSH
NumParms 2
Nexemplary 30
sd .45 .57 .65
alpha .05 .01
NTotal 15 21 33
effects
"ET vs SP on VHchange"   1  1.29066667
%tables;
```

copied from GLM output

Scenario: exemplary SSH

			Standard Deviation								
			0.45			0.57			0.65		
			Total N			Total N			Total N		
			15	21	33	15	21	33	15	21	33
			Power	Power	Power	Power	Power	Power	Power	Power	Power
Test	Alpha	Type									
ET vs SP on VHchange	0.05	2-tail t	.380	.518	.727	.257	.353	.526	.209	.284	.427
		1-tail t	.519	.652	.828	.379	.485	.655	.318	.407	.559
	0.01	2-tail t	.156	.260	.472	.091	.147	.276	.068	.108	.201
		1-tail t	.235	.358	.581	.145	.219	.372	.112	.167	.283

Same results as the easy way!

```
*-------------------------------------------------------------------------;
* General linear models: one-way analysis of covariance, with contrasts ;
*-------------------------------------------------------------------------;
title2 "Dr. Mindy Bowdy: 4-Group Design + Covariate";
title3 "(D) Dominators, (R) Regulars {Ordinary or Loner}, (F) Friendlies";
/*
Three groups, one covariate (LESI), unequal slopes. LESI is the Life Events
Stress Index:  -2 = low stress, -1 = below average, 0 = average, 1 = above
average, 2 = high stress.

Order of exemplary data:
  D = Dominator
  R = Regulars = {Ordinary or Loner}
  F = Friendly
*/
data; input lysis0 DRFgrp $ LESI n;

/* create exemplary data */
if DRFgrp = 'D' then beta = -.03;
if DRFgrp = 'R' then beta = -.01;
if DRFgrp = 'F' then beta =  .00;
lysis = lysis0 + beta*LESI;

*  lysis at                      ;
*   LESI=0  DRF  LESI    n       ;   datalines;
    .3350    D    -2    02
    .3350    D    -1    03
    .3350    D     0    04
    .3350    D     1    05
    .3350    D     2    06
    .5033    R    -2    12
    .5033    R    -1    12
    .5033    R     0    12
```

Model for i-th case in group j

$$lysis_{ij} = \tau_j + \beta_j(LESI)_{ij} + \varepsilon_{ij}$$

(Different intercepts and slopes for the 3 groups.)


```
    .5033    R     1    12
    .5033    R     2    12
    .6000    F    -2    04
    .6000    F    -1    04
    .6000    F     0    04
    .6000    F     1    04
    .6000    F     2    04
;
```

Proc Print of exemplary data set:

OBS	LYSIS	DRFGRP	LESI	N_EXEMP
1	0.3950	D	-2	2
2	0.3650	D	-1	3
3	0.3350	D	0	4
4	0.3050	D	1	5
5	0.2750	D	2	6
6	0.5233	R	-2	12
7	0.5133	R	-1	12
8	0.5033	R	0	12
9	0.4933	R	1	12
10	0.4833	R	2	12
11	0.6000	F	-2	4
12	0.6000	F	-1	4
13	0.6000	F	0	4
14	0.6000	F	1	4
15	0.6000	F	2	4

```
/* "analyze" exemplary data to get SSHe values */
proc glm order=data; class DRFgrp; freq n;      ← regular data analysis on exemplary data
model lysis = DRFgrp DRFgrp*LESI/noint solution;
contrast 'DvRvF main | LESI=0' DRFgrp 1 -1 0, DRFgrp 0 1 -1;
contrast 'Ave LESI slope' DRFgrp*LESI .333 .333 .333;
contrast 'DvRvF x LESI' DRFgrp*LESI 1 -1 0, DRFgrp*LESI 0 1 -1;
contrast 'DvR | LESI=0' DRFgrp 1 -1 0;
contrast 'RvF | LESI=0' DRFgrp 0 1 -1;
contrast 'Slopes: DvR' DRFgrp*LESI 1 -1 0;
contrast 'Slopes: RvF' DRFgrp*LESI 0 1 -1;

/*    Key parts of the output:
--------------------------------------------------------------------------
                Number of observations in data set = 100

Contrast                DF   Contrast SS  Mean Square  F Value    Pr > F

DvRvF main | LESI=0      2     0.6722149    0.3361074 99999.99    0.0001
Ave LESI slope           1     0.0258462    0.0258462 99999.99    0.0001
DvRvF x LESI             2     0.0175385    0.0087692 99999.99    0.0001
DvR | LESI=0             1     0.3837566    0.3837566 99999.99    0.0001
RvF | LESI=0             1     0.1402634    0.1402634 99999.99    0.0001
Slopes: DvR              1     0.0108387    0.0108387 99999.99    0.0001
Slopes: RvF              1     0.0030000    0.0030000 99999.99    0.0001
--------------------------------------------------------------------------
```


This information is then used in UnifyPow as follows.

```
%include "&UnifyPow";
datalines;
Exemplary SSH
Nexemplary 100
alpha .05
SD .12 .15
Ntotal 200 300 500
NumParms 6
tails 2
effects
"DvRvF main | LESI=0"      2     0.6722149
"Ave LESI slope"           1     0.0258462          copied from GLM output
"DvRvF x LESI"             2     0.0175385
;
data PowData1; set PowData;  /* store these .05 results; */
/* now run .025 tests (Bonferroni adjustment) */
%include "&UnifyPow";
datalines;
Exemplary SSH
Nexemplary 100
alpha .025
SD .12 .15
Ntotal 200 300 500
NumParms 6
tails 2
effects
"DvR | LESI=0"             1     0.3837566
"RvF | LESI=0"             1     0.1402634          copied from GLM output
"Slopes: DvR"              1     0.0108387
"Slopes: RvF"              1     0.0030000
;
/* concantenate datasets and produce one table */
data PowData;  set PowData1 PowData; %tables;
```

Scenario: Exemplary SSH

			Std Dev					
			0.12			0.15		
			Total N			Total N		
			200	300	500	200	300	500
			Power	Power	Power	Power	Power	Power
Test	Alpha	Type						
DvRvF main | LESI=0	0.05	Regular F	.999	.999	.999	.999	.999	.999
Ave LESI slope	0.05	2-tail t	.470	.638	.848	.326	.456	.667
DvRvF x LESI	0.05	Regular F	.264	.380	.588	.182	.256	.404
DvR | LESI=0	0.025	2-tail t	.999	.999	.999	.999	.999	.999
RvF | LESI=0	0.025	2-tail t	.984	.999	.999	.897	.981	.999
Slopes: DvR	0.025	2-tail t	.154	.228	.380	.103	.148	.244

Remember this?

Design and Scenario

	Lived	Died	
Placebo	72%of n_1	28%of n_1	$n_1 = N/3$
DCA	79%of n_2	21%of n_2	$n_2 = 2N/3$

```
*------------------------------------------------------------------------;
* Logit analysis (intro.): comparing 2 indep. proportions the hard way  ;
*------------------------------------------------------------------------;
/*
   A trick to handle any situation involving likelihood ratio tests
   in categorical modeling.
*/
title2 "Lactic Acidosis in Children with Malaria (revisited)";
title3 "28% die untreated. What if DCA cuts this by 25%?";
/*
   Input the groups (here, Trtment), their conjectured mortality rates
   (PrDeath), and give sample sizes that reflect the cell weights
   (n_exemp). Then compute exemplary frequencies for each outcome
   category (ExempFrq).  This creates an exemplary data set
   (i.e., sample proportions = "true" proportions).

   The theory justifying this approach is given in O'Brien (1986,
   SUGI-11 Proceedings, 778-784) and summarized in Agresti (1990,
   Analysis of Categorical Data, Wiley, p 243).
*/
data;
input Trtment $ PrDeath  n_exemp;
if Trtment = "Placebo" then DCA = 0;
else DCA = 1;
died = 1; PopFreq = PrDeath*n_exemp; output;
died = 0; PopFreq = (1-PrDeath)*n_exemp; output;
/* Trtment PrDeath  n_exemp */  datalines;
   Placebo   .28      100
     DCA     .21      200
;
```


Proc Print of exemplary data set:

OBS	TRTMENT	DCA	DIED	PRDEATH	N_EXEMP	EXEMPFRQ
1	Placebo	0	1	0.28	100	28
2	Placebo	0	0	0.28	100	72
3	DCA	1	1	0.21	200	42
4	DCA	1	0	0.21	200	158

```
proc logistic;
 weight PopFreq; model died = DCA;
/*
```

← **regular data analysis on exemplary data**

```
Key part of output
------------------
  Criterion       Only       Covariates    Chi-Square for Covariates

  -2 LOG L        325.964       324.173       1.790 with 1 DF
*/

/* Use UnifyPow to compute and table powers or sample sizes */
%include "&UnifyPow";
datalines;
/#
Nobody should use UnifyPow this way to test two independent
proportions. It is only done here to introduce the use of the
"exemplary chi**2" method to compute and table powers or
sample sizes for complex logistic regression or log-linear models,
or similar methods using -2lnL chi**2 test statistics.
#/
exemplary chi**2
Nexemplary 300
alpha .01 .045
NTotal 750 999 1500
effects
"DCA vs. Placebo"  1  1.790 ◄—copied from LOGISTIC output
%tables;
```


Scenario: exemplary chi**2

		ALPHA					
		0.01			0.045		
		Total N			Total N		
		750	999	1500	750	999	1500
		Power	Power	Power	Power	Power	Power
Effect	Type						
DCA vs. Placebo	2-tail Z	.323	.447	.661	.544	.669	.838
	1-tail Z	.416	.546	.747	.663	.772	.903

Same results as the easy way!

```
*-------------------------------------------------------------------------;
* Logit analysis                                                          ;
*-------------------------------------------------------------------------;
data babies;
title2 "Get Exemplary -2LL values for Study of Lyon's Nonstress Test";
title3 "See O'Brien (1986, SUGI-11 Proceedings, 778-784)";
Title4
"Summarized in Agresti (1990, Analysis of Categorical Data, Wiley, p 243).";

/*
"PrNRcare" stands for Probability that "Non-Routine" cardiac care is
given at birth.  The question is whether a new diagnostic test
("Lyons") provides additional value in predicting NRcare beyond that
povided by using only the Standard test.

      Outcome at birth
        "NonRoutn"         0 = no non-routine care was given
                           1 = non-routine care was given

      Predictors before birth
        "Standard" test    1 = worrisome result
                           2 = reassuring result

        "Lyons" test       1 = very worrisome result
                           2 = somewhat worrisome result
                           3 = reassuring result
*/
```



```
      input Standard Lyons CellProb PrNRcare;

      * expected number of routine outcomes;
      NonRoutn=0;
        ExempFrq = (1-PrNRcare)*CellProb*1000;
        output;
      * expected number of nonroutine outcomes;
      NonRoutn=1;
        ExempFrq = PrNRcare*CellProb*1000;
        output;

      *Standard  Lyons  CellProb  PrNRcare  ;     datalines;
           1       1      .04       .40
           1       2      .08       .32
           1       3      .04       .27
           2       1      .02       .30
           2       2      .18       .22
           2       3      .64       .15
      data babies; set babies;
      *build dummy vars for PROC LOGIST;
      XStd = (Standard=2) - (Standard=1);
      XLyons1 = (Lyons=1) - (Lyons=3);
      XLyons2 = (Lyons=2) - (Lyons=3);
      XSL1 = XStd*XLyons1;
      XSL2 = XStd*XLyons2;
      proc print; var Standard Lyons CellProb PrNRCare NonRoutn ExempFrq;
```

Proc Print of exemplary data set:

OBS	STANDARD	LYONS	CELLPROB	PRNRCARE	NONROUTN	EXEMPFRQ
1	1	1	0.04	0.40	0	24.0
2	1	1	0.04	0.40	1	16.0
3	1	2	0.08	0.32	0	54.4
4	1	2	0.08	0.32	1	25.6
5	1	3	0.04	0.27	0	29.2
6	1	3	0.04	0.27	1	10.8
7	2	1	0.02	0.30	0	14.0
8	2	1	0.02	0.30	1	6.0
9	2	2	0.18	0.22	0	140.4
10	2	2	0.18	0.22	1	39.6
11	2	3	0.64	0.15	0	544.0
12	2	3	0.64	0.15	1	96.0

```
/* "analyze" exemplary data to get -2lnL values */
proc logistic;  weight ExempFrq;
  model NonRoutn = Xstd;
proc logistic;  weight ExempFrq;
  model NonRoutn = Xstd XLyons1;
proc logistic;  weight ExempFrq;
  model NonRoutn = Xstd XLyons1 XLyons2;
proc logistic; weight ExempFrq;
  model NonRoutn = Xstd XSL1 XSL2;
proc logistic;  weight ExempFrq;
  model NonRoutn = Xstd XLyons1 XLyons2 XSL1 XSL2;
/*
  These analyses provide the -2LnL values for UnifyPow's
  EFFECTS specifications. */
```


Key parts of proc logistic output

Criterion	Intercept Only	Intercept and Covariates	
-2 LOG L	983.943	964.437	← Model: Standard main only
-2 LOG L	983.943	956.319	← Model: Standard main + Lyons(lin)
-2 LOG L	983.943	956.277	← Model: Standard main + Lyons(2 df) main
-2 LOG L	983.943	960.910	← Model: Standard main + Standard*Lyons(2 df) interaction
-2 LOG L	983.943	955.990	← Model: Standard + Lyons(2 df) main + Standard*Lyons(2 df) interaction

```
%include "&UnifyPow";
datalines;
Exemplary chi**2
Nexemplary 1000
alpha .01 .05 .
Ntotal  400 600 1000 .
effects
"Standard Only [LR]"  1  983.943  964.437
"Lyons(lin) Given Standand [LR]"  1  964.437  956.319
"Lyons Given Standard [LR]"  2  964.437  956.277
"Lyons Main [LR, Sat Mdl]"  2  960.910  955.990
;
/* save results for likelihood ratio tests */
data PowData2;  set PowData;
```



```
%include "&UnifyPow";
datalines;
pi .40 .32 .27 .30 .22 .15
weight .04 .08 .04 .02 .18 .64
alpha .01 .05
Ntotal  400 600 1000
NoOverall
contrasts
"Lyons Main [Wald, Sat Mdl]" 1 -1  0  1 -1  0
>                            0  1 -1  0  1 -1
"Lyons(lin) SubMain [Wald, Sat Mdl]" 1 0 -1 1 0 -1
;

/* Concatenate LRT and Wald-type results and make one table
*/
data PowData; set PowData2 PowData;
ProbStmt =
 "pi = {.40 .32 .27 .30 .22 .15}, CellProbs = {.04 .08 .04
.02 .18 .64}";
%tables;
```

		ALPHA					
		0.01			0.05		
		Total N			Total N		
		400	600	1000	400	600	1000
		Power	Power	Power	Power	Power	Power
Effect	Type						
Standard Only [LR]	2-tail Z	.586	.801	.967	.798	.928	.993
	1-tail Z	.680	.863	.982	.875	.962	.997
Lyons(lin) Given Standand [LR]	2-tail Z	.220	.356	.608	.437	.598	.813
	1-tail Z	.300	.452	.699	.562	.713	.886
Lyons Given Standard [LR]	Chi-square	.156	.266	.498	.347	.495	.727
Lyons Main [LR, Sat Mdl]	Chi-square	.083	.137	.268	.222	.317	.497
Lyons Main [Wald, Sat Mdl]	Chi-square	.084	.138	.270	.224	.319	.500
Lyons(lin) SubMain [Wald, Sat Mdl]	2-tail Z	.109	.177	.327	.270	.378	.567
	1-tail Z	.163	.249	.421	.382	.501	.686

LR test (better)

same hypothesis

Wald test ("cheaper")

Want more on the exemplary data method?

See lecture notes for Shieh-O'Brien invited paper given at the 1998 Joint Statistical Meetings in Dallas:

"A Simpler Method to Compute Power for
Likelihood Ratio Tests in
Generalized Linear Models"

These notes are on the UnifyPow website as an Acrobat (*.pdf) file.

Thanks for coming!

Go Tribe!