The Generative "Social Circles" Feedback Network as Explanatory/Statistical Model: Questions and a Conjecture[1]

Douglas R. White   Excursis2.doc

Discussion draft for the weekly "open problems" network group meeting, UCI

Abstract. How do we explain 1-to-1 mapping between three parameters well-fitted by MLE estimates to a single function that describes all the degree distributions of a particular set of social networks and that also describes all the degree-distribution outcomes of a generative network model? The generative model is known as the *social circles* or feedback networks model with parameters α, β, γ that govern the simulation and outcome parameters δ, κ and q for the shapes and scales of the degree distributions of the networks it generates (WKTFW: eqn. 4). The 1-1 mapping from model parameters (α, β, γ) ↔ (δ, κ, q) has already been described and weakly tested as 1-1 from extensive simulation results. Can we describe through appropriate MLE statistics a clear inferential route to explanatory modeling with the social circles model?

The conditions for a quasi-maximum likelihood estimator (QMLE) (e.g., H. White (1982, 1994)). describe the problem posed. The equation that fits the degree distributions passes most of the tests that qualify for a quasi-maximum likelihood estimator, which converges under conditions given by H. White (1982,1994) to the solution to the problem

$$\max_{\delta,\kappa,q} \ell\,(\delta,\kappa,q;\alpha,\beta,\gamma) \equiv \int \ln p(\kappa;\,\delta,\kappa,q)(\,\kappa;\alpha,\beta,\gamma)\,\mathrm{d}\,\eta(\kappa),$$

where η is the counting measure on the integers.

When, as is typically the case, the above maximization problem satisfies the hypotheses of the theorem of the maximum (e.g., Berge, 1963), the 1-to-1 mapping condition is satisfied. H. White and Chalak (2006, Theorem 4.2) give conditions for application of the theorem of the maximum in a related context.

Network theory is in need of generative models that can serve as explanatory paradigms. Barabási (2002) attempted to do so with the single parameter-constant ($\alpha=1$) in $p(k) \sim k^{-\alpha}$, the scale-free generative network model, where $p(k)$ is the probability of the next link being chosen to a node with degree k in an evolving network, $\alpha=1$ being a "universal" constant for "preferential attraction."[2] The explanatory theory offered by Barabási and colleagues for large networks would require that for nodes to link preferentially to hubs they much possess "perfect detectors" of the degree of other nodes in the network. Google, for example, might serve as such a detector, combining with "preferential attraction" to explain the observed power-law tendencies for degree distributions of number of incoming links for nodes (web sites) on the internet. In models of probabilistic network evolution, adding nodes and requiring randomly chosen existing nodes to link with others according to their degree and $p(k) \sim 1/k$, the resultant degree distributions will have a power-law slope that approaches 3 in the limit as the number of nodes approaches infinity. Alternatively, one could utilize a generating function to model an activity bias $p(k) \sim k^{-\alpha}$ for a node with degree k to link to a randomly chosen node. From a degree distribution alone it might be impossible to distinguish "preferential attraction" and its mechanism from "preferential activity" with nothing but randomness operating in choice of an alter. In addition, there are localized processes that when summed over localities form degree distributions that are power-law even though the distributions for local regions are not. In his popularization of scale-free networks as explanatory models Barabási (2002) makes the mistake of inferring a generative network model such as $p(k) \sim 1/k$ and preferential attachment from a power-law tendency exhibited by a degree distribution.

The generative feedback network model of White, Kejžar, Tsallis, Farmer, and White (WKTFW) (2006) provides a plausible account of network formation for social processes that reflects three commonly observed tendencies in both the generating model and the degree distributions of simulated networks: an activity bias with probabilistic $p(k) \sim k^{-\alpha}$ where α is an activity exponent that might vary from 0 to 1; an exponent γ that might vary from 0 to 1 to

---

[2] Outcome distributions of cumulative $P(x) \sim x^{-q}$ have a valid probability density, provided that $q>1$ (for $q=1$ the "density" doesn't integrate to one, which is not generally a problem since in the range of the Zipfian $q \sim 2$) and yield MLE estimates that are asymptotically unbiased, consistent and asymptotically normal, but not unbiased for small samples.

parameterize p(k)~k$^{-\gamma}$ to describe probabilistic (by parameter γ) how "small world" instructions might be phrased when the activated node sends a token in a parameterized search process varying from random (γ = 0) to fully hub-to-hub (γ = 0) traversal of that node's network (i.e., "please forward to the greatest hub in your network until a target is found that fits description x"), and a distance decay bias β for probabilistic selection of a distance d that the token will travel before sufficient time to elapse that the active nodes recruits a partner from outside the network.

This "social circles" model mimics new-tie formation that results from attempts to find suitable partners within a personal network with whom to form a new (feedback) link and, failing that, to recruit a new partner from outside. It encapsulates processes that may tend (depending on the parameter settings) to lead to power-laws in the tails of degree distributions ("scale-free" tails), tendencies for such degree distributions to approach power-law tails only asymptotically, tendencies for there to be fewer members in networks than the number expected from power-law extrapolations of the power-law tails, tendencies for social circles and local clusters to form even in large networks ("small-world" tendencies), tendencies for reciprocity to occur in social networks, and tendencies for new nodes to be recruited into a social network at a self-regulated rate that would account for low network density and the absence of network saturation into local clusters that are much more dense than needed for social cohesion to operate. Further detail may be found in the description of the model.

Several features of this model are of special interest. One is that it has the potential to represent, through its three parameters (α,β,γ), a wide range of different mixes of the tendencies of different observed social networks. We call this property "broad potential verisimilitude."

Second, all the degree distributions of the networks produced by the simulation are fit by a single equation which itself has three parameters: δ, q, and κ. As shown in Table 1 (WKTFW 2006:5), fit to this equation is tested by Kolmogorov-Smirnoff (KS) and Wilcoxon rank sum (W) tests. In this model, in addition to a normalizing constant p_0 needed for a probability distribution, the probability of a node occurring with degree k is $p(k)=p_0 k^{\delta} e_q^{k/\kappa}$, examined after motivating use of a 1-1 (α,β,γ) ↔ (δ, κ, q) mapping from theoretical model to empirical data.

TABLE I. Parameters for the best fit to a q-exponential function for networks with different parameters. The first three columns are the parameters of the network model, and the next three columns are the parameters for the fit to the q-exponential. The parameter b represents the exponent in the limiting (when $k \to \infty$) Pareto distribution. It is defined by $b \equiv 1/(q-1)-\delta$ (see the text). The last two columns are p values for nonparametric statistical Kolmogorov-Smirnov (KS) and Wilcoxon rank sum (W) tests. The standard acceptance criterion is to have $p > 0.05$, i.e., less than one failure in 20. The asterisk depicts the one case where the null hypothesis was rejected. Consequently, if we demand that both KS and W tests are satisfied, we obtained failure in only one among the 12 cases that we have analyzed.

Network model			Fitted parameters			p values for nonparametric tests [21]		
α	β	γ	q	κ	δ	b	KS test	W test
0	1.2	0	1.08	1.7	0	12.5	0.90	0.54
0.5	1.2	0	1.2	2.1	−0.6	5.6	0.91	0.50
1	1.2	0	1.65	2.75	−1.4	2.94	1.0	0.80
0	1.2	1	1.21	1.5	0	4.76	0.80	0.429
0.5	1.2	1	1.38	1.8	−0.6	3.23	0.15	0.096
1	1.2	1	2.1	2.8	−1.5	2.41	0.76	0.65
0	1.4	0	1.16	1.91	0	6.25	1.0	0.83
0.5	1.4	0	1.31	2.35	−0.6	3.83	1.0	0.95
1	1.4	0	1.85	3.2	−1.4	2.58	0.07	0.03[*]
0	1.4	1	1.16	5.4	0	6.25	0.96	0.92
0.5	1.4	1	1.42	4.5	−0.6	2.98	0.73	0.44
1	1.4	1	2.9	3	−1.5	2.03	0.24	0.35

Motivation

Findings of empirical studies done after development of the feedback network simulation motivate the questions addressed here. Certain empirical network studies are found where the degree distributions of networks fit the equation for networks simulated by the social-circles model better than any other family of distributions than we can identify (power-law, exponential, loglinear, etcetera). This is true, for example, for a study of groups of people in different communities were the data consist of complete nominations of intra-community links on a number of different dimensions, and relevant attribute and contextual data are collected as well. For 5 groups and data on seven network in each, the networks represent salient activities or problem-solving ties within the context of the community, and lend themselves to interpretations from the $(\alpha, \beta, \gamma) \leftrightarrow (\delta, \kappa, q)$ mapping of parameters from the social circles model.

The Simulation Networks and its Generative Parameters

Networks with N=250 and N=5000 nodes were simulated from the feedback model for 12 combinations of parameter values [$(\alpha=0, .5, 1)$ x $(\beta = 1.4, 1.6)$) x $(\gamma = 0, 1)$]. Each parameter

combination was realized 10 times, for a total of 120 networks. Degree distributions were summed for each parameter combination (WKTFW 2006:4, FIG. 3).

As shown in FIG. 3, parameter α governs the shape of the degree distribution, whether power-law (α=1) or more exponential (α=0), or a mixture of the two (α=.5). γ tends to govern the power-law slopes of the tails of the distribution, but other variables affect this as well. β has an effect on the bend in the curves.
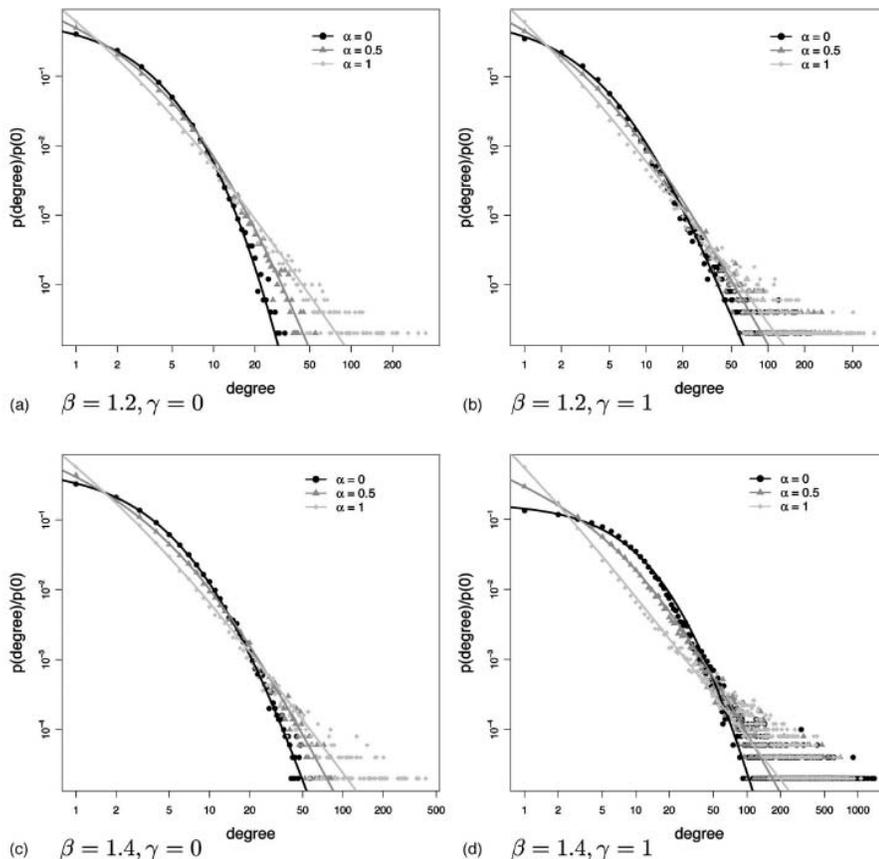

FIG. 3. Degree distributions and fits to a q-exponential for simulations of networks with N=5000 and 10 realizations. The dots correspond to the empirically observed frequency of each degree; the lowest row of dots in each case corresponds to observing one node with that degree. The solid curves represent the best fit to a q-exponential. In each case α has the three values $\{0, 0.5, 1\}$, corresponding to black, dark gray, and light gray, respectively. (a) $\beta=1.2, \gamma=0$, (b) $\beta=1.2, \gamma=1$, (c) $\beta=1.4, \gamma=0$, and (d) $\beta=1.4, \gamma=1$. Note that the scale of the x axes changes. The parameters of the fitted generalized q-exponential functions are given in Table I.

Key statistical aspects of the simulation of concern here are these (WKTFW 2006:5):

> "Despite that fact that network formation in our model depends purely on local information, i.e., each step only depends on information about nodes and their nearest neighbors, the probability of cycle formation is strongly dependent on the global properties of the graph, which evolve as the network is being constructed. In our model there is a competition between successful searches, which increase the degree of two nodes and leave the number of nodes unaltered, and unsuccessful searches, which increase the degree of an existing node but also create a new node with degree 1.

Successful searches lower the mean distance of a node to other nodes, and failed searches increase this distance. This has a stabilizing effect—a nonzero rate of failed searches is needed to increase distances so that future searches can succeed. Using this mechanism to grow the network ensures that local connectivity structures, in terms of the mean distance of a node to other nodes, are somewhat similar across nodes thus creating long-range correlations between nodes. Because these involve long-range interactions, we check whether the resulting degree distributions can be described by the form

$$p(k) = p_0 k^\delta e_{q-}^{k/\kappa} \qquad (4)$$

where the *q-exponential* function e_q^x is defined as

$$e_q^x = [1+(1-q)x]^{1/(1-q)} \qquad (e_q^x = e_q^x),$$

if $1+(1-q)x > 0$ and zero otherwise."

This function "arises naturally as the solution of the equation" $dk/dt = k^q$, which is $k = e_q^{at} = [1+(1-q)\,a\,t]^{1/(1-q)}$. If you take q=1, then you have $k = e^{at}$. All 12 simulated degree distributions were found to fit this probability distribution function, which for degree k combines a power-law term k^δ with a stretched exponential term $e_q^{-k/\kappa}$.

Parameter q>1 measures departure from an exponential (degree) distribution toward a Pareto I tail, and $1/(q-1)$ estimates the asymptote to a power-law at the tail, as seen in FIG. 3. The Pareto II or *q*-exponential is the mix of Pareto I and exponential. The scale parameter κ governs the cross-over in the *q*-exponential deviation from Pareto I to a Pareto II distribution in degree, seen as curvature in FIG. 3.

FIG. 5 (WKTFW 2006:5) shows how outcome parameters q and κ are related to the generating parameters. The data in figure 5 and 4 (linear relation between α and δ) are used to map the generating parameters to output parameters for degree distributions.

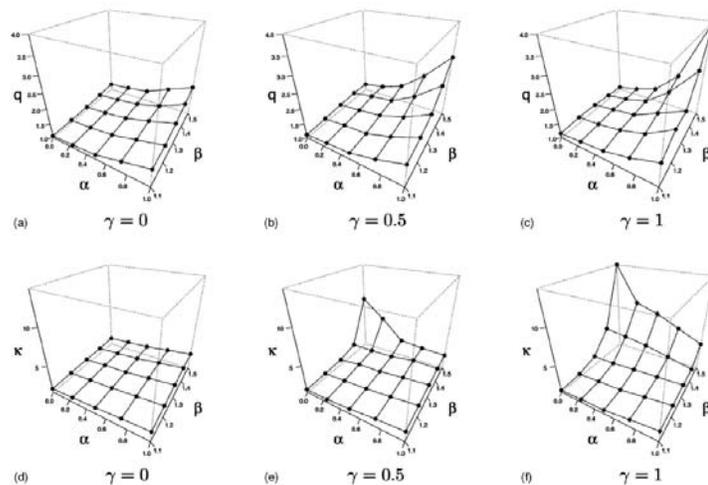
FIG. 5. Dependencies of *q*-exponential parameters *q* and κ that were fitted to network models.

Requirements for 1-1 Mappings: Halbert White (2007, personal communication):

"Let $\psi(\kappa; \alpha,\beta,\gamma)$ define the true degree density implied by the generative model for feedback networks introduced by D. White, Kejžar, Tsallis, Farmer, and S. White (2006) (WKTFW). Let $p(\kappa; \delta,\kappa,q)$ define the q-exponential density proposed by WKTFW to approximate ψ. Maximizing the sample likelihood with respect to the parameters δ,κ,q when p is not necessarily identical to ψ yields a quasi-maximum likelihood estimator (QMLE) (e.g., H. White (1982, 1994)). The quasi-maximum likelihood estimator converges under conditions given by H. White (1982,1994) to the solution to the problem

$$\max_{\delta,\kappa,q} \ell\,(\delta,\kappa,q;\alpha,\beta,\gamma) \equiv \int \ln p(\kappa; \delta,\kappa,q)(\,\kappa;\alpha,\beta,\gamma)\,\mathrm{d}\,\eta(\kappa),$$

where η is the counting measure on the integers.

When, as is typically the case, the above maximization problem satisfies the hypotheses of the theorem of the maximum (e.g., Berge, 1963), the solution can be represented as

$$\delta^* = D(\alpha,\beta,\gamma)$$
$$\kappa^* = K(\alpha,\beta,\gamma)$$
$$q^* = Q(\alpha,\beta,\gamma),$$

where $F \equiv (D,K,Q)'$ represents a non-empty, compact-valued upper hemi-continuous correspondence. (White and Chalak, 2006, Theorem 4.2 gives conditions for application of the theorem of the maximum in a related context.)

If in addition ψ is such that F' obeys the conditions of the inverse function theorem (among other things, F is continuously differentiable with non-zero Jacobian at $(\alpha_o,\beta_o,\gamma_o)$, ($|J(\alpha_o,\beta_o,\gamma_o)|$), then there exists a function F^{-1} such that

$$(\alpha,\beta,\gamma)\,' = F^{-1}(\delta_o,\kappa_o.q_o)$$

for all (α,β,γ) in an open neighborhood of $(\alpha_o,\beta_o,\gamma_o)$. These conditions can be verified to hold for a broad class of functions ψ."

Mapping the Simulated Networks to the Generative Parameters

One plan for estimating the 1-1 functions involve three stages is: first, estimate the two trend surfaces, second, identify and solve convergent simultaneous equations, third, check the White-Chalak (2006), H.White (1994), and Berge (1963) conditions for 1-1 mappings:

1a. Estimate (δ,κ,q) as functions of (α,β,γ) from the published trend surfaces (creating a complete dataset for the six variables from FIG. 5 (WKTFW 2006:5).

1b. Estimate (α,β,γ) as functions of (δ,κ,q) from dataset as a trend surface.

2. Checking results in terms of solving for each one set of functions by solving simultaneous equations, one from the other.

3. Checking conditions.

An alternate plane by Scott White is to configure the 3-dimensional representations of the trend surfaces for δ, κ, q, as f(α,β,γ), and vice versa and solve the simultaneous equations by reconfiguration of the 3D space.

At a much cruder level of approximation, preliminary estimations of three outcome parameters of the fitted parameters were done by various forms of regression from the generating parameters as independent variables (WKTFW: figures 4 and 5). These are only provisional or indicative experiments:

$\delta = -1.5\,\alpha$ (95% adj. R^2)   1

$q = \gamma$ (66% R^2) (quadratic)   2

$q-1 = .25 + .3\gamma$ (26% adj. R^2)   2

$q = 1 + .9\alpha - 6\alpha\gamma(\beta-1)^2$ (84% adj. R^2)   2

    $\gamma = q$ (60% R^2) (not comparable)   2

    $\gamma = .41 + .9(q-1)$ (26% adj. R^2)   2

$\kappa = 246\gamma(1-\alpha)(\beta-1)^4$ (87% adj. R^2)   3

Reversing the functional relations, these equations may be solved for α, β, γ as a function of outcome parameters and re-estimated by regression analysis:

$\alpha = -\delta/1.5$ 3

$\beta = 1 + \kappa/246\gamma(1+\delta/1.5))^{1/4}$ (87% adj. R^2)   4

    $^{\wedge\wedge}(k-1.5)$?

$\gamma = (1-q + \alpha/1.1)/ 6\alpha(\beta-1)^2$ substituting β and δ into γ   5

$\gamma = ((1-q -\delta/1.5/1.1)/-6\delta/1.5(\kappa/246 (1+\delta/1.5))^{1/2})^{2/3}$ (now a function of fitted parameters)

Thus

$\beta = 1 - 1.5\kappa/246((1-q -\delta/1.5/1.1/6)/ \delta (\kappa/246 (1+\delta/1.5))^{1/2})^{2/3}(1+\delta/1.5))^{1/4}$ 4

$\beta = 1 - 1.5\kappa/246((1-q -4.4\delta)/ \delta (\kappa/246 (1+\delta/1.5))^{1/2})^{2/3}(1+\delta/1.5))^{1/4}$ 4

Interpretation

Because these functions are approximate estimates by regression, they may be improved. They help to show the effects of each process in the simulation models, and the interactive effects between them.

The activity bias α has a 1-to-1 relation with an outcome parameter for degree distributions, so it may be recovered from an empirical network from parameter δ. The

distance decay β and traversal γ parameters may be recovered but complex nonlinear equations.

As shown in FIG. 5 (WKTFW: 6), γ produces both high q and high κ. This needs to be taken into account in the equation for κ. Retrodicting the generating parameter γ seems to involve the most complex nonlinear regression estimate.

Conjecture for Deriving an MLE of p(k)= $p_0 k^{\delta} e_{q-}^{k/\kappa}$ for use in estimating its parameters from Empirical Networks

In the empirical network, using Spss nonlinear regression, equation (4) seems to characterize the empirical degree distributions. The question is whether we can do MLE estimation of equation 4. Shalizi (2007) provided an MLE solution for Pareto II which provides unbiased and efficient estimates of parameters for the *q*-exponential (the two are equivalent under 1-1 parameter remappings).

The conjecture, because the two components of the equation, p(k)= $p_0 k^{\delta} e_{q-}^{k/\kappa}$, have MLE solutions (Pareto I and II), is that their product has an MLE solution. This would involve following out Shalizi's (2007:2) derivation for Pareto II, starting with

$$P(X \geq x) = x^{\delta}(1+x/\sigma)^{-\Theta}$$

Derivation of the appropriate formula to optimize for an appropriate MLE solution to this equation seems an open problem, since a prior solution has not yet surfaced..

Interpretation in the Application Context

In a study of Chinese migrant networks involving of 7x5 or 35 networks on 200 people in each of 5 urban migrant communities, each responding to how they relate to each other migrant in their group, all 35 networks have degree distributions that fit the *q*-exponential quite well. This should entail an even better fit to eqn. 4 since if $\delta=0$ the two are equivalent, do adding δ can only improve fit (but: at the cost of an extra parameter.

The empirical observation predicted is that the three fitted parameters (δ, k, q) in the 35 networks map to (δ, k, q) in the simulated networks and these map back to the α, β, γ in the generative feedback network model.

We have examined this mapping without MLE and found it to be substantively meaningful: the generative feedback network model seems to represent in quantitative terms the type of process that generated these networks, and the interpretation of the α, β, γ

parameters of the generative model are substantively meaningful for (and correlated with features of) the 5 populations and 35 networks (in the 7x5 factor design).

Starting from the Simulation itself

An alternative suggested by H. White is to express the generative process in terms of Fokker-Planck equations or something similar, and solve for the stationary distribution.

Concluding questions

Can we regard the *feedback network* or *social circles* model in the same light that Barabási asked us to regard the scale-free model? This model has not only "broad potential verisimilitude" but actual verisimilitude to the full range of all the networks observed in an empirical study, under assumptions such as $dx/dt = x^q$," which acts like a variable parameter exponent in $p(k) \sim k^{-\alpha}$, generalizing the scale-free model to a special case of a scale-dependent model? If so, can we also use this as a *modular* component of more general set of models, such as *adding* MLE testing of the *social circles* and, for example, the triads census, simultaneously? (A suggestion of Katie Faust).

Although the social circles model has broader verisimilitude than the scale-free model, we must be careful about what might be concluding by fitting the model, or this model plus others. As in loglinear we can say we accept the null hypothesis of the *social circles* model fitting the data although we cannot reject that there might be some better alternative model. But can the *social circles* model then provide one of the general frameworks in which we can examine theoretical and comparative statements about the parameters for basic processes that might explain the differences in properties of our data, network by network, community by community, and tracing the empirical attributional, contextual or observed processes that correlated with variations in the inferred *social circles* model parameters. With MLE the estimates of the empirical degree distributions in the survey data are unbiased. If these values are within the ranges of those of the simulated networks of the model, then with the mean value parameterization theorem we are assured of a valid 1-1 mapping from these simulated parameters back to the parameters of the model. The existence of such a 1-1 mapping is already evident in the figures in the original publication of the model that describe the generative results.

References

Barabási, Laszlo. 2002. *Linked: The New Science of Networks*. New York NY: Perseus Press.

Berge, Claude. 1963. *Topological Spaces*. Translated by E. M. Patterson. New York: Macmillan.

Brown, Lawrence D. 1986. *Fundamentals of Statistical Exponential Families*. Institute of Mathematical Statistics Lecture Notes-Monograph Series.

Kiefer, N. M. Econ 719. Exponential Families. http://instruct1.cit.cornell.edu/courses/econ719/Exponential_Families.doc

Lauritzen, Steffen. 2004. *Maximum Likelihood in Exponential Families*. BS2 Statistical Inference, Lecture 6 Michaelmas Term November 7, 2004. University of Oxford. http://www.stats.ox.ac.uk/~steffen/teaching/bs2siMT04/si6c.pdf

Shalizi, Cosma. 2007. Maximum Likelihood Estimation for q-Exponential (Tsallis) Distributions. http://arxiv.org/PS_cache/math/pdf/0701/0701854.pdf

Walker, Mark. 1979. A Generalization Of The Maximum Theorem. *International Economic Review* 20(1): 267-271. doi:10.2307/2526431 http://links.jstor.org/sici?sici=0020-6598(197902)20%3A1%3C267%3AAGOTMT%3E2.0.CO%3B2-X

White, Douglas, Nataša Kejžar, Constantino Tsallis, Doyne Farmer, and Scott White (WKTFW). 2006. Generative model for feedback networks. *Physical Review E* 73, 016119 http://eclectic.ss.uci.edu/~drwhite/pw/WhiteKejzarTsallisFarmerWhite06.pdf

White, Halbert. 1994. *Estimation, Inference and Specification Analysis*. Econometric Society Monographs.

--- 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometric* 50:1-25.

White, Halbert, and Karim Chalak. 2006. A Unified Framework for Defining and Identifying Causal Effects. http://weber.ucsd.edu/%7Embacci/white/pub_files/hwcv-sub005.pdf