

## Paper 208-30

# Estimation of Type III Error and Power for Directional Two-Tailed Tests Using PROC POWER

Cam-Loi Huynh, Ph.D., Department of Psychology, University of Manitoba,  
Winnipeg, MB, Canada

### ABSTRACT

By imposing directional decisions on traditional two-tailed tests, you will overestimate power, underestimate sample size, and ignore the risk of Type III error. You can avoid these problems by using the *directional two-tailed* test. This paper demonstrates how to use PROC POWER to estimate the prospective power and Type III error for one-sample, two-sample and paired t tests for means, and binomial tests for proportions, using the *directional two-tailed* testing procedure.

Examples are given with SAS<sup>®</sup> codes accompanied by selected results. Power and Type III error for the *directional two-tailed* sign test with unknown direction is also discussed and illustrated with a SAS macro written by the author.

### INTRODUCTION

In developing a test for deciding whether one of the  $k$  populations means is larger (smaller) than the rest, under the null hypothesis that all populations are continuous and identical, Mosteller (1948) identified three types of error for a statistical decision: (i) Type I error = the probability of "rejecting the null hypothesis when it is true", (ii) Type II error = the probability of "failing to reject the null hypothesis when it is false", and (iii) Type III error = the probability of "correctly rejecting the null hypothesis for the wrong reason" (i.e., the risk that both the (rejected) null and (accepted) alternative hypotheses are false. See Mosteller, 1948, p. 63). Type III error exists when the false null hypothesis is rejected but the sample having the largest (smallest) sample mean does not actually contain the largest (smallest) population mean. In other words, in testing the null hypothesis that the means discrepancy is zero, Type III error represents the risk of correctly rejecting the null hypothesis in supporting the "wrong direction" of the mean difference (most often because the sample with the larger (smaller) sample mean does not come from the population having the larger (smaller) population mean). For a  $t$  test of any means difference among  $k$  population means, two immediate consequences of Mosteller's Type III error can be recognized: (i) It is encountered only when the null hypothesis is false but the predicted direction based on the alternative hypothesis does not represent the sign (i.e., direction) of the true population means difference, and (ii) If Type III error is possible, or encountered, then the conventional definition of statistical power for the test should be modified. For example, the conventional definition of power as "the probability that we reject the null hypothesis, say, because the rightmost population yields a sample with too many large observations" (Mosteller, 1949, p.61) can be revised to be "the probability of both correct rejection and correct of rightmost population, when it exists." (Mosteller, 1948, p. 63).

Mosteller's concept of Type III error renders important implications for statistical decisions. First, in acknowledging the possibility of Type III error, we should prefer a (*directional*) two-tailed test over the one-tailed alternative. Given the same data, the one-tailed test tends to yield higher power than the two-tailed test if the assumed direction is correct. However, if the supported one-tailed alternative is false, its power is misleading because, instead of power, it may represent the probability of Type III error. Secondly, it is important to realize that Type III error could exist in all statistical tests (e.g., test of means, variances, correlations, proportions, etc.) as well as tests with any set of  $k$  samples ( $k \geq 1$ ). Finally, a procedure to estimate the probability of Type III error for two-tailed tests is surely desired. In responding to this need, Kaiser (1960) developed a *directional two-tailed* test. In this paper, we illustrate how to calculate Type III error rate by means of PROC POWER, an experimental procedure available in SAS/STAT<sup>®</sup>, Version 9.1 (SAS Institute Inc. 2004). In this procedure, the prospective power for Kaiser's test can be determined. In this study, applications for several t tests and an extension for the case of unknown direction in binomial tests are discussed.

## TYPE III ERROR AND POWER FOR t TESTS

### GENERAL THEORY

In the conventional two-tailed test, one is restricted to the choice of two hypotheses: Null ( $H_0$ ):  $\delta = 0$  versus Alternative ( $H_1$ ):  $\delta \neq 0$ , where  $\delta$  represents differences in one of the following parameters: means ( $\delta = \mu_X - \mu_Y$ ) of variables X and Y, proportions ( $\delta = \pi_A - \pi_B$ ), or correlation coefficients ( $\delta = \rho_A - \rho_B$ ) of groups A and B. For a given level of Type I error ( $\alpha$ ), the conventional test power is defined as

$$\Psi_C(\alpha) = 1 - \beta,$$

where  $\beta$  denotes Type II error (which is defined as “the probability of failing to reject a false null hypothesis”). Kaiser (1960) proposed a test that involves three (mutually exclusive) true states of nature ( $H_1$ ,  $H_0$  and  $H_2$ ) for which the three corresponding hypotheses are specified as:

$$\begin{aligned} \text{Left-tailed alternative } ((H_1): \delta < 0, \\ \text{Null } (H_0): \delta = 0, \\ \text{Right-tailed alternative } (H_2): \delta > 0, \end{aligned}$$

where  $\delta$  represents one of the differences in population parameters mentioned above. This approach is often called the “*directional two-tailed* test.” For fuller discussions of Kaiser’s test, see Kaiser (1960), Harris (1997), Leventhal and Huynh (1996), and Leventhal (1999). Its test power is defined as,

$$\Psi_K(\alpha) = 1 - \beta - \gamma,$$

where  $\gamma$  represents the Type III error. The maximum value of Type III error for *directional two-tailed* test is equal to  $\alpha / 2$  (Kaiser, 1960, p.164). This is also equal to the “power” of one-tailed test evaluated at  $\alpha / 2$  but the direction specified in the alternative hypothesis is “wrong.” Note that the conventional and Kaiserian test powers share the same values of  $\alpha$  and  $\beta$ . However, the latter is smaller by the presence of Type III error ( $\gamma$ ). The conservative nature of the Kaiserian approach is the strength of this method since it may result in larger required sample size for prospective sample size determination. Zumbo and Hubley (1998) and Knapp (1999) maintained that prospective power (an estimate of test power for planning purposes) is the more meaningful concept than the retrospective, or observed, power (computed after data collection).

Kaiser (1960) and Shaffer (1972) showed that the *directional two-tailed* test at a predetermined  $\alpha$  level is equivalent to the testing of two simultaneous one-tailed hypotheses, each evaluated at  $\alpha / 2$ . Therefore, by conducting two one-tailed tests, each at the size  $\alpha / 2$ ,  $\Psi_K$  is equal to the power of the one-tailed test for which the null hypothesis is rejected, and  $\gamma$  is equal to the power of the other test for which the null hypothesis is retained. In using SAS /STAT software, there are two equivalent ways to calculate Type III error and power for the *directional two-tailed* test

### METHOD 1: APPLYING BOTH PROC TTEST AND PROC POWER

Type III error, test power and planned sample size for prospective design of *directional two-tailed* tests can be obtained as follows:

(a) Estimating prospective power and Type III error for t tests for a given sample size (n):

- (i) Choosing an alpha level ( $\alpha$ ), say  $\alpha = 0.01$ .
- (ii) Evaluating the right-tailed test at  $\alpha^* = \alpha / 2$ , i.e.,  $\alpha^* = 0.005$ , by using PROC

TTEST (for statistical significance evaluation) and PROC POWER (for sample size determination).

(iii) Repeat (ii) for the left-tailed test.

(iv) In (ii) and (iii), if in the t test of PROC TTEST, it is found that "Prob>|T|"  $< \alpha / 2$  then  $\Psi_K$  = the power value obtained in PROC POWER. On the other hand, if "Prob>|T|"  $\geq \alpha / 2$ , then  $\gamma$  = Type III error = the power value obtained in PROC POWER.

(b) Determining the required sample size and Type III error for a desired level of power:

(i) Choosing an alpha level ( $\alpha$ ), say  $\alpha = 0.01$  and a desired value of power ( $\Psi_K$ ), say  $\Psi_K = .90$ .

(ii) Evaluating the *directional two-tailed* test at  $\alpha^* = \alpha$ , i.e.,  $\alpha = 0.01$  by using PROC POWER with a specified value for power to be  $\Psi_K$  (= .90, say). The resulting value of Ntotal is the required sample size (n) for the Kaiserian test.

(iii) Evaluating the one-tailed test at  $\alpha^* = \alpha / 2$  (= 0.005, say) by using PROC POWER with a specified value for Ntotal = n = the required sample size above. Then,  $\Psi_C$  = the resulting value of power. Finally, Type III error =  $\gamma = \Psi_C - \Psi_K$ .

#### METHOD 2: APPLYING ONLY PROC POWER

PROC POWER yields the conventional power for one-tailed and (non-directional) two-tailed tests (See Chapter 57, SAS/STAT User's Guide, Version 9.1, SAS Institute Inc. 2004). The default for TEST =, DIST =, NULLDIFF =, ALPHA =, and SIDES = options specify a two-sided t test of group means difference equal to 0, assuming a normal distribution with a significance level of  $\alpha = 0.05$ . Thus, in applying PROC POWER, you may need to enter appropriate values for these options. Because PROC POWER can be used to evaluate one-tailed test in the same direction as the given effect by specifying SIDES= 1 (i.e., it performs a left-tailed test for negative effect, and a right-tailed test for positive effect), a much simpler procedure than Method 1 above can be used to estimate Type III error and test power as follows:

(a) Estimating prospective power and Type III error for t tests for a given sample size (n):

(i) Choosing an alpha level ( $\alpha$ ), say  $\alpha = 0.01$ .

(ii) Obtaining the Kaiserian power ( $\Psi_K$ ) by using PROC POWER with option SIDES=1 at  $\alpha^* = \alpha / 2$

(iii) Obtaining the conventional power ( $\Psi_C$ ) by using PROC POWER with option SIDES=2 at  $\alpha$ .

(iv) Then, calculating Type III error =  $\gamma = \Psi_C - \Psi_K$ .

(b) Determining the required sample size and Type III error for a desired level of power (same as under Method 1 above).

#### SAS<sup>®</sup> PROGRAM CODES FOR TYPE III AND POWER OF T TESTS

The SAS codes and selected results for some examples involving one-sample, two-sample (with equal and unequal variances), paired t tests and binomial test for proportions using PROC POWER to estimate test power, Type III error and required sample size are given below. In each case, to save space and computing time, both one-tailed and two-tailed hypotheses are invoked in each run (by specifying "Sides=1 2 Alpha = .025 .05"), a macro program (%MACRO GAMMA) is used to sort the one-tailed from two-tailed is used (given in the APPENDIX).

Note that when SIDES=1 is used for one-tailed tests, this condition implies that the population direction is assumed "known" and is the same as that of the specified effect. Therefore, Type III error would be very

small. For example, consider the one-sample t test. Suppose in a study of leadership, the mean and standard deviation of the measure "Inner Guidance" was expected to be at least 6 and at most 40, respectively. The researcher would like to estimate the statistical power if a sample of 150 U.S. military officers would be drawn. Upon applying PROC POWER, from the results reported for one-sample t test below, the power for  $N = 150$  would be quite low (about 0.42 for both conventional and Kaiserian approaches), with Type III error about 0.01% for  $\alpha = .05$ . It appears that the sample size of 356 or more is required for a decent power of .80.

### 1. One-sample t test.

```
* Determining Type 3 error and power for directional tests;
Title 'One-sample t Test';
Data One;
  PROC POWER Plotonly;
  Ods Output Plotcontent=Plotdata;
  Onesamplemeans Sides=1 2 Alpha = .025 .05
  Mean = 6
  Stddev = 40
  Ntotal = 150
  Power = .;
  plot X=n min=140 max=500;
  Run;
Data Plotdata;
Set Plotdata;
*****;
  %Gamma(Plotdata);
*****;
RUN;
Proc Delete Data=Plotdata;
Run;
```

Selected results for one-sample t test:

Obs	NTotal	PowerC	PowerK	Type3
1	140	0.42185	0.42176	.000098581
2	158	0.46578	0.46572	.000063041
3	176	0.50765	0.50761	.000040895
4	194	0.54735	0.54732	.000026852
5	212	0.58479	0.58477	.000017816
6	230	0.61996	0.61995	.000011929
7	248	0.65287	0.65286	.000008052
8	266	0.68353	0.68353	.000005474
9	284	0.71202	0.71202	.000003745
10	302	0.73841	0.73840	.000002577
11	320	0.76277	0.76277	.000001783
12	338	0.78520	0.78520	.000001239
13	356	0.80581	0.80581	.000000865

### 2. Two-sample t test with equal variances.

```
Title 'Two-sample t Test with Equal Variances';
PROC POWER Plotonly;
Ods Output Plotcontent=Plotdata;
```

```

Twosamplemeans Test=diff Sides=1 2 Alpha =.025 .05
Groupmeans = 40|45
Stddev = 16
Ntotal = 60
Power = .;
plot X=n min=50 max= 500;
Run;
Data Plotdata;
Set Plotdata;
*****;
      %Gamma(Plotdata);
*****;
RUN;
Proc Delete Data=Plotdata;
Run;

```

Selected results for two-sample t test with equal variances:

Obs	NTotal	PowerC	PowerK	Type3
1	50	0.19138	0.19021	.001170645
2	72	0.25765	0.25711	.000541829
3	94	0.32272	0.32245	.000270861
4	116	0.38556	0.38542	.000142378
5	138	0.44543	0.44535	.000077589
6	160	0.50181	0.50177	.000043461
7	182	0.55439	0.55437	.000024883
8	204	0.60301	0.60300	.000014504
9	226	0.64763	0.64762	.000008582
10	248	0.68831	0.68830	.000005144
11	270	0.72518	0.72517	.000003117
12	292	0.75841	0.75841	.000001908
13	314	0.78822	0.78822	.000001178
14	336	0.81485	0.81485	.000000733

### 3. Two-sample t test with unequal variances.

```

Title 'Two-sample t Test with Unequal Varainces';
PROC POWER Plotonly;
Ods Output Plotcontent=Plotdata;
Twosamplemeans Test=diff_satt Sides=1 2 Alpha =.025 .05
Nfractional
Groupmeans = 40|45
Groupstddevs = 14|18
Ntotal = 60
Groupweights = (1 1.5)
Power = .;
plot X=n min=50 max=500;
Run;
Data Plotdata;
Set Plotdata;
*****;
      %Gamma(Plotdata);
*****;

```

```

RUN;
Proc Delete Data=Plotdata;
Run;

```

Selected results for two-sample t test with unequal variances:

Obs	NTotal	PowerC	PowerK	Type3
1	50	0.18982	0.18865	.001177782
2	73	0.25894	0.25841	.000530647
3	96	0.32660	0.32635	.000259131
4	119	0.39175	0.39162	.000133357
5	142	0.45359	0.45352	.000071258
6	165	0.51160	0.51156	.000039180
7	188	0.56544	0.56542	.000022037
8	211	0.61498	0.61497	.000012626
9	234	0.66021	0.66020	.000007348
10	257	0.70121	0.70121	.000004333
11	280	0.73816	0.73816	.000002585
12	303	0.77126	0.77126	.000001557
13	326	0.80077	0.80077	.000000947

4. Paired t test of dependent means:

```

Title 'Paired t Test';
Data One;
  PROC POWER Plotonly;
  Ods Output Plotcontent=Plotdata;
  Pairedmeans test=diff Sides=1 2 Alpha = .025 .05
pairedmeans = 83.2294 | 90.4941
Corr = 0.351
pairedstddevs = (5.0167 8.4751)
npairs = 7
Power = .;
plot X=n min=5 max=20 Step=1;
Run;
Data Plotdata;
Set Plotdata;
Ntotal = npairs;
*****;
  %Gamma(Plotdata);
*****;
RUN;

```

Selected results for paired t test with dependent means :

Obs	Ntotal	PowerC	PowerK	Type3
1	5	0.33123	0.33107	.000163312
2	6	0.42025	0.42018	.000069486
3	7	0.50333	0.50330	.000030868
4	8	0.57873	0.57871	.000014185

Obs	Ntotal	PowerC	PowerK	Type3
5	9	0.64575	0.64574	.000006700
6	10	0.70437	0.70437	.000003236
7	11	0.75497	0.75496	.000001593
8	12	0.79815	0.79815	.000000797

## TYPE III ERROR AND POWER FOR BINOMIAL TESTS

### POPULATION DIRECTION IS ASSUMED KNOWN

The same procedures in Methods 1 and 2 above can be applied to estimate planned sample size, power and Type III error for testing hypotheses about proportions upon replacing the codes for t tests by those for binomial tests. As an example, in a double-blind listening study, suppose  $n$  listeners were asked to identify differences between sound components in a (well-designed and well-conducted) experiment. Let  $\pi$  and  $p$  represent the population and observed proportions of correct identification, respectively. Suppose a *directional two-tailed* sign test was used, given  $n = 10$ ,  $\pi = .50$  and  $p = .30$ . The hypotheses can be specified as:

Left-tailed alternative ( $H_1$ ):  $\pi < .50$   
 Null ( $H_0$ ):  $\pi = .50$ ,  
 Right-tailed alternative ( $H_2$ ):  $\pi > .50$ .

In a single-listener study,  $\pi$  can be interpreted as the proportion of correct choices the listener is capable of making over an infinite number of trials, from which the trials in one's study constitute a random sample. In a test with multiple listeners, each receiving multiple trials, then there are three possibilities for data analysis: each listener can constitute a separate experiment, each trial can constitute a separate experiment, or one can combine listeners and trials to produce a single experiment.

First, the Kaiserian power and Type III error can be calculated by using the option SIDES=1 for one-tailed tests as in t tests above (i.e., the population direction is assumed to be the same as that specified in the alternative hypothesis). The SAS code and relevant PROC POWER results are given below.

```
Data Rate;
*Two-tailed test at Alpha = 0.05;
PROC POWER Plotonly;
Ods Output Plotcontent=Plotdat2;
Onesamplefreq test=Exact Sides=2 Alpha = .05
Nullproportion = 0.50
Proportion = 0.30
Ntotal = 10
Power = .;
plot X=n min=8 max=20 Step=2;
Run;
Data Plotdat2;
Set Plotdat2;
Count+1;
PowerC = Power;
Data Two;
Merge Plotdat1 Plotdat2;
by Count;
Type3 = PowerC - PowerK;
Ods rft file="Power.rtf";
Proc Print;
Var Sides Alpha Ntotal PowerC PowerK Type3;
Run;
Ods rft close;
```

Obs	Alpha	NTotal	PowerC	PowerK	Type3
1	0.00781	8	0.05771	0.05765	.000065610
2	0.02148	10	0.14945	0.14931	.000143686
3	0.03857	12	0.25302	0.25282	.000206376
4	0.01294	14	0.16086	0.16084	.000025307
5	0.02127	16	0.24589	0.24586	.000033601
6	0.03088	18	0.33269	0.33265	.000039496
7	0.04139	20	0.41641	0.41637	.000042940

Note that as N increases, the power values (under PowerC = conventional definition and PowerK= Kaiserian definition) and type III demonstrate a saw-tooth pattern (See also Example 56.2, pp. 3248-3255 in SAS/STAT User Guide, SAS Institute Inc., 2004).

#### POPULATION DIRECTION IS UNKNOWN

In all the examples so far, because SIDES=1 is used for one-tailed tests in computing the Kaiserian power, Type III error would be quite small. However, this is not the case when the population direction is assumed unknown as it will be illustrated for the following binomial tests.

A macro (%POWBIN) is written by the author to calculate sample size (N), critical values (cv), probabilities of type I error ( $\alpha$ ), Type III error ( $\gamma$ ) and power for the binomial tests that do not use the option SIDES=1 for one-tailed tests. The critical values (cv) are the sample sizes needed for statistical significance (i.e., for rejecting a null hypothesis in favor of an alternative one). For example, cv = (R, M) = (9, 1) means that nine or more correct choices (R) are needed to reject  $H_0$  in favor of  $H_2$ , and one or fewer correct choices (M) are needed to reject  $H_0$  in favor of  $H_1$ .

The macro %POWBIN compute all critical values for N=8(2)20. For example, selected results for N = 10 and N=20 are listed below.

#### N=10

Obs	N	R	M	Alpha	Power6	Gamma6	Power7	Gamma7	Power8	Gamma8	Power9	Gamma9
9	10	10	0	.0020	.0060	.0001	.0282	.0000	.1074	.0000	.3487	.0000
10	10	9	1	.0215	.0464	.0017	.1493*	.0001*	.3758	.0000	.7361	.0000
11	10	8	2	.1094	.1673	.0123	.3828	.0016	.6778	.0001	.9298	.0000
12	10	7	3	.3438	.3823	.0548	.6496	.0106	.8791	.0009	.9872	.0000
13	10	6	4	.7539	.6331	.1662	.8497	.0473	.9672	.0064	.9984	.0001

#### N=20

Obs	N	R	M	Alpha	Power6	Gamma6	Power7	Gamma7	Power8	Gamma8	Power9	Gamma9
49	20	18	2	.0004	.0036	.0000	.0355	.0000	.2061	.0000	.6769	.0000
50	20	17	3	.0026	.0160	.0000	.1071	.0000	.4114	.0000	.8670	.0000
51	20	16	4	.0118	.0510	.0003	.2375	.0000	.6296	.0000	.9568	.0000
52	20	15	5	.0414	.1256	.0016	.4164*	.0000*	.8042	.0000	.9887	.0000
53	20	14	6	.1153	.2500	.0065	.6080	.0003	.9133	.0000	.9976	.0000
54	20	13	7	.2632	.4159	.0210	.7723	.0013	.9679	.0000	.9996	.0000



For these tables, the following notations are used: R and M = critical values needed to reject the null hypothesis ( $R + M = N$ ), Power6 = Kaiserian power for  $\pi = (0.6, 0.4)$ , Power7 = Kaiserian power for  $\pi = (0.7, 0.3)$ , Power8 = Kaiserian power for  $\pi = (0.8, 0.2)$ , Power9 = Kaiserian power for  $\pi = (0.9, 0.1)$ . Similarly, Gamma6 = Type III error for  $\pi = (0.6, 0.4)$ , etc. The "ALPHA" column provides the actual  $\alpha$  (the actual type I probability) produced by the critical values (R, M). The remaining columns list power and  $\gamma$  for a number of  $\pi$  values. The Kaiserian power and type III results obtained with SIDES=1 option above are reproduced and identified by the asterisk (\*). Note that for a given configuration of power, (e.g. Power7,  $\pi = (0.7, 0.3)$ , and  $N = 10$ ), the values of power and type III estimates increase steadily as the critical value R decreases (or M increases).

## DISCUSSION AND CONCLUSIONS

Suppose an experimenter conducts a right-tailed test at the .05 level and finds that the sample mean is not large enough to be statistically significant. Worst, the sample mean is in fact negative. She now finds herself sorely tempted to conduct a left-tailed test to determine whether the sample mean is small enough to be statistically significant. Suppose she yields to the temptation and carries out the left-tailed test. The biggest problem with this procedure is that both the t statistic and the nominal alpha level will become invalid. Meeks and D'Agostino (1983) argued that in a strict two-stage inference procedure (in which the hypothesis of interest is tested first and if it results in a rejection decision then proceed with estimating the relevant parameter) then the conventional confidence limits and associated test statistics are no longer valid. Instead, she should have conducted one *directional two-tailed* test. Although this is equivalent to simultaneously performing two one-tailed tests, each at  $\alpha / 2$ , both the form of the test statistic and the size of the nominal alpha level are protected. Although this test has just begun to stir up some interests in the educational and behavioral literature (Hand, McCarter & Hand, 1985; Harris, 1997; Hopkins, 1973), it is hoped that it will become more popular due to its straightforward procedure and practical implications.

## REFERENCES

- Hand, J., McCarter, R. E., & Hand, M. R. (1985), "The procedures and justification of a two-tailed directional test of significance," *Psychological Report*, 56, 495-498.
- Harris, R. J. (1973), "Significance tests have their place," *Psychological Science*, 8, 8-11.
- Hopkins, B. (1973), "Educational research and Type III errors," *The Journal of Experimental Education*, 41, 31-32.
- Kaiser, H. F. (1960), "Directional statistical decision," *Psychological Review*, 67, 160-167.
- Knapp, I. R. (1999), "Letter to the Editor," *Nursing Research*, 48, 6.
- Leventhal, L. & Huynh, C-L. (1996), "Directional decisions for two-tailed tests: Power, error rates, and sample size," *Psychological Methods*, 1(3), 278-292.
- Leventhal, L. (1999), "Updating the debate on one- versus two-tailed test with the *directional two-tailed* test.," *Psychological Reports*, 84, 707-718.
- Mosteller, F. (1948), "A k-sample slippage test for an extreme population," *Annals of Mathematical Statistics*, 19, 58-65.
- Meeks, S.L., & D'Agostino, R.B. (1983), "A note on the use of confidence limits following rejection of a null hypothesis," *The American Statistician*, 57(2), 134-136.
- SAS Institute Inc. (2004). *SAS/STAT User's Guide: Version 9.1*. Cary NC: SAS Institute Inc.
- Shaffer, J. P. (1972), "Directional statistical hypotheses and comparisons among means," *Psychological Bulletin*, 77, 195-197.
- Zumbo, B. D. & Hubley, A. M. (1998), "A note on misconceptions concerning prospective and retrospective power," *The Statistician*, 47, Part 2, 385-388.

## ACKNOWLEDGEMENTS

This research is partially supported by a grant from the Social Science and Humanity Research Council of Canada (SSHRC Acct No. 410-2004-1213)

## CONTACT INFORMATION

The SAS codes for %POWBIN are available from the author upon request. Your comments and questions are valued and encouraged. Contact the author at:

Cam-Loi Huynh  
 Department of Psychology, University of Manitoba  
 Winnipeg, Manitoba, Canada R3T 2N2  
 Phone: (204) 474-7400  
 Fax: (204) 474-7599  
 E-mail: [huynh@cc.umanitoba.ca](mailto:huynh@cc.umanitoba.ca).  
 Web: <http://home.cc.umanitoba.ca/~huynh/>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. In the USA and other countries, ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

## APPENDIX

### THE MACRO PROGRAM %GAMMA FOR CALCULATING POWER AND TYPE III ERROR

```
%MACRO GAMMA(Plotdata);
  Data Side1;
  Set Plotdata;
  If Sides=2 then delete;
  Count+1;
  If Sides=1 and Alpha > .03 then delete;
  PowerK = Power;
Data Side2;
  Set Plotdata;
  If Sides=1 then delete;
  If Sides=2 and Alpha < .04 then delete;
  PowerC = Power;
Data Side2;
  Set Side2;
  Count+1;
Data Two;
  Merge Side1 Side2;
  by Count;
  Type3 = PowerC - PowerK;
  If Ttype = 4 then Ntotal = Npairs;

Proc Print;
  Var Sides Ntotal PowerC PowerK Type3;
Proc Delete Data=Plotdata One;
Run;
%Mend;
```