AbstractID: 9817 Title: Models in Medicine. III. Introduction to Generalized Linear and Nonlinear Models.

Advances in statistical theory, numerical methods, and computer software and hardware have allowed investigators to extend regression modeling well beyond the GLM to two types of important problems in which either the *stochastic* or the *deterministic parts*, respectively, differ from the GLM: (1) the response variables y have a probability distribution other than Normal – the response variables may even be *categorical* as well as *continuous*; (2) the relationship, $y = f(x, \beta) + \varepsilon$, between response, y, and explanatory, x, variables and error, ε, is not the additive form $y = x^T\beta + \varepsilon$ of the GLM in which the predictor is a linear function of the unknown parameters. Non-Normal responses are addressed by the Generalized Linear Model; nonlinear relationships by Nonlinear Regression (NLR). Unlike the GLM models, which are typically empirical interpolation models, the Generalized Linear and Nonlinear Regression Models are mechanistic and therefore more reliable in extrapolation (though not necessarily in interpolation). Mechanistic models typically arise in nonlinear differential equations (either ordinary or stochastic), the solutions to which yield one of these two classes of models (usually NLR). The Generalized Linear Model consists of three elements: (i) a *deterministic* part, the linear predictor η, which, as in the GLM, is a linear combination of the parameters, $\eta = x^T\beta$; (ii) a *stochastic part*, ε, which has a distribution from the *exponential family*, a large class of distributions that includes the Binomial and Poisson, as well as the Normal;
(iii) a *link function*, $g(\mu)$, which links a function of the *expected value*, $\mu = E(y)$, of the response variable, y, to the linear predictor, η, $g(\mu) = \eta$. If ε has a Normal distribution with constant variance then the link function is the *identity*, $g(\mu) = \mu$ and $\mu = x^T\beta$ i.e., the regression model. If ε has a Binomial distribution then $\mu = \pi$, where $0 \leq \pi \leq 1$, and the link function is the logit, $\log[\pi/1-\pi]$; e.g., a dose-response model, $\log \pi/(1-\pi) = x^T\beta$ or $\pi = [\exp(x^T\beta)]/[1 + \exp(x^T\beta)]$. Inference is based on maximum likelihood theory. Estimates of β and goodness-of-fit are typically obtained by iterated reweighted least squares (IRLS) procedures. In Nonlinear Regression (NLR) models the predictor, $f(x, \beta)$, is a nonlinear function of the parameters, β; i.e., $\partial f/\partial\beta$ is a function of β. A typical example is the chemical reaction-rate model, $y = \beta_0 x/(\beta_1 + x) + \varepsilon$, where ε is Normally distributed with constant variance, σ^2. In NLR modeling the deterministic and stochastic terms may combine either *additively*, $f(x, \beta) + \varepsilon$, or *multiplicatively*, $f(x, \beta)\varepsilon$. Note that NLR mechanistic models are generally more *parsimonious* than an empirical regression model of the same data. Estimates of β, σ^2 and goodness-of-fit are obtained by an iterated least squares procedure, e.g., Gauss-Newton. In both the Generalized Linear Model and the Nonlinear Regression Model, Regression Diagnostics are key elements in model evaluation and interpretation. Medical applications of both models are described. **Educational Objectives:** Understanding of 1) The Generalized Linear Model. 2) The Nonlinear Regression Model.