Novel Use of SAS® Software in Industrial and Biomedical Consulting, Research and Teaching

Timothy E. O'Brien
Loyola University Chicago

ABSTRACT

Basic courses in applied biostatistics, statistical methods, experimental design and applied regression methods often focus on hypothesis testing and estimation for linear and nonlinear models by providing students and decision-makers the methodology (i.e., test statistics and confidence intervals) to reach conclusions by narrowly focusing on individual t-tests and global F-tests. These courses leave students and managers with an overly simplistic view of how informed statistical decisions are made in practice in the real world. This paper focuses on the new pedagogical ideas of exposing students to underlying likelihood methods and treating these specific (t- and F-) tests as special cases embedded in this larger structure. Key to this better decision-making process is powerful statistical software: our focus is here on the use of the NLMIXED and IML procedures available in SAS® software to provide the means to make some of these important decisions. This approach enables practitioners to pose and examine more meaningful queries. For example, the techniques discussed here allow practitioners to focus on the estimation of important model parameters in the presence of serially correlated errors rather than on the detection of the specific time-series error structure (and treating the model function and parameters as secondary). Numerous additional practical examples of the applicability of likelihood methods are provided and discussed; specifically, the provided illustrations include novel approaches useful in statistical modelling, drug synergy, relative potency and optimal experimental design.

Keywords: decision-making; likelihood; modelling; optimal design; statistical education.

INTRODUCTION

Decision-makers and real-world practitioners often underscore the rift between statistical techniques taught in introductory methods courses and those used in practice on the job. Whereas in the past, great emphasis was placed on adapting a real-world problem to classical problems, nowadays – using sophisticated statistical techniques and SAS® software – statistical consultants can adapt the statistical methods to better address the actual situation. The following illustrations have been chosen to demonstrate these situations.

SOME MOTIVATING EXAMPLES

The following illustrations provide relevant examples of situations in which practitioners can use SAS® software to fit meaningful models and answer relevant research questions. Clearly, without the needed statistical software, these results remain merely theoretical; these statistical techniques empower researchers to test their hypotheses and implement the scientific method.

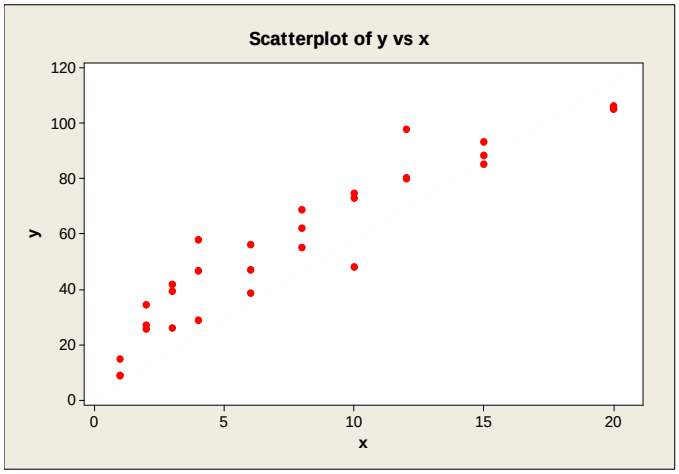
Example 1. The Lack of Fit (LOF) Test. By way of highlighting the usefulness of SAS/IML, consider the usual LOF test, which checks for departures from the assumed base model function. Typically, we check

for departures from a straight-line fit. This test is useful when the chosen n-point design includes replicates of k support points, and uses the usual Full and Reduced F test statistic,

$$F_{(dfF - dfR),dfF} = \frac{[(SSE_R - SSE_F) / (df_F - df_R)]}{[SSE_F / df_F]} \quad (1)$$

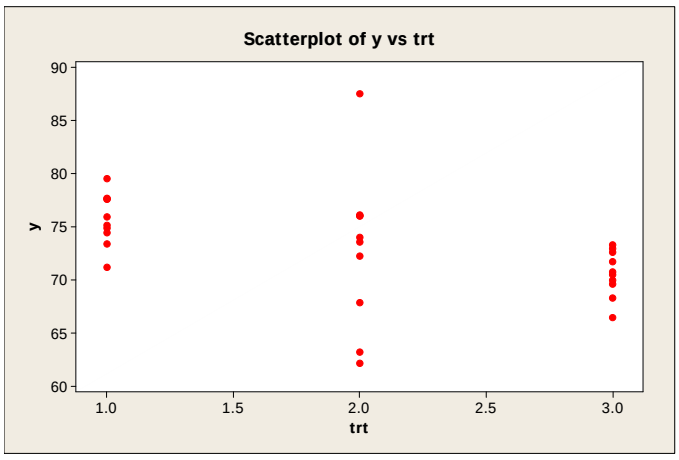
The above LOF test – including its advantages and limitations – is discussed in O'Brien & Berg (2009) and O'Brien, Chooprateep & Funk (2009); in the latter article, it is pointed out that it lacks power to detect lack-of-fit departures for intermediate models.

To illustrate, consider the data plotted in the following graph where the sample size is n = 30 and the design comprises three replicates of k = 10 design support points.



In spite of the above nonlinear pattern, it is surprising that the usual LOF test here fails to reject the assumption that a line fits these data ($F_{8,20} = 1.9965$, $p = 0.1003$). This seeming contradiction is resolved in the next section with the use of the IML procedure in SAS®. ■

Example 2. Heteroskedastic One-Way ANOVA. It is simple to handle the homoskedastic one-way ANOVA problem, but things can get somewhat more challenging for the novice when variances are believed to differ from group to group. To illustrate, consider the data plotted in the graph below.

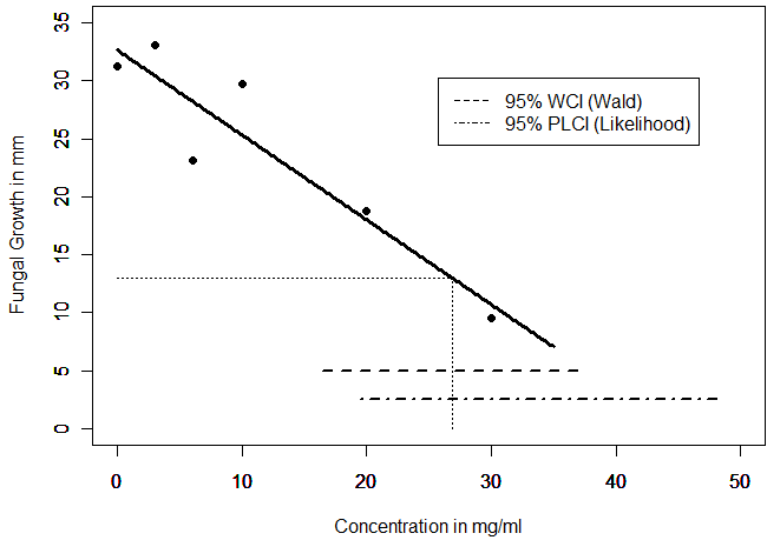


These data, taken from O'Brien and Berg (2009), are yield data for three treatments. It is reasonable here to assume normality of the respective yield measurements. It is also clear here that the variances for the three groups are not the same, but if this fact is ignored and the three treatment means are compared using the usual one-way ANOVA test, the claim of equal means (H_0: $\mu_1 = \mu_2 = \mu_3$) is retained ($F_{2,27} = 3.1038$, $p = 0.0612$). The unequal-variance means test can be accomplished using the chi-square counterpart of the Full-and-Reduced F test in Equation (1) as outlined in Agresti (2007:12). This test is easily implemented in SAS/IML as demonstrated below. ■

Example 3. Calibration in Linear Regression. In this example, adapted from Samuels & Witmer (2003:538), the concentration of laetisaric acid (independent variable, in μg/mLi) is related to fungal growth (dependent variable, in mm) using simple linear regression. The data are plotted in the following graph. Since we want the value of x (denoted γ_{13}) such that $E(Y) = y_k = 13$, the complication here is that instead of writing the linear model in the usual manner, we write this line as

$$\eta(x,\boldsymbol{\theta}) = E(Y) = y_k + \beta(x - \gamma_k) = 13 + \beta(x - \gamma_{13}) \qquad (2)$$

The parameters to be estimated here are β and γ_{13} so the parameter vector is $\boldsymbol{\theta} = \begin{pmatrix} \beta \\ \gamma_{13} \end{pmatrix}$.



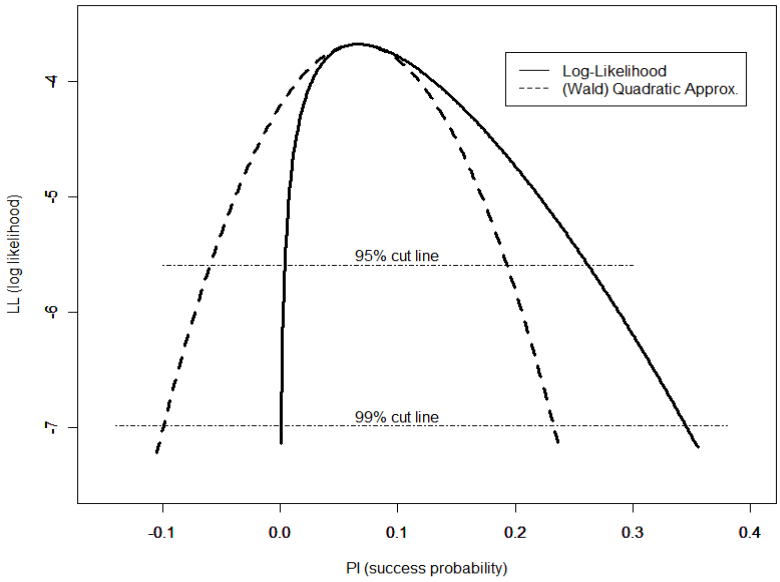
Note that the reparameterization given in Equation (2) is a nonlinear model ("nonlinear" in the parameters) since at least one of the partial derivatives with respect to the parameters involves model parameters; in fact this is true for both parameters since $\frac{\partial \eta}{\partial \beta} = x - \gamma_{13}$ and $\frac{\partial \eta}{\partial \gamma_{13}} = -\beta$.

This model is easily fit using the NLIN procedure in SAS, but problems associated with the Wald confidence intervals (provided in the NLIN output) have been underscored in Donaldson and Schnabel (1987), Haines *et al* (2004), and elsewhere. In the next section, we provide the means to use PROC IML to obtain the more reliable profile likelihood confidence intervals. The 95% Wald (WCI) and Profile

Likelihood (PLCI) intervals for the parameter of interest (γ_{13}) are plotted above in the above figure; the WCI is symmetric whereas for this illustration the PLCI is skewed to the right. ■

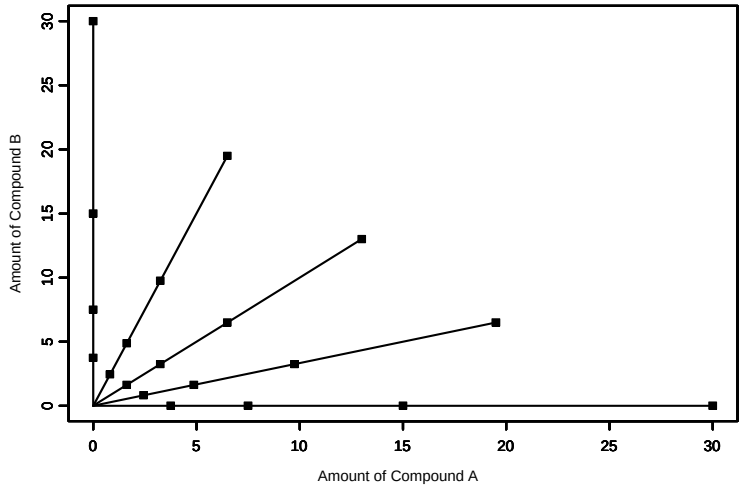
Example 4. Confidence Intervals for a Single Binomial Proportion. In the setting where fifteen coin tosses results in just one Head (i.e., one success out of 15 so n = 15 and y = 1), the usual 95% Wald confidence interval for the success probability (π), $p \pm z_{0.025}SE$, here extends from –0.0596 to 0.1929. This interval is clearly nonsensical since this success probability must be non-negative. This Wald interval is based on a quadratic approximation (dashed curve plotted below) to the log-likelihood expression (solid curve plotted below),

$$LL(\pi) = y\log(\pi) + (n-y)\log(1-\pi) = \log(\pi) + 14\log(1-\pi) \qquad (3)$$



Obtaining more reliable confidence intervals – for example, based on this log-likelihood – given in the next section, can be easily obtained using PROC IML. These intervals are obtained by finding the intersection of the above (log-likelihood or quadratic approximation) curves with the respective "cut lines", also plotted above. ■

Example 5. Bioassay and the Assessment of Drug Synergy. Data provided in Giltinan *et al* (1988) resulted from an experiment designed to investigate how two insecticides (A and B) may act in combination. The data, plotted below, correspond to the number of dead insects corresponding to a fixed number of insects tested (around 30) for 20 combinations of the insecticides. The chosen design is a so-called 2+3[4] ray design: the two one-at-a-time rays plus three interior rays (with slopes of 3, 1 and 1/3), and with 4 design points per ray. Of interest here is whether the insecticides interact to produce enhanced performance (termed "synergy"), a reduction in performance ("antagonism") or neither ("independent action").



For these data, we fit the so-called Finney$_4$ model. For this model, similar compounds A and B in respective amounts x_1 and x_2 are related to the binomial response Y by first calculating the effective dose,

$$z = x_1 + \theta_4 x_2 + \theta_5 \sqrt{\theta_4 x_1 x_2} \tag{4}$$

In this expression, θ_4 is the relative potency parameter and θ_5 is the coefficient of synergy. If $\theta_5 < 0$, compounds A and B exhibit antagonism; if $\theta_5 > 0$, synergy is indicated; and if $\theta_5 = 0$, then compounds A and B behave independently. The binomial response variable and effective dose in Equation (4) are related using a dose-response model function such as the 2-parameter log-logistic (LL2) function

$$\eta(\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{1 + \left(z/\theta_2\right)^{\theta_3}} \tag{5}$$

In the next section, we use the NLMIXED procedure to fit this generalized nonlinear model. ■

Example 6. Obtaining and Verifying Optimal Designs. For a given (Normal) linear or nonlinear model function $\eta(\mathbf{x},\boldsymbol{\theta})$, an experiment is often undertaken in which a chosen design is used to efficiently estimate the p model parameters $\boldsymbol{\theta}$. An n-point design (or probability measure), denoted by ξ, is written

$$\xi = \begin{Bmatrix} x_1, x_2, \ldots, x_n \\ \omega_1, \omega_2, \ldots, \omega_n \end{Bmatrix} \tag{6}$$

Here the ω_k (k = 1,2 … n) are non-negative weights that sum to one, and the x_k (which could be vectors) belong to the relevant design space. Note that the design points (x_k) are not necessarily distinct). For the chosen model function $\eta(x,\boldsymbol{\theta})$, the Jacobian matrix is $\mathbf{V} = \partial\boldsymbol{\eta}/\partial\boldsymbol{\theta}$ (of dimension n × p) and the p × p Fisher information matrix is

$$\mathbf{M}(\xi,\boldsymbol{\theta}) = \mathbf{V}^{\mathrm{T}}\boldsymbol{\Omega}\mathbf{V} \tag{7}$$

where $\boldsymbol{\Omega} = \text{diag}\{\omega_1, \omega_2, \ldots \omega_n\}$. The first-order (and asymptotic) variance of the least-squares estimator $\hat{\vartheta}$ is proportional to $\mathbf{M}^{-1}$, so designs are often chosen to minimize some convex function of $\mathbf{M}^{-1}$. For example, designs which minimize its determinant are called D-optimal.

The (first-order) variance of the predicted response at X = x is given by

$$d(x,\xi,\theta) = \frac{\partial \eta(x,\theta)}{\partial \theta^T} \mathbf{M}^{-1}(\xi) \frac{\partial \eta(x,\theta)}{\partial \theta^T} \quad (8)$$

Designs that minimize (over ξ) the maximum (over x) of $d(x,\xi,\theta)$ are called G-optimal. The General Equivalence Theorem demonstrates that D- and G-optimal designs are equivalent, and that the variance function evaluated using the D-/G-optimal design does not exceed the line (or hyper-plane) y = p (i.e., the number of model parameters) – but that it will exceed this line/plane for all other designs. A corollary establishes that the maximum of the variance function (i.e., p) is achieved for the D-/G-optimal design at the support points of this design.

These concepts are discussed more fully in O'Brien and Funk (2003). In the next section, we show how to use SAS/IML to obtain and verify optimal designs. ■

Clearly, one of the limitations encountered by non-statistical practitioners is the unavailability of sophisticated software to handle situations and datasets for which statistical methods beyond very basic analysis is needed. Thankfully, the IML and NLMIXED procedures in SAS® can easily handle these situations, and it is incumbent upon us to educate students and practitioners as to the proper analysis and to demonstrate use of the relevant software. As such, we discuss with our students several of the analyses and results given in the next section.

RESOLUTION USING SAS PROC IML AND PROC NLMIXED

In this section, we reconsider the above illustrations and show how key SAS procedures, modules and functions can be used to obtain meaningful solutions. Relevant output is given in the Appendix; corresponding SAS programs may be obtained from the author.

Example 1 continued. The Lack of Fit (LOF) Test. SAS Output 1 given in the Appendix shows that the usual LOF test retains the assumed linear fit for these data ($F_{8,20} = 1.9965$, $p = 0.1003$). On the other hand, given the above plot, one might feel that a model function which might better fit these data is the quadratic model function

$$\eta(x,\theta) = \beta_0 + \beta_1 x + \beta_2 x^2 \quad (9)$$

Since the test of H_0: $\beta_2 = 0$ is rejected here ($F_{1,27} = 7.6061$, $p = 0.0103$), this quadratic model is deemed appropriate for these data and thus the assumed line does indeed show significant lack of fit. This apparent contradiction points out that the usual LOF test sometimes lacks power to detect departures from the assumed model (i.e., the line) since it compares the line with the highest order polynomial (a ninth order polynomial in this case). Thus, it misses the intermediate (quadratic) model.

As pointed out in O'Brien, Chooprateep & Funk (2009), we write the full model here as

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon = [\mathbf{X}_1 \mid \mathbf{X}_2]\beta + \varepsilon = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \varepsilon \quad (10)$$

In this expression, $\mathbf{X}$ is of dimension $n \times p$, $\mathbf{X}_1$ is $n \times (p\text{-}q)$, and $\mathbf{X}_2$ is $n \times q$. $\mathbf{X}_1$ here corresponds to the assumed linear model. Let $\mathbf{P}_1$ denote the projection matrix associated with $\mathbf{X}_1$, and we note that it is important to understand the nature of the matrix

$$\mathbf{R}_2 = (\mathbf{I} - \mathbf{P}_1)\mathbf{X}_2 \quad (11)$$

This matrix targets is the subspace of $C(\mathbf{X}_2)$ that is orthogonal to $C(\mathbf{X}_1)$. If we let $\mathbf{P}_2$ denote the projection matrix associated with $\mathbf{R}_2$, then it is straightforward to show that the numerator of the test statistic in Equation (1) can be written $SSE_R - SSE_F = \mathbf{y}^T\mathbf{P}_2\mathbf{y} = \|\mathbf{P}_2\mathbf{y}\|^2$; see Seber and Lee (2003:100). Using SAS/IML, it is seen that this squared norm is equal to 1254.53 for these data. Then using SAS/IML, we next sequentially decompose $\mathbf{R}_2$ one column at a time in an orthogonal manner, obtaining the corresponding projection matrix and the squared norm. This task is indeed akin to finding sequential sums of squares. These eight values here are 621.01, 51.39, 121.13, 258.98, 12.69, 55.20, 69.86, 64.27, and orthogonality is confirmed since these eight values sum to $\|\mathbf{P}_2\mathbf{y}\|^2 = 1254.53$. Since the largest of these values (621.01) – corresponding to the quadratic component – is so large, it is not surprising that the usual LOF test lacks power in this instance. ■

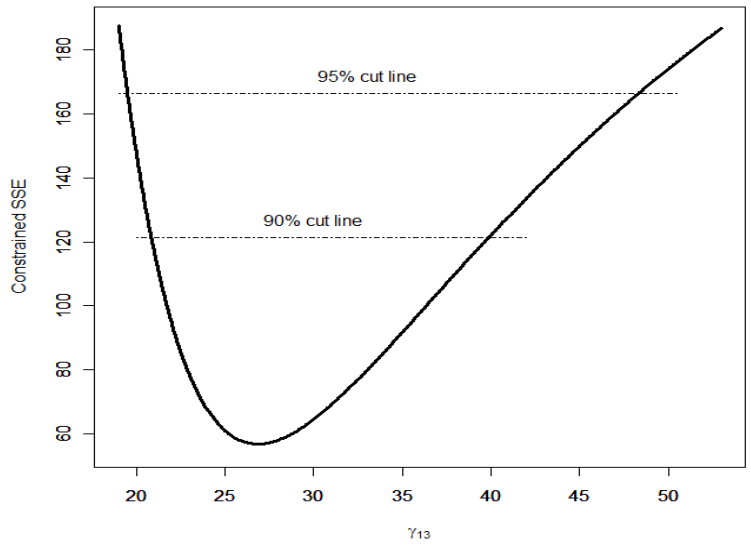
Example 2 continued. Heteroskedastic One-Way ANOVA. As pointed out above, the extension of the Full-and-Reduced F test to other than homoskedastic Normal settings is the chi-square test which is based on twice the change in the log-likelihood. Obtaining the log-likelihood values is relatively simple using SAS/IML, and the exercise of writing such programs is an excellent learning tool for statistics students.

For these data, the full model contains six parameters (a mean and variance for each of the three treatments) and the reduced model contains only four parameters (three variances and a common mean). The SAS program given in Output 2 in the Appendix gives these parameter estimates as well as the values $-2LL_{FULL} = 99.9954$ and $-2LL_{REDUCED} = 116.7110$, from which we obtain the test statistic $\chi_2^2 = 116.7110 - 99.9954 = 16.7156$ ($p = 0.0002$). Thus, in contrast with the above incorrect result, this result indicates that at least one of the treatment means differs from the others. ■

Example 3 continued. Calibration in Linear Regression. As highlighted in Donaldson and Schnabel (1987), O'Brien and Wang (1996) and Haines *et al* (2004), Wald confidence intervals (WCIs) for nonlinear models can be problematic in the sense that even though the nominal coverage of the WCI might be 95%, the actual coverage may be quite different from 95%. Likelihood or profile likelihood confidence intervals (obtained by profiling out nuisance parameters), on the other hand, are typically quite reliable. Also, Clarke (1987) and Haines *et al* (2004) show that important connections exist between departures between Wald and likelihood intervals, on one hand, and marginal curvature measures, on the other.

As shown in O'Brien and Wang (1996), marginal curvature measures are easily obtain using SAS/IML, and these are given in Output 3b of the Appendix. The PROC NLIN 95% Wald confidence interval for γ_{13}, (16.45 , 37.29), is given in Output 3a. This interval is also repeated in Output 3b. The profile likelihood interval, on the other hand, is (19.50 , 48.36). The latter interval is obtained from the following profile likelihood plot, where the horizontal cut lines correspond to 90% (lowest horizontal line) and 95%

(middle cut line). The Wald interval corresponds to a profile likelihood plot with a quadratic shape since Wald intervals are symmetric. For these data, the profile likelihood plot is far from quadratic in shape, and this explains the big difference between the Wald and likelihood confidence intervals. Output 3b also shows that the overlap (ratio of intersection to union) of the 95% Wald to the Likelihood interval is only 55.75%; this output also shows that the overlap of the 95% marginal curvature-adjusted interval (MCCI) proposed in Clarke (1987), (18.53 , 44.88), is 85.10%; this indicates that the MCCI provides a better approximation to the likelihood interval.



One reason the profile likelihood curve increases quickly on the left and gently on the right is that five of the original six design points are to the left of $\hat{\gamma}_{13}$ = 26.87 and only one point is to the right. This demonstrates that a better design (at least in terms of estimating γ_{13}) would be to choose a more equitable number of design points on either side of $\hat{\gamma}_{13}$ = 26.87. ■

Example 4 continued. Confidence Intervals for a Single Binomial Proportion. The IML results given in Output 4 are used to obtain the likelihood-based 95% CI for π, which, in this case extends from 0.0039 to 0.2621. Due to the likelihood results given in Agresti (2007:12), these values are obtained by solving the expression

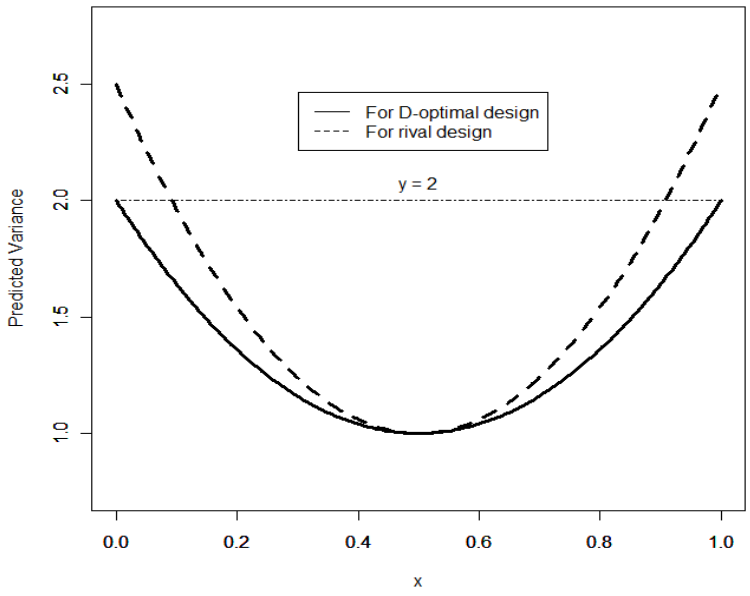
$$LL(\pi) = LL(p) - \tfrac{1}{2}\chi_{1,\alpha}^{2} \qquad (12)$$

Here, $LL(\pi)$ is given above in Equation (3) and p = y/n. Solving this nonlinear equation is equivalent to finding the roots of an equation, and this task is easily performed using one of the "NLP" (nonlinear optimization) routines in SAS/IML. As pointed out in the previous example, the simulation results given in Donaldson and Schnabel (1987) underscore the reliability and preference of these likelihood-based intervals over Wald intervals. ■

In comparing Examples 3 and 4, note that Example 3 involves a homoskedastic Normal model whereas Example 4 involves a generalized linear model as the Binomial distribution is assumed. Further, since Example 3 involves multiple parameters, profiling is used to remove the additional (so-called "nuisance") parameters.

Example 5 continued. Bioassay and the Assessment of Drug Synergy. The generalized nonlinear Finney$_4$ model given in Equations (4) and (5) is easily fit to the data given in Giltinan *et al* (1988) using the NLMIXED procedure; the relevant output is given in the Appendix. For these two compounds, antagonism is indicated since the estimate of the coefficient of synergy, $\hat{\vartheta}_5$ = -1.0349, is negative. To test whether independent action is observed here ($\theta_5 = 0$), instead of using the Wald results given the NLMIXED output, we again advocate the use of the likelihood-based test. Thus, we fit the reduced model with the condition $\theta_5 = 0$ imposed; this results in the value –2LL = 110.9 (results not shown), and the test statistic $\chi_1^2 = 110.9 - 80.6 = 30.3$ ($p < 0.0001$). Clearly, these compounds appear to interact antagonistically. ■

Example 6 continued. Obtaining and Verifying Optimal Designs. For the homoskedastic Normal linear model in one independent variable defined over the design space X = [0 , 1] – i.e., for $0 \le x \le 1$ – the D-optimal design (ξ_1) assigns the weight $\omega = ½$ with each of the two design points x = 0 and x = 1. This design, obtained using SAS/IML and the NLPNRA nonlinear optimization subroutine, is given in the Appendix. The corresponding determinant of the information matrix is equal to ¼, and the corresponding variance function, $d(x,\xi_1) = 4(x - ½)^2 + 1$, is graphed below as the solid curve (parabola). D-/G-optimality of this design is verify by noting that the variance function $d(x,\xi_1)$ does not exceed the line y = 2 (the number of model parameters here), and also that $d(x_1 = 0,\xi_1) = d(x_2 = 1,\xi_1) = 2$.



In contrast with the above D/G-optimal design, consider the design ξ_2 which assigns the weight $\omega = ⅓$ with each of the three design points $x_1 = 0$, $x_2 = ½$ and $x_3 = 1$. That this design is not D-optimal is established by noting first that the corresponding determinant of the information matrix is 1/6 (less than the value of ¼ associated with the D-optimal design). The second way to establish non-optimality is to note that the variance function, $d(x,\xi_2) = 6(x - ½)^2 + 1$, graphed above as the dashed curve (parabola), does exceed the cut-line y = 2. These designs and determinant values are obtained using SAS/IML and graphs can easily be plotted using SAS/GRAPH. ■

CONCLUSION

As highlighted above, it is incumbent upon the practicing statistician to more directly adapt the statistical techniques to better address the practitioner's research queries, and using the SAS® NLMIXED and IML procedures (in addition to others) can often give the consultant the means to do so.

REFERENCES

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*, Wiley: New York.

Atkinson, A.C., Donev, A.N. and Tobias, R.D. (2007). *Optimum Experimental Designs, with SAS*, Oxford: Oxford University Press

Clarke, G.P.Y. (1987). Marginal curvatures and their usefulness in the analysis of nonlinear regression models. *Journal of the American Statistical Association*, 82(399), 844-850.

Donaldson, J.R and Schnabel, R.B. (1987). Computational experience with confidence regions and confidence intervals for nonlinear least squares, *Technometrics*, 29(1), 67-82.

Giltinan, D.M., Capizzi, T.P. and Malani, H. (1988), Diagnostic tests for similar action of two compounds, *Applied Statistics*, 37(1), 39-50.

Haines, L.M., O'Brien, T.E. and Clarke, G.P.Y. (2004). Kurtosis and curvature measures for nonlinear regression models, *Statistica Sinica*, 14(2), 547-570.

O'Brien, T.E. (2002). Hypothesis tests, confidence intervals and common sense, *Proceedings of the International Conference on Teaching Statistics*, Cape Town, S.A.

O'Brien, T.E. (2008). Teaching statistical concepts, fundamentals and modelling, *Teaching Statistics*, 30(1), 81-85.

O'Brien, T.E. and Berg, M.B. (2009). Getting the most from data – maximizing information and power by using appropriate and modern statistical methods, *Journal of Data Science*, 7, 537-550.

O'Brien, T.E., Chooprateep, S. and Funk, G.M. (2009). Encouraging students to think critically: Regression modeling and goodness-of-fit, *Journal of Data Science*, 7, 235-253.

O'Brien, T.E., Chooprateep, S. and Homkham, N. (2009). Efficient geometric and uniform design strategies for sigmoidal regression models, *S. African Statistical Journal*, 43, 49-83.

O'Brien, T.E. and Funk, G.M. (2003). A gentle introduction to optimal design for regression models, *The American Statistician*, 57(4), 265-267.

O'Brien, T.E. and Wang, Y. (1996). Using SAS software to assess and adjust for nonlinearity in nonlinear regression models, *SUGI 21 Proceedings*, Cary, NC: SAS Institute, 1274-1283.

Pawitan, Y. (2001). *In All Likelihood*, Oxford: Clarendon Press.

Samuels, M.L. and Witmer, J.A. (2003). *Statistics for the Life Sciences*, 3rd Edition, Upper Saddle River, N.J.: Prentice Hall.

Seber, G.A.F. and Lee, A.J. (2003). *Linear Regression Analysis,* 2nd edition. New York: Wiley.

Wolfinger, R. (1999). Fitting nonlinear mixed models with the new NLMIXED procedure, *SUGI 24 Proceedings* (http://www2.sas.com/proceedings/sugi24/Stats/p287-24.pdf)

APPENDIX

Output 1. Linear LOF Tests

F_LOF1	PVAL1	F_LOF2	PVAL2
1.9965096	0.1003176	7.6061499	0.0103079

Output 2. Heteroskedastic One-Way ANOVA

THETAA						N2LLA
5.829999	72.939999	70.689999	2.3220895	6.9381841	2.0549695	99.995372

THETAB				N2LLB
71.864014	4.595775	7.0211213	2.3666872	116.71103

TESTSTAT	PVAL
16.715658	0.0002346

Output 3a. Calibration, PLCI's and Curvature – PROC NLIN

The NLIN Procedure

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	1	348.1	348.1	24.49	0.0078
Error	4	56.8451	14.2113		
Corrected Total	5	404.9			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits		Skewness
b	-0.7309	0.1477	-1.1410	-0.3209	-104E-18
g13	26.8686	3.7519	16.4517	37.2854	1.0035

Approximate Correlation Matrix

	b	g13
b	1.0000000	0.8276845
g13	0.8276845	1.0000000

Output 3b. Calibration, PLCI's and Curvature – IML

```
          TH        SIG         WALD
-0.730929 3.7697852     -1.14099 -0.320868
26.868578              16.451585  37.28557

       ADJGAM     ADJBET       FUNC
    -2.65E-17 -7.98E-17          1
    0.1672435 0.0343997 0.6597623

             MCBETACI
          -1.14099 -0.320868
          18.526309 44.884949

                PL
          -1.14099 -0.320868
          19.498871  48.35563

              OVER
            1          1
          0.5575061 0.8510444
```

Output 4. Binomial Likelihood Confidence Interval

```
              Optimization Results
              Parameter Estimates
                                    Gradient
                                   Objective
N Parameter         Estimate        Function
1 X1                0.003931    -0.000005051

Value of Objective Function = 3.269083E-12


              Optimization Results
              Parameter Estimates
                                    Gradient
                                   Objective
N Parameter         Estimate        Function
1 X1                0.262078    1.615404E-13

 Value of Objective Function = 2.031197E-14
```

Output 5. Bioassay and the Assessment of Drug Synergy

```
                        The NLMIXED Procedure

                          Specifications
Data Set                                    WORK.ONE
Dependent Variable                          y
Distribution for Dependent Variable         Binomial
Optimization Technique                      Dual Quasi-Newton
Integration Method                          None
```

Dimensions	
Observations Used	20
Parameters	4

Fit Statistics	
-2 Log Likelihood	80.6

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th2	9.9531	1.0590	20	9.40	<.0001	0.05	7.7441	12.1621
th3	1.8018	0.1612	20	11.18	<.0001	0.05	1.4655	2.1381
th4	0.9002	0.1170	20	7.69	<.0001	0.05	0.6562	1.1442
th5	-1.0349	0.1390	20	-7.44	<.0001	0.05	-1.3248	-0.7449

Output 6. Obtaining and Verifying Optimal Designs

Optimization Results
Parameter Estimates

N	Parameter	Estimate	Gradient Objective Function	Active Bound Constraint
1	X1	1.477225E-18	0.250000	Lower BC
2	X2	1.000000	-0.250000	Upper BC
3	X3	0.500002	-0.249999	
4	X4	0.499998	-0.250001	

Value of Objective Function = -0.0625

CONTACT INFORMATION

Contact the author at:

Professor T.E. O'Brien
Department of Mathematics and Statistics
Loyola University Chicago
6525 N. Sheridan Road
Chicago, IL 60626 USA
Work Phone: 1-773-508-2129
Fax: 1-773-508-2123
Email: tobrie1@luc.edu
Webpage: http://webpages.math.luc.edu/~tobrien/home.html