THE NOMINATE MODEL

Legislator i’s (i=1,...,p) utility for the Yea outcome on roll call j (j=1,...,q) is:

\[ U_{ijy} = u_{ijy} + \varepsilon_{ijy} = \beta \left( \frac{1}{2} \sum_{k=1}^{s} d_{iky}^2 \right) + \varepsilon_{ijy} \]  

(1)

where \( u_{ijy} \) is the deterministic portion of the utility function, \( \varepsilon_{ijy} \) is the stochastic portion, and \( d_{iky}^2 \) is the distance of legislator i to the Yea outcome on the \( k^{th} \) \((k=1,...,s)\) dimension for roll call j:

\[ d_{iky}^2 = (X_{ik} - O_{jky})^2 \]  

(2)

and \( X_{ik} \) is the \( i^{th} \) legislator’s ideal point on dimension \( k \), and \( O_{jky} \) is the “Yea” outcome location for the \( j^{th} \) roll call on the \( k^{th} \) dimension.

The same level of utility, \( U_{ijy} \), can be produced in two different ways. First, fixing the scale of the deterministic utility, \( u_{ijy} \), and varying the standard deviation of the error, \( \varepsilon_{ijy} \), that is:

\[ \varepsilon_{ijy} \sim N(0, \frac{\sigma^2}{2}) \]  

so that \( \varepsilon_{ijn} - \varepsilon_{ijy} \sim N(0, \sigma^2) \)
Second, \( \sigma^2 \) can be fixed and the relative weight of \( u_{ijy} \) in the overall utility, \( U_{ijy} \), can be adjusted by increasing/decreasing \( \beta \). In other words, without loss of generality we can assume that:

\[
\varepsilon_{ijn} - \varepsilon_{ijy} \sim N(0, 1) \quad (5)
\]

This implies that the distribution of the difference between the latent utilities for Yea and Nay is normal with mean \( u_{ijy} - u_{ijn} \) and variance 1; that is

\[
y_{ij}^* = U_{ijy} - U_{ijn} = u_{ijy} - u_{ijn} + \varepsilon_{ijn} - \varepsilon_{ijy} = \beta \left\{ e^{-\frac{1}{2} \sum_{k=1}^d d_{ijy}} - e^{-\frac{1}{2} \sum_{k=1}^d d_{ijn}} \right\} + \varepsilon_{ijn} - \varepsilon_{ijy}
\sim N(u_{ijy} - u_{ijn}, 1) \quad (6)
\]

where \( y_{ij}^* \) is the difference between the latent utilities.

The probability that legislator \( i \) votes Yea on the \( j^{th} \) roll call is:

\[
P_{ijy} = P(U_{ijy} > U_{ijn}) = P(\varepsilon_{ijn} - \varepsilon_{ijy} < u_{ijy} - u_{ijn}) = \\
\Phi(u_{ijy} - u_{ijn}) = \\
\Phi \left[ \beta \left\{ e^{-\frac{1}{2} \sum_{k=1}^d d_{ijy}} - e^{-\frac{1}{2} \sum_{k=1}^d d_{ijn}} \right\} \right] \quad (7)
\]
Let \( Y \) be the \( p \) by \( q \) matrix of observed Yea/Nay choices and let \( Y^* \) be the \( p \) by \( q \) matrix of unobserved latent utility differences. From a classical perspective the joint probability distribution of the sample is:

\[
f(Y^* | u_{ijy} - u_{ijn}) = \prod_{i=1}^{p} \prod_{j=1}^{q} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_{ij}^*(u_{ijy} - u_{ijn}))^2}
\]  

(8)

Where \( Y^* \) is the \( p \) by \( q \) matrix of latent utility differences. Note that equation (8) is not a typical joint p.d. of the sample. Technically, a random sample is a set of independent and identically distributed random variables so that the joint p.d. of the sample is (DeGroot, 1986, p. 316):

\[
f_n(X_1, X_2, \ldots, X_n | \Theta) = f(X_1 | \Theta) f(X_2 | \Theta) \ldots f(X_n | \Theta)
\]

where \( f(X | \Theta) \) is the distribution from which the random sample, \( X_1, X_2, \ldots, X_n \), is drawn. In contrast each of the \( pq \) elements of \( Y^* \) is a random sample of size one from the corresponding \( N(u_{ijy} - u_{ijn}, 1) \) distribution. The joint p.d. is a \( pq \)-variate normal distribution with variance-covariance matrix \( I_{pq} \):

\[
f(Y^* | u_{ijy} - u_{ijn}) =
\frac{1}{(2\pi)^{pq/2}} e^{-\frac{1}{2} \sum_{i=1}^{1} \sum_{j=1}^{q} (y_{ij}^* - (u_{ijy} - u_{ijn}))^2}
\]

(8)
To see that equation (8) is indeed a legal probability distribution note that:

\[ f(Y^* | u_{ijy} - u_{ijn}) \geq 0 \text{ for all } y_g^*, \quad -\infty < y_g^* < +\infty \]

and

\[
\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(Y^* | u_{ijy} - u_{ijn}) dy^*_1 dy^*_2 \cdots dy^*_pq =
\]

\[
\left\{ \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_{ij1} - (u_{ijy} - u_{ijn}))^2} dy^*_1 \right\} \cdots \left\{ \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_{ij2} - (u_{ijy} - u_{ijn}))^2} dy^*_2 \right\} \cdots \left\{ \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_{pq} - (u_{pqy} - u_{pqn}))^2} dy^*_pq \right\} = 1 \]

In the joint p.d. of the sample, \( f(Y^* | u_{ijy} - u_{ijn}) \), the \( y_g^* \) are the random variables and the \( ps + 2qs + 1 \) parameters -- \( X_{i1}, X_{i2}, \ldots, X_{is} \), the \( qs \) Yea outcome coordinates -- \( O_{j1y}, O_{j2y}, \ldots, O_{jsy} \), the \( qs \) Nay outcome coordinates -- \( O_{j1n}, O_{j2n}, \ldots, O_{jsn} \), and \( \beta \) -- are fixed constants. Following DeGroot (1986, p. 317), if we regard \( f(Y^* | u_{ijy} - u_{ijn}) \) as a function of the parameters for given values of the \( y_g^* \) then it is a likelihood function; that is

\[
L^*( u_{ijy} - u_{ijn} | Y^* ) =
\]

\[
\frac{1}{(2\pi)^{pq}} e^{-\frac{1}{2} \left\{ (y_{ij1} - (u_{ijy} - u_{ijn}))^2 + (y_{ij2} - (u_{ijy} - u_{ijn}))^2 + \cdots + (y_{pq} - (u_{pqy} - u_{pqn}))^2 \right\}}
\]

(9)
Which is identical to equation (8) only now the $pq{y^*_ij}$ are observed and the $ps+2qs+1$ parameters are variables (but not random variables), and the problem is to find values of the parameters that maximize equation (9).

Equation (8) is a probability distribution over the $pq$ dimensional hyperplane with dimensions corresponding to the $y^*_ij$. Equation (9) is a function defined over the $ps+2qs+1$ dimensional hyperplane with dimensions corresponding to the $ps$ legislator coordinates -- $X_{i1}, X_{i2}, \ldots, X_{is}$, the $qs$ Yea outcome coordinates -- $O_{j1y}, O_{j2y}, \ldots, O_{jsy}$, the $qs$ Nay outcome coordinates -- $O_{j1n}, O_{j2n}, \ldots, O_{jsn}$, and $\beta$. Although equation (9) is not a probability distribution it is the case that it is above zero over the $ps+2qs+1$ hyperplane; that is:

$$L^*(u_{ijy} - u_{ijn} | Y^*) \geq 0, \quad 0 < \beta < +\infty,$$

$$-\infty < X_{ik}, O_{jky}, O_{jkn} < +\infty$$

In addition, the hypervolume underneath the function is almost certainly finite:

$$\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} L^*(\beta \left\{ e^{\left( \frac{1}{2} \sum_{j=1}^{qs} \beta_j Y_{jy}^2 \right)} - e^{\left( \frac{1}{2} \sum_{j=1}^{qs} \beta_j Y_{jn}^2 \right)} \right\}) |Y^*| d\beta dX_{i1} \ldots dX_{ps} dO_{i1y} \ldots dO_{qsy} dO_{i1n} \ldots dO_{qsn} = K^* < +\infty$$

(10)
Because as $\beta \to +\infty$ clearly $L^*(u_{ijy} - u_{ijn} \mid Y^*) \to 0$; and as the absolute value of any of the legislator and roll call parameters becomes large the likelihood function goes to zero; that is, as $|X_{ik}| \to +\infty$ clearly $L^*(u_{ijy} - u_{ijn} \mid Y^*) \to 0$. $L^*$ is shaped like a multivariate normal in that it is quasi-concave along each dimension and asymptotes towards zero fairly quickly. However, I have no formal proof that the hypervolume is finite.

Unfortunately, the latent utility differences are not observed and we do not have any simple expression for the joint probability distribution for the sample of discrete choices -- $f(Y \mid u_{ijy} - u_{ijn})$. However, it is easy to write down the distribution corresponding to any particular choice, that is, $f_{ij}(y_{ij} \mid u_{ijy} - u_{ijn})$. The product of these $pq$ distributions is proportional to the joint p.d. of the sample and the likelihood function. Specifically, let:

$$y_{ij} = \begin{cases} 1 \text{ (Yea) if } y^*_{ij} > 0 \\ 0 \text{ (Nay) if } y^*_{ij} \leq 0 \end{cases}$$

so that

$$P(y^*_{ij} > 0) = \Phi(u_{ijy} - u_{ijn})$$
$$P(y^*_{ij} \leq 0) = 1 - \Phi(u_{ijy} - u_{ijn})$$

If the $y_{ij}$ are independent Bernoulli random variables, that is:

$$f_{ij}(y_{ij} \mid u_{ijy} - u_{ijn}) \sim \text{Bernoulli}(\Phi(u_{ijy} - u_{ijn}))$$

then
where $\tau$ is the index for Yea and Nay, $P_{ij\tau}$ is the probability of voting for choice $\tau$, and $C_{ij\tau}=1$ if the legislator's actual choice is $\tau$ and zero otherwise.

Note that $f(Y \mid u_{ijy}-u_{ijn})$ is a discrete distribution with $2^{pq}$ possible outcomes. By definition, the $ps+2qs+1$ parameters are fixed constants and the $y_{ij}$ are the random variables. Hence, we can apply standard probability theory to find the proportionality constant:

$$
\sum_{y_{i1}=0}^{1} \sum_{y_{j1}=0}^{1} \cdots \sum_{y_{pq}=0}^{1} \left\{ \prod_{i=1}^{p} \prod_{j=1}^{q} f_{ij}(y_{ij} \mid u_{ijy}-u_{ijn}) \right\} = K
$$

So that

$$
f(Y \mid u_{ijy}-u_{ijn}) = \frac{1}{K} \left\{ \prod_{i=1}^{p} \prod_{j=1}^{q} f_{ij}(y_{ij} \mid u_{ijy}-u_{ijn}) \right\}
$$

Fortunately, knowing the value of $K$ is not important and does not affect the analysis of the likelihood function, $L(u_{ijy}-u_{ijn} \mid Y)$. Technically, the likelihood function has exactly the same expression as $f(Y \mid u_{ijy}-u_{ijn})$. In this case the
1/K is missing. This has no effect because it is as if the true likelihood function were multiplied by K. When gradient methods are applied to the likelihood function all that matters is the relative heights of the function. In addition, when logs are taken the proportionality constant becomes an additive constant and plays no role in the estimation.

The joint p.d. of the sample, \( f(\mathbf{Y} | u_{ijy} - u_{ijn}) \), is a discrete probability distribution with \( 2^{pq} \) possible outcomes. The likelihood function, \( L(u_{ijy} - u_{ijn} | \mathbf{Y}) \), is a continuous distribution over the \( ps+2qs+1 \) dimensional hyperplane with dimensions corresponding to the \( ps \) legislator coordinates, the \( 2qs \) outcome coordinates, and \( \beta \). Although equation (11) is not a probability distribution it is the case that it is above zero over the \( ps+2qs+1 \) hyperplane; that is:

\[
L(\ u_{ijy} - u_{ijn} | \mathbf{Y} ) \geq 0, \ 0 < \beta < +\infty, \\
-\infty < X_{ik}, O_{jky}, O_{jkn} < +\infty
\]

Unfortunately the hypervolume underneath the function is not finite; that is:

\[
\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} L(\beta) \left[ e^{\left[-\frac{1}{2} \sum_{k=1}^{ps} d_{ik}^2\right]} - e^{\left[-\frac{1}{2} \sum_{k=1}^{2qs} d_{ijk}^2\right]} \right] \mathbf{Y} d\beta dX_1 \ldots dX_{ps} dO_{i1} \ldots dO_{qy} \ldots dO_{i1n} \ldots dO_{qyn} = +\infty
\]

(12)
The likelihood function in equation (11) is the product of the \( pq \) probabilities of the observed choices. The value of the function is a maximum of 1 and a minimum of zero. Note that if all the legislators are voting correctly, that is, \( P_{ijc} > .5 \) (or \( u_{ijc} > u_{ijb} \)) for all \( i \) and \( j \) where “c” means “correct choice” and “b” means “incorrect choice”, then as \( \beta \rightarrow +\infty \) clearly \( L( u_{ijc} - u_{ijb} \mid Y ) \rightarrow 1 \). Conversely, if for at least one choice a legislator votes “incorrectly”, \( u_{ijc} < u_{ijb} \) so that \( P_{ijc} < .5 \), then as \( \beta \rightarrow +\infty \) clearly \( L( u_{ijc} - u_{ijb} \mid Y ) \rightarrow 0 \). With voting error \( L( u_{ijy} - u_{ijn} \mid Y ) \) asymptotes very quickly to zero because \( \Phi(u_{ijc} - u_{ijb}) \) goes to zero very quickly as \( \beta \) increases.

Now consider the effect of the legislator and outcome coordinates. Suppose \( |X_{ik}| \rightarrow +\infty \) then a simple inspection of equation (7) shows that:

\[
\Phi \left[ \beta \left\{ e^{\left( \frac{-\sum_{k=1}^{q} d_{ijkc}}{2} \right)} - e^{\left( \frac{-\sum_{k=1}^{q} d_{ijkb}}{2} \right)} \right\} \right] \rightarrow \Phi \left[ \beta \left\{ e^{(-\infty)} - e^{(-\infty)} \right\} \right] = \Phi \left[ \beta \{0\} \right] = .5
\]

The portion of the likelihood function corresponding to legislator \( i \) is:

\[
\prod_{j=1}^{q} \Phi(u_{ijc} - u_{ijb}) \quad (13)
\]

So that this converges to \( .5^q \) as \( |X_{ik}| \rightarrow +\infty \). This shows that
\( \mathbf{L}( u_{ijy} - u_{ijn} \mid \mathbf{Y} ) \) does not asymptote to zero along the dimensions corresponding to legislator coordinates so that the hypervolume underneath \( \mathbf{L}( u_{ijy} - u_{ijn} \mid \mathbf{Y} ) \) is infinite.

The fact that \( \mathbf{L}( u_{ijy} - u_{ijn} \mid \mathbf{Y} ) \) has an infinite hypervolume has no practical effect on a standard maximum likelihood analysis. This is so because at a great distance from the center of the space defined by the legislator and outcome points the likelihood function is a flat, featureless vista. That is, the maxima are towards the interior of the function and are easily found by conventional gradient and quasi-gradient methods. However, the use of simulation methods is inappropriate because \( \mathbf{L}( u_{ijy} - u_{ijn} \mid \mathbf{Y} ) \) cannot be treated as if it were a probability distribution.

The Bayesian approach avoids the problem of infinite volume through the judicious choice of prior distributions that when multiplied through the likelihood function produce a distribution that is proportional to a probability distribution. In a standard Bayesian approach the prior distribution for a legislator ideal point is a multivariate normal distribution with variance-covariance matrix \( \sigma^2 \mathbf{I}_s \):

\[
\xi(\mathbf{x}_i) = \frac{1}{(2\pi\sigma)^{\frac{s}{2}}} e^{-\frac{1}{2\sigma^2}(X_i^2+X_2^2+\ldots+X_s^2)}
\]  

(14)
Similarly, assume that the prior distributions for the outcome points are also multivariate normal distributions with variance-covariance matrices $\sigma^2 I_s$:

$$
\xi(O_{jy}) = \frac{1}{(2\pi\sigma)^{\frac{s}{2}}} e^{-\frac{1}{2\sigma^2} (O^2_{jy} + O^2_{j2} + \ldots + O^2_{jn})}
$$

and

$$
\xi(O_{jn}) = \frac{1}{(2\pi\sigma)^{\frac{s}{2}}} e^{-\frac{1}{2\sigma^2} (O^2_{jn} + O^2_{j2n} + \ldots + O^2_{jnn})}
$$

(15)

The posterior distribution for the NOMINATE model is:

$$
\xi(\beta, O_y, O_n, X | Y) \propto \prod_{i=1}^{p} \prod_{j=1}^{q} \left\{ f(y_{ij} | u_{ij} - u_{ijm}) \xi(X_i) \xi(O_{jy}) \xi(O_{jn}) \right\} =
$$

$$
\prod_{i=1}^{p} \prod_{j=1}^{q} \left\{ [\Phi(u_{ij} - u_{ijm})]^y_i [1 - \Phi(u_{ij} - u_{ijm})]^t_i \xi(X_i) \xi(O_{jy}) \xi(O_{jn}) \right\} =
$$

$$
\prod_{i=1}^{p} \prod_{j=1}^{q} \xi(X_i) \xi(O_{jy}) \xi(O_{jn}) \prod_{t=1}^{C_{ijm}} p_{ijt}
$$

(16)

By definition

$$
\xi(\beta, O_y, O_n, X | Y) \geq 0, \quad 0 < \beta < +\infty,
$$

$$
-\infty < X_{ik} , O_{jky} , O_{jkn} < +\infty
$$
and the hypervolume underneath $\xi(\beta, O_y, O_n, X \mid Y)$ is finite. Specifically, as the legislator ideal points and/or the outcome points go to $+\infty$ then $\xi(\beta, O_y, O_n, X \mid Y)$ goes to zero. For example, as $|X_{ik}| \to +\infty$ then $\xi(X_i) \to 0$ so that

$$\xi(\beta, O_y, O_n, X \mid Y) \to 0.$$ 

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \xi(\beta, O_y, O_n, X \mid Y)d\beta dX_{i1} \cdots dX_{iK} dO_{1y} \cdots dO_{qy} dO_{1n} \cdots dO_{qn} = K < +\infty$$

Theoretically, simulation methods can be applied to the posterior distribution expressed in equation (16) but not to the likelihood distribution expressed in equation (11). The irony here is that if “believable” standard errors are to be obtained for the basic model either the parametric bootstrap (L&P) or the simulation approach are the best methods to do so because inverting the information matrix is problematic. I say “ironic” because one could favor the use of (16) over (11) on these practical grounds rather than the philosophical grounds of the “Bayesians” (non-“Frequentists”).

12