Bayesian Analysis of Choice Data

Simon Jackman

Stanford University
http://jackman.stanford.edu/BASS

February 3, 2012
Discrete Choice

- binary (e.g., probit model; we looked at with data augmentation)
Discrete Choice

- binary (e.g., probit model; we looked at with data augmentation)
- ordinal (ordinal logit or probit)
Discrete Choice

- binary (e.g., probit model; we looked at with data augmentation)
- ordinal (ordinal logit or probit)
- multinomial models for unordered choices: e.g., multinomial logit (MNL), multinomial probit (MNP).
Discrete Choice

- Binary (e.g., probit model; we looked at with data augmentation)
- Ordinal (ordinal logit or probit)
- Multinomial models for unordered choices: e.g., multinomial logit (MNL), multinomial probit (MNP). We won’t consider models for “tree-like” choice structures (nested logit, GEV, etc).
Binary Choices: logit or probit

- For “standard” models (e.g., no “fancy” hierarchical structure, no concerns re missing data etc), other avenues besides BUGS/JAGS
  - E.g., MCMCpack
  - `dbern` or `dbin` and sample from the conditional distributions using Metropolis-within-Gibbs, slice sampling
Voter turnout example.

```jags
model{
  for (i in 1:N) {
    y[i] ~ dbern(p[i]) ## binary outcome
    logit(p[i]) <- ystar[i] ## logit link
    ystar[i] <- beta[1] ## regression structure for covariates
    + beta[2]*educ[i]
    + beta[3]*(educ[i]*educ[i])
    + beta[4]*age[i]
    + beta[5]*(age[i]*age[i])
    + beta[6]*south[i]
    + beta[7]*govelec[i]
    + beta[8]*closing[i]
    + beta[9]*(closing[i]*educ[i])
    + beta[10]*(educ[i]*educ[i]*closing[i])
  }

  ## priors
  beta[1:10] ~ dnorm(mu[], B[, ] ) # diffuse multivariate Normal prior
  # see data file
}
```
Binary Data Is Binomial Data when Grouped (§8.1.4)

- big, micro-level data sets with binary data (e.g., CPS)
- MCMC gets slow
- collapse the data into covariate classes, treat as binomial data; much smaller data set, much shorter run-times
- $y_i | x_i \sim \text{Bernoulli}(F[x_i \beta])$, where $x_i$ is a vector of covariates.
- **Covariate classes**: a set $C = \{ i : x_i = x_C \}$ i.e., the set of respondents who have covariate vector $x_C$.
- probability assignments over $y_i \forall i \in C$ are conditionally exchangeable given their common $x_i$ and $\beta$.
- binomial model $r_C \sim \text{Binomial}(p_C; n_C)$, where $p_C = F(x_C \beta)$, $r_C = \sum_{i \in C} y_i$ is the number of “successes” in $C$ and $n_C$ is the cardinality of $C$. 
Example 8.5; binomial model for grouped binary data

Form covariate classes, and groupedData object; original data set \(n \approx 99000\); only 636 unique covariates classes.

---

R code

```r
# collapse by covariate classes
X <- cbind(nagler$age, nagler$educYrs)
X <- apply(X, 1, paste, collapse = ":")
covClasses <- match(X, unique(X))
covX <- matrix(unlist(strsplit(unique(X), ":")), ncol = 2, byrow = TRUE)
r <- tapply(nagler$turnout, covClasses, sum)
n <- tapply(nagler$turnout, covClasses, length)
groupedData <- list(n = n, r = r,
                      age = as.numeric(covX[, 1]),
                      educYrs = as.numeric(covX[, 2]),
                      NOBS = length(n))
```

Simon Jackman (Stanford)
We can then pass the `groupedData` data frame to JAGS. We specify the binomial model $r_i \sim \text{Binomial}(p_i; n_i)$ with $p_i = F(x_i \beta)$ and vague normal priors on $\beta$ with the following code:

```jags
model{
  for (i in 1:NOBS){
    logit(p[i]) <- beta[1] + age[i]*beta[2]
    + pow(age[i],2)*beta[3]
    + educYrs[i]*beta[4]
    + pow(educYrs[i],2)*beta[5]
    r[i] ~ dbin(p[i],n[i])  ## binomial model for each covariate class
  }

  beta[1:5] ~ dmnorm(b0[,],B0[,])
}
```
Ordinal Responses

e.g., 7-point scale when measuring party identification in the U.S., assigning the numerals $y_i \in \{0, \ldots, 6\}$ to the categories \{“Strong Republican”, “Weak Republican”, \ldots, “Strong Democrat”\}.

Censored, latent variable representation:

$$y_i^* = x_i \beta + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2), \quad i = 1, \ldots, n.$$  

$y_i = 0 \iff y_i^* < \tau_1$  

$y_i = j \iff \tau_j < y_i^* \leq \tau_{j+1}, \quad j = 1, \ldots, J - 1$  

$y_i = J \iff y_i^* > \tau_J$

Threshold parameters obey the ordering constraint  

$\tau_1 < \tau_2 < \ldots < \tau_J$.

The assumption of normality for $\epsilon_i$ generates the probit version of the model; a logistic density generates the ordinal logistic model.

Bayesian analysis: we want $p(\beta, \tau | y, X) \propto p(y | X, \beta, \tau) p(\beta, \tau)$.  

Simon Jackman (Stanford)  
Bayesian Analysis of Choice Data  
February 3, 2012  
8 / 1
Ordinal responses, $y_i^* \sim N(x_i \beta, \sigma^2)$

$$
Pr[y_i = j] = \Phi[(\tau_{j+1} - x_i \beta)/\sigma] - \Phi[(\tau_j - x_i \beta)/\sigma]
$$
Identification

\[ y_i^* = x_i \beta + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2), \quad i = 1, \ldots, n. \]

\[ y_i = 0 \iff y_i^* < \tau_1 \]
\[ y_i = j \iff \tau_j < y_i^* \leq \tau_{j+1}, \quad j = 1, \ldots, J - 1 \]
\[ y_i = J \iff y_i^* > \tau_J \]

- Model needs identification constraints
- Set one of the \( \tau \) to a point (zero); set \( \sigma \) to a constant (one)
- Drop the intercept and fix \( \sigma \)
- Fix two of the \( \tau \) parameters.
Priors on thresholds

\[ \tau_j \sim N(0, 10^2), \text{ subject to ordering constraint } \tau_j > \tau_{j-1}, \forall j = 2, \ldots, J. \]

In JAGS only, use nifty `sort` function:

```
for(j in 1:4){
    tau0[j] ~ dnorm(0,.01)
}
tau[1:4] <- sort(tau0)  ## JAGS only, not in WinBUGS!
```

BUGS:

\[
\begin{align*}
\tau_1 & \sim N(t_1, T_1) \\
\delta_j & \sim \text{Exponential}(d), \quad j = 2, \ldots, J, \\
\tau_j & = \tau_{j-1} + \delta_j, \quad j = 2, \ldots, J,
\end{align*}
\]

```
tau[1] ~ dnorm(0,.01)
for(j in 1:3){
    delta[j] ~ dexp(2)
    tau[j+1] <- tau[j] + delta[j]
}
```
Example 8.6, interviewer ratings of respondents

- 5 point rating scale used by interviewers in assessing respondents’ levels of political information

- In 2000 ANES:

<table>
<thead>
<tr>
<th>Label</th>
<th>y</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Low</td>
<td>0</td>
<td>105</td>
<td>6</td>
</tr>
<tr>
<td>Fairly Low</td>
<td>1</td>
<td>334</td>
<td>19</td>
</tr>
<tr>
<td>Average</td>
<td>2</td>
<td>586</td>
<td>33</td>
</tr>
<tr>
<td>Fairly High</td>
<td>3</td>
<td>450</td>
<td>25</td>
</tr>
<tr>
<td>Very High</td>
<td>4</td>
<td>325</td>
<td>18</td>
</tr>
</tbody>
</table>

- covariates: education, gender, age, home-owner, public sector employment
model{
  for(i in 1:N) {
    ## loop over observations
    ## form the linear predictor (no intercept)
    mu[i] <- x[i,1]*beta[1] +
              x[i,2]*beta[2] +
              x[i,3]*beta[3] +
              x[i,4]*beta[4] +
              x[i,5]*beta[5] +
              x[i,6]*beta[6]

    ## cumulative logistic probabilities
    logit(Q[i,1]) <- tau[1]-mu[i]
    p[i,1] <- Q[i,1]
    for(j in 2:4) {
      logit(Q[i,j]) <- tau[j]-mu[i]
      ## trick to get slice of the cdf we need
      p[i,j] <- Q[i,j] - Q[i,j-1]
    }
    p[i,5] <- 1 - Q[i,4]
    y[i] ~ dcat(p[i,1:5]) ## p[i,] sums to 1 for each i
  }

  ## priors over betas
  beta[1:6] ~ dmnorm(b0[1],B0[1,1])

  ## thresholds
  for(j in 1:4) {
    tau0[j] ~ dnorm(0, .01)
  }
  tau[1:4] <- sort(tau0) ## JAGS only not in BUGS!
}
exploit lack of identification
run the MCMC algorithm deployed in the space of unidentified parameters
post-processing: map MCMC output back mixes better than the MCMC algorithm in the space of the identified parameters
get a better mixing Markov chain
in ordinal model case, exploit lack of identification between thresholds and intercept parameters
take care!
Interviewer heterogeneity in scale-use

- Different interviewers use the rating scale differently: e.g., interviewer \( k \) is a tougher grader than interviewer \( k' \).
- We tap this with a set of interviewer terms, varying over interviewers \( k = 1, \ldots, K \).
- We augment the usual ordinal model as follows:

\[
\begin{align*}
\Pr(y_i \geq j) &= F(\tau_j - \mu_i), \quad j = 0, \ldots, J - 1 \\
\Pr(y_i = J) &= 1 - F(\tau_{j-1} - \mu_i) \\
\mu_i &= x_i \beta + \eta_k \\
\eta_k &\sim N(0, \sigma^2) \quad k = 1, \ldots, K
\end{align*}
\]

- A positive \( \eta_k \) is equivalent to the thresholds being shifted down (i.e., interviewer \( k \) is an easier-than-average grader).
- Zero-mean restriction on \( \eta_k \): why?
- Alternative model: each interviewer gets their own set of thresholds, perhaps fit these hierarchically.
model{
  for(i in 1:N) {  ## loop over observations
    ## form the linear predictor

    ## cumulative logistic probabilities
    logit(Q[i,1]) <- tau[1]-mu[i]
    p[i,1] <- Q[i,1]
    for(j in 2:4){
      logit(Q[i,j]) <- tau[j]-mu[i]
      p[i,j] <- Q[i,j] - Q[i,j-1]
    }
    p[i,5] <- 1 - Q[i,4]
    y[i] ~ dcat(p[i,1:5])  ## p[i,] sums to 1 for each i
  }
  ## priors over betas
  beta[1:6] ~ dmnorm(b0[], B0[,])
  ## hierarchical model over etas, note zero mean restriction
  for(k in 1:NID){
    eta[k] ~ dnorm(0.0, eta.tau)
  }
  eta.tau <- 1/pow(sigma,2)  ## convert stddev to precision
  sigma ~ dunif(0,2)
  ## priors over thresholds
  for(j in 1:4){
    tau0[j] ~ dnorm(0,.01)
  }
  tau[1:4] <- sort(tau0)  ## JAGS only, not in WinBUGS!
}
since $\eta_k \sim N(0, \sigma^2)$, if we set $\sigma$ to its posterior mean of .77, then half of the interviewer effects will lie more than 1.35 $\sigma \approx 1.04$ “logits” away from zero.
Tabular summary of results

<table>
<thead>
<tr>
<th></th>
<th>Non-Hierarchical</th>
<th>Hierarchical</th>
</tr>
</thead>
<tbody>
<tr>
<td>College Degree</td>
<td>1.46 (0.10)</td>
<td>1.61 (0.10)</td>
</tr>
<tr>
<td>Female</td>
<td>-0.66 (0.09)</td>
<td>-0.76 (0.09)</td>
</tr>
<tr>
<td>log(Age)</td>
<td>0.47 (0.12)</td>
<td>0.42 (0.13)</td>
</tr>
<tr>
<td>Home Owner</td>
<td>0.45 (0.10)</td>
<td>0.48 (0.10)</td>
</tr>
<tr>
<td>Government Employee</td>
<td>0.17 (0.14)</td>
<td>0.16 (0.14)</td>
</tr>
<tr>
<td>log(Interview Length)</td>
<td>1.13 (0.15)</td>
<td>1.45 (0.18)</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.00</td>
<td>0.77 (0.08)</td>
</tr>
</tbody>
</table>

Threshold parameters:

<table>
<thead>
<tr>
<th>( \tau_0 )</th>
<th>3.85 (0.67)</th>
<th>4.69 (0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_1 )</td>
<td>5.66 (0.67)</td>
<td>6.60 (0.75)</td>
</tr>
<tr>
<td>( \tau_2 )</td>
<td>7.37 (0.68)</td>
<td>8.46 (0.76)</td>
</tr>
<tr>
<td>( \tau_3 )</td>
<td>8.83 (0.69)</td>
<td>10.08 (0.77)</td>
</tr>
</tbody>
</table>

Models for Multinomial Choices, §8.3

- multinomial logit (MNL), §8.3.1
- multinomial probit (MNP) §8.3.2
Multinomial logit (MNL), §8.3.1

- Random utility rationale: utility to decision-maker \( i \) of choice \( j \) is linear in some predictors, plus a random component,

\[
U_{ij} = x_i \beta_j + \varepsilon_{ij}, j = 0, \ldots, J
\]

- \( \varepsilon_{ij} \) are drawn a distribution whose cumulative distribution function is a Type-1 extreme value distribution with functional form \( F(\varepsilon_{ij}) = \exp[-\exp(-\varepsilon_{ij})] \) and hence \( \varepsilon_{ij} \) has density

\[
p(\varepsilon_{ij}) = \exp(\varepsilon_{ij})\exp[-\exp(-\varepsilon_{ij})].
\]

- Decision-maker \( i \) chooses option \( j \) with probability

\[
\pi_{ij} = \Pr(y_i = j) = \Pr[U_{ij} > U_{ik}], \quad \forall \ k \neq j.
\]
Consider a choice set with 3 elements, {"0", "1", "2"}. Suppose we observe $y_i = 2$:

$$\Pr(y_i = 2) = \Pr(U_{i2} > U_{i1}, U_{i2} > U_{i0})$$

$$= \Pr[x_i\beta_2 + \epsilon_{i2} > x_i\beta_1 + \epsilon_{i1}, x_i\beta_2 + \epsilon_{i2} > x_i\beta_0 + \epsilon_{i0}],$$

$$= \Pr[\epsilon_{i2} + x_i\beta_2 - x_i\beta_1 > \epsilon_{i1}, \epsilon_{i2} + x_i\beta_2 - x_i\beta_0 > \epsilon_{i0}],$$

$$= \int_{-\infty}^{\infty} f(\epsilon_2) \left[ \int_{-\infty}^{\epsilon_{i2}+x_i\beta_2-x_i\beta_1} f(\epsilon_1) d\epsilon_1 \cdot \int_{-\infty}^{\epsilon_{i2}+x_i\beta_2-x_i\beta_0} f(\epsilon_0) d\epsilon_0 \right] d\epsilon_2,$$

$$= \int_{-\infty}^{\infty} f(\epsilon_2) \times \exp[-\exp(-\epsilon_{i2} - x_i\beta_2 + x_i\beta_1)] \times \exp[-\exp(-\epsilon_{i2} - x_i\beta_2 + x_i\beta_0)]$$

$$= \frac{\exp(x_i\beta_2)}{\exp(x_i\beta_0) + \exp(x_i\beta_1) + \exp(x_i\beta_2)}.$$

Thus:

$$\pi_{ij} = \Pr(y_i = j) = \frac{\exp(x_i\beta_j)}{\sum_{k=0}^{J} \exp(x_i\beta_k)}.$$
Multinomial Logit (MNL), §8.3.1

\[ \pi_{ij} = \Pr(y_i = j) = \frac{\exp(x_i \beta_j)}{\sum_{k=0}^{J} \exp(x_i \beta_k)}. \]

- Identification by normalizing on a "baseline outcome", e.g., \( \beta_0 = 0 \).
- Independence of irrelevant alternatives §8.3.2
Vote choice in the 1992 U.S. Presidential election
ANES data; choices are Clinton, George H.W. Bush, Perot. \( n = 909 \).
Original analysis by Alvarez and Nagler (1995), who used MNP.
Predictors: dummies for Dem or Rep party-id, dummy for gender, retrospective evaluations of the national economy \((-1, 0, 1)\), and \(z_{ij}\), square of the distance of respondent \(i\) from candidate \(j\).

\[
\Pr(U_{ij} > U_{ik}) = \Pr(x_i \beta_j + z_{ij} \gamma + \epsilon_{ij} - x_i \beta_k - z_{ij} \gamma - \epsilon_{ik} > 0)
= \Pr(x_i [\beta_j - \beta_k] + [z_{ij} - z_{ik}] \gamma > \epsilon_{ik} - \epsilon_{ij}).
\]
Example 8.7, using dcat

JAGS code

model{
  for(i in 1:NOBS){
    for(j in 1:3){  ## loop over choices
      mu[i,j] <- beta[j,1]
      + beta[j,2]*dem[i]
      + beta[j,3]*ind[i]
      + beta[j,4]*rep[i]
      + beta[j,5]*female[i]
      + beta[j,6]*natlecon[i]
      + gamma*dist[i,j]
      emu[i,j] <- exp(mu[i,j])
      p[i,j] <- emu[i,j]/sum(emu[i,1:3])
    }
    y[i] ~ dcat(p[i,1:3])
  }

  ## priors
  for(k in 1:6){
    beta[1,k] <- 0  ## identifying restriction
  }
  for(j in 2:3){
    beta[j,1:6] ~ dmnorm(b0,B0)  ## b0, B0 passed as data from R
  }
  gamma ~ dnorm(0,.01)

  ## plus code for mapping to identified parameters, see book
}
same random utility rationale:

\[ U_{ij} = r_{ij} \beta + v_{ij}, \quad j = 0, 1, \ldots, J; i = 1, \ldots, n \]

MNP for MVN model for un-modelled sources of utility:

\[ v_i = (v_{i1}, \ldots, v_{ij})' \overset{iid}{\sim} N(0, V) \]

where \( V \) is a \((J + 1)\)-by-\((J + 1)\) covariance matrix.

But probabilities are difficult to compute:

\[ \pi_{ij} = \Pr(y_i = j) = \Pr(U_{ij} > U_{ik}), \quad \forall \ k \neq j \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{U_{ij}} \cdots \int_{-\infty}^{U_{ij}} f(U_0, U_1, \ldots, U_J) \, dU_0 \, dU_1 \cdots \, dU_j \]
if choice $j$ is observed for person $i$, we know that $U_{ij} - U_{ik} > 0 \forall j \neq k$.

Without loss of generality choose a “baseline” outcome, $j = 0$, and define the utility differences $w_i = (w_{i1}, \ldots, w_{ij})'$ with $w_{ij} = U_{ij} - U_{i0}$, $j = 1, \ldots, J$:

$$w_{ij} = (r_{ij} - r_{i0})\beta + v_{ij} - v_{i0} = x_{ij}\beta + \epsilon_{ij}.$$ 

where $\epsilon_{ij} \overset{iid}{\sim} N(0, \Sigma)$

mapping from latent variables to observed choices:

$$y_i = h(w_i) \equiv \begin{cases} 0 & \text{if max} (w_i) < 0 \\ j & \text{if max} (w_i) = w_{ij} > 0 \end{cases}$$

Identification: the distribution of $y|X, \beta, \Sigma$ is the same as the distribution of $y|X, c\beta, c^2\Sigma$

solution: set $\sigma_{11} = 1$. 
posterior density $p(\beta, \Sigma | y, X)$

1. sample $w_i^{(t)}$ from $p(w_i | \beta^{(t-1)}, \Sigma^{(t-1)}, y, X)$, $i = 1, \ldots, n$, the data-augmentation step
2. sample $\beta^{(t)}$ from $p(\beta | \Sigma^{(t-1)}, W^{(t)}, y, X)$.
3. sample $\Sigma^{(t)}$ from $p(\Sigma | \beta^{(t)}, W^{(t)}, y, X)$.

Conditional on the latent $w_i$, we have a very simple multivariate normal regression (McCulloch and Rossi 1994; Chib and Greenberg 1997; McCulloch, Polson and Rossi 1998).

For step 3, the prior and the conditional distribution for $\Sigma$ is complicated by the identifying constraint $\sigma_{11} = 1$.

Implemented in MNP package in R (Imai and van Dyk 2005).
Example 8.8, 1992 U.S. Presidential election

- $n = 909$, $j \in \{\text{Perot, Bush, Clinton}\}$
- Mix of individual (party-id, gender, evaluations of the economy) and choice-specific covariates (squared ideological distance from candidates)
- MNP in $\mathbb{R}$, 1.5M iterations, extremely inefficient exploration of the posterior densities for some parameters
Example 8.8, 1992 U.S. Presidential election

Autocorrelations: $\sigma_{22}$

Effective sample size: 2,412

Autocorrelations: $\sigma_{12}$

Effective sample size: 2,540

Autocorrelations: $\rho$

Effective sample size: 2,575
### Example 8.8, 1992 U.S. Presidential election

1.5 million iterations:

<table>
<thead>
<tr>
<th></th>
<th>$z$</th>
<th>$p$</th>
<th>$\rho$</th>
<th>$N$</th>
<th>$l$</th>
<th>EffSamp</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{11}$, Intercept, Perot</td>
<td>-0.82</td>
<td>0.62</td>
<td>0.37</td>
<td>303,375</td>
<td>81.00</td>
<td>5,908</td>
</tr>
<tr>
<td>$\beta_{21}$, Intercept, Clinton</td>
<td>-0.36</td>
<td>0.36</td>
<td>0.46</td>
<td>547,875</td>
<td>146.00</td>
<td>6,211</td>
</tr>
<tr>
<td>$\beta_{12}$, Dem Id, Perot</td>
<td>-1.12</td>
<td>0.49</td>
<td>0.48</td>
<td>320,025</td>
<td>85.40</td>
<td>4,076</td>
</tr>
<tr>
<td>$\beta_{22}$, Dem Id, Clinton</td>
<td>-0.27</td>
<td>0.69</td>
<td>0.73</td>
<td>860,850</td>
<td>230.00</td>
<td>3,175</td>
</tr>
<tr>
<td>$\beta_{13}$, Repub Id, Perot</td>
<td>0.66</td>
<td>0.72</td>
<td>0.34</td>
<td>659,850</td>
<td>176.00</td>
<td>6,599</td>
</tr>
<tr>
<td>$\beta_{23}$, Repub Id, Clinton</td>
<td>0.42</td>
<td>0.72</td>
<td>0.75</td>
<td>958,400</td>
<td>256.00</td>
<td>2,763</td>
</tr>
<tr>
<td>$\beta_{14}$, Female, Perot</td>
<td>-0.07</td>
<td>0.32</td>
<td>0.01</td>
<td>94,025</td>
<td>25.10</td>
<td>58,544</td>
</tr>
<tr>
<td>$\beta_{24}$, Female, Clinton</td>
<td>1.04</td>
<td>0.51</td>
<td>0.02</td>
<td>99,850</td>
<td>26.70</td>
<td>50,304</td>
</tr>
<tr>
<td>$\beta_{15}$, Econ Retro, Perot</td>
<td>0.51</td>
<td>0.97</td>
<td>0.18</td>
<td>198,950</td>
<td>53.10</td>
<td>11,272</td>
</tr>
<tr>
<td>$\beta_{25}$, Econ Retro, Clinton</td>
<td>0.15</td>
<td>0.57</td>
<td>0.59</td>
<td>607,750</td>
<td>162.00</td>
<td>3,709</td>
</tr>
<tr>
<td>$\gamma$, Ideological Distance</td>
<td>0.21</td>
<td>0.57</td>
<td>0.72</td>
<td>967,950</td>
<td>258.00</td>
<td>2,778</td>
</tr>
<tr>
<td>$\sigma_{12}$</td>
<td>-1.46</td>
<td>0.09</td>
<td>0.90</td>
<td>975,375</td>
<td>260.00</td>
<td>2,540</td>
</tr>
<tr>
<td>$\sigma_{22}$</td>
<td>-0.78</td>
<td>0.98</td>
<td>0.88</td>
<td>902,500</td>
<td>241.00</td>
<td>2,412</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-1.16</td>
<td>0.33</td>
<td>0.89</td>
<td>1,159,350</td>
<td>309.00</td>
<td>2,575</td>
</tr>
</tbody>
</table>


