simultaneously cluster people, social predicates, and demographic attributes.

Formally, suppose that the observed data are \( m \) relations involving \( n \) types. Let \( R^i \) be the \( i \)th relation, \( T^j \) be the \( j \)th type, and \( z^j \) be a vector of cluster assignments for \( T^j \). Our task is to infer the cluster assignments, and we are ultimately interested in the posterior distribution \( P(z^1, \ldots, z^n | R^1, \ldots, R^m) \). We specify this distribution by defining a generative model for the relations and the cluster assignments:

\[
P(R^1, \ldots, R^m, z^1, \ldots, z^n) = \prod_{i=1}^{m} P(R^i | z^1, \ldots, z^n) \prod_{j=1}^{n} P(z^j)
\]

where we assume that the relations are conditionally independent given the cluster assignments, and that the cluster assignments for each type are independent. To complete the generative model we first describe the prior on the cluster assignment vectors, \( P(z^j) \), then show how the relations are generated given a set of these vectors.

**Generating clusters**

To allow the IRM the ability to discover the number of clusters in type \( T \), we use a prior that assigns some probability mass to all possible partitions of the type. A reasonable prior should encourage the model to introduce only as many clusters as are warranted by the data. Following previous work on nonparametric Bayesian models (Rasmussen 2002, Antoniak 1974), we use a distribution over partitions induced by a Chinese Restaurant Process (CRP, Pitman 2002).

Imagine building a partition from the ground up: starting with a single cluster containing a single object, and adding objects until all the objects belong to clusters. Under the CRP, each cluster attracts new members in proportion to its size. The distribution over clusters for object \( i \), conditioned on the cluster assignments of objects \( 1, \ldots, i-1 \) is

\[
P(z_i = a | z_1, \ldots, z_{i-1}) = \begin{cases} \frac{n_a}{i-\gamma} & n_a > 0 \\ \frac{\gamma}{i-\gamma} & a \text{ is a new cluster} \end{cases}
\]

where \( n_a \) is the number of objects already assigned to cluster \( a \), and \( \gamma \) is a parameter. The distribution on \( z \) induced by the CRP is exchangeable: the order in which objects are assigned to clusters can be permutated without changing the probability of the resulting partition. \( P(z) \) can therefore be computed by choosing an arbitrary ordering and multiplying conditional probabilities as specified above. Since new objects can always be assigned to new clusters, the IRM effectively has access to a countably infinite collection of clusters, hence the first part of its name.

A CRP prior on partitions is mathematically convenient, and consistent with the intuition that the prior should favor partitions with small numbers of clusters. Yet it is not a universal solution to the problem of choosing the right number of clusters. In some settings we may have prior knowledge that is not captured by the CRP: for instance, we may expect that the clusters will be roughly equal in size. Even so, the CRP provides a useful starting point for structure discovery in novel domains.

**Generating relations from clusters**

We assume that relations are binary-valued functions, although extensions to frequency data and continuous data are straightforward. Consider first a problem with a single type \( T \) and a single two-place relation \( R : T \times T \rightarrow \{0,1\} \). Type \( T \), for example, could be a collection of people, and \( R(i,j) \) might indicate whether person \( i \) likes person \( j \). The complete generative model for this problem is:

\[
z | \gamma \sim \text{CRP}(\gamma) \quad 
\eta(a, b) | \beta \sim \text{Beta}(\beta, \beta) \quad 
R(i, j) | z, \eta \sim \text{Bernoulli}(\eta(z_i, z_j)) \tag{1}
\]