Exchangeability: Variance components model

A *variance components* model is a form of 'random effects' model for observable quantities that takes the form

$$Y_i = M + \epsilon_i$$

for $i = 1, \ldots, n$, where

- M is a random variable
- $\epsilon_1, \ldots, \epsilon_n$ are identically distributed random variables with

$$\mathbb{E}[\epsilon_i] = 0$$

with *all variables independent*.

Exchangeability: Variance components model

Suppose

- $M \sim Normal(0, \tau^2)$,
- $\epsilon_1, \ldots, \epsilon_n \sim Normal(0, \sigma^2)$,

so that *conditional* on $M = m$

$$Y_1, \ldots, Y_n | M = m \sim Normal(m, \sigma^2)$$

are independent.

Exchangeability: Variance components model

Then, by a standard marginalization calculation,

$$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = \int_{-\infty}^{\infty} \prod_{i=1}^{n} f_{Y_i|M}(y_i|m) f_M(m)\, dm$$

$$= \int_{-\infty}^{\infty} \left\{ \prod_{i=1}^{n} \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma^2}(y_i - m)^2\right\}\right\} \left(\frac{1}{2\pi\tau^2}\right)^{1/2} \exp\left\{-\frac{1}{2\tau^2} m^2\right\} dm$$

$$= \left(\frac{1}{2\pi}\right)^{(n+1)/2} \left(\frac{1}{\sigma^2}\right)^{n/2} \left(\frac{1}{\tau^2}\right)^{1/2} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} \underbrace{\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - m)^2 + \frac{1}{\tau^2} m^2\right]}\right\} dm.$$

We need to integrate with respect to m.

Exchangeability: Variance components model

By a standard sums-of-squares decomposition

$$\sum_{i=1}^{n}(y_i - m)^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 + n(\bar{y} - m)^2.$$

Also, recall the completing-the-square formula

$$A(t-a)^2 + B(t-b)^2 = (A+B)\left(t - \frac{Aa+Bb}{A+B}\right)^2 + \frac{AB}{A+B}(a-b)^2.$$

Exchangeability: Variance components model

Therefore we may rewrite

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - m)^2 + \frac{1}{\tau^2}m^2$$

as

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{n}{\sigma^2}(m - \bar{y})^2 + \frac{1}{\tau^2}m^2$$

and then to combine the second and third terms, take

$$A = \frac{n}{\sigma^2} \quad a = \bar{y} \qquad B = \frac{1}{\tau^2} \quad b = 0$$

in the above formula.

We have

$$\frac{n}{\sigma^2}(m-\bar{y})^2+\frac{1}{\tau^2}m^2=\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}\right)\left(m-\frac{n\bar{y}/\sigma^2}{(n/\sigma^2)+(1/\tau^2)}\right)^2+\frac{n/(\sigma^2\tau^2)}{(n/\sigma^2)+(1/\tau^2)}\bar{y}^2$$

which we may rewrite as

$$\frac{1}{\lambda^2}(m-\mu)^2+\frac{n}{n\tau^2+\sigma^2}\bar{y}^2$$

where

$$\mu=\frac{n\bar{y}/\sigma^2}{(n/\sigma^2)+(1/\tau^2)}=\frac{n\tau^2\bar{y}}{n\tau^2+\sigma^2}\qquad\lambda^2=\left(\frac{\sigma^2\tau^2}{n\tau^2+\sigma^2}\right)$$

Therefore, for the integral

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - m)^2 + \frac{1}{\tau^2}m^2\right]\right\} dm$$

$$= \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{n}{n\tau^2 + \sigma^2}\bar{y}^2\right]\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\lambda^2}(m - \mu)^2\right\} dm$$

$$= \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{n}{n\tau^2 + \sigma^2}\bar{y}^2\right]\right\} \left(2\pi\lambda^2\right)^{1/2}$$

as the integrand is proportional to a Normal pdf.

Exchangeability: Variance components model

Thus for $(y_1, \ldots, y_n) \in \mathbb{R}^n$,

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)$$

$$= \left(\frac{1}{2\pi}\right)^{n/2} \left(\frac{1}{\sigma^2}\right)^{n/2} \left(\frac{1}{\tau^2}\right)^{1/2} \lambda \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 + \frac{n}{n\tau^2 + \sigma^2}\bar{y}^2\right]\right\}$$

which also relies only upon the summary statistics

$$s_1 = \bar{y} \qquad s_2 = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

and so we may deduce *exchangeability*, as the statistics are invariant to the indexing of the ys.

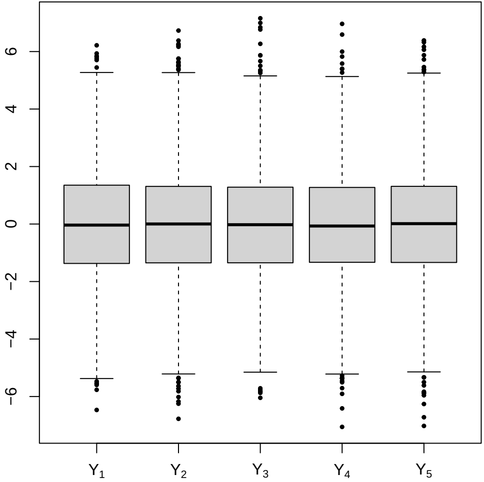
Exchangeability: Variance components model

Simulation: $n = 5$, $\tau^2 = 3$, $\sigma^2 = 1$

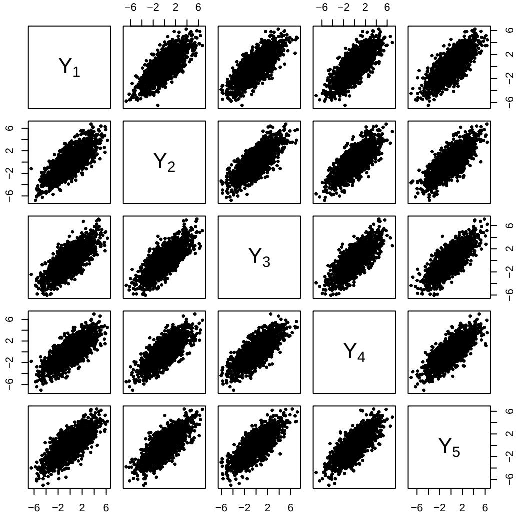
```
set.seed(213)
n<-5
sim.exch01<-function(nv,tauv,sigv){  #Sample the exchangeable variables.
    Mv<-rnorm(1)*tauv
    Yv<-rnorm(nv,Mv,sigv)
}
tau<-sqrt(3)
sig<-sqrt(1)

Ymat<-t(replicate(2000,sim.exch01(n,tau,sig)))  #2000 replicate draws of Y
```

Exchangeability: Variance components model



Exchangeability: Variance components model



Exchangeability: Variance components model

We have that for $i = 1, \ldots, n$,

$$\mathbb{E}_{Y_i}[Y_i] = \mathbb{E}_M[M] + \mathbb{E}_{\epsilon_i}[\epsilon_i] = 0$$

and by independence

$$\mathbb{V}ar_{Y_i}[Y_i] = \mathbb{V}ar_M[M] + \mathbb{V}ar_{\epsilon_i}[\epsilon_i] = \tau^2 + \sigma^2 = 3 + 1 = 4.$$

```
apply(Ymat,2,mean)

+ [1] -0.022506382  0.005840274 -0.025417308  0.001990556 -0.015283486

apply(Ymat,2,var)

+ [1] 4.102284 4.084988 4.076868 4.062566 4.040734
```

Exchangeability: Variance components model

For the covariances, using iterated expectation we have

$$\begin{aligned}
\mathbb{C}ov_{Y_i,Y_j}[Y_i, Y_j] &\equiv \mathbb{E}_{Y_i,Y_j}[Y_i Y_j] \\
&= \mathbb{E}_M \left[\mathbb{E}_{Y_i,Y_j|M}[Y_i Y_j|M]\right] \\
&= \mathbb{E}_M \left[\mathbb{E}_{Y_i|M}[Y_i|M]\mathbb{E}_{Y_j|M}[Y_j|M]\right]
\end{aligned}$$

as Y_i and Y_j have expectation zero, and are conditionally independent given M.

Exchangeability: Variance components model

Thus, as $\mathbb{E}_{Y_i|M}[Y_i|M] = M$ for each i, we have

$$\mathbb{C}ov_{Y_i,Y_j}[Y_i, Y_j] = \mathbb{E}_M[M^2] = \mathbb{V}ar_M[M] + \{\mathbb{E}_M[M]\}^2 = \tau^2$$

and hence

$$\mathbb{C}orr_{Y_i,Y_j}[Y_i, Y_j] = \frac{\mathbb{C}ov_{Y_i,Y_j}[Y_i, Y_j]}{\sqrt{\mathbb{V}ar_{Y_i}[Y_i]\mathbb{V}ar_{Y_j}[Y_j]}} = \frac{\tau^2}{\tau^2 + \sigma^2} = \frac{3}{4}.$$

```
round(cor(Ymat),3)

+       [,1]  [,2]  [,3]  [,4]  [,5]
+ [1,] 1.000 0.754 0.758 0.755 0.756
+ [2,] 0.754 1.000 0.760 0.754 0.763
+ [3,] 0.758 0.760 1.000 0.759 0.768
+ [4,] 0.755 0.754 0.759 1.000 0.768
+ [5,] 0.756 0.763 0.768 0.768 1.000
```

Note: if $\tau \longrightarrow 0$,

$$\mathbb{C}orr_{Y_i,Y_j}[Y_i, Y_j] \longrightarrow 0$$

and the Ys are *uncorrelated*. In fact, in this case the Ys are *independent* as

$$M = 0$$

with probability 1, so $Y_i \equiv \epsilon_i$.

In vector form, we have

$$\mathbf{Y} = M\mathbf{1}_n + \epsilon$$

where

- $\mathbf{1}_n = (1, 1, \ldots, 1)^\top$ is the $n \times 1$ vector of 1s.
- $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)^\top$.

That is

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \ldots \\ Y_n \end{bmatrix} = M \begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \ldots \\ \epsilon_n \end{bmatrix}.$$

Exchangeability: Variance components model

By properties of vector random variables, we have that marginally

$$\mathbf{Y} \sim Normal_n(\mathbf{0}, \Sigma)$$

where

$$\Sigma = \tau^2 \mathbf{1}_n \mathbf{1}_n^\top + \sigma^2 \mathbf{I}_n.$$

with

- $\mathbf{1}_n \mathbf{1}_n^\top$ an $n \times n$ matrix of 1s
- $\mathbf{I}_n$ the $n \times n$ identity matrix