Variance Partition Coefficients (Simulation and Approximation Methods)

Variance partition coefficients (VPCs) are not straightforward for beta-distributed mixed models, for much the same reasons that they are problematic in the binomial or logistic mixed models. Here, we follow the model linearization (approximation) and simulation methods proposed in Goldstein, Browne, and Rasbash (2002) and extended in Browne, Subramanian, Jones, and Goldstein (2005). VPC results are in the Excel file VPV_1x4_1.xls.

The approximation method, as explained in Goldstein, Browne and Rasbash (2002), uses a first-order Taylor expansion of the link function at the mean of the distribution of the level 2 random effects. Browne et al. (2005) extend their approach to mixed models with more than two levels. We adapt their approach for the logistic link by incorporating the dispersion effect parameter ϕ. In a 2-level model, for one or more appropriate combinations of values for the covariates x_j, their formulation yields

$\sigma^2_{\mu j|xj} = [\sigma_u\mu_{(j)}/\exp(\Sigma\beta_j x_j)]^2$, and

$\sigma^2_{\mu(j)} = \mu_{(j)}(1 - \mu_{(j)})/(\phi + 1)$,

where $\mu_{(j)}$ denotes the predicted value for a combination of x_j values. Finally, the approximate VPC is

$\tau_j = \sigma^2_{\mu j|xj}/(\sigma^2_{\mu j|xj} + \sigma^2_{\mu(j)})$.

The simulation approach includes the following steps.

1. We start with an estimated location submodel $\mu_{ij} = \exp(\Sigma\beta_j x_{ij} + u_i)/[1 + \exp(\Sigma\beta_j x_{ij} + u_i)]$ with $u_i \sim N(0,\sigma_u)$, and dispersion submodel $\phi = \exp(-\delta_0)$.
2. Generate a large number (say, M = 5000) of random level-2 residual values $u_m \sim N(0,\sigma_u)$, for m from 1 to M.
3. Select one or more appropriate combinations of values for the covariates x_j. For each such combination, use sample estimates of the relevant parameters to compute M predicted values $\mu_{m(j)} = \exp(\Sigma\beta_j x_j + u_m)/[1 + \exp(\Sigma\beta_j x_j + u_m)]$. In our 1x4 example, we do this for each of the four combinations of 0-1 values for the dummy x_j variables.
4. Compute the variance $\sigma^2_{u(j)}$ of the μ_m.
5. Compute $\sigma^2_{\mu(j)} = \mu_{(j)}(1 - \mu_{(j)})/(\phi + 1)$, where $\mu_{(j)}$ denotes the predicted value for a combination of x_j values.
6. The approximate VPC is $\tau_j = \sigma^2_{u(j)}/(\sigma^2_{u(j)} + \sigma^2_{\mu(j)})$.

First sample of 1x4 Data with VPC Approximations

<u>Model and Parameter Estimates</u>

The BUGs code below obtains parameter estimates from the 1x4 simulation data (in the file named Simulation.doc).

```
model
{
## Main routine
for(i in 1:N){
  for(j in 1:4){
m[i,j] <- beta1*x[1,j] + beta2*x[2,j] + beta3*x[3,j] + beta4*x[4,j] + u[i]
disp[i,j] <- delta1*x[1,j] + delta2*x[2,j] + delta3*x[3,j] + delta4*x[4,j]
E[i,j] <- exp(m[i,j])/(1+exp(m[i,j]))
phi[i,j] <- exp(disp[i,j])
a[i,j] <- E[i,j]*phi[i,j]
b[i,j] <- (1-E[i,j])*phi[i,j]
p[i,j] ~ dbeta(a[i,j], b[i,j])
# Get the log-likelihoods and compute CPO
g[i,j] <-  exp(loggam(phi[i,j]) - loggam(E[i,j]*phi[i,j])-loggam(phi[i,j]-E[i,j]*phi[i,j]) + E[i,j]*phi[i,j]*log(p[i,j]) + (phi[i,j]-
E[i,j]*phi[i,j])*log(1-p[i,j])-log(p[i,j])-log(1-p[i,j]))
h[i,j] <- 1/g[i,j]
  }
u[i] ~ dnorm(0,taub1)
}
## location and dispersion
mu1 <- exp(beta1)/(1+exp(beta1))
mu2 <- exp(beta2)/(1+exp(beta2))
mu3 <- exp(beta3)/(1+exp(beta3))
mu4 <- exp(beta4)/(1+exp(beta4))
phi1 <- exp(delta1)
phi2 <- exp(delta2)
phi3 <- exp(delta3)
phi4 <- exp(delta4)
## Priors
beta4 ~ dnorm(0.0, 1.0E-6)
beta1 ~ dnorm(0.0, 1.0E-6)
beta2 ~ dnorm(0.0, 1.0E-6)
beta3 ~ dnorm(0.0, 1.0E-6)
delta4 ~ dnorm(0.0, 1.0E-6)
delta1 ~ dnorm(0.0, 1.0E-6)
delta2 ~ dnorm(0.0, 1.0E-6)
delta3 ~ dnorm(0.0, 1.0E-6)
sigb1 ~ dunif(0,100)
taub1 <- 1/(sigb1*sigb1)
#taub1 ~ dgamma(0.001,0.001)
#sigb1 <- 1/sqrt(taub1)
}
```

The parameter estimates are as follows:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu1	0.1807	0.02487	0.001036	0.1386	0.1791	0.2335	4001	8000
mu2	0.5136	0.03145	0.001018	0.4518	0.5143	0.5734	4001	8000
mu3	0.5106	0.03164	0.001097	0.4496	0.5105	0.5727	4001	8000
mu4	0.7974	0.02277	9.16E-4	0.7511	0.798	0.8401	4001	8000
phi1	1.438	0.2235	0.00889	1.072	1.419	1.936	4001	8000
phi2	2.019	0.2819	0.01012	1.538	1.989	2.625	4001	8000
phi3	1.77	0.2358	0.00866	1.357	1.754	2.264	4001	8000
phi4	2.787	0.4352	0.01631	1.981	2.764	3.667	4001	8000
sigb1	0.5538	0.08995	0.005033	0.3778	0.5521	0.7385	4001	8000

These are the sample parameter estimates that are input into the excel file (VPC_1x4_1.xls) for the VPC approximations.

Variance Partition Coefficients

The VPC results are shown in the table below, for both the approximation and simulation (a typical run) methods. They are reasonably close, with the simulation method returning a somewhat lower VPC for the 2nd and 3rd categories than the approximation method does.

	$\sigma^2_{\mu(1)}$	$\sigma^2_{\mu(2)}$	$\sigma^2_{\mu(3)}$	$\sigma^2_{\mu(4)}$
	0.0607	0.0827	0.0902	0.0427
	$\sigma^2_{u(1)}$	$\sigma^2_{u(2)}$	$\sigma^2_{u(3)}$	$\sigma^2_{u(4)}$
Approximation	0.0067	0.0191	0.0192	0.0080
Simulation	0.0072	0.0165	0.0165	0.0081
	τ_1	τ_2	τ_3	τ_4
Approximation	0.0997	0.1879	0.1751	0.1580
Simulation	0.1060	0.1666	0.1550	0.1604