

Statistical Analysis of Single Relational Networks

in Wasserman, Stanley, and Katherine Faust (1994), Social Network Analysis: Methods and Applications, New York: Cambridge University Press, 605-674.

by Dawn Iacobucci

We now turn our attention to stochastic models for social network data. The methodology described here continues the development of statistical methods for network data begun in Chapter 13. We begin in Chapter 15 by considering a (very special) class of statistical distributions for random directed graphs, which, as we will show, is a special case of the uniform random directed graph distributions presented in Chapter 13. This class is more interesting than the distributions of Chapter 13, and contains substantively meaningful parameters which reflect a wide variety of graph properties. Further, the parameters can actually be estimated from data. The basic model has many generalizations and extensions, some of which are described in Chapter 16.

In Chapter 16 we turn to the last question raised in Chapter 9 concerning methodology for studying a positional analysis. We want to measure the adequacy of a representation of a positional analysis. We stated that there are four tasks that have to be undertaken in a positional analysis:

- (i) Define equivalence
- (ii) Measure how closely the actors adhere to this definition
- (iii) Represent the equivalences of the actors
- (iv) Measure the adequacy of this representation

Two of the necessary tasks are measurement-oriented. These tasks are the second and fourth. The second task requires the analyst to determine how equivalent the actors are, for a given set of relations; that is, one must find which actors are equivalent, and which ones are not, using some measurement device(s). After such an examination, one then turns to the

third task in order to represent the discovered equivalences (and non-equivalences) mathematically. In Chapter 16, we focus on the adequacy of this mathematical model.

The statistical models described in Chapter 15 allow a researcher to perform significance tests — a formal evaluation of the statistical significance of various substantive effects based on null hypotheses. For example, an outdegree used as a descriptive measure of an actor's expansiveness cannot be evaluated as absolutely large or as significantly larger than other outdegrees, but such inferences can be made with statistical models. Furthermore, parameters that quantify the "structural effects" present in a network, such as reciprocity and tendencies toward differential indegrees, can be estimated simultaneously; for example, we can model actor expansiveness while controlling for differential actor popularity. The models described here are *dyadic interaction* models, which use the (natural) log of probabilities as their basic modeling unit. The models posit a structural form for the (natural) logarithm of the probability that actor i "chooses" actor j at one strength while actor j "chooses" actor i at a possibly different strength.

Chapter 16 first describes goodness-of-fit indices for positional analyses not based on statistical models. Next, it presents methods that assume that a statistical model is actually operating, so that the index considered arises naturally from the underlying model. We call both types of indices goodness-of-fit indices, because both attempt to measure the fit of a model to a data set, but note that there is this fundamental distinction between them. And the last section in the chapter describes generalizations and extensions of the models presented in Part VI.

Statistical network analyses allow the researcher to assess a model by measuring the fit of the model to data. In addition, statistical approaches yield flexible probabilistic models that can be generalized by using random directed distributions based on network characteristics. These distributions allow comparisons of the observed effects to hypothesized effects, as well as significance tests to determine whether an effect is due to sampling variability.

We begin this chapter by presenting models for a network with measurements on a single, directional relation for one set of actors. We then describe and demonstrate the interpretation and fitting of a basic statistical network model. Attribute variables measured on the actors can also be incorporated into the models, so that we have the flexibility to model network structure among individual actors or among subsets of actors in situations in which the subsets are defined *a priori* based on actor

attributes. We also describe models that focus entirely on the relation, to the exclusion of the individual actors or the subsets to which they belong. Modifications of the basic statistical model are also described that allow for ordinal, rather than just nominal or dichotomous, relations. Lastly, we briefly discuss recent research on related statistical models for single relations.

Toward the end of the chapter, we present models for networks with two sets of actors in which a single relation is measured. We describe models for one set of actors in greater detail, but we also hope to encourage researchers to consider more applications to networks with two sets of actors.

This chapter does require the reader to have some background in categorical data analysis. Specifically, a knowledge of log linear models, and the methods for fitting such models to three- and four-dimensional contingency tables is needed. Those desiring more background in log linear models should study the excellent texts of Fienberg (1980), Kennedy (1983), Wickens (1989), or Agresti (1990).

The end of this chapter gives the "commands" needed to fit these models to network data using several computer packages. We give specific details on how to fit the models described in this chapter using the standard packages. Some of the computations for the basic models presented in this chapter are included in the latest release of *UCINET IV*.

More statistically knowledgeable readers may find sections of this chapter rather elementary, and possibly boring — we suggest to such readers that the elementary, discursive parts of the chapter, which explain likelihood functions and maximum likelihood estimation of the parameters in log linear models, can be skipped.

15.1 Single Directional Relations

In this section, we first describe the construction and modeling of the *Y-array*, a contingency table basic to our models which is derived from the relational data in *X*. This array focuses on dyads, and is descriptive of individual actors' ties to other actors.

We demonstrate these methods in detail on the hypothetical set of second-grade children. We use this fabricated social network as an illustrative example because of its small size, which makes the analyses easier to follow. We also present the application of these methods to Krackhardt's friendship relation measured on managers in a high-tech

organization, and Padgett's Florentine marital and business relations. Analyses of these data display different aspects of the methodology discussed in this chapter.

15.1.1 The Y-array

The models for single relational networks are not easily fit to the X -matrix, so we reorganize the network data into a different contingency table, to which the models are more easily fit. We first illustrate the construction of this new table using the small hypothetical social network of second-grade children.

Data Review and the Definition of Y . We begin by describing a model for a single, directional relation measured for a single set of actors, \mathcal{N} . Recall that a dyad with measurements on a directional relation consists of two actors, i and j , and the possible ties between these two actors.

The ties between the actors may be viewed from the perspective of either actor i or actor j . First, take the perspective of n_i . The relational variable X_{ij} records the possible "choice" of n_j by n_i , while the relational variable X_{ji} records the possible "choice" received by n_i from n_j . Now, take the perspective of actor j . The relational variable X_{ji} records the possible "choice" of actor i by actor j , while the relational variable X_{ij} records the possible "choice" received by actor j from actor i . Both of these perspectives are incorporated into our modeling.

Recall that a social network consisting of g actors contains $\binom{g}{2}$ dyads. In a statistical model, each dyad consists of information represented by two random variables, X_{ij} and X_{ji} . We will let D_{ij} denote the dyadic variable. With g actors and a single relation, we have $g(g - 1) = 2\binom{g}{2}$ dyadic random variables to consider. We wish to model all the dyadic ties in a network simultaneously and as parsimoniously as possible.

Consider a pair of actors, a single dichotomous relation, and the dyad D_{ij} . The ties in the dyad, for both actors, can be presented in a 2×2 array. The two variables of this array, both of which have just two levels, are rather novel. The first, with two levels, which we index with a k and which can be either 0 or 1, codes the value of the tie sent by the row actor i to the column actor j . The second, also with just two levels and which we index with an l , codes the value of the tie sent by the column actor j to the row actor i . So, the ties for each and every dyad can be

presented in one of these 2×2 arrays. The new indices k and l equal either 0 or 1, depending on the state of the dyad.

Consider now all dyads and this single, dichotomous relation. If we take the original $g \times g$ X binary matrix, and replace each entry with the appropriate 2×2 table, we obtain a new contingency table. Since there are $\binom{g}{2}$ dyads, which can be indexed by the $g \times g$ pairs of actors involved, the new contingency table will be of size $g \times g \times 2 \times 2$.

We can consider valued, as well as dichotomous, relations. The restriction to dichotomous relations common to the statistical methods presented in Chapters 13 and 14 is relaxed here. To model all dyads on a single, valued relation simultaneously, we create a four-dimensional contingency table of size $g \times g \times C \times C$. The first two dimensions of this table are indexed by the actors in \mathcal{N} . The size of the third and fourth dimensions is C , the number of integer values the measured relation can take on. For dichotomous data that are coded $k, l = 0$ or 1 , C equals 2. For relational data coded as $k = 0, 1, 2$ and $l = 0, 1, 2$, C equals 3.

We call the $g \times g \times C \times C$ matrix Y , and define its entries as follows:

$$\begin{aligned} Y_{ijkl} &= 1 \text{ if the dyad } D_{ij} \text{ takes on the values} \\ &\quad (X_{ij} = k, X_{ji} = l) \\ &= 0 \text{ otherwise.} \end{aligned} \tag{15.1}$$

The Y -array is a cross-classification of four variables and thus, its entries have four subscripts: The actors as senders (i), the actors as receivers (j), and the relational variables X_{ij} (indexed by the third subscript, k) and X_{ji} (indexed by the fourth subscript, l). The structure of Y is similar to a sociomatrix, where rows represent sending actors and columns represent receiving actors. The entry in the (i, j) th cell of a sociomatrix is x_{ij} . The (i, j) th cell of Y is not a single quantity, but rather a $C \times C$ submatrix. In this $C \times C$ submatrix, there will be a single 1 found in the (k, l) th cell. The remaining $C^2 - 1$ elements will be 0. Thus, one can view these submatrices as simply indicator matrices, giving the "state" of each dyad. The Y -array has a special symmetry, $Y_{ijkl} = Y_{jilk}$ for all (i, j) and (k, l) pairs, due to the fact that the dyad may be viewed from either the perspective of actor i or the perspective of actor j . The Y -array was created so that the models we are about to describe could be fit to discrete-valued relations using standard log linear modeling procedures that exist in the widely available statistical computing packages.

An Example of Y . As an example, refer to the small fabricated social network of second-grade children first introduced in Chapter 3.

Table 15.1. Sociomatrix for the second-grade children

Friendship at Beginning of Year						
	Allison	Drew	Eliot	Keith	Ross	Sarah
n_1 Allison	-	1	0	0	1	0
n_2 Drew	0	-	1	0	0	1
n_3 Eliot	0	1	-	0	0	0
n_4 Keith	0	0	0	-	1	0
n_5 Ross	0	0	0	0	-	1
n_6 Sarah	0	1	0	0	0	-

In this network, the actors are labeled as follows: $n_1 = \text{Allison}$, $n_2 = \text{Drew}$, $n_3 = \text{Eliot}$, $n_4 = \text{Keith}$, $n_5 = \text{Ross}$, and $n_6 = \text{Sarah}$. To focus on one dyad in particular, we might observe the data for $n_1 = \text{Allison}$ and $n_5 = \text{Ross}$ on the relation of friendship at the beginning of the school year. The data show that Ross does not name Allison as a child he likes, but Allison nominates Ross. From Allison's perspective, the relational variable sent is $x_{15} = 1$, implying that Allison likes Ross as a friend, and the relation received is $x_{51} = 0$, implying that Allison is not liked as a friend by Ross. From Ross's perspective, the relation sent is $x_{51} = 0$, Ross does not choose Allison, and the relation received is $x_{15} = 1$, Ross is chosen by Allison. The recorded data for actors 1 and 5 in this pair $\langle n_1, n_5 \rangle$ would be $D_{15} = (x_{15}, x_{51}) = (1, 0)$, so that $y_{1510} = 1$, while $y_{1500} = y_{1501} = y_{1511} = 0$.

We can build the Y that corresponds to the network describing friendship choices among these six children at the beginning of a school year. We first present these data as a sociomatrix in Table 15.1.

Remember that it is common statistical practice to use capital, bold-faced letters (such as Y) to denote random variables, while actual realizations (such as the y given here) have lowercase, boldfaced letters.

In Table 15.2, we present the y -array for these data. The size of this array is $6 \times 6 \times 2 \times 2$ because the contingency table is actors ($i = 1, 2, \dots, 6$) by partners ($j = 1, 2, \dots, 6$) by strength of choices sent ($x_{ij} = 0, 1$) by strength of choices received ($x_{ji} = 0, 1$), where $C = 2$.

Note the other stated properties of y hold in the example: In each 2×2 submatrix, there is one 1 and $(C^2 - 1)/2 = 3$ 0's. The submatrices along the main diagonal are filled entirely with —'s, because no reflexive ties ("self-choices") are measured for this relation. Finally, note that y is symmetric as described earlier ($y_{ijkl} = y_{jilk}$).

Table 15.2. y for the second-grade children

		Allison	Drew	Eliot	$l = x_{ji}$	Keith	Ross	Sarah
i	k	0	1	0	1	0	1	0
n_1 Allison	$x_{ij} = 0$	-	-	0	0	1	0	0
	$x_{ij} = 1$	-	-	1	0	0	0	1
n_2 Drew	$x_{ij} = 0$	0	1	-	-	0	0	1
	$x_{ij} = 1$	0	0	-	-	0	1	0
n_3 Eliot	$x_{ij} = 0$	1	0	0	0	-	1	0
	$x_{ij} = 1$	0	0	0	1	-	0	0
n_4 Keith	$x_{ij} = 0$	1	0	1	0	-	0	0
	$x_{ij} = 1$	0	0	0	0	-	1	0
n_5 Ross	$x_{ij} = 0$	0	1	1	0	0	1	-
	$x_{ij} = 1$	0	0	0	0	0	-	1
n_6 Sarah	$x_{ij} = 0$	1	0	0	1	0	0	-
	$x_{ij} = 1$	0	0	0	1	0	0	-

The margins of the y -array are quite important to the estimation of parameters for various models. These margins are sums over the elements of y , and are denoted with subscripts including "+" signs. A + used as a subscript on various y terms indicates that one sums over the subscripts replaced by the +'s. For example, y_{++k+} denotes the sum of the entries of y over i , j , and l , for each k . These sums form a one-way table with one cell for each level of k . This margin, $\{y_{++k+}\}$ gives the number of ties on the relation at the various strengths $k = 0, 1, \dots, C - 1$. It is aggregated over actors (i), their partners (j), and the choices received (l).

For the example of the fabricated network of second-grade children (the y -array appears in Table 15.2), the $\{y_{++k+}\}$ margin is:

$$y_{++0+} = 22$$

$$y_{++1+} = 8$$

These numbers tell us that 22 ties have strength $k = 0$ (that is, 22 possible ties are absent), and 8 have strength $k = 1$ (that is, 8 ties are present). Another example is the y_{i1k+} margin, which gives the numbers of ties that are present ($k = 1$) and ties that are absent ($k = 0$) for each actor:

$$y_{i1+0} = 3$$

$$y_{i1+1} = 2$$

$$\begin{aligned}
 y_{2+0+} &= 3 \\
 y_{2+1+} &= 2 \\
 y_{3+0+} &= 4 \\
 y_{3+1+} &= 1 \\
 y_{4+0+} &= 4 \\
 y_{4+1+} &= 1 \\
 y_{5+0+} &= 4 \\
 y_{5+1+} &= 1 \\
 y_{6+0+} &= 4 \\
 y_{6+1+} &= 1
 \end{aligned}$$

For example, n_3, n_4, n_5 , and n_6 , all have one tie — these four children choose just one child as a friend.

15.1.2 Modeling the Y-array

We now present statistical models for the analysis of a single, directional relation, whose data we represent by a 4-dimensional Y-array. Before presenting the mathematical model statement itself, we will motivate the model and describe its utility by explaining the substantive effects that the parameters of the model are designed to reflect. For a single, directional relation, we focus on effects that represent the “expansiveness” of actors, the “popularity” of their partners, and the “reciprocation” of the ties within the dyads.

Description of the Key Model Parameters. The basic model consists primarily of three sets of parameters: one set of parameters describes the actors’ sending behavior, one set describes the actors’ receiving behavior, and one set describes the interactions between pairs of actors within a dyad. The first set of parameters are called *expansiveness* effects. In a children’s friendship network, these effects reflect the propensity of each child to nominate others as friends. The second set of parameters are called *popularity* effects. In the children’s friendship example, popularity reflects the tendency for a child to be nominated by others as a friend.

Patterns of friendship choices among children are described in terms of the expansiveness and popularity of the individual children. While these terms, “expansiveness” and “popularity,” might apply equally well to other network data sets, particularly when actors are people and relations

measure positive affect or evaluation (for example, the expansiveness and popularity of employees as measured on a communication relation), they apply less well in other applications. For example, if the network is one describing children taking toys from other children, one would not necessarily describe as “popular” a child whose toys are frequently taken. Nevertheless, the terms “expansiveness” and “popularity” have become commonplace in the literature. We use these terms in this context to mean precisely this: parameters representing the propensities for actors to have ties to and from the other actors. Positive values of the parameters increase the probability of having ties.

The final set of parameters are those that reflect the reciprocation, or mutuality, between two actors, independent of the expansiveness or popularity of either actor. This set is not all that different from the measures of reciprocity described in Chapter 13. However, the parameters described here are not limited to dichotomous data, and are probabilistic in nature. Further, these reciprocity effects describe interactive behaviors unique to the dyad, above and beyond the probabilistic tendencies for expansiveness and popularity of the actors who comprise the dyad.

Reciprocity is the extent to which a dyad exhibits mutual, as opposed to asymmetric, ties. With respect to the statistical models discussed here, positive reciprocity parameters increase the likelihood that the dyad is mutual. The model we present for a single relation includes parameters to measure the probabilistic tendencies of all of these substantive effects: expansiveness, popularity, and reciprocity.

We estimate these parameters using log linear modeling techniques. Log linear models are the standard statistical method for studying discrete-valued data organized as counts in multi-way contingency tables (see Agresti 1984, 1990; Bishop, Fienberg, and Holland 1975; Fienberg 1980; Goodman 1979; Haberman 1978, 1979; Kennedy 1983; Wickens 1989). The vast majority of social network data are discrete, and almost always C is small. Social network data that are not discrete can often be categorized without losing important information in the data. For example, we might take a continuous measure of *time spent talking* and code it as *high*, *medium*, or *low*. Thus, our concentration on network models that can be fit to discrete-valued relations using log linear models seems appropriate.

The Basic Model for Dichotomous Relations. We begin by discussing the modeling of a dichotomous relation. After presenting the

models for dichotomous relations, we extend the model to the more general case of discrete relations ($C > 2$).

The basic model, introduced and termed " p_1 " by Holland and Leinhardt (1977, 1981), is expressed in four statements. Each of the four statements represents one of the four possible states of any given dyad: the null dyad ($X_{ij} = X_{ji} = 0$, or $Y_{ij00} = 1$), the mutual dyad ($X_{ij} = X_{ji} = 1$, or $Y_{ij11} = 1$), and two cases of asymmetric dyads ($X_{ij} = 1, X_{ji} = 0$, or $Y_{ij10} = 1$, and $X_{ij} = 0, X_{ji} = 1$, or $Y_{ij01} = 1$). We represent the (natural) log of the probabilities of each of these four dyadic states as a function of several parameters, in order to specify p_1 :

$$\begin{aligned}\log P(Y_{ij00} = 1) &= \lambda_{ij} \\ \log P(Y_{ij10} = 1) &= \lambda_{ij} + \theta + \alpha_i + \beta_j \\ \log P(Y_{ij01} = 1) &= \lambda_{ij} + \theta + \alpha_j + \beta_i \\ \log P(Y_{ij11} = 1) &= \lambda_{ij} + 2\theta + \alpha_i + \alpha_j + \beta_i + \beta_j + (\alpha\beta).\end{aligned}\quad (15.2)$$

This model is log linear. It can be viewed as an analogue of the linear models arising in analysis of variance. Log-linear models begin multiplicatively, but once the log of the response variable is taken, the model is additive, or linear, in the parameters. Thus, p_1 begins with a probability of a dyadic state as a response variable, equated to an expansiveness parameter (actually, e raised to the power of the expansiveness parameter) multiplied by a popularity parameter. When the model and response "probabilities" are transformed to the log scale, p_1 shows an expansiveness parameter *added* to a popularity parameter. The log-linear form of the model is simple to fit and to understand. The log of the probability that n_i has ties to and from n_j becomes an additive function of terms that include the expansiveness of n_i and n_j , the popularity of both actors, and the reciprocal effects between the two. When a parameter is positive, it contributes to (or increases) the (log) probability that n_i has a tie to n_j , and if it is negative, the probability decreases.

The $\{\lambda_{ij}\}$ parameters are mathematical necessities included in the model to insure these four probabilities sum to one for each dyad. Thus, these parameters appear in all four statements, regardless of the state of the dyad. The θ parameter is interpreted as an overall choice effect (analogous to a grand mean), reflecting the overall volume of choices sent and received. If one tie is present in the dyad, one θ 's appear in the statement; when the tie is reciprocated, two θ 's appear.

Note that θ does not appear in the model statement when ties are not present, and $(\alpha\beta)$ is present only when the dyad is mutual. No substantive parameters appear in the first statement of the model which represents a null dyad. For asymmetric dyads, the log probabilities depend on parameters reflecting only one of the two possible ties in the dyad: dyads in which actor i chooses actor j without reciprocation (so an α_i but not an α_j is relevant, and a β_j but not a β_i is included) and dyads in which actor j chooses actor i with no reciprocated choice (so the relevant parameters are α_j and β_i , but not α_i or β_j or $(\alpha\beta)$). All the parameters appear together only for mutual dyads (the last statement of the model).

The $(\alpha\beta)$ (sometimes denoted by ρ in the literature), is also called a *mutuality* parameter. When choices on some relation, such as friendship, tend to be mutual in some network, the parameter will be positive and large. In this sense, the parameter is a measure of association between ties sent and received (analogous to a correlation coefficient for continuous data). For some relations like friendship, one would expect reciprocity to be present. However, we might not expect reciprocity for other relations, such as "assigns work to" or "asks for advice." Although a superior might ask a subordinate for advice, we might expect this to occur less frequently than the subordinate asking the superior for advise. With dichotomous data, such patterns on these relations would be modeled with a large negative, $(\alpha\beta)$ parameter indicating that actors who choose others tend not to be those chosen by those others. When reciprocity is not an important factor in a network, the reciprocity parameter would equal 0. We can view $(\alpha\beta)$ as a model-based measure of reciprocity, so that it can be compared and contrasted to the indices for reciprocity discussed in Chapter 13.

Constraints are necessary to estimate the parameters in this model. We use the standard analysis of variance-like constraints in which the parameters and their estimates sum to 0 across their subscripts. We have $\sum_i \alpha_i = 0$, and $\sum_j \beta_j = 0$.

These constraints determine the degrees of freedom (df) associated with each set of parameters. The df associated with any set of parameters is the number of parameters that are independent and free to vary. The expansiveness parameters $\{\alpha_i\}$ have a subscript of i , which ranges from 1 to g , the number of actors. There are g α_i parameters, but they are constrained to sum to 0. Thus, the df for this set of parameters is $(g - 1)$ because we can calculate α_g from the other $(g - 1)$ parameters. Similarly, the popularity parameters $\{\beta_i\}$ also require $(g - 1)$ degrees of freedom. Lastly, the reciprocity parameter $(\alpha\beta)$ requires a single degree

of freedom. Estimation of the parameters of this model is discussed in detail by Fienberg and Wasserman (1981a) and Wasserman and Weaver (1985); we also describe how to estimate these parameters shortly.

We now consider the more general form of the p_1 model, which allows us to study single relational variables that are discrete and not necessarily dichotomous (Wasserman and Iacobucci 1986). We assume that the relational variable can take on values $0, 1, \dots, C - 1$, and that $Y_{ijkl} = 1$, when $X_{ij} = k$ and $X_{ji} = l$. The following model statement generalizes the four statements of p_1 :

$$\begin{aligned}\log P(Y_{ijkl} = 1) = & \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} \\ & + \beta_{j(l)} + \beta_{i(l)} + (\alpha\beta)_{kl}.\end{aligned}\quad (15.3)$$

Note the actor-level parameters in this model. The parameter $\alpha_{i(k)}$ measures the tendency for actor i to send ties at strength k , while $\beta_{j(l)}$ measures the tendency for actor j to receive ties at strength l . We will sometimes refer to such parameters as *actor-level*, because of their dependence on the individual actors.

The parameters are subject to the following constraints:

$$\begin{aligned}\theta_0 &= 0 \\ \alpha_{i(0)} &= 0, \text{ for all } i \\ \sum_i \alpha_{i(k)} &= 0, \text{ for all } k \\ \beta_{j(0)} &= 0, \text{ for all } j \\ \sum_j \beta_{j(l)} &= 0, \text{ for all } l \\ (\alpha\beta)_{k0} &= 0, \text{ for all } k \\ (\alpha\beta)_{0l} &= 0, \text{ for all } l \\ (\alpha\beta)_{kl} &= (\alpha\beta)_{lk}\end{aligned}$$

In words, we constrain the parameters to equal zero when a choice is made at the lowest strength ($k = 0$ or $l = 0$). This generalization of the p_1 model thus becomes equivalent to p_1 when $C = 2$. In p_1 , α_i is defined only when a choice is sent ($k = 1$). Here, $\alpha_{i(k)} = 0$ when $k = 0$, but $\alpha_{i(k)}$ can be non-zero (and usually is) when choices are made at any strength ($k = 1, 2, \dots, C - 1$). For every $k > 0$, the $\alpha_{i(k)}$ sum to 0 across actors. Because the estimates sum to zero across actors, relative comparisons among actors (at each strength) are easily made. The constraints on the

Table 15.3. Constraints on the $\{\alpha_{i(k)}\}$ parameters in model (15.3)

	$k = 0$	$k = 1$	\dots	$k = C - 1$
$i = 1$	0	$\alpha_{1(1)}$	\dots	$\alpha_{1(C-1)}$
$i = 2$	0	$\alpha_{2(1)}$	\dots	$\alpha_{2(C-1)}$
$i = 3$	0	$\alpha_{3(1)}$	\dots	$\alpha_{3(C-1)}$
\vdots	\vdots	\vdots	\ddots	\vdots
$i = g$	0	$\alpha_{g(1)}$	\dots	$\alpha_{g(C-1)}$
Total	0	0	\dots	0

α parameters are depicted in Table 15.3. The constraints stated above are consistent with the fact that the $\{\theta_k\}$, $\{\alpha_{i(k)}\}$, $\{\beta_{j(l)}\}$, and $\{(\alpha\beta)_{kl}\}$ require $(C - 1)$, $(g - 1)(C - 1)$, $(g - 1)(C - 1)$, and $C(C - 1)/2$ degrees of freedom, respectively.

As before, the $\{\alpha_{i(k)}\}$ parameters are the expansiveness parameters, and the $\{\beta_{j(l)}\}$ are the popularity parameters. The $\alpha_{i(k)}$ represents the tendency (or the additive effect on the logarithm of the probability) of actor i to send ties at strength k . Similarly, $\beta_{j(l)}$ represents the tendency (or the additive effect on the logarithm of the probability) of actor j to receive ties at strength l . Actor i 's expansiveness is reflected by $\alpha_{i(k)}$ and actor i 's popularity by $\beta_{i(l)}$. Actor j 's expansiveness is reflected by $\alpha_{j(l)}$ and actor j 's popularity by $\beta_{j(k)}$.

The $\{(\alpha\beta)_{kl}\}$ parameters are the reciprocity effects. Note that these parameters do not depend on the specific actors being modeled (there is no i or j subscript). The model assumes these effects are constant across all pairs of actors. The reciprocity parameters are symmetric in their indices, $(\alpha\beta)_{kl} = (\alpha\beta)_{lk}$, so there is $C(C - 1)/2$ degrees of freedom.

The model for dichotomous data contains just a single reciprocity parameter (because $C(C - 1)/2 = 2(2 - 1)/2 = 1$), as specified by p_1 . The single $(\alpha\beta)$ parameter for modeling dichotomous relations is analogous to a measure of association. When $C > 2$, the $C \times C$ matrix of $(\alpha\beta)_{kl}$ parameters is analogous to an entire matrix of such measures. For example, $(\alpha\beta)_{15} = (\alpha\beta)_{51}$ measures the positive or negative association between ties sent at the weak strength of $k = 1$ and ties received at the much greater strength of $l = 5$.

It is important to note that when ties are valued, the α and β estimates (derived from (15.3), not (15.2)) depend on the number of possible values. For every level $k = 1, \dots, (C - 1)$, there are g alpha's and g beta's. For a

Table 15.4. p_1 parameter estimates for the second-graders

Node label	Actor	$\hat{\alpha}_i$	$\hat{\beta}_i$
n_1	Allison	1.414	$-\infty$
n_2	Drew	0.817	0.867
n_3	Eliot	-0.474	-1.223
n_4	Keith	0.197	$-\infty$
n_5	Ross	-0.977	0.178
n_6	Sarah	-0.977	0.178

$$(\hat{\alpha}\hat{\beta}) = 3.077$$

$$\hat{\theta} = -1.437$$

fixed k , these parameters measure how likely it is that an actor has ties (sent or received) at that particular strength.

An Example — Fitting p_1 to the Fabricated Network. To illustrate, we fit the model to the fabricated network of second-grade children and study its parameters. The parameter estimates resulting from fitting model (15.2) to the y -array based on the friendship choices among the 6 children are presented in Table 15.4.

Note that these parameters are on a logarithmic scale. Thus, if an α increases by 1 unit, say, from 1 to 2, the logarithm of the probability of a choice increases by 1 unit. Or, the actual probability increases by $\exp(1) = 2.718$.

The alpha estimates tell the following story: Actor 1 (Allison) has the largest expansiveness parameter. She was far more likely to have friends (at the beginning of the school year) than were any of the other children. Actor 2 (Drew) was next most likely, and actors 5 and 6 (Ross and Sarah) were least likely.

The beta estimates quantify the tendencies with which each of these children is chosen as a friend by the other children in the network. And these estimates are quite different from the alphas; specifically, two of the parameter estimates are infinite. We will discuss this situation technically later in the chapter, but for now, we note that Allison and Keith are not chosen by any other children as friends; hence, they have 0 indegrees. This forces the beta parameters to be $-\infty$ for these two children.

Whereas child 1 (Allison) was most likely to choose others, she was least likely to be chosen by others, since her $\hat{\beta}$ is the smallest (in fact, she was not chosen so her $\hat{\beta}$ is infinite). Child 4 (Keith) is similar. The other

15.1 Single Directional Relations

children (n_2, n_5, n_6 , and n_3) can be chosen as friends, with Drew, Ross, and Sarah exhibiting positive tendencies.

The reciprocity parameter gives additional information about this relation. With dichotomous data, the analogy between the $(\alpha\beta)$ parameter and a measure of association is especially easy to see. Here, the estimate is positive and large, indicating tendencies for positive association or mutual friendships — if child i nominates another j as a friend, that friend j in turn tends to reciprocate the friendship. Similarly, if n_i does not nominate n_j , n_j also tends to not nominate n_i . If this parameter estimate were negative, it would indicate a relation for which there are many non-reciprocated friendships (or asymmetric dyads).

15.1.3 Parameters

We now discuss theoretical issues and the practical means for calculating the parameter estimates of model (15.3). We also describe how to test the statistical significance of each set of parameters to see which effects in the network are statistically large.

Parameter Estimation — Theory and Practice. In this section, we discuss several issues regarding the estimation of the parameters in model (15.3). We begin with some theoretical issues (such as maximum likelihood estimation and the likelihood-ratio goodness-of-fit statistic), and then explain how to analyze network data using these methods (including significance testing of the parameters). Finally, we describe the statistical analyses of several social network data sets.

⊗ Distributions and Maximum Likelihood Estimation. In this section, we discuss the statistical theory underlying model (15.3), including the form of the likelihood function, the statistical function from which the maximum likelihood parameter estimates are derived. Alternative estimation techniques are described later in this chapter.

Holland and Leinhardt (1981) describe model (15.3) as belonging to an *exponential family of distributions*, which means parameters can be estimated via the *maximum likelihood* estimation procedure. With maximum likelihood, estimated parameters are those that give the best fit to the data. By “fitting the data the best,” we mean that the best-fitting parameters maximize the joint probability distribution of the data. It is as if many estimates of each parameter, say $\hat{\beta}_{1(i)} = 0.05$, or 0.12, or 0.73, were tried out in the model, and the value that “fits” the data

the best is chosen as the “best” estimate. In practice, several values are not actually tried out; the optimal estimates for parameters are usually computed numerically through various exact (or, if necessary, approximate) mathematical procedures. These procedures are used in statistical computing packages to maximize the likelihood function (or really, the logarithm of this function, which is easier computationally).

For p_1 , the goal of maximum likelihood estimation is, “Find the best (maximum likelihood) estimates of all the parameters in the model (the λ ’s, θ ’s, α ’s, β ’s, and $(\alpha\beta)$ ’s) that could have produced the given dyadic interaction data, represented by our y -array.” The likelihood function is the joint probability distribution of the data. Maximum likelihood estimation strives to find parameters that maximize this function. The log likelihood function explicitly tells which functions of the data — which “margins” of the y -array — are needed to estimate the parameters. These margins must be specified when using a statistical computing package. We will discuss these margins at length shortly.

The log likelihood function for model (15.3) assumes independence of dyads (an assumption we discuss at the end of this chapter) and is as follows (Wasserman and Iacobucci 1986):

$$\begin{aligned} & \sum_{i,j} \lambda_{ij} + \frac{1}{2} \sum_k \theta_k y_{++k+} + \frac{1}{2} \sum_l \theta_l y_{+++l} \\ & + \frac{1}{2} \sum_i \sum_k \alpha_{i(k)} y_{i+k+} + \frac{1}{2} \sum_j \sum_l \alpha_{j(l)} y_{+j+l} \\ & + \frac{1}{2} \sum_j \sum_k \beta_{j(k)} y_{+jk+} + \frac{1}{2} \sum_i \sum_l \beta_{i(l)} y_{i++l} \\ & + \sum_k \sum_l (\alpha\beta)_{kl} \sum_{i < j} y_{ijkl}. \end{aligned} \quad (15.4)$$

The terms in the log likelihood function that depend on the data y have subscripts which include “+”’s.

The margins of y which are needed to estimate the parameters are those arising in the log likelihood function (15.4). These margins are *sufficient statistics* for the parameters in the model. These are the only summaries of the data (y) that are needed to maximize the likelihood and obtain the parameter estimates. We later describe these margins at length.

Maximum likelihood estimation not only yields estimates of the parameters of a model (such as p_1), but also estimates of the random

variables being modeled. In this instance, we are modeling the elements of the Y -array; and we will denote the maximum likelihood estimates of these quantities using “^”’s: $\{\hat{Y}_{ijkl}\}$. These maximum likelihood estimates, which are often referred to as *fitted values*, usually do not have a simple equation to calculate them. They are calculated iteratively by the algorithm chosen to solve the likelihood equations arising from the maximum likelihood problem.

Goodness-of-Fit Statistic. Once we have fit model (15.3) to the data and obtained maximum likelihood estimates of the parameters, we need to be able to evaluate how well the model fits and which of the parameters are statistically large (that is, statistically different from 0). The statistic used in maximum likelihood estimation for these purposes is usually the *likelihood-ratio statistic*. For the log-linear models discussed here, this statistic is G^2 , which, for our application, is:

$$G^2 = 2 \sum_{i < j} \sum_{k,l} y_{ijkl} \log (y_{ijkl}/\hat{y}_{ijkl}). \quad (15.5)$$

This statistic is well-known in categorical data analysis; further. The likelihood-ratio statistic is a function of the observed data, y_{ijkl} , and the calculated model predictions (fitted values), \hat{y}_{ijkl} , which are standard output from log-linear model computing packages. These fitted values are the predictions for the network data arising from the model p_1 . The parameter estimates can also be combined, as stated by the model, to give us the fitted values (the \hat{y} ’s). The statistic is calculated as the sum over all cells in a contingency table of quantities that compare the observed values to the fitted values.

The fitted values, \hat{y}_{ijkl} , for the friendship relation for the fabricated network of second-graders are given in Table 15.5.

The goodness-of-fit statistic for our statistical models for social network data (in 15.5) has a similar interpretation. The only difference is that the sum is not taken over all cells of the y -array, but over all dyads ($i < j$), since the basic unit in these models is the dyad. We want to compare each dyad to its fitted value only once, so we sum over one-half of the y -array (remember that y is symmetric).

The statistic G^2 , defined in equation (15.5), has the property that as we add parameters to a model and make the model more complicated, it will stay the same or go down, indicating better fit, or closer agreement of the y ’s and \hat{y} ’s. This property is especially important when

we compare two hierarchically nested models. For example, a model with parameters λ, θ, α , and β is hierarchically nested in a model with parameters $\lambda, \theta, \alpha, \beta$, and $(\alpha\beta)$. The models are identical except that one model contains one more set of parameters (the $\{(\alpha\beta)_{kl}\}$). The "larger" model has all the parameters of the first, plus an additional set. All of the models we consider here will be *hierarchical* — if one model can be obtained from a second model by setting some of the parameters to zero, then we say that the first model is hierarchically nested within the second.

In general, we can test hypotheses about the model parameters by comparing the test statistics for hierarchically nested models. The difference between the G^2 statistics for two hierarchically nested models is approximately asymptotically distributed as a χ^2 random variable. Thus, this difference can be compared to tabled values of χ^2 (with certain degrees of freedom) to establish whether the model with more parameters fits significantly better than the simpler model. If it does, the extra parameters are statistically important. We can compare the fit of model (15.3) to the fit of a model that is similar except that it contains no α parameters, for example. The difference between the two G^2 's would be tested on $(g - 1)(C - 1)$ degrees of freedom in order to evaluate the statistical significance of the $\{\alpha_{ik}\}$.

UCINET IV can fit p_1 . One can also use a standard statistical computing package, such as *GLIM*, *SPSS^X*, or *SYSTAT*, to fit the model and to calculate the fitted values, $\{\hat{y}_{ijkl}\}$, which are used to compute the goodness-of-fit statistic G^2 . It is important to note that the G^2 and degrees of freedom obtained directly from such standard packages are incorrect, and need modifications that we describe shortly. The fitted values and the residuals (the differences between the observed table entries and the fitted table entries $\{y_{ijkl} - \hat{y}_{ijkl}\}$) can also be diagnostic of cells that are fit particularly well or particularly badly. Holland and Leinhardt (1981) discuss how to use the residuals from fitting p_1 to network data. Particularly poorly fit cells or dyads may be of special concern to a researcher. One should investigate possible reasons for the large discrepancies. When models do not fit data well, one means of improving the fit is to add parameters to the model. For example, some more complicated models have been investigated by Wasserman and Galaskiewicz (1984), Wasserman and Anderson (1987), and Wang and Wong (1987), who add terms that depend on the blocks or positions into which actors can be classified (more on this in Chapter 16).

Table 15.5. y fitted values for p_1 fit to the second-grade children

		Allison	Drew	Eliot	Keith	Ross	Sarah
		x_{ij}	x_{ji}	$x_{ji} : 0$	$x_{ji} : 1$	n_j	n_i
Allison	0	—	0.30	0.00	0.78	0.00	1.00
	1	—	0.70	0.00	0.22	0.00	0.00
Drew	0	0.30	0.70	—	0.37	0.13	0.59
	1	0.00	0.00	—	0.06	0.45	0.00
Eliot	0	0.78	0.22	0.37	0.06	—	0.92
	1	0.00	0.00	0.13	0.45	—	0.00
Keith	0	1.00	0.00	0.59	0.00	—	—
	1	0.00	0.00	0.41	0.00	0.08	0.00
Ross	0	0.46	0.54	0.21	0.13	0.77	0.14
	1	0.00	0.00	0.04	0.62	0.02	0.08
Sarah	0	0.46	0.54	0.21	0.13	0.77	0.14
	1	0.00	0.00	0.04	0.62	0.02	0.08

Practical Guide to Fitting These Models to Data. Once the data are arranged into a y -array, model (15.3) can be fit (and parameters estimated) using the log-linear modeling procedures that are contained in any of the major statistical computing packages.

Comments on parameter estimation and model fitting using several of the major packages, including *SPSS^X*, *BMDP*, *GLIM*, and others can be found in the literature (Fienberg and Wasserman 1981b; Wasserman and Weaver 1985; Wasserman and Iacobucci 1986).

Fitting the Models. Maximum likelihood estimation of the parameters of log linear models for categorical data is discussed at length in Appendix II of Fienberg (1980). Specifically, each set of parameters has associated with it sufficient statistics, which are margins of the data array being modeled. The maximum likelihood equations, which must be solved to obtain the maximum likelihood estimates of the expected values of the cell counts (or the “fitted” values), all have the form: observed margin = fitted margin. Thus, the appropriate model parameters will be estimated, and the fitted values obtained, when we constrain the fitted margins to be equal to the observed margins. Philosophically, this is like saying, if forty men and sixty women responded to a sample survey, any modeling of the data that includes the variable gender must still reproduce the given figures of 40 and 60. Or, referring to network data, if we include β 's in our model, and if actor 1 is chosen by four of the eight actors, then fitted probabilities that n_1 is chosen must sum to $(4/8) = 0.50$. This equating is the critical computation that will produce maximum likelihood estimates of cell expected values, and parameters.

Our models for single relational network data can be fit by following the theory discussed above: specifically, one focuses on the model parameters, and their sufficient statistics, which are margins of the Y -array. All six of the two-dimensional margins of the four-dimensional Y -array are the sufficient statistics for the parameters in (15.3). The same is true for p_1 , since it is equivalent to the basic model when $C = 2$.

The four variables that define Y comprise the four dimensions of the contingency table to be modeled. We have been referring to these variables and their subscripts as:

- (i) Sending actor (i)
- (ii) Receiving actor (j)
- (iii) Strength of choice made ($k = X_{ij}$)
- (iv) Strength of choice received ($l = X_{ji}$)

We label these variables as 1, 2, 3, 4, respectively. In order to fit model (15.3), we fit the following log-linear model to a Y -array:

$$[12][13][24][23][14][34]. \quad (15.6)$$

The numbers in brackets in (15.6) are the margins of Y , which are sufficient statistics for the parameters in the basic model. This notation is common for log-linear models, and was first introduced by Fienberg (1980), based on the notation in Bishop, Fienberg, and Holland (1975). We will call it *Fienberg* notation.

A collection of variable numbers in brackets implies that we have included the parameter associated with the sufficient statistics given by the variable numbers in a model. Because the log-linear models considered here are hierarchical, all lower-order terms (for example, main effects) are also necessarily fit. For example, [12] is the interaction between variables 1 and 2 (that index initiating and receiving actors) and corresponds to the margin $\{Y_{ij++}\}$. This margin includes the 1-dimensional margins for these two variables $\{Y_{i+++}\}$ and $\{Y_{+j++}\}$.

The correspondence between the margins and the parameter sufficient statistics is as follows. The [12] margin is fit in order to include the $\{\lambda_{ij}\}$ parameters in the model. Because the λ parameters must be included in all our models in order to constrain the probabilities properly, the margin [12] must always be included.

Margins [13] and [24] are sufficient statistics for the $\{\alpha_{i(k)}\}$ parameters. There is only one set of $(g-1)(C-1)\{\alpha_{i(k)}\}$ expansiveness parameters, but there are two sets of margins, [13] and [24]. These margins are actually equal, due to the symmetry of the Y -array, and thus do not produce superfluous parameters. Both margins [13] and [24] must be included in any model for which we desire estimates of the alpha's.

Similarly, margins [23] and [14] are sufficient statistics for the $\{\beta_{j(l)}\}$ parameters. The margin [23] is equal to the margin [14], because of the symmetry of the Y -array. Both must be included in any model for which we desire estimates of the beta's.

The final margin, [34], is the sufficient statistic for the reciprocity parameters, $\{(\alpha\beta)_{kl}\}$. The only remaining parameters given in model (15.3) are the elements of the set of $\{\theta\}$'s. The sufficient statistics for these parameters are the margins [3] and [4], but since our models are hierarchical, these are automatically “fit” whenever higher-order margins [13] and [24], or [23] and [14], or [34], are fit.

Infinite Parameters. One problem in this modeling and computation process can arise when an actor has an indegree or outdegree equal to either 0 or $(g - 1)$. In the first instance, the actor neither sends nor receives any ties; while in the second, the actor either has ties to or from every other actor. We discuss each of these situations in turn. When a relation is valued, the problem also arises when all ties sent (or received) by a particular actor have exactly the same strength.

Let us focus just on dichotomous relations. Suppose first that an actor has an indegree equal to 0. Statistically, this actor has no tendency to receive ties. Thus, the fitted probabilities that this actor receives ties must all be 0's. If an actor has an outdegree equal to 0, this actor has no tendency to send ties. Thus, the fitted probabilities that this actor makes choices must all be 0's.

If a fitted probability is to be 0, then the logarithm of this probability is $-\infty$. And to make this logarithm $-\infty$, the appropriate parameter must also be $-\infty$. A 0 row sum in the sociomatrix forces the associated alpha for this actor to be $-\infty$. A zero column sum in the sociomatrix forces the associated beta for this actor to be $-\infty$. For example, two children in the fabricated second-grader network received no friendship nominations; hence, there are two betas fitted with p_1 equal to $-\infty$.

Now suppose that an actor has an indegree equal to $(g - 1)$. Statistically, this actor is certain to receive ties from all other actors. Thus, the fitted probabilities that this actor receives any ties must all be 1's. If an actor has an outdegree equal to $(g - 1)$, this actor is certain to send ties to all other actors. Thus, the fitted probabilities that this actor sends ties must all be 1's.

If a fitted probability is to be unity, then the logarithm of this probability is ∞ . And to make this logarithm infinity, the appropriate parameter must also be ∞ . A row sum in the sociomatrix equal to $(g - 1)$ forces the associated alpha for this actor to be ∞ . A column sum equal to $(g - 1)$ forces the associated beta for this actor to be ∞ .

Lastly, we note that these infinite parameters are not counted when tallying up degrees of freedom. For the second-graders, we have a full set of alphas, and hence 5 degrees of freedom for these parameters, but only estimate four betas. Hence, there is only 3 degrees of freedom for the betas. Further, the infinite parameters are not considered when centering the finite parameters to sum to 0. Note once again from the second-graders that the four finite estimated betas sum to 0.

Model Comparisons and Statistical Tests for Parameters. We have presented (15.3) with three sets of meaningful parameters — representing effects of expansiveness, popularity, and reciprocity. We can conduct hypothesis tests to determine whether the actors demonstrate any or all of these effects on a relation. We can postulate and subsequently fit alternative, simpler models that might be able to reproduce the network data just as well, or almost as well, as the more complicated model (15.3). Comparing the fit of (15.3) to simpler versions of the model allows us to determine whether each set of parameters is statistically different from 0.

We use likelihood-ratio hypothesis tests to study each of the sets of parameters. The test statistic for these hypotheses compares the goodness-of-fit statistic of model (15.3) to the goodness-of-fit statistics of the simpler models. These statistics are all likelihood-ratio statistics, where the $\{\hat{y}_{ijkl}\}$ fitted values are computed from different competing models.

To be more concrete, there are seven models, shown below, that are simpler than (15.3), derived by dropping one or more sets of the parameters $\{\alpha_{i(k)}\}$, $\{\beta_{j(l)}\}$, or $\{(\alpha\beta)_{kl}\}$ from the model. A model is more *parsimonious* than (15.3) if it contains fewer parameters. Model (15.3) is listed again, and then the seven simpler models follow:

$$\begin{aligned}
 (15.3) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} \\
 &\quad + \beta_{j(k)} + \beta_{i(l)} + (\alpha\beta)_{kl} \\
 a) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + \beta_{j(k)} + \beta_{i(l)} \\
 &\quad + (\alpha\beta)_{kl} \\
 b) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} \\
 &\quad + (\alpha\beta)_{kl} \\
 c) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} \\
 &\quad + \beta_{j(k)} + \beta_{i(l)} \\
 d) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} \\
 e) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + \beta_{j(k)} + \beta_{i(l)} \\
 f) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l + (\alpha\beta)_{kl} \\
 g) \quad \log P(Y_{ijkl} = 1) &= \lambda_{ij} + \theta_k + \theta_l
 \end{aligned} \tag{15.7}$$

In Table 15.6 we list these models, (15.3) and (15.7a) — (15.7g), along with the list of margins that specify each model. Recall from the definition of p_1 that if the relation under study is dichotomous, these parameters

Table 15.6. p_1 parameters, models, and associated margins.

Model	Parameters included in the model	Margins
15.3	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_{ik}\}, \{\beta_{jl}\}, \{(\alpha\beta)_{kl}\}$	[12][13][24][23][14][34]
15.7a	$\{\lambda_{ij}\}, \{\theta_k\}, \{\beta_{jl}\}, \{(\alpha\beta)_{kl}\}$	[12][23][14][34]
15.7b	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_{ik}\}, \{(\alpha\beta)_{kl}\}$	[12][13][24][34]
15.7c	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_{ik}\}, \{\beta_{jl}\}$	[12][13][24][23][14]
15.7d	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_{ik}\}$	[12][13][24]
15.7e	$\{\lambda_{ij}\}, \{\theta_k\}, \{\beta_{jl}\}$	[12][23][14]
15.7f	$\{\lambda_{ij}\}, \{\theta_k\}, \{(\alpha\beta)_{kl}\}$	[12][34]
15.7g	$\{\lambda_{ij}\}, \{\theta_k\}$	[12][3][4]

appear in the models only when choices are made (that is, for example, $\alpha_i(0) = 0$ and $\alpha_i(1) = \alpha_i$).

Consider one of these models in more detail. In model (15.7b), there are no β 's. This model assumes that all β 's equal 0; thus, there are no differential popularity effects among the actors. By definition, we would expect this model to fit single relational social network data reasonably well only when the popularity of the actors is constant. If it were true that all β 's were 0, we would expect (15.7b) to "fit" the observed dyadic interactions just as well as the full model (15.3). To test the null hypothesis,

$$H_0 : \beta_{jl} = 0 \text{ for all } j \text{ and } l,$$

we compare the goodness-of-fit statistics for the two models (15.7b) and (15.3).

Similarly, a comparison of the fit statistics for model (15.7a) or (15.7c) to (15.3) would test the null hypotheses

$$H_0 : \alpha_{ik} = 0 \text{ for all } i \text{ and } k,$$

or

$$H_0 : (\alpha\beta)_{kl} = 0 \text{ for all } k \text{ and } l, \text{ respectively.}$$

Model (15.7b) is hierarchically nested in model (15.3), so that the first model is a special case of the second. Thus, the fit statistic for (15.7b) will be equal to or larger than the fit statistic for (15.3). The comparison of G^2 statistics indicates whether the parameters are statistically different from zero. The difference between the two G^2 's is a *conditional likelihood-ratio test statistic*, testing a null hypothesis that the parameters are indeed equal to 0. We statistically condition on the more complicated model and test to see whether we can simplify it by "dropping" terms.

In practice, to test these conditional hypotheses regarding sets of parameters, we take the G^2 for model (15.3), and subtract it from the G^2 for the model that does *not* contain the parameters being tested. For example, to test whether the β 's are 0, we calculate two G^2 's, one for the full model, and one for (15.7b). We get a difference score (G^2 for (15.7b) minus the G^2 for model (15.3)), a ΔG^2 that quantifies whether the additional parameters are significantly different from 0. This new statistic, ΔG^2 , is approximately asymptotically distributed as a χ^2 random variable, with a certain number of degrees of freedom. Just as the statistic ΔG^2 is equal to the difference in G^2 for the two models, the degrees of freedom, Δdf , is equal to the difference in degrees of freedom for the two models. The model that contains more parameters is the alternative hypothesis model (H_A), and the model with fewer terms is the null hypothesis model (H_0). The Δdf is equal to the number of independent parameters being tested.

For example, when evaluating the null hypothesis, $H_0 : (\alpha\beta)_{kl} = 0$ (for all k, l), we reject H_0 if ΔG^2 is greater than the upper percentage point of the χ^2 distribution with $C(C - 1)/2$ degrees of freedom. If ΔG^2 is statistically large compared to the χ^2 distribution on $C(C - 1)/2$ degrees of freedom, we would conclude that the $(\alpha\beta)_{kl}$ parameters are statistically large and should be included in any model truly descriptive of the given network.

We should note that one reason the "asymptotic" distributions of G^2 or ΔG^2 are only approximations to χ^2 is due to the difficulty in achieving the "asymptotic" property itself. As the number of actors in a set of actors, g , increases, so do the number of α_{ik} and β_{jl} parameters to be estimated ($(g - 1)(C - 1)$), so even at the asymptote (g tending toward infinity), the number of parameters in p_1 and model (15.3) also gets infinitely large. We comment on these problems a bit later in this chapter.

The tests that can be conducted for parameters that model a single relation are listed in Table 15.7. The first test listed in Table 15.7 compares the fit of model (15.7a) with the fit of model (15.3). Model (15.7a) has no α 's, so in conducting this comparison, we are testing the null hypothesis, $H_0 : \alpha_{ik} = 0$ for all i, k . The degrees of freedom for this test is $(g - 1)(C - 1)$. If ΔG^2 is not large, we cannot reject the null hypothesis, which means the actors are statistically equivalent with respect to their expansiveness (sensible perhaps in networks in which the number of ties is restricted by design to a constant value). The second test is the counterpart for testing the null hypothesis, $H_0 : \beta_{jl} = 0$ (for all partners j and all strengths l).

Table 15.7. Tests of significance for parameters in model (15.3)

Test	ΔG^2	Δdf	Null hypothesis
1	15.7a - 15.3	$(g-1)(C-1)$	$H_0 : \alpha_{i(k)} = 0$ for all i, k
2	15.7b - 15.3	$(g-1)(C-1)$	$H_0 : \beta_{j(l)} = 0$ for all j, l
3	15.7c - 15.3	$C(C-1)/2$	$H_0 : (\alpha\beta)_{kl} = 0$ for all k, l

The second test in Table 15.7 compares the fits of models (15.7b) and (15.3). It is a test of the null hypothesis, $H_0 : \text{all } \beta_{j(l)} = 0$. If ΔG^2 is large, we would conclude that the actors exhibit differential popularity, thus, the $\{\beta_{j(l)}\}$ parameters need to be included in any model that would adequately represent the data. If the statistic is not large, we would conclude the data may be described adequately without including effects for popularity.

The alpha, beta, and reciprocity parameters are independent sets of parameters in the sense that for any particular set of actors with measurements on a single relation, any of the three effects may be statistically large. Knowing the characteristics of the $\hat{\alpha}$'s for some network data set, for example, does not inform us about the characteristics of the $\hat{\beta}$'s, or the reciprocation parameter(s) present in that same network.

Examples. We have fit model (15.3) and its relatives, models (15.7a), (15.7b), and (15.7c), to a variety of data sets. We first demonstrate the model fits for the fabricated network of second-grade children, and then discuss the analysis of Krackhardt's friendship and advice relations at length. The second-graders and Krackhardt's network both consist of dichotomous relations, so we fit p_1 directly. Freeman's EIES communications network, which we also would like to model, has valued relations; thus, we must fit the more general model to it.

Fabricated Network of Second-Graders. Table 15.8 gives the goodness-of-fit statistics for each of the four models applied to the dichotomous friendship relation for the second-graders.

The information in this table indicates which parameters are statistically important for these fabricated data. We see that the expansiveness parameters are not very important (ΔG^2 is quite small). Knowing this, we would not bother to study the individual actor α_i parameter estimates in an attempt to distinguish among the children. On the other hand, the popularity parameters (the β 's) are of interest in this network because

15.1 Single Directional Relations

Table 15.8. Goodness-of-fit statistics for the fabricated network

Second-graders — friendship at the beginning of the year

Model	Null hypothesis	G^2	ΔG^2	Δdf
15.3		20.630		
15.7a	$H_0 : \alpha_{i(k)} = 0$ for all i and k	23.072	2.442	3
15.7b	$H_0 : \beta_{j(l)} = 0$ for all j and l	31.956	11.326	5
15.7c	$H_0 : (\alpha\beta) = 0$	22.564	1.934	1

Table 15.9. Goodness-of-fit statistics for Krackhardt's network

Krackhardt's high-tech managers — advice relation

Model	Null hypothesis	G^2	ΔG^2	Δdf
15.3		322.564		
15.7a	$H_0 : \alpha_{i(k)} = 0$ for all i and k	506.778	184.214	20
15.7b	$H_0 : \beta_{j(l)} = 0$ for all j and l	440.471	117.907	20
15.7c	$H_0 : (\alpha\beta) = 0$	339.632	17.068	1

Krackhardt's high-tech managers — friendship relation

Model	Null hypothesis	G^2	ΔG^2	Δdf
15.3		288.303		
15.7a	$H_0 : \alpha_{i(k)} = 0$ for all i and k	421.965	133.662	20
15.7b	$H_0 : \beta_{j(l)} = 0$ for all j and l	337.068	48.765	20
15.7c	$H_0 : (\alpha\beta) = 0$	312.455	24.152	1

$\Delta G^2 = 31.956 - 20.630 = 11.326$, which is large. We reported and discussed these parameter estimates earlier in the chapter (Table 15.4). Also noteworthy here is the fact that reciprocity does not seem to be a strong force for this friendship relation.

Krackhardt's Network. Now look at the analyses of Krackhardt's data, shown in Table 15.9. We analyzed both the advice relation and the friendship relation (separately). All three sets of parameters appear to be large when modeling either of the relations.

The α and β estimates are listed in Table 15.10. The expansiveness parameters indicate that, for the advice relation, actors 3, 5, and 18 are likely to give advice, while actors 2, 6, and 12 are unlikely to. Note also that actor 15 gives advice to all other actors. The popularity parameters for this relation indicate that actor 2 is very likely to receive advice, while actors 9 and 15 are very unlikely to receive advice. For friendship, the

Table 15.10. Parameter estimates for Krackhardt's high-tech managers

Actor	Advice		Friendship	
	$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\alpha}_i$	$\hat{\beta}_i$
n_1	-0.98	1.75	-0.37	1.40
n_2	-2.64	4.54	-1.46	2.36
n_3	2.63	-2.09	-1.36	0.66
n_4	1.40	-0.63	0.41	0.03
n_5	2.63	-2.09	0.60	0.34
n_6	-3.75	1.10	0.86	-1.84
n_7	-0.28	1.56	$-\infty$	0.04
n_8	-0.01	0.42	-2.22	0.83
n_9	2.02	-2.44	$-\infty$	1.39
n_{10}	1.95	-0.44	1.29	-3.17
n_{11}	-2.20	1.25	2.39	-0.41
n_{12}	-2.47	-0.08	-0.76	1.53
n_{13}	-0.13	-1.79	-0.58	-2.30
n_{14}	-1.58	0.79	-1.36	0.66
n_{15}	$+\infty$	-2.91	1.16	-0.74
n_{16}	-1.38	0.07	-1.18	0.19
n_{17}	-1.03	0.33	4.59	-0.96
n_{18}	2.55	1.48	-2.04	0.38
n_{19}	1.42	-2.27	1.33	-0.36
n_{20}	1.40	-0.63	-0.99	0.38
n_{21}	0.46	2.07	-0.33	0.33

expansiveness parameters indicate that actor 17 has many other actors as friends, while actors 7 and 9 (who make no choices) do not. Actor 2 is a relatively popular friend, while actor 10 is not. Note the dual role played by actor 2. This actor gives very little advice, but is a friend to many other actors.

The reciprocity parameters for the advice ($(\hat{\alpha}\hat{\beta})=2.233$) and friendship ($(\hat{\alpha}\hat{\beta})=2.937$) relations are both positive and large (relative to the actor-level parameters), indicating substantial tendencies toward mutuality.

The α and β estimates for the advice relation can also be studied by using attribute information on the tenure of each actor (the number of years of service in the organization). One way to use such information is to correlate the parameter estimates with this attribute variable.

The $\hat{\alpha}_i$'s are negatively correlated with tenure while the $\hat{\beta}_j$'s are positively correlated, indicating that more experienced workers seek advice less frequently than others and are the source of advice more frequently than others. Tenure has a very small correlation with the $\hat{\alpha}_i$'s and the $\hat{\beta}_j$'s for the friendship relation.

15.1 Single Directional Relations

We note that there are other ways to incorporate attribute variables directly into our statistical models. We discuss another approach in the next section.

Other Analyses. In Table 15.8, we presented test statistics for the second-graders that told us (in general) whether the α , β , and $(\alpha\beta)$ parameters were statistically large. We could stop our statistical analyses at this point. For example, for the second-graders, we would summarize by saying that the network structure indicates strong differential popularity effects among the actors and that there is a small tendency for reciprocated friendship ties. Alternatively, we might proceed to study these results in more detail, by conducting other analyses to learn, for example, exactly which actors are significantly more or less popular than which others. To do so, we can use the methods reported in Wasserman and Weaver (1985), based on statistical theory for log-linear models for categorical data (see Bishop, Fienberg, and Holland 1975). These methods involve contrasts, as in the analysis of variance.

For example, the magnitude of each individual $\beta_{j(l)}$ parameter may be studied in turn, by calculating

$$t = (\hat{\beta}_{j(l)} - \beta^*) / S.E.(\hat{\beta}_{j(l)}),$$

where β^* is an a priori constant, specified in the null hypothesis (often $\beta^* = 0$), and $S.E.(\hat{\beta}_{j(l)})$ is the standard error of the parameter estimate, a value obtained as a function of information reported by statistical computing packages.

15.1.4 \otimes Is p_1 a Random Directed Graph Distribution?

Given our discussion of random directed graph distributions in Chapter 13, one important question is how p_1 compares to these distributions. Specifically, is p_1 a conditional uniform distribution for random graphs?

The answer to this question is yes, with a bit of explanation needed. First, note that p_1 is an exponential family of distributions, with minimal sufficient statistics consisting of the indegrees, outdegrees, and the number of mutual dyads. Thus, any two random digraphs with equal values of these statistics have identical fitted values from p_1 . Thus, if we condition on specific values of $\{X_{i+}\}$, $\{X_{+j}\}$, and M , all random digraphs with these values have exactly the same probability of occurring. Thus, p_1 is identical to the random digraph distribution $U|\{X_{i+}\}, \{X_{+j}\}, M$, as discussed by Snijders (1991b).

But, p_1 is also a Bernoulli digraph distribution (see Karoński 1982; Palmer 1985; Bollobás 1985; Frank 1985, 1989; and, of course, Chapter 13). Such distributions focus on a set of p_{ij} probabilities, specifying the probability that i chooses j , for all pairs of actors i and j . Since digraphs are representations of dichotomous relations, one can define θ_{ij} as the logit for p_{ij} :

$$\theta_{ij} = \log \left(\frac{p_{ij}}{1 - p_{ij}} \right).$$

If we assume that all ties are statistically independent (as with a general Bernoulli distribution), then the likelihood function can be written down.

Further, one can postulate an additive model for the logits, quite similar to p_1 . The main difference between such a logit model and p_1 is that the off-diagonal elements of the sociomatrix, x_{ij} and x_{ji} , are not assumed by p_1 to be statistically independent. Indeed, the basic modeling unit of p_1 is the dyad, not the individual tie or arc. Thus, one can view p_1 as a Bernoulli dyad distribution, rather than a Bernoulli arc distribution, where all dyads, not ties or arcs, are assumed to be independent.

15.1.5 Summary

We have presented a class of models designed to study how ties from a single relation vary across actors. The class is comprehensive, and allows one to posit substantively meaningful parameters, all of which can be estimated and studied via significance tests.

But, there are a number of statistical problems with this class. These problems are caused by the lack of an asymptotic statistical theory — the number of parameters does not remain constant as the number of actors in the set increases. Even so, we note two facts. First, as g increases, the number of $(\alpha\beta)_{kl}$ parameters to be estimated ($C(C - 1)/2$) remains constant. Thus, the test of reciprocity is not affected by the main statistical problem, and we can be confident that the asymptotic properties hold for this test. Secondly, we note that these problems are non-existent when incorporating attributes of actors into the models (which we describe in detail shortly). For now we note that subsets of actors can be formed on the basis of their attributes, such as gender (male and female). Then as g increases, the number of subsets remains fixed. Thus, if parameters depend on subset memberships, rather than the individual actors themselves, we can be confident that asymptotic distributions hold. We note that these uncertainties regarding the

asymptotic distribution of G^2 force us to evaluate hypothesis tests more tentatively. For example, we do not attach p -values to these comparisons in general, but can do so when the modeling includes attribute variables.

Fortunately, one does not need an asymptotic statistical theory for the significance tests described here. There is an alternative approach. One can use permutation tests to compare matrices of fitted values, as discussed by Hubert and Baker (1978). Specifically, one can determine how close an observed data matrix is to a predicted, fitted matrix, by looking at a large number of permutations of the original data matrix, and comparing each to the “target.” Details on such tests, as applied to social network data, can be found in Baker and Hubert (1981) and Krackhardt (1987b). We apply such technology to blockmodels in Chapter 16.

As a brief aside, we mention some researchers who have investigated models more complicated than (15.3), such as stochastic blockmodels (Wasserman and Anderson 1987; Wang and Wong 1987). Fienberg and Wasserman (1981a) also extended p_1 to allow for “differential reciprocity”; the $\{(\alpha\beta)_{kl}\}$ parameters are replaced with $\{(\alpha\beta)_{i(kl)}\}$ parameters, allowing the rates of reciprocation to depend on the sending or receiving actors. (The model that adds these parameters includes the margins: [12][134][234].) Further generalizations of this idea can be found in Wasserman and Galaskiewicz (1984).

15.2 Attribute Variables

We now wish to analyze data on the attributes for the g actors contained in \mathcal{N} in conjunction with the relation measured between actors. We thus build models that use both network composition and network structure. We first discuss why information about the characteristics of the individual actors is helpful in modeling the relational information. We then modify our Y -array into a new contingency table (the W -array) that incorporates such attribute information. Once again, we use the small fabricated network of second-graders to illustrate. We will then modify the statistical model (15.3) designed for a single relational variable to allow for the simultaneous analysis of one or more actor attribute variables. Lastly, we offer more examples.

15.2.1 Introduction

Studying \mathbf{Y} using statistical models for dyadic interactions is one way to summarize the information contained in a single relational variable measured on the $\binom{8}{2}$ pairs of actors from the set \mathcal{N} . In some applications, we might be interested in studying the patterns of dyadic interactions within and between subsets of persons who share similar characteristics. For example, we might be interested in understanding how a classroom of thirty children interacts. We would create \mathbf{Y} based on a relation among the thirty actors. However, even simple theories might predict that boys and girls differ in the way they interact with others. For example, we might at least suspect that in a classroom of kindergarteners, children would be most likely to nominate as friends children of their same gender. We might be less interested in modeling individual differences in the interactions of thirty children than in understanding how gender affects the childrens' interactions. Thus, we might be led by theoretical concerns to model subset differences — the interactions of these thirty children in conjunction with the attribute of gender. We would form two subsets, boys and girls, and study the between- and within-subset interactions.

Attribute variables can be of interest in many different social networks. For example, network researchers might use variables such as gender, age, and socioeconomic status. Researchers studying dyadic interactions in married couples might suspect that communication patterns depend on attributes such as gender of the speaker, or whether the couple was in marital distress (Gottman 1979a). Researchers interested in modeling interorganizational behavior might wish to include the size of the organization or its industry as predictive attribute variables (Galaskiewicz 1979; Galaskiewicz and Marsden 1978; Wasserman and Galaskiewicz 1984).

Before we proceed with the methodology, it is important to note that the attribute variables, and hence the categorization of actors into subsets, must be chosen independently of the relation under study. Subsets must not be formed using the relational information; otherwise, the error rates of the tests discussed here will be affected. The use of attribute variables described here is *a priori*; *a posteriori* analyses will be described in Chapter 16.

For the small network of second-graders nominating friends, we use the attribute of age to distinguish between the 7- and 8-year-olds. We might hypothesize that children prefer to play with others of the same

15.2 Attribute Variables

age, so 7-year-olds would choose other 7-year-olds as friends, 8-year-olds would choose 8-year-olds, but few 7-year-olds would have 8-year-olds as friends. Perhaps a competing hypothesis would be that children gravitate toward older children, so that the 7-year-olds nominate other 7-year-old children as well as 8-year-old children, whereas the 8-year-old children only nominate each other.

In our analysis of our other examples, we use the following attribute information: for Krackhardt's management network, there are two attribute variables of interest to us: age of actors, and seniority or tenure of actors in the organization. We categorized the employees into four subsets on the basis of their age and tenure with the organization:

- (i) Those with 10 or fewer years of tenure and younger than 40 years of age
- (ii) Those with 11+ years of tenure and younger than 40
- (iii) Those with 10 or fewer years of tenure and older than 40 years of age
- (iv) Those with 11+ years of tenure and older than 40

15.2.2 The W-array

Attribute variables were first introduced into models based on Holland and Leinhardt's p_1 by Fienberg and Wasserman (1981a). Individuals are placed into subsets using relevant actor characteristics, and actors within a subset are assumed to behave similarly (at least with regard to the measured relations being modeled). Specifically, p_1 (or model (15.3)) is postulated, and it is assumed that all actors within a subset have equal parameters; for example, if we partition actors based on gender, then every male actor shares a common expansiveness parameter α and a common popularity parameter β (as would the females). We give a formal mathematical definition of this assumption below.

Stochastic Equivalence. This assumption of comparable behavior of actors within subsets (with respect to some statistical model) has been viewed as a generalization of structural equivalence. Wasserman and Weaver (1985) termed this assumption *stochastic equivalence*, and Wasserman and Anderson (1987) developed this idea in more detail. Holland, Laskey, and Leinhardt (1983) and Wang and Wong (1987) proposed alternative versions of stochastic actor equivalence, all using generalizations of p_1 for their statistical models. We will discuss them in

more detail in Chapter 16, along with the idea of a *stochastic blockmodel*, which follows naturally from the notion of stochastic equivalence.

Based on the assumption of equal parameters within subsets, our basic model can be greatly simplified. Rather than postulating and estimating separate expansiveness and popularity parameters for each individual actor, a single expansiveness parameter and a single popularity parameter is assumed for each subset. This simplification, enabled by the assumption of stochastic equivalence, allows us to estimate far fewer parameters.

Definition. A set of attribute variables is used to classify each of the g actors into one of S subsets. The size of S depends on how many attribute variables are used, and how many levels the individual attributes take on. With just a single, dichotomous attribute (such as gender) we need only $S = 2$ subsets. In other social network data sets, we may have more attribute variables with considerably more levels. For example, gender ("males" and "females") crossed with race ("Whites," "Blacks," and "Asians") would result in $S = 6$ subsets. We denote the subset to which the sending actor (i) belongs by $s(i)$, and the subset in which the receiving actor (j) belongs by $s(j)$. Thus, we have a mapping function $s(\bullet)$, where this function is defined a priori (based on attribute variables) and not on the basis of relational information.

We use subscripts for parameters which index the actors, and superscripts for subsets of actors. Thus, we will replace $\alpha_{i(k)}$ and $\beta_{j(l)}$ in model (15.3) by $\alpha_k^{[s(i)]}$ and $\beta_l^{[s(j)]}$, respectively, where the superscript indexes the subset modeled by that parameter.

The assumption of stochastic equivalence may be stated more precisely with this new notation. We assume that:

$$\alpha_{i(k)} = \alpha_k^{[s(i)]}$$

for all $n_i \in$ subset $s(i)$, and

$$\beta_{j(l)} = \beta_l^{[s(j)]}$$

for all $n_j \in$ subset $s(j)$.

We modify \mathbf{Y} so that all actors with similar characteristics are aggregated into a single subset. To accomplish this, we aggregate over the elements in \mathbf{Y} , collapsing over actors with identical attributes. In doing so, we obtain \mathbf{W} , with elements $\{W_{s(i)s(j)kl}\}$, defined formally as

$$W_{s(i)s(j)kl} = \sum_{i \in s(i)} \sum_{j \in s(j)} Y_{ijkl}. \quad (15.8)$$

15.2 Attribute Variables

That is, we simply sum the relational data, coded as the elements of \mathbf{Y} , over pairs (n_i, n_j) where actor i is a member of subset $s(i)$ and actor j is a member of subset $s(j)$.

The \mathbf{Y} -array is of size $g \times g \times C \times C$, but the \mathbf{W} -array can be much smaller. In general, the \mathbf{W} -array (the aggregated \mathbf{Y} -array), is of size $S \times S \times C \times C$, where S , the number of subsets, is usually much smaller than g . For example, in Krackhardt's network, twenty-one managers' friendship ($g=21$), aggregating on the basis of age (categorized into those younger than 40, and those older than 40) would give us only two subsets ($S = 2$). Thus, \mathbf{W} for this example is $2 \times 2 \times C \times C$. The number of levels of the relational variable, C (the different strengths the relation can take on), remains the same for both \mathbf{Y} and \mathbf{W} .

An Example \mathbf{W} -array. The \mathbf{W} -array is symmetric in the same way that \mathbf{Y} is symmetric, $W_{s(i)s(j)kl} = W_{s(j)s(i)lk}$. Unlike \mathbf{Y} , however, the entries in \mathbf{W} are usually not simply 0's and 1's, since they are sums of the 0's and 1's contained in the elements of \mathbf{Y} . The entries in \mathbf{W} are counts of the frequencies of the different dyadic states. In Table 15.11, we present the \mathbf{w} -array obtained by collapsing over the \mathbf{y} -array for the six second-graders, using the attribute of age. The children in the first subset are 7 years old (Eliot, Keith, and Sarah), and the children in the second subset are 8 years old (Allison, Drew, and Ross). This partitioning gives us two subsets of actors: $\{n_3, n_4, n_6\}$ and $\{n_1, n_2, n_5\}$. Note how the array is symmetric, and that entries in the table are integer counts, some of which are greater than 1. For example, consider the two values of "5" in this array. The lower-left entry of 5, cell (2,1,0,0), indicates there were five dyads with the following characteristics: The actor (i) was 8 years old and thus belonged in the second subset ($s(i) = 2$), the chosen partner (j) was 7 years old, and so belonged in the first subset ($s(j) = 1$), and neither actor chose the other — a null dyad. Due to the symmetry, these five null dyads also appear in cell (1,2,0,0), where the actor is a 7-year-old ($s(i) = 1$) and the partner is an 8 year-old ($s(j) = 2$).

Table 15.12 contains some of the \mathbf{w} -arrays for Krackhardt's network, for both of the relations, using one or two attribute variables (tenure in the organization and age of actor). The first tenure subset consists of the managers with ten or fewer years of service, and the second subset consists of the managers with more than ten years of service. We also dichotomize age into two subsets: those actors forty years and younger, and those actors older than forty years. We will model these data sets shortly, but for now simply note that these \mathbf{w} -arrays are of

Table 15.11. The W-array for the second-graders using friendship and age (the first subset consists of the 7-year-old children, Eliot, Keith, and Sarah, and the second subset consists of the 8-year-old children, Allison, Drew, and Ross.)

		$s(j) = 1$: 7-year-olds		$s(j) = 2$: 8-year-olds	
		$l: x_{ji} = 0$	$x_{ji} = 1$	$x_{ji} = 0$	$x_{ji} = 1$
$s(i) = 1$	$k: x_{ij} = 0$	6	0	5	1
	$x_{ij} = 1$	0	0	1	2
$s(i) = 2$	$x_{ij} = 0$	5	1	2	2
	$x_{ij} = 1$	1	2	2	0

more manageable sizes ($2 \times 2 \times 2 \times 2$ if we use one attribute variable or $4 \times 4 \times 2 \times 2$ if we use two attribute variables) than their corresponding y-arrays ($21 \times 21 \times 2 \times 2$ for both of the relations).

15.2.3 The Basic Model with Attribute Variables

The model we fit to this new contingency table (the W-array defined in equation (15.8)) is a special case of the basic model (15.3), subject to the constraints placed on the parameters (defined above), which arise through the use of the actor attribute variables and the assumption of stochastic equivalence. This version of model (15.3) follows:

$$\log P(Y_{ijkl} = 1) = \lambda_{ij} + \theta_k + \theta_l + \alpha_k^{[s(i)]} + \alpha_j^{[s(j)]} + \beta_k^{[s(j)]} + \beta_l^{[s(i)]} + (\alpha\beta)_{kl}, \quad (15.9)$$

where n_i is a member of subset $s(i)$ and n_j is a member of subset $s(j)$ (for $s(i)$ and $s(j)$ between 1 and S). Note that the model is still postulated for dyads, so that the basic modeling unit remains unchanged. The use of attribute variables is taken into account in the parameter structure on the right-hand side of the equation above, where the α_i 's are replaced with $\alpha^{[s(i)]}$'s, and so on. If each actor belonged to a unique subset of size 1, we would have $g = S$, and this model would be equivalent to model (15.3).

The parameter structure in (15.9) is simpler than in the basic model (15.3). There are fewer parameters to estimate. When fitting the model (15.3) to Y , we must estimate:

- $(g - 1)(C - 1)$ α 's
- $(g - 1)(C - 1)$ β 's
- $C(C - 1)/2$ $(\alpha\beta)$'s

15.2 Attribute Variables

Table 15.12. The W-arrays for Krackhardt's high-tech managers, using tenure, and age and tenure

		$s(j) = 1$		$s(j) = 2$	
		$l: x_{ji} = 0$	$x_{ji} = 1$	$x_{ji} = 0$	$x_{ji} = 1$
$s(i) = 1$	$k: x_{ij} = 0$	32	20	39	9
	$x_{ij} = 1$	20	38	48	14
$s(i) = 2$	$x_{ij} = 0$	39	48	20	23
	$x_{ij} = 1$	9	14	23	24

		$s(j) = 1$		$s(j) = 2$	
		$l: x_{ji} = 0$	$x_{ji} = 1$	$x_{ji} = 0$	$x_{ji} = 1$
$s(i) = 1$	$k: x_{ij} = 0$	74	12	68	15
	$x_{ij} = 1$	12	12	14	13
$s(i) = 2$	$x_{ij} = 0$	68	14	52	15
	$x_{ij} = 1$	15	13	15	8

		$s(j) = 1$	$s(j) = 2$	$s(j) = 3$	$s(j) = 4$
		$l: x_{ji} = 0$	$x_{ji} = 1$	$x_{ji} = 0$	$x_{ji} = 1$
$s(i) = 1$	$(\text{Less tenure, Younger})$	18 11	11 32	7 3	5 3
	$(\text{Less tenure, Older})$	7 5	3 3	0 1	1 0
$s(i) = 2$	$(\text{More tenure, Younger})$	12 7	15 11	6 1	3 0
	$(\text{More tenure, Older})$	18 1	23 1	3 0	7 0
$s(i) = 3$	$(\text{Less tenure, Younger})$	6 1	3 0	2 5	6 8
	$(\text{Less tenure, Older})$	12 5	15 8	14 8	5 5
$s(i) = 4$	$(\text{More tenure, Younger})$	3 0	7 0	6 14	6 14
	$(\text{More tenure, Older})$	18 1	23 0	6 5	4 6

		$s(j) = 1$	$s(j) = 2$	$s(j) = 3$	$s(j) = 4$
		$l: x_{ji} = 0$	$x_{ji} = 1$	$x_{ji} = 0$	$x_{ji} = 1$
$s(i) = 1$	$(\text{Less tenure, Younger})$	44 8	8 12	14 3	1 0
	$(\text{Less tenure, Older})$	14 1	3 0	2 0	0 0
$s(i) = 2$	$(\text{More tenure, Younger})$	25 7	7 6	1 0	3 2
	$(\text{More tenure, Older})$	7 5	1 6	3 2	5 3
$s(i) = 3$	$(\text{Less tenure, Younger})$	27 5	7 6	7 1	2 1
	$(\text{Less tenure, Older})$	14 3	5 3	9 0	7 0
$s(i) = 4$	$(\text{More tenure, Younger})$	12 4	4 6	14 3	12 4
	$(\text{More tenure, Older})$	4 4	5 3	6 3	4 0

When using attribute variables, and postulating S subsets of actors, we fit model (15.9) to \mathbf{W} and estimate only $(S - 1)(C - 1)$ each of α 's and β 's, but still $C(C - 1)/2$ $(\alpha\beta)$'s, for all the combinations of a pair of relational tie levels. In network data sets where the number of actors (g) is large, this simplification can mean the estimation of many fewer parameters. For example, in Krackhardt's network with twenty-one actors and a simple dichotomous relation, when we model individual differences among actors, we would have $(21 - 1)(2 - 1) = 20$ α 's, 20 β 's, 1 θ , and 1 $(\alpha\beta)$, for a total of 42 independent parameters. If we aggregate the actors into four age and tenure subsets ($S = 4$), and model the resulting \mathbf{W} -array, we estimate only $(4 - 1)(2 - 1) = 3$ α 's and 3 β 's, 1 $(\alpha\beta)$, and 1 θ , for a total of just 8 parameters.

This simplification is due to our stochastic equivalence assumption. All actors in a subset relate to all other actors statistically similarly and are also related to by all others statistically similarly. Statistical similarity implies that the probabilities of such interactions are constant. This similarity in the behavior of social interactions is assumed on theoretical grounds, using the attribute variables at hand.

The importance of this special case of our basic model to social network analysis is that it allows researchers to go beyond using only relational variables to understand network structure. With these models, information about the actors themselves, such as their age, gender, size, status, and so on, can also be used to help understand the network structure. One can study and test the association between the attribute variables and the relational variables by comparing model (15.3) fit to \mathbf{y} with model (15.9) fit to \mathbf{w} , as we will discuss and demonstrate shortly.

Parameter Estimation and Testing. Parameter estimation and testing proceeds just as in fitting model (15.3) to \mathbf{y} . The correspondence between the margins and the parameters discussed earlier still applies. So, for example, to include the α_{ikl} 's in a model for \mathbf{y} , we fit the margins [13] and [24]. To include the $\alpha_k^{[s(i)]}$'s in a model for \mathbf{w} , we still fit the margins [13] and [24]. The difference is simply that in the first, the margins [13] and [24] are fit to the $\hat{\mathbf{y}}$ -array (and are $g \times C$ in size) and in the second, the margins [13] and [24] are fit to the $\hat{\mathbf{w}}$ -array (and are $S \times C$ in size). The margins themselves are quite different, since the sums are over different arrays. In Table 15.13, we list some special cases of model (15.9), alternatively dropping the α , β , and $(\alpha\beta)$ parameters, and give the rules for testing hypotheses regarding the various model parameters.

Table 15.13. Parameters, models, and associated margins for models for attribute variables

Model	Parameters included in the model			Margins
Test	ΔG^2	Δdf	Null hypothesis	
15.9a	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_k^{[s(i)]}\}, \{\beta_l^{[s(j)]}\}, \{(\alpha\beta)_{kl}\}$	$(S - 1)(C - 1)$	$H_0 : \alpha_k^{[s(i)]} = 0$ for all i, k	[12][13][24][23][14][34]
15.9b	$\{\lambda_{ij}\}, \{\theta_k\}, \{\beta_l^{[s(j)]}\}, \{(\alpha\beta)_{kl}\}$	$(S - 1)(C - 1)$	$H_0 : \beta_l^{[s(j)]} = 0$ for all j, l	[12][23][14][34]
15.9c	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_k^{[s(i)]}\}, \{(\alpha\beta)_{kl}\}$	$(S - 1)(C - 1)$	$H_0 : (\alpha\beta)_{kl} = 0$ for all k, l	[12][13][24][34]
15.9d	$\{\lambda_{ij}\}, \{\theta_k\}, \{\alpha_k^{[s(i)]}\}, \{\beta_l^{[s(j)]}\}$	$C(C - 1)/2$	$H_0 : \alpha_k^{[s(i)]} = 0$ for all k, l	[12][13][24][23][14]

The model fitting strategy for (15.9) is parallel to that for model (15.3), from the general logic to the specific margins fit and tests made. The arrays being modeled are usually quite a bit smaller, so the computations are easier to handle.

There is one important computational difference which concerns the likelihood-ratio statistics. When fitting models to a $\hat{\mathbf{y}}$ -array using statistical packages designed for standard data, rather than just for social network data, one needs to divide the likelihood-ratio statistic G^2 obtained as output from the statistical package by 2 (to adjust for the fact that these arrays count each dyad twice). However, when fitting the models given in Table 15.13 to a $\hat{\mathbf{w}}$ -array, the G^2 's on the printouts must be adjusted in a slightly more complicated way. The reason for this adjustment is that while the "subset" models are fit to the aggregated array (defined in (15.8)), the model itself uses the dyad as its basic unit (in (15.9)). The adjustment that is needed takes the parameter estimates for a model with attribute variables (which are based on $S(S - 1)/2$ pairs of subsets), and calculates fitted values for all $g(g - 1)/2$ pairs of actors, then compares these fitted values to the relational data contained in the original sociomatrix.

The FORTRAN program GSQUARE (Iacobucci and Wasserman 1990) takes as input the fitted values for \mathbf{w} as calculated by the standard log-linear model statistical packages, and the observed \mathbf{y} -array. The program then calculates the correct, adjusted G^2 .

Other than the adjustments that must be made to get the correct test statistics, the models fit to \mathbf{y} and to \mathbf{w} are analogous. The important dis-

tinction revolves around what is most interesting to the researcher. That is, a researcher wishing to model dyadic interactions at the individual actor level would analyze y . The researcher who has attribute variables, and is less interested in individual differences than in subset differences, would analyze w .

Testing the Statistical Importance of Attributes. Suppose that two (or more) attribute variables are available, along with the single, relational variable. We now discuss how one can use the above methodology to test directly that each attribute variable is statistically important.

If an attribute variable is statistically important, then actors within subsets defined by the categories of the attribute variable really do have equal parameter values. Thus, the partitioning of the actors based on the attribute variable has not only theoretical meaning but also statistical importance. With respect to model fits, the subset model will fit as well as an equivalent model not based on subsets.

To test the statistical importance of an attribute variable, one must fit two models: one that uses the attribute variable under study, and one that does not. These two models must otherwise be identical. Due to the hierarchical nature of the models described here, the model with the attribute variable includes the model without the attribute variable as a special case. Because these models are nested within each other (the model with the attribute variable being a null hypothesis, and the model without, an alternative hypothesis), standard likelihood-ratio tests can be used. One must fit these models to different w -arrays, and obtain the G^2 's, whose difference becomes the likelihood-ratio statistic for the test of the importance of the attribute.

The details of this strategy are as follows. Use all the attribute variables to categorize the actors in S subsets. For example, if the network data set measured for $g = 30$ actors contains two attributes (gender — Males and Females — and race — White, Black, Asian, Other), then there are eight subsets that can be formed from the cross-classification of these two: White Males, White Females, Black Males, . . . , Other Females. Thus, $S = 8$, and the $30 \times 30 \times C \times C$ y -array can be compressed through aggregation (summing over all actors in the eight subsets) into a $8 \times 8 \times C \times C$ w -array. Such aggregation assumes stochastic equivalence, so that actors differ only at the level of subsets.

One now fits models to this w -array, and finds the one that fits best. We will call this w -array the "big w ," since one next forms another w , of smaller size. The best-fitting model found here is the alternative

hypothesis model for the test of the importance of the attribute variable. The G^2 for this model should be recorded.

One now has to fit a model which does not use the attribute variable in question. One should form a second, smaller w -array, using all the *other* attribute variables (that is, exclude the one under study). For the above example, this means there will be a new w : a w -array based just on race (of size $4 \times 4 \times C \times C$). One then fits the "best-fitting" model found for the big w matrix to the smaller w , and obtains the G^2 .

Then, to test that the attribute variable under study has no effect on the relational variable, we assume that the parameters from the model using all the attributes do not depend on this first attribute variable. If so, then the difference between the G^2 for the big w and the G^2 for the smaller w , which does not use the attribute variable, is a conditional likelihood-ratio statistic for the test of this hypothesis. We will illustrate this methodology shortly, using Krackhardt's high-tech managers, on which we have measurements on two attribute variables (age and tenure).

Some Difficulties Solved by Fitting Subset Models. We should mention some technical difficulties that can arise when fitting models to y -arrays. The first problem is practical and centers on the size of y . The second problem is theoretical, concerning the likelihood-ratio test statistics discussed in this chapter.

The practical issue is that the y -arrays, which have g^2C^2 cells, can get quite large and become difficult to work with computationally. For Krackhardt's data set, the y -array is of size $21 \times 21 \times 2 \times 2$, with 1724 cells. Consider one of the attribute variables, age, categorized into just 2 categories (young and old). The w -array for these data is smaller, of size $2 \times 2 \times 2 \times 2$, or just 16 cells. It is wise to keep the data matrices small, since some computer packages for fitting log-linear models rely on the inversion of matrices, which can tax moderately sized computers.

The theoretical problem in fitting models to y is also related to the size of the array. This problem also arose in this chapter when we discussed the asymptotic distributions of G^2 and ΔG^2 , which are very approximately χ^2 . For the standard asymptotic statistical theory to apply here, g must increase. However, as g increases, the dimensions of the y -array grow. For example, adding one actor turns a $g \times g \times C \times C$ array into a table of size $(g+1) \times (g+1) \times C \times C$. This implies that the number of parameters and the size of the table both grow as g does, which violates one of the basic assumptions of standard asymptotic statistical theory for the tests in conjunction with model (15.3).

Tests made on typical categorical data analyzed with log-linear models are appropriate for large numbers of observations classified into the cells of a contingency table which is fixed in size. Larger numbers of observations in a network data set can only be obtained by increasing g . Unfortunately, doing so enlarges the \mathbf{Y} -array. As we add actors to \mathcal{N} , and rows and columns to \mathbf{Y} , we must also estimate more parameters. This problem was recognized by Holland and Leinhardt (1981), Fienberg and Wasserman (1981a, 1981b), Haberman (1981), and Reitz (1982).

Fortunately, this problem does not arise when modeling \mathbf{W} . The size of this contingency table depends not on g , but on S (a function of the attribute variables) and C (a function of the relational variable), and hence is fixed. Adding actors to the network increases the counts in the cells of this array, but not the size of the table or the number of parameters. Of course, the choice of attributes and the magnitude of S should be made independently of the number of actors.

For this reason, researchers have cautioned networkers about the assumption that G^2 is asymptotically distributed as a χ^2 random variable when evaluating the fit of models to \mathbf{Y} -arrays (see Fienberg and Wasserman 1981a; Haberman 1981; Wong and Yu 1989). In addition, these contingency tables are usually quite large and sparse (that is, they contain many 0's). We suggest that conclusions about model fit based on G^2 statistics for \mathbf{Y} -arrays be stated cautiously. Statistical evaluations of model comparisons, based on ΔG^2 , conditional likelihood-ratio statistics, are more sound because such statistics compare the fits of different models applied to the same data.

We emphasize that these problems (both the computational and statistical ones) do not arise when evaluating the fit of models applied to \mathbf{W} (via G^2 statistics), or the comparison of fits to evaluate the statistical significance of parameters from models fit to \mathbf{W} (via ΔG^2 statistics). Consequently, we recommend the use of actor attribute variables whenever possible.

15.2.4 Examples: Using Attribute Variables

Just as we fit model (15.3) and its variants to test for the statistical significance of the sets of $\{\alpha\}$'s, $\{\beta\}$'s, and $\{(\alpha\beta)\}$'s to the \mathbf{y} -arrays under study, we have also fit model (15.9) and its variants to the corresponding \mathbf{w} -arrays. In Table 15.14, we list the goodness-of-fit statistics for the fabricated network of second-graders, and indicate which sets of parameters are statistically important.

Table 15.14. Goodness-of-fit statistics for the fabricated network, using attribute variables

Second-graders — friendship at the beginning of the school year and age					
Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		32.024			
15.9b	$H_0 : \alpha_{i(j)}^{(s(i))} = 0$ for all i and k	32.639	0.615	1	Not small
15.9c	$H_0 : \beta_k^{(s(j))} = 0$ for all j and k	32.639	0.615	1	Not small
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	33.044	1.020	1	Not small

For the fabricated network of second-graders, we have aggregated over the ages of the children (two categories), so that $S = 2$. We can see from the table that none of the parameters ($\{\alpha\}$, $\{\beta\}$, $\{(\alpha\beta)\}$) are statistically large. Note that there is just one parameter in each set, because there are just two levels for the relational variable and just two subsets (thus, $S = C = 2$, so $(S - 1)(C - 1) = 1$ and $C(C - 1)/2 = 1$). Prior to aggregating over age, the popularity parameters were statistically important (see Table 15.8). Based on these analyses, we would choose a model that has just a θ parameter. There appears to be no difference between the two age groups with respect to expansiveness, popularity, or reciprocity.

We can study the hypothesis of no age effect further by comparing the G^2 obtained by fitting p_1 to \mathbf{y} ($G^2 = 20.630$) to the G^2 obtained by fitting the analogous model to \mathbf{w} ($G^2 = 32.024$). The first test statistic can be found in Table 15.8, and the second, in Table 15.14. The conditional likelihood-ratio statistic $\Delta G^2 = 11.39$ for this test, with $\Delta df = 8$, is not large. For this test, the null hypothesis is that (15.9) fits \mathbf{w} as well as (15.3) fits \mathbf{y} . Given the relatively small ΔG^2 , we cannot reject this hypothesis; the simpler model described the data as well as the model with more parameters. This result means that the assumption of stochastic equivalence is reasonable — all 7-year-olds share common propensities for sending and receiving, as do all 8-year-olds. Rather than describe each individual child's behavior, we can describe the average behavior of a 7-year-old child and that of an 8-year-old without losing descriptive power.

Note that we can use the same logic to test the importance of multiple actor attribute variables. For example, if this included the attribute variable gender, we could test for the impact of gender and age (and

Table 15.15. Parameter estimates for children's friendship and age

Subset	$\hat{\alpha}$	$\hat{\beta}$
1 (7-year-olds)	-0.353	-0.353
2 (8-year-olds)	0.353	0.353

logically, their interaction) by comparing conditional likelihood-ratio statistics (ΔG^2 's) as we have described in the previous section.

Because none of the α , β , or $(\alpha\beta)$ parameters are statistically important here, we would not usually report and interpret the estimates. However, to illustrate how much simpler the parameter structure is (compared to modeling all actors via y), we list the parameter estimates in Table 15.15. The alpha and beta estimates suggest the 8-year-old children ($s(i) = 2$) are both more likely to nominate others ($\alpha_{k=1}^{[s(i)=2]}$ is positive) and to be nominated by others ($\beta_{k=1}^{[s(j)=2]}$ is positive) than are the 7-year-old children ($s(i) = 1$). The reciprocity parameter is positive, indicating some tendency for mutual ties. In conclusion, we note that this small network displays differential popularity tendencies, implying that the actors are chosen with differing rates. Aggregating the actors based on their age is important here — the children are not distinguishable once we put all the 7-year-olds and all the 8-year-olds together.

We now look at an analysis of Krackhardt's network of high-tech managers (shown in Tables 15.16 and 15.17). Here, we look at both the advice and friendship relations, but use the tenure and age attribute variables. We do three sets of models: one set using tenure, one using age, and one using both tenure and age.

First, we look at advice. Table 15.16 gives the goodness-of-fit statistics for four models for each of the three attribute variable collections, all fit to the advice relation. Actors differ substantially on advice seeking and receiving when classified by tenure, and when classified by age and tenure. However, actors differ only on advice seeking when classified by age; that is, younger actors are just as likely to receive advice as older actors. From these models, there also appears little tendency for the advice to be reciprocated. So, age and tenure interact with the advice relation, but in rather different ways.

Next, we look at friendship. Table 15.17 gives the goodness-of-fit statistics for four models for each of the three attribute variable collections, all fit to the friendship relation. The friendship relation is quite different (when studied with these models) from advice. From Table 15.17, one

Table 15.16. Goodness-of-fit statistics for Krackhardt's managers and the advice relation, with attribute variables

Krackhardt's high-tech managers, the advice relation and tenure

Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		547.121			
15.9b	$H_0 : \alpha_k^{[s(i)]} = 0$ for all i and k	564.804	17.683	1	$p < 0.001$
15.9c	$H_0 : \beta_k^{[s(j)]} = 0$ for all j and k	561.792	14.671	1	$p < 0.001$
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	549.809	2.688	1	$p = 0.101$

Krackhardt's high-tech managers, the advice relation and age

Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		547.265			
15.9b	$H_0 : \alpha_k^{[s(i)]} = 0$ for all i and k	576.188	28.92	1	$p < 0.001$
15.9c	$H_0 : \beta_k^{[s(j)]} = 0$ for all j and k	549.436	2.171	1	$p = 0.141$
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	548.465	1.200	1	Not small

Krackhardt's high-tech managers, the advice relation, and age and tenure

Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		508.541			
15.9b	$H_0 : \alpha_k^{[s(i)]} = 0$ for all i and k	553.820	45.279	3	$p < 0.001$
15.9c	$H_0 : \beta_k^{[s(j)]} = 0$ for all j and k	537.624	29.083	3	$p < 0.001$
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	513.838	5.297	1	$p = 0.0214$

can see that regardless of the attribute variable(s) used to classify the actors, the receipt of friendship ties is constant from subset to subset. But there is a difference in friendship nominations between the age groups. The younger and older actors have different numbers of friends. Their friendship expansiveness differs, as can be seen from the parameter estimates from model (15.9), fit to the friendship relation, using just age as the actor attribute variable. One finds that: $\hat{\alpha}^{[younger]} = 0.134$ and $\hat{\alpha}^{[older]} = -0.376$. Thus, younger actors are more likely to nominate others as friends, while older actors are considerably less likely to do so. Friendship nominations seem quite age-specific. Further, unlike the advice relation, there are large tendencies for these friendship ties to be reciprocated (in fact, $(\hat{\alpha}\hat{\beta}) = 1.360$).

15.3 Related Models for Further Aggregated Data

We now describe several related models. The first class of models is similar to the models just described for dyadic data with subset-level

Table 15.17. Goodness-of-fit statistics for Krackhardt's managers and the friendship relation, with attribute variables

Krackhardt's high-tech managers, the friendship relation and tenure

Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		450.691			
15.9b	$H_0 : \alpha_k^{(s(i))} = 0$ for all i and k	450.936	0.245	1	Not small
15.9c	$H_0 : \beta_k^{(s(j))} = 0$ for all j and k	450.734	0.043	1	Not small
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	465.224	14.533	1	$p < 0.001$

Krackhardt's high-tech managers, the friendship relation and age

Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		447.096			
15.9b	$H_0 : \alpha_k^{(s(i))} = 0$ for all i and k	450.985	3.889	1	$p = 0.0486$
15.9c	$H_0 : \beta_k^{(s(j))} = 0$ for all j and k	447.139	0.043	1	Not small
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	461.688	14.592	1	$p < 0.001$

Krackhardt's high-tech managers, the friendship relation, and age and tenure

Model	Null hypothesis	G^2	ΔG^2	Δdf	p-value
15.9		438.390			
15.9b	$H_0 : \alpha_k^{(s(i))} = 0$ for all i and k	449.738	11.348	3	$p = 0.010$
15.9c	$H_0 : \beta_k^{(s(j))} = 0$ for all j and k	438.768	0.378	3	Not small
15.9d	$H_0 : (\alpha\beta)_{kl} = 0$ for all k and l	451.885	13.495	1	$p < 0.001$

parameters. It differs in that one now assumes that all actors belong to one and only one subset. In other words, we assume that all alphas are equal for all actors (or subsets), as are all betas.

The relevant array, which we call V , is a contingency table like Y or W , but aggregates over all actors or over all subsets of actors. In a sense, we have just a single subgroup, so that $S = 1$. The models we fit to such tables inform us about the network structure as given by the ties, without regard to the identities of the particular actors or their attributes. The models in the class contain parameters which do not depend on the actors or on their subsets; thus, these models focus strictly on the relations, not on the actors or their attributes.

The second class of models described here takes the models for Y and W and modifies them by allowing for attribute information and/or relational data that are ordinal (for example, "big," "medium," and "small" corporations, or "high," "medium," and "low" frequencies of

communications). With these models, one can handle valued relational variables as well as valued attribute variables, such as size or socio-economic status.

15.3.1 Strict Relational Analysis — The V-array

In the same way that we aggregated individuals into a W -array so that we could fit models with subset-level parameters, and hence study stochastic equivalence, we can further aggregate W . We now want to consider models with neither actor-level nor subset-level parameters.

The models that we have presented in this chapter allow researchers to study patterns of ties for a single relational variable among individual actors (as coded by the entries of Y), or among subsets (as coded by the entries of W). We might wish to ignore the actors and their subsets altogether. Models postulated for such data focus on the relations alone, without consideration of who sent or received the ties, or of any attributes of these actors. Substantively, these models assume that all expansiveness (or popularity) parameters are constant across all actors or all subsets of actors. In other words, these parameters no longer depend on the actors. Thus, the α 's (or β 's), measuring differential actor expansiveness (or popularity) are 0; that is, there are no such differential tendencies.

We first define a new array, the V -array, with entries $\{v_{kl}\}$, which we can obtain by aggregating the entries in either Y or W :

$$\begin{aligned} V_{kl} &= \sum_{i=1}^g \sum_{j=1}^g Y_{ijkl} \\ &= \sum_{s(i)=1}^S \sum_{s(j)=1}^S W_{s(i)s(j)kl}. \end{aligned} \quad (15.10)$$

This array is of size $C \times C$, and its entries give counts of the various types of dyads. For a dichotomous relation, the V -array is 2×2 . For such a dichotomous relation, the count in the (0, 1) and (1, 0) cell, for example, equals the number of asymmetric dyads. In general, the off-diagonal cells of V , such as the (k, l) th cell ($k \neq l$) give the counts of the number of dyads for which actors send ties at level k and receive ties at level l , distinct from k . The diagonal cells, because of "double counting" when summing over dyads, give twice the number of dyads for which actors send at level k and also receive at level k .

Most social network data sets are probably better modeled using parameters that allow for individual or subset differences, but there may be

Table 15.18. The V-array constructed from the Y-array for the second-graders and friendship

i :	$x_{ji} = 0$	$x_{ji} = 1$
k :	$x_{ij} = 0$	18
	$x_{ij} = 1$	4
		4

circumstances in which modeling a V-array might be appropriate. Frequently, the decision of whether to model individuals (using Y) or subsets (using W) or to aggregate over both (V) is primarily driven by theoretical issues. A V-array can be viewed as appropriate for statistical models with only one subset of actors. If stochastic equivalence is appropriate, then we can go from the Y- to the W-array. When considering whether to go from the W- to the V-array, we are simply further assuming all subsets, or all actors, are homogeneous with respect to the dyadic interactions on the relational variable under study. In other words, all alphas are equal, as are all betas, and are all equal to 0 (as we will see from the formal model given below). If this is a tenable substantive hypothesis, then such modeling is proper. And we can certainly study such hypotheses, by first fitting p_1 and then fitting special cases of it without the α 's and without the β 's.

The v-array for the fabricated network of second-graders is given in Table 15.18. The entries in this table do not depend on actors i and j , or on subsets $s(i)$ and $s(j)$. Instead, the $g(g - 1) = 30$ ordered pairs of actors (each of the fifteen dyads is viewed from both perspectives) are classified only according to the strengths of the relational variable between the actors. The elements in this small array are the counts from the dyad census: there are 9 (18/2) null dyads, 2 (4/2) mutual dyads, and the off-diagonal entry (4) gives the number of asymmetric dyads. Note the table is symmetric. And also note, as mentioned above, that because the summation here is over all ordered pairs of actors, rather than over all unordered pairs of actors, dyads are counted twice; hence, the diagonal entries in the array are doubled dyad counts.

A statistical model for dyadic interactions that have been organized into a V-array follows:

$$\log P(Y_{ijkl} = 1) = \lambda_{ij} + \theta_k + \theta_l + (\alpha\beta)_{kl}. \quad (15.11)$$

Here, k and l take all possible integer values between 0 and $C - 1$. As noted, the α and β parameters of model (15.3) are set equal to 0. And we usually assume that the θ 's and $(\alpha\beta)$'s sum to 0 across their

respective subscripts; further, these latter parameters are symmetric: $(\alpha\beta)_{kl} = (\alpha\beta)_{lk}$.

As is proper, we give the statistical model in terms of the dyadic variables (Y_{ijkl}). Since the model is postulated at the level of the individual actors in a dyad, but the table has been aggregated over individuals, the goodness-of-fit statistics given as output of common statistical programs must be adjusted (as was necessary for fitting models to W).

Note the simplicity of this model — there are only two interesting parameters. The more substantively interesting parameters are the reciprocation parameters, the $\{(\alpha\beta)_{kl}\}$. The thetas are indicators of volume of ties sent and received at each relational strength. No parameters appear in the model that depend on the actors, i or j , or the subsets, $s(i)$ or $s(j)$, because we have aggregated over all these possibilities, and formed a table that cross-classifies only the levels of the relational variables.

To illustrate, let us look at the fabricated second-grade network. Recall that we have already demonstrated (see Table 15.8) that the children do differ with respect to their popularity effects; that is, the β 's are statistically different from each other. Thus, strictly speaking, these parameters should not be all equated to each other. Nevertheless, we use this example for illustrative purposes. To fit model 15.11, we simply fit the saturated model [12] to the 2×2 v of Table 15.18.

The model clearly fits "perfectly"; that is, a 2×2 table has a total of 3 degrees of freedom (only the grand total of the table is fixed), and we have a model with three parameters, as follows: one for θ , one for $(\alpha\beta)$, and one for λ . The parameter estimates are: $\theta = 0.385$ and $(\alpha\beta) = 0.02$ (we have no interest in λ , since it appears in the model simply to insure that the probabilities sum to unity). Clearly, there is very little tendency for ties to be sent, but when ties are sent, there is some tendency for these ties to be reciprocated.

We have found these relational v-arrays most useful when modeling multirelational network data sets. In fact, the first use of such arrays was by Davis (1968a), who proposed methods for the analysis of two relations, and Galaskiewicz and Marsden (1978), who analyzed a social network data set containing three relations. In Chapter 16, we discuss statistical models for multiple relational social network data sets which incorporate associations or interactions for the relational variables under study. Such models can also be applied to the multirelational versions of the Y-, W-, or V-arrays.

15.3.2 Ordinal Relational Data

We now consider models for relations that not only are valued, but have ordered categories. For example, the strength of a relation might be measured with $C = 3$ levels, coded as 0, 1, or 2, indicating "no," "sometimes," and "often" frequency of phone calls between actors, or "strangers," "acquaintances," "friends," and "best friends," which might be coded as 0, 1, 2, and 3. Thus, we assume that the levels of the relational variable are ordered, with smaller values indicating weaker ties, and larger values representing stronger ties.

Log-linear models for discrete data have been extended to categorical variables whose categories are ordered (for example, Agresti 1984, 1990; Goodman 1979). For social network data, relational variables and attribute variables can be ordinal. Here, we modify the models of this chapter to incorporate the possibly ordinal measurement properties inherent in the data. The models discussed in this section are described in detail in Wasserman and Iacobucci (1986).

The statistical models (15.3) and (15.9) can be generalized to incorporate not only ordinal relational variables, but also ordinal attribute variables. For example, we might classify actors into five ordered subsets based on their university rank: lecturer, instructor, assistant professor, associate professor, and professor.

We begin by assuming that scores can be assigned to the ordered strengths of the relational variable. We label these scores $u_k : 0, 1, 2, \dots, (C - 1)$, and center them to have a mean of zero. That is, we calculate the average score, $\bar{u} = \sum_{k=0}^{C-1} u_k / C$, and then subtract it from each of the u_k 's. For example, with $C = 3$ strengths, we might assign the u_k scores of $u_1 = 0, u_2 = 1, u_3 = 2$, and center them to $-1, 0$, and 1 , so that the centered scores have a mean of 0. If we had $C = 4$ strengths, we might assign the scores $u_1 = 1, u_2 = 2, u_3 = 4, u_4 = 8$, and center them to obtain $-2.75, -1.75, 0.25, 4.25$.

We use these scores to estimate a set of regression-like slope parameters. In model (15.3) there are $(g - 1)(C - 1)$ expansiveness parameters. We change these parameters by using the ordinal information. Specifically:

$$\alpha_{i(k)} = \alpha_i(u_k - \bar{u}).$$

The $(u_k - \bar{u})$'s are the assigned, centered scores. Because the u_k 's are known, there are fewer parameters to estimate (assuming that C exceeds two categories). The α_i parameters are the "slope" parameters mentioned above. There are g alphas, one for each actor. The $\{\alpha_i\}$ effects sum to

15.3 Related Models for Further Aggregated Data

0. Thus, instead of estimating $(g - 1)(C - 1)$ independent quantities, we estimate only $(g - 1)$ parameters. The general expansiveness parameters $\{\alpha_{i(k)}\}$ are assumed to be linearly related to the scores assigned to the relational strengths.

Consider an example. Suppose the relation is frequency of interactions among children, which we have coded as "low," "medium," and "high." In model (15.3), we would estimate $(g - 1)(3 - 1) = 2(g - 1) \alpha_{i(k)}$ parameters. With the ordinal version of the model discussed here, we have only $(g - 1) \alpha_i$'s. These new α_i 's do not depend on k .

Let us focus on the regression-nature of these parameters. The difference between the low category ($k = 0$) and the medium category ($k = 1$) is (for the i th actor) equal to $\alpha_i(u_1 - u_0)$. The difference between the medium category and the high category ($k = 2$) is $\alpha_i(u_2 - u_1)$. If the two differences $(u_1 - u_0)$ and $(u_2 - u_1)$ are equal, which would be the case if the categories are equally spaced on the score variable, then $\alpha_{i(1)} - \alpha_{i(0)} = \alpha_{i(2)} - \alpha_{i(1)}$. Thus, the effect of "going from" the medium category ($k = 1$) to the high category ($k = 2$) is the same as the effect of going from the low category to the medium category. This constancy is identical to the change in a response variable in a regression, with a linear regression slope.

Similarly, the popularity parameters are simplified by taking the set of $\{\beta_{j(l)}\}$ effects, for a nominal relation, and replacing it with the set of $\{\beta_j\}$ effects for an ordinal relation. Specifically,

$$\beta_{j(l)} = \beta_j(u_l - \bar{u}).$$

In general, any parameters in models (15.3) and (15.9) can be revised to ordinal relational versions by replacing parameters that depend on the index k , with a simpler parametric structure using the known scores $\{u_k\}$. In addition to the alphas and betas, $\{\theta_k\}$ and $\{(\alpha\beta)_{kl}\}$ can be modified, giving rise to the following model:

$$\begin{aligned} \log P(Y_{ijkl} = 1) = & \lambda_{ij} + \theta(u_k - \bar{u}) + \theta(u_l - \bar{u}) + \alpha_i(u_k - \bar{u}) \\ & + \alpha_j(u_l - \bar{u}) + \beta_j(u_k - \bar{u}) + \beta_i(u_l - \bar{u}) \\ & + (\alpha\beta)(u_k - \bar{u})(u_l - \bar{u}), \end{aligned} \quad (15.12)$$

where, as usual, k and l take on integer values between 0 and $C - 1$. Note that there is only a single $(\alpha\beta)$ parameter in this model. Its interpretation is analogous to a measure of association between sending strengths and receiving strengths, much like in model (15.3) for dichotomous data.

If we use attribute variables to categorize the actors into $S < g$ subsets, then we can simplify model (15.12) even further:

$$\begin{aligned}\log P(Y_{ijkl} = 1) = & \lambda^{[s(i)s(j)]} + \theta(u_k - \bar{u}) + \theta(u_l - \bar{u}) \\ & + \alpha^{[s(i)]}(u_k - \bar{u}) + \alpha^{[s(j)]}(u_l - \bar{u}) \\ & + \beta^{[s(j)]}(u_k - \bar{u}) + \beta^{[s(i)]}(u_l - \bar{u}) \\ & + (\alpha\beta)(u_k - \bar{u})(u_l - \bar{u}).\end{aligned}\quad (15.13)$$

Finally, the attribute variable(s) might also be ordered, so that we could assign scores to subset categories. For example, distinguishing actors by gender or race would result in dichotomous and discrete attribute variables. But such attribute variables are not ordinal. However, if we have measurements on some measure of socioeconomic status, we could form subsets on some categorization of this variable (for example, lower, middle, and upper class). More details and examples illustrating the application of these models for ordinal relational and attribute variables can be found in Wasserman and Iacobucci (1986).

15.4 ○Nondirectional Relations

Consider now a single, nondirectional relation. We now show how model (15.3) can be modified for such a relation. As an example, we examine Padgett's Florentine families.

There are (at least) two main differences between directional and nondirectional relations. First, the indegrees and outdegrees are equal for a nondirectional relation. Thus, there can be no difference between expansiveness and popularity parameters in a model. Second, because the states of the dyads for a nondirectional relation are either "on" or "off," there is no reciprocity. We cannot study the tendency for i to send a tie to j and not for j to send a tie to i since there are no asymmetric dyads. Thus, the models for nondirectional relations cannot contain reciprocity parameters.

15.4.1 A Model

We now present a model designed to reflect individual actor differences. Thus, it should be fit to a \mathbf{Y} -array. Models for actors in subsets, which should be fit to a \mathbf{W} -array, are similar: actors i and j would simply be replaced by subsets $[s(i)]$ and $[s(j)]$. We first must define the correct

15.4 ○Nondirectional Relations

\mathbf{Y} -array, designed to reflect the dyadic states that are possible with a nondirectional relation:

$$\begin{aligned}Y_{ijk} &= 1 \text{ if } X_{ij} = k \\ &= 0 \text{ otherwise.}\end{aligned}\quad (15.14)$$

The \mathbf{Y} -array is a three-dimensional array, since the \mathbf{X} matrix is symmetric by definition. Our model is

$$\log P(Y_{ijk} = 1) = \lambda_{ij} + \theta_k + \gamma_{i(k)} + \gamma_{j(k)}. \quad (15.15)$$

In this model, there are no α 's and no β 's. Rather, they are replaced by γ 's. Each actor has a single set of γ parameters, indexed by the choosing strength (k). Note that there is no $(\alpha\beta)$ term.

This model is equivalent to [12][13][23], the no three-factor interaction model, fit to the \mathbf{Y} -array. Log-linear model computer packages can be used to obtain fitted values. Various special cases of model (15.15) can be fit by dropping parameters. In these cases, corresponding margins of \mathbf{Y} are added or deleted from the fitted log-linear model.

15.4.2 An Example

We fit several models to Padgett's Florentine network. The basic model (15.15) has goodness-of-fit statistics equal to 51.83 (for the business relation) and 87.97 (for the marriage relation). To study whether the families differ with respect to business or marriage, we fit the special case of the basic model without the γ parameters. Comparing the basic model to the model without these actor-level parameters, we found that the γ effects are large for both marital ties ($\Delta G^2 = 108.13 - 87.97 = 20.16$) and business ties ($\Delta G^2 = 90.42 - 51.83 = 38.59$). Degrees of freedom are 14 for marriage and 10 for business (since several of the families have infinite parameters). The conditional likelihood ratio statistic for the marriage relation is small, while that for business is not. Hence, we can conclude that the families are indeed different with respect to the volume and patterns of their marital and business ties to others. The $\hat{\gamma}_i$'s for the business and marital relations are given in Table 15.19. We focus just on the business relation, since these $\hat{\gamma}$'s appear to be more statistically important than those for the marital relation.

The largest negative $\hat{\gamma}$'s for business are for the families Pazzi, Salvati, and Tornabuoni. The families with the largest positive $\hat{\gamma}$'s are Medici, Barbadori, and Lamberteschi. As can be seen, five families have no business ties with the others, and hence have $-\infty$ parameter estimates.

Table 15.19. Parameter estimates for Padgett's Florentine families

Family	Actor	Business $\hat{\gamma}_i$	Marriage $\hat{\gamma}_i$
Acciaiuoli	n_1	$-\infty$	-1.106
Albizzi	n_2	$-\infty$	0.310
Barbadori	n_3	0.904	-0.265
Bischeri	n_4	0.290	0.310
Castellani	n_5	0.290	0.310
Ginori	n_6	-0.399	-1.106
Guadagni	n_7	-0.399	0.779
Lamberteschi	n_8	0.904	-1.106
Medici	n_9	1.444	1.539
Pazzi	n_{10}	-1.313	-1.106
Peruzzi	n_{11}	0.904	0.310
Pucci	n_{12}	$-\infty$	$-\infty$
Ridolfi	n_{13}	$-\infty$	0.310
Salvati	n_{14}	-1.313	-0.265
Strozzi	n_{15}	$-\infty$	0.779
Tornabuoni	n_{16}	-1.313	0.310

We also modeled these relations using wealth as an attribute of each actor. Families with wealth less than or equal to 40,000 lira formed one group, and families with more formed another. Half of the families fall into the wealthier group.

Aggregating families into these two wealth categories has a very large effect on the analysis. After aggregation, we fit models to $2 \times 2 \times 2$ three-dimensional y -arrays. For the analyses of the w 's, the \hat{y} 's for marriage are not large ($\Delta G^2 = 1.21$ with $\Delta df = 1$) but they are for business ($\Delta G^2 = 7.59$ again with $\Delta df = 1$). These results suggest that wealth is quite important in distinguishing families who have business ties, but not for marital arrangements. That is, wealthy families enter into business relationships at different rates than less wealthy families, but wealth is not an important influence on marital ties.

15.5 \otimes Recent Generalizations of p_1

In this section, we briefly introduce some interesting developments and recent research on other generalizations of the p_1 model. These include Bayesian estimation of p_1 parameters as described in Wong (1987), and the pseudo-likelihood estimation described in Strauss and Ikeda (1988) designed for the Markov random graphs of Frank and Strauss (1986).

15.5 \otimes Recent Generalizations of p_1

The Bayesian ideas offered by Wong (1987) allow a priori information about the α 's and β 's to be used in an effort to improve their estimation. The Bayesian approach assumes that the α 's and β 's are no longer fixed constants, but are random, possibly associated, variables. Wong assumes that these p_1 parameters are multivariate normal random variables, with some structured covariance matrix.

A Bayesian approach can help "smooth" estimation of parameters when fitting models to large, sparse contingency tables. On the down side, the algorithm developed by Wong is slow and has problems if the α 's and β 's are highly correlated (indicating that parameter redundancy and singular matrices can arise in the computation process). He focuses his attention on the α and β parameters, but presumably his approach could incorporate prior information on $(\alpha\beta)$ also.

Wong compares α and β estimates from the Bayesian approach to α and β estimates using maximum likelihood (ML) estimation, as we have described here. Wong states that the sets of estimates (Bayesian and ML) are usually different, and that even the relative ordering among the estimated α 's and among the estimated β 's may differ. (In particular, he notes that Bayesian methods cause estimates to "shrink" toward 0, compared to the ML estimates, and that more shrinkage occurs for the more extreme values of α or β .) Nevertheless, the correlations between the ML and Bayesian estimates for the example he analyzed were quite high. The correlation between the two sets of estimated α 's was 0.891, and the correlation for the estimated β 's was 0.966. Even so, prior information might change or even improve parameter estimates in other network data sets.

Another development to note is the work reported in Strauss and Ikeda (1990). They investigate a *pseudo-likelihood* estimation procedure, a generalization of maximum likelihood, that uses an approximate likelihood function which does not assume dyadic independence (see comments in Iacobucci and Wasserman 1990). The theoretical foundation of Strauss and Ikeda's work is found in Frank and Strauss (1986). Strauss and Ikeda derive a pseudo-likelihood as a function of each data point (x_{ij}), conditional on the rest of the data. Any interdependencies in the data can be directly modeled by this statistical conditioning, so no assumptions need to be made that the data points are all independent.

Strauss and Ikeda compared the performance of standard maximum likelihood estimates to their maximum pseudo-likelihood (MP) estimates both in a simulation study, and by analyzing the "like" relation measured on the monks in the monastery studied by Sampson (1967). In the

simulations, they looked at the performance of the estimates in five replicated networks containing fifteen, twenty, or thirty actors. They found that MP and ML estimates performed equally well, as evaluated by a root mean squared error measure. The estimates had greater standard errors for the networks with fewer actors, but this would be true of any procedure — better precision usually occurs with larger data sets.

Under all the conditions for which both ML and MP estimates could be estimated, the two performed similarly. The main advantage in the use of MP estimation is that there are conditions under which MP estimates exist, but ML estimates do not. The MP approach further expands the applicability of p_1 because it can be used to fit models that do not assume dyadic independence, such as those described by Frank and Strauss (1986).

Strauss and Ikeda's comparisons also address the issue of how well the maximum likelihood estimation of p_1 parameters performs even under conditions where the assumption of dyadic independence is known to be violated. The fact that the ML estimates are as good as MP estimates is good news. We can proceed to use the relatively simple methods described here without much concern that violation of the assumption of dyadic independence will greatly affect the results. In addition, the ML and MP $\hat{\alpha}$'s were highly correlated (in fact, almost equal) in the reported analysis of the Sampson data, as were the ML and MP $\hat{\beta}$'s.

Another approach to dyadic independence models comes from the multilevel models common in medical and educational studies. Rosner's (1989) statistical work was motivated not by social network concerns, but by concerns in modeling ophthalmological data. The left and right eye of a person could be viewed as analogous to an actor and a partner in a dyad. Measures on the left or right eye (or actor or partner) clearly cannot be treated as independent; consequently, Rosner develops a logistic regression model for multilevel data, essentially a categorical data version of a linear model incorporating an intraclass correlation coefficient (see also Kraemer and Jacklin 1979).

Still other methods have been pursued by Frank and his colleagues (Frank, Hallinan, and Nowicki 1985; Frank, Komanska, and Widaman 1985; Frank, Lundquist, Wellman, and Wilson 1986; Wellman, Frank, Espinoza, Lundquist, and Wilson 1991). This research uses stochastic equivalence (or, as stated by these authors, "certain homogeneity assumptions") in conjunction with models for dyads and triads. All possible states of a dyad (X_{ij}, X_{ji}) become categories of a discrete random

variable. All possible cross-classifications of actor attribute information become another variable. For example, using race and gender as attribute variables, one variable would have categories consisting of all pairs of the following subsets: white males, black males, white females, and black females.

The elements in each cell of the cross-classification of these two variables are the frequencies with which a specific type of actor (such as black females) interacts with other types of actors (such as white females) at the various dyadic states. Each row, then, describes a distribution of dyadic states for each pair of subsets of actors. The information in such a cross-classification is identical to that in a W-array. The similarities between all possible pairs of subsets are summarized by χ^2 statistics, and then the subsequent matrix of similarities is cluster analyzed (see Frank, Komanska, and Widaman 1985). The cluster analysis yields subsets that are distinct, and possibly separable, and subsets that are similar, and possibly aggregable.

Stochastic equivalence and comparisons of within-subset to between-subset relations are also examined by Marsden (1981, 1986, 1987, and 1989). A two-way table is created in which both the rows and columns represent the subsets (white males, black males, and so on). The elements of the matrix are densities of ties, for which several models can be developed. Such density matrices are like w-arrays, in that the rows and columns are indexed by the $[s(i)]$'s and $[s(j)]$'s. However, the density matrix essentially collapses over sending and receiving strengths k and l , so that the directions and reciprocation of the ties are lost.

Ties within and between subsets are also analyzed by Yamaguchi (1990). "Homophily" or "inbreeding" describe the expected tendencies for actors to have friends of the same attribute category as themselves. For example, this structural hypothesis would lead us to expect that white males would be more likely to chose each other as friends, than to choose men of other races or women. Yamaguchi also examines the links among the friends chosen by any given actor. Thus, the paper looks at both the ties between n_i and n_j 's (in conjunction with the attribute information), and also the ties among the chosen n_j 's.

Finally, Strauss and Freeman (1989) translate several classic substantive theories or hypotheses into statistical models that use network descriptors as parameters. For example, one model created is based on "small world" studies. Strauss and Freeman create a three-dimensional sociomatrix with elements x_{iji} , equaling 1 if n_i passes the message to n_j at step i (and otherwise equaling 0). Elements of matrices such as

this one can be studied using Markov models. A very similar model, focusing not on the "step" at which a message is sent but on its strength, was discussed by Elsas (1990) and commented upon by Iacobucci and Hopkins (1991). It is based on the social interaction theory and model of Scheiblechner (1971, 1972), and is derived from the common Rasch model in psychometrics (Scheiblechner 1977).

Another model discussed by Strauss and Freeman is the random and biased net paradigm of Rapoport (1953, 1957, 1963, 1979) and Rapoport and Horvath (1961), which began as a model of animal sociology (Rapoport 1949a, 1949b, 1950). This paradigm has been the subject of recent theory by Fararo and Skvoretz (1987). A variety of structural theories, such as Blau's (1977) social differentiation theory and Granovetter's (1973) strength-of-weak-ties principle, have been studied using the paradigm. Fararo (1981, 1983), Fararo and Skvoretz (1984), and Skvoretz (1983, 1985, 1990) have done extensive work on the mathematical aspects of this model, particularly approximations and simulations.

15.6 \otimes Single Relations and Two Sets of Actors

We now turn our attention to statistical models for two-mode social networks. We assume that we have two sets of actors, and a relation that is directed from actors in one set to actors in the other. Details of these models can be found in Iacobucci and Wasserman (1990), Wasserman and Iacobucci (1990), and Iacobucci and Hopkins (1992). The models and results presented here for one-mode networks are easily modified for this generalization. Because of this commonality, we will just briefly present these modifications.

15.6.1 Introduction

Define two actor sets: $\mathcal{N} = \{n_1, n_2, \dots, n_g\}$, $\mathcal{M} = \{m_1, m_2, \dots, m_h\}$. A tie originates at n_i and is directed toward the receiver m_j . The relational data recorded on the ordered pair (n_i, m_j) is denoted by x_{ij} .

We can study the tendency for actors in \mathcal{N} to initiate ties and actors in \mathcal{M} to receive ties. But since it is not possible for the tie $m_j \rightarrow n_i$ to occur, we cannot study tendencies for actors in \mathcal{M} to initiate and tendencies for actors in \mathcal{N} to receive. Thus, for this two-mode relation, we cannot study reciprocity.

With g actors in \mathcal{N} , and h actors in \mathcal{M} , there are gh dyads to consider, or gh relational variables to be modeled. We create a \mathbf{Y} -array and fit

15.6 \otimes Single Relations and Two Sets of Actors

various log-linear models. In this new situation, we create not a four-dimensional array, but a three-dimensional contingency table of size $g \times h \times C$, defined as follows:

$$\begin{aligned} Y_{ijk} &= 1 \text{ if the ordered pair } \langle n_i, m_j \rangle \text{ takes on} \\ &\quad \text{the value } x_{ij} = k \\ &= 0 \text{ otherwise.} \end{aligned} \tag{15.16}$$

The \mathbf{Y} -array is thus defined by three variables: the actors in \mathcal{N} , the actors in \mathcal{M} , and the value of the tie between the two.

15.6.2 The Basic Model

Model (15.3) is a version of p_1 generalized to discrete-valued relations. It can be extended in a straightforward theoretical manner to two-mode networks. To begin, we note that only actors in \mathcal{N} can send ties, while only actors in \mathcal{M} can receive. Thus, a specific actor will have only one parameter: expansiveness parameters are estimated for the actors in \mathcal{N} , while popularity parameters are estimated for the actors in \mathcal{M} .

The basic model for two-mode networks simplifies further because reciprocity cannot occur; thus, we do not include the $\{(\alpha\beta)_{kl}\}$ parameters in our models. There can be only one tie within each dyad.

The basic model is:

$$\log P(Y_{ijk} = 1) = \lambda_{ij} + \theta_k + \alpha_{i(k)} + \beta_{j(k)}. \tag{15.17}$$

We assume that the parameters in this model sum to 0 across k for each i or j .

The details of obtaining maximum likelihood estimates of the parameters of (15.3), such as the log likelihood function and the formula for the goodness-of-fit statistic G^2 , are straightforward generalizations of the results for p_1 . We refer the interested reader to Iacobucci and Wasserman (1990) and Wasserman and Iacobucci (1990).

We can simplify model (15.17) in only three ways. We can drop the α 's, the β 's, or both, from the basic model. These alternative models are listed below.

$$\begin{aligned} a) \log P(Y_{ijk} = 1) &= \lambda_{ij} + \theta_k + \beta_{j(k)} \\ b) \log P(Y_{ijk} = 1) &= \lambda_{ij} + \theta_k + \alpha_{i(k)} \\ c) \log P(Y_{ijk} = 1) &= \lambda_{ij} + \theta_k \end{aligned} \tag{15.18}$$

When we compare the goodness-of-fit statistics for models (15.17) and

(15.18a), we have the conditional likelihood-ratio test statistic for the null hypothesis,

$$H_0 : \alpha_{i(k)} = 0 \text{ for all } i, k;$$

that is, there are no differences among the actors in \mathcal{N} with respect to their expansiveness. The alternative hypothesis is (as usual) that some of these parameters are not 0. When comparing models (15.17) and (15.18b), we obtain the conditional likelihood-ratio test statistic for the null hypothesis

$$H_0 : \beta_{j(k)} = 0 \text{ for all } j, k.$$

If we were not able to reject either of these hypotheses, we would conclude that model (15.18c) is the best-fitting model. Such a null model is rather uninteresting because it states the actors in \mathcal{N} have no differential expansiveness and actors in \mathcal{M} have no differential popularity effects. Thus, ties are just functions of the strength at which those ties are made, not of the actors involved.

15.6.3 Aggregating Dyads for Two-mode Networks

We can use actor attribute variables, now for both sets of actors, to create a **W**-array, and equate parameters for all actors belonging to a subset.

For example, focusing on Galaskiewicz's CEOs and clubs network, we might classify the CEOs by the size of their firm (based perhaps on after-tax income for a recent year). We could also classify the clubs by some measure of their prestige: high; very high; and very, very high.

Using the notation first introduced in Chapter 3, we assume that we have Q_1 attribute variables for the actors in \mathcal{N} , and Q_2 attribute variables for the actors in \mathcal{M} . From these, we categorize the actors appropriately, to obtain S subsets for the actors in \mathcal{N} , and T subsets for the actors in \mathcal{M} . We need two mapping functions: $s(i)$, which maps the g actors $\{n_i\}$ in \mathcal{N} to their subsets, and another function $t(j)$ to map the h actors $\{m_j\}$ in the second set \mathcal{M} to their respective subsets. Thus, the subset memberships of the actors in \mathcal{N} can be denoted by $s(n_1), s(n_2), \dots, s(n_g)$, and the subsets to which the actors in \mathcal{M} belong can be denoted by $t(m_1), t(m_2), \dots, t(m_h)$.

To obtain a **W**-array we aggregate over actors in a subset, just as we did in the one-mode case:

$$W_{s(n_i)t(m_j)k} = \sum_{n_i \in s(n_i)} \sum_{m_j \in t(m_j)} Y_{ijk}. \quad (15.19)$$

15.7 Computing for Log-linear Models

Because the two sets of actors differ, we might use attribute variables for one set of actors, but not the other. Aggregating over actors in \mathcal{N} yields:

$$W_{s(n_i)jk} = \sum_{n_i \in s(n_i)} Y_{ijk}. \quad (15.20)$$

Aggregating over actors in \mathcal{M} yields:

$$W_{it(m_j)k} = \sum_{m_j \in t(m_j)} Y_{ijk}. \quad (15.21)$$

Thus, the **W**-array can be of size $S \times T \times C$, if we use attribute variables for both sets of actors, or $S \times h \times C$, or $g \times T \times C$, if we use attributes for just one of the sets of actors.

We can also aggregate completely over actors, and postulate models that do not contain any actor- or subset-level parameters. Aggregating over all actors gives a very simple one-dimensional **V**-array:

$$V_k = \sum_i \sum_j Y_{ijk} \quad (15.22)$$

describing only the relational data, not distinguishing among the sending or receiving actors. Models fit to this array would contain only λ and θ parameters. These latter parameters measure tendencies toward choices at the various strengths, but are independent of the natures of the sending and receiving actors.

We note that the goodness-of-fit statistics for models fit to either the **W**- or **V**-arrays require special computations just as in the one-mode situation. Standard output from statistical packages is not correct for G^2 statistics, even though the fitted probabilities are correct.

We can also consider ordinal relational variables, just as we did in the one-mode situation. We would model the relational data as follows:

$$\begin{aligned} \log P(Y_{ijk} = 1) = & \lambda_{ij} + \theta(u_k - \bar{u}) + \alpha_i(u_k - \bar{u}) \\ & + \beta_j(u_k - \bar{u}), \end{aligned} \quad (15.23)$$

where the u 's are the assigned, known scores. All theoretical and practical details of modeling and testing hypotheses about two-mode networks carries over from the procedures described in detail for one-mode networks.

15.7 Computing for Log-linear Models

Following are the computing steps necessary to fit the models described in this chapter and the next. Commands are listed and explained for each

of the following statistical packages: *GLIM* (Baker and Nelder 1978), *SPSS^X* (Norusis 1985), *BMDP* (Dixon 1983), and *SYSTAT* (Wilkinson 1987). We highly recommend the use of the program *GLIM* or *SPSS^X*. Other statistical packages will work similarly to these two. For example, *BMDP* and *SYSTAT* are very similar to *SPSS^X*. (*SAS* is a very powerful and flexible package, but its *CATMOD* procedure is rather peculiar and has difficulties with contingency tables that have many 0's, which characterize *y*-arrays. Parameter estimates can disappear and become confounded as functions of each other, and it is up to the researcher to disentangle them.) The *Y*- and *W*-arrays used to illustrate each package are those data from the fabricated second-grader network.

15.7.1 Computing Packages

After the commands for these packages are presented, we describe how one obtains the model parameter estimates from the *u*-terms that are given on the computer printouts. Some of the packages use different constraints than others, and we note the differences here.

GLIM Commands.

```
$C TO FIT MODELS TO THE Y-ARRAY:
$ECHO
$OUTPUT 3
$UNITS 144
$FACTOR I 6 J 6 K 2 L 2
$CALC I=%GL(6,24): J=%GL(6,4): K=%GL(2,2): L=%GL(2,1)
$DATA Y
$READ
 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0
 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 1
 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 0
 1 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0
 0 1 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0
 1 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0
$C END DATA
$YVAR Y $LINK L $ERROR P
$FIT I*I + I*K + J*L + J*K + I*L + K*L $DISPLAY MDAR
$FIT I*I +           J*K + I*L + K*L $DISPLAY MDAR
$FIT I*I + I*K + J*L +   K*L $DISPLAY MDAR
```

```
$FIT I*I + I*K + J*L + J*K + I*L           $DISPLAY MDAR
$STOP

$C TO FIT MODELS TO THE W-ARRAY:
$ECHO
$OUTPUT 3
$UNITS 16
$FACTOR R 2 S 2 K 2 L 2
$CALC R=%GL(2,8): S=%GL(2,4): K=%GL(2,2): L=%GL(2,1)
$DATA W
$READ
 6 0 0 0 5 1 1 2 5 1 1 2 2 2 2 0
$C END DATA
$YVAR W $ERROR P $LINK L
$FIT R*S + R*K + S*L + S*K + R*L + K*L $DISPLAY MDAR
$FIT R*S +           S*K + R*L + K*L $DISPLAY MDAR
$FIT R*S + R*K + S*L +           K*L $DISPLAY MDAR
$FIT R*S + R*K + S*L + S*K + R*L           $DISPLAY MDAR
$STOP
```

An explanation of the commands follows. "ECHO" gets the commands entered onto the log of the computational session for review later. "OUTPUT 3" tells *GLIM* to write all results to a file called *TAPE3* (input and output instructions will vary with computer site). "UNITS x" tells *GLIM* the contingency table has *x* cells. "FACTOR" gives the variable name for each dimension of the table as well as the number of levels the discrete variable takes. "CALC" forms the factors needed for modeling and tells *GLIM* which data point goes in which cell of the table. The commands:

*CALC I=%GL(G,G*C*C): J=%GL(G,C*C): K=%GL(C,C): L=%GL(C,1)*

are used when reading in the *y*- or *w*-array with the last subscript changing fastest. (Compare the *y*-arrays and the *w*-arrays shown earlier.)

"DATA z" tells *GLIM* you are about to enter some data and you want that variable named *z*. "READ" then initiates the procedure and the data follow. There must be *x* numbers — the same number as in the units statement. "C" always denotes a comment ignored by *GLIM*. "YVAR q" tells *GLIM* the frequencies to be modeled are stored in the variable called *q*. "ERROR P" and "LINK L" are the commands signifying that *GLIM* should fit log-linear models. (Other error distributions and link

functions give *GLIM* great flexibility.) "FIT" specifies the model. The first of the four models gives the full p_1 model and the remaining three models will provide significance tests for alpha, beta, and reciprocity parameters, respectively. "STOP" ends the *GLIM* session.

SPSS^X Commands. The data file is called "MY DATA A" and it contains the relations as follows:

0000000001111...
Column of Data File: 1234567890123...

Variable: i j r s k l

1 2 2 2 1 0
2 1 2 2 0 1
1 3 2 1 0 0
3 1 1 2 0 0
1 4 2 1 0 0
4 1 1 2 0 0
1 5 2 2 1 0
5 1 2 2 0 1
1 6 2 1 0 0
6 1 1 2 0 0
2 3 2 1 1 1
3 2 1 2 1 1
2 4 2 1 0 0
4 2 1 2 0 0
2 5 2 2 0 0
5 2 2 2 0 0
2 6 2 1 1 1
6 2 1 2 1 1
3 4 1 1 0 0
4 3 1 1 0 0
3 5 1 2 0 0
5 3 2 1 0 0
3 6 1 1 0 0
6 3 1 1 0 0
4 5 1 2 1 0
5 4 2 1 0 1
4 6 1 1 0 0

6 4 1 1 0 0
5 6 2 1 1 0
6 5 1 2 0 1

To fit models to a y-array, use the following program file (called "RUN-JOB SPSSX A"):

```
TITLE 'MY ANALYSIS'
FILE HANDLE A1 / NAME= 'MY DATA A'
DATA LIST FILE=A1
/I 1 J 3 R 5 S 7 K 9 L 11
HILOGLINEAR I J (1,6) K L (0,1)
/PRINT = ALL ASSOCIATION
/DESIGN I*J I*K J*L J*K I*L K*L
/DESIGN I*J J*K I*L K*L
/DESIGN I*J I*K J*L K*L
/DESIGN I*J I*K J*L J*K I*L
LOGLINEAR I J (1,6) K L (0,1)
/PRINT = ESTIM
/DESIGN I J K L I BY J I BY K J BY L J BY K
I BY L K BY L
```

To fit models to a w-array, use these commands:

```
TITLE 'MY W ANALYSIS'
FILE HANDLE A2 / NAME = 'MY DATA A'
DATA LIST FILE=A2
/I 1 J 3 R 5 S 7 K 9 L 11
HILOGLINEAR R S (1,2) K L (0,1)
/PRINT = ALL ASSOCIATION
/DESIGN R*S R*K S*L S*K R*L K*L
/DESIGN R*S S*K R*L K*L
/DESIGN R*S R*K S*L K*L
/DESIGN R*S R*K S*L S*K R*L
LOGLINEAR R S (1,2) K L (0,1)
/PRINT = ESTIM
/DESIGN R S K L R BY S R BY K S BY L S BY K
R BY L K BY L
```

The commands "TITLE," "FILE HANDLE," and "DATA LIST" initiate *SPSS^X* and read in the data file. "HILOGLINEAR" generates hierarchical log-linear models and the fitting procedures. It is a fast

procedure, both for typing in compact model statements, and in terms of computational running time. However, it does not give estimated u -terms. "LOGLINEAR" is a slower and more cumbersome procedure, but it will calculate estimated u -terms for small networks.

BMDP Commands. To fit models to a y-array (contained in a data file, like that needed for $SPSS^X$):

```
/PROBLEM TITLE IS 'MY ANALYSIS'.
/INPUT VARIABLES ARE 6.
CASES ARE 30.
FORMAT IS '(I1,5I2)'.
TABLE IS 6,6,2,2.
/VARIABLE NAMES ARE I,J,R,S,K,L.
/TABLE INDICES ARE L,K,J,I.
/FIT MODEL IS IJ,IK,JL,JK,IL,KL.
MODEL IS IJ, JK, IL, KL.
MODEL IS IJ, IK, JL, KL.
MODEL IS IJ, IK, JL, JK, IL.
/PRINT EXPECTED. LAMBDA.
/END
```

To fit models to a w-array, replace i 's and j 's with r 's and s 's in the statements "TABLE" and "FIT," and replace the 6's with 2's on the fifth line.

BMDP outputs its fitted values in a somewhat awkward form. Rather than listing the fitted values as a column vector with the subscripts $ijkl$ changing in the "last is fast" standard manner, **BMDP** creates L columns and then strings the subscripts kji in column vector format where the first subscript changes fastest.

SYSTAT Commands. To fit models to a y-array, enter data in a file called "MY.DAT" which resembles the input to $SPSS^X$ given above. Then, commands entered on a personal computer follow:

```
SYSTAT
TABLES
USE MY.DAT
TABULATE I*J*K*L
MODEL I*J + I*K + J*L + J*K + I*L + K*L /FITTED, RESIDUALS
MODEL I*J + J*K + I*L + K*L /FITTED, RESIDUALS
MODEL I*J + I*K + J*L + K*L /FITTED, RESIDUALS
```

```
MODEL I*J + I*K + J*L + J*K + I*L /FITTED, RESIDUALS
QUIT
QUIT
```

To fit models to a w-array, replace i 's and j 's with r 's and s 's. Unfortunately **SYSTAT** does not calculate estimates of u -terms.

15.7.2 From Printouts to Parameters

The values found in the printouts can be translated in simple ways to obtain estimates of the parameters of the models. We demonstrate here for α and $(\alpha\beta)$ of p_1 , and note that the translation for β is analogous to that for α .

The table that follows contains some of the u -terms that $SPSS^X$ produces from fitting p_1 to the fabricated network of second-graders.

	K:	0	1
I=1	*	-0.707*	0.707
I=2	*	-0.408*	0.408
I=3	*	0.237*	-0.237
I=4	*	-0.098*	0.098
I=5	*	0.488*	-0.488
I=6		0.488	-0.488

Having named the actors variable " I " and the relation sent variable " K ," the u -terms corresponding to the IK interaction are used to derive the α_{ik} . The key is to look at the subscripts. Because the IK margin is identical to the JL margin, the JL u -terms are also identical to the IK u -terms. (Similarly, just as the IL margin is equal to the JK margin, so are the JK and IL u -terms equal.)

We have noted several of the values in the table above with asterisks to indicate those values that $SPSS^X$ actually prints. We filled in the remainder of the table by subtracting the estimates in a given row from 0, since u -terms have the property that they sum to 0 across all indices (or in all directions).

The IK u -terms are translated to α_{ik} 's by two simple steps. First, make all entries in the first column 0. Second, for dichotomous relations, multiply all the entries in the second column by 2 (Wasserman and Weaver 1985). These simple steps will result in the alpha estimates we reported earlier.

When C , the number of levels of the relational variable, is greater than two, the translation is as follows. First, for each row, subtract the value in the first column from every column (this will result in the first column equaling 0 for all rows, upon completion, and the remaining columns will also have been altered). Second, center Columns 2 through C .

For the reciprocity parameter, *SPSS^X* calculated the estimated value as 0.769. The u -terms for this KL interaction can be written in a 2×2 table, as follows:

	L=0	L=1
K=0	0.769	-0.769
K=1	-0.769	0.769

To obtain the estimate of the model parameter ($\alpha\beta$), place 0's in the first row and column, and then multiply the value 0.769 in the (2, 2) cell by 4. This translation yields the parameter estimate for the reciprocity parameter reported earlier.

For valued relations ($C > 2$), take the following steps. First, for all cells (k, l) where k or l exceeds 2, subtract the value in the first column in row k , and subtract the value in the first row in column l , and add back the value in the (1, 1) (upper leftmost) cell. After all cells for which k or l exceed 2 have been modified, simply change all the values in the first row and column to 0.

The translation is slightly different for users of *GLIM*. The estimated u -terms that are produced for the IK interaction are those in the following table:

	K=0	K=1
I=1	0	0
I=2	0	-0.597
I=3	0	-1.888
I=4	0	-1.217
I=5	0	-2.391
I=6	0	-2.391

Notice that *GLIM* uses a constraint that the first row and the first column of any set of u -terms is defined to be 0. We maintain the first column of 0's, in accordance with standard model constraints, and we simply adjust the remaining columns. Thus, the second column in the table above needs to be "recentered." We compute the mean: $((0) + (-0.597) + (-1.888) + (-1.217) + (-2.391) + (-2.391))/6 = -1.414$, and subtract this value (-1.414) from each element in the second column,

to obtain the following estimates (which are the estimated α 's reported earlier):

	K=0	K=1
I=1	0	(0 - 1.414) = 1.414
I=2	0	(-0.597 - 1.414) = 0.817
I=3	0	(-1.888 - 1.414) = -0.474
I=4	0	(-1.217 - 1.414) = 0.197
I=5	0	(-2.391 - 1.414) = -0.977
I=6	0	(-2.391 - 1.414) = -0.977

When the relational variable is valued, the translation to obtain β 's is the same as that to obtain α 's. The first column remains 0, and all the remaining columns are recentered.

Finally, we note that the constraints *GLIM* uses are exactly those we use for the $(\alpha\beta)$ parameters, whether the relation is dichotomous or valued. That is, *GLIM* will print the following estimated u -terms for the KL interaction:

	L=0	L=1
K=0	0	0
K=1	0	3.077

Clearly, no further adjustment is required.

The standard errors of the model parameter estimates are also derived from slight adjustments of the standard errors of the u -terms. Drawing from Wasserman and Weaver (1985), the standard errors of the model parameter estimates for dichotomous relations are computed as follows:

$$\begin{aligned} SE[\hat{\alpha}_{i(1)}] &= 2SE[\hat{u}_{13(i1)}] \\ SE[\hat{\beta}_{j(1)}] &= 2SE[\hat{u}_{23(j1)}] \\ SE[\widehat{(\alpha\beta)_{11}}] &= 4SE[\hat{u}_{34(11)}] \end{aligned}$$

15.8 Summary

In this chapter, we began by presenting a model for one-mode network data that included effects for the expansiveness of sending actors, the popularity of receiving actors, and the reciprocity of the actors. We discussed models for dichotomous relations, and extended the models to

include general discrete, and ordinal data. We discussed how to model individual actors, and then showed how to model subsets of actors. Thus, we can focus on the dyadic interactions in a network, or we can model both the ties and the actor attributes. We also showed how we might model the patterns in the relational data without considering the actors at all. We briefly described some of the current research on extending these modeling procedures. We demonstrated most of these methods on several examples to give the reader a better understanding of the practice and interpretation of the models and methods. We also described analogous models and methods for two-mode networks.

In the next chapter, we extend these models to multiple relations measured on a network. These multivariate models will allow testing somewhat different hypotheses about network structure. For example, many researchers have been interested in the problem of how one measures the similarity between network representations (Hubert and Arabie 1985; Hubert and Baker 1978; Katz and Powell 1953; Schultz and Hubert 1976; Wasserman 1987). The conformity between two sociomatrices can be measured using a "symmetric" index (like a measure of association), which would answer questions such as, "How similar are these two social networks?" For example, we would expect the relations such as "like" and "dislike" to be negatively associated, but relations such as "like" and "respect" to be positively associated. Alternatively, conformity can be measured in an "asymmetric" way, such as with regression models, to address such questions as, "How well can we predict the structure in one network knowing the structure in another?"