Models In Epidemiology And Biostatistics
Gordon Hilton Fick

A Very Brief Set of Results with Vectors and Matrices

Suppose that $\boldsymbol{y}$ is multivariate normal with mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$.

We will write $\boldsymbol{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with

$\boldsymbol{\mu} = E(\boldsymbol{y})$ and $\boldsymbol{\Sigma} = VAR(\boldsymbol{y}) = E(\boldsymbol{y}-\boldsymbol{\mu})(\boldsymbol{y}-\boldsymbol{\mu})' = E\,\boldsymbol{y}\,\boldsymbol{y}' - \boldsymbol{\mu}\,\boldsymbol{\mu}'$

Now, lets us consider $\boldsymbol{w} = \boldsymbol{a} + \boldsymbol{C}\boldsymbol{y}$

Then, we get these 2 key findings $\boldsymbol{E}\boldsymbol{w} = \boldsymbol{\nu} = \boldsymbol{a} + \boldsymbol{C}\boldsymbol{\mu}$ and $\boldsymbol{w} - \boldsymbol{\nu} = C(\boldsymbol{y} - \boldsymbol{\mu})$

so that $VAR(\boldsymbol{w}) = E(\boldsymbol{w}-\boldsymbol{\nu})(\boldsymbol{w}-\boldsymbol{\nu})' = E(C(\boldsymbol{y}-\boldsymbol{\mu})(\boldsymbol{y}-\boldsymbol{\mu})'C') = C\,VAR(\boldsymbol{y})C'$

Now write $\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix}$

Now consider the conditional distribution for $\boldsymbol{y}_2$ given $\boldsymbol{y}_1$:

We denote the mean and variance-covariance matrix for this distribution as:

$E(\boldsymbol{y}_2 : \boldsymbol{y}_1)$ often called the regression of $\boldsymbol{y}_2$ on $\boldsymbol{y}_1$

and $VAR(\boldsymbol{y}_2 : \boldsymbol{y}_1)$ called the variance matrix about regression on $\boldsymbol{y}_1$

Two more key findings:

$E(\boldsymbol{y}_2) = E(E(\boldsymbol{y}_2 : \boldsymbol{y}_1))$:
The mean of the conditional mean is the marginal mean. This innocuous looking statement turns out be quite crucial and very specific to the conditional mean. There is no comparable result of the log odds, for example.

$VAR(\boldsymbol{y}_2) = E(VAR(\boldsymbol{y}_2 : \boldsymbol{y}_1)) + VAR(E(\boldsymbol{y}_2 : \boldsymbol{y}_1))$:
The marginal variance is the mean variance about regression plus the variance of regression. Again, very specific to conditional means.

For all the details, see Chapter 5 in D.A.S. Fraser 'Probability and Statistics : Theory and Applications' - Duxbury This book is highly recommended. Do seek it out. GHF was a [small] part its development.

The Conditional Approach

$E(\boldsymbol{y} : \boldsymbol{u}) = X\boldsymbol{\beta} + Z\boldsymbol{u}$

and, in particular

$\boldsymbol{y}$ given $\boldsymbol{u} \sim N(X\boldsymbol{\beta}+Z\boldsymbol{u}, \sigma_e^2\boldsymbol{I})$ and

$\boldsymbol{u} \sim N(\boldsymbol{0}, \sigma_u^2\boldsymbol{I})$

Marginal from Conditional

We see that:

$$E(\boldsymbol{y})=E(E(\boldsymbol{y}:\boldsymbol{u}))=E(X\boldsymbol{\beta}+Z\boldsymbol{u})=X\boldsymbol{\beta}+EZ\boldsymbol{u}=X\boldsymbol{\beta}+ZE\boldsymbol{u}=X\boldsymbol{\beta}$$

and

$$VAR\,\boldsymbol{y}=E(VAR(\boldsymbol{y}:\boldsymbol{u}))+VAR(E(\boldsymbol{y}:\boldsymbol{u}))=\sigma_e^2\boldsymbol{I}+VAR(X\boldsymbol{\beta}+Z\boldsymbol{u})=\sigma_e^2\boldsymbol{I}+Z\sigma_u^2\boldsymbol{I}Z'=\sigma_e^2\boldsymbol{I}+\sigma_u^2 ZZ'$$

An Example : The Dental Study

$\boldsymbol{y}$ is a 108 x 1 column vector. There 27 children each with 4 measurements. 108=27 x 4

X is a 108 x 4 matrix. $X=(\boldsymbol{1}:\boldsymbol{A}:\boldsymbol{G}:\boldsymbol{AG})$

$\boldsymbol{u}$ is, for now, a 27 x 1 column vector

Z is a 108 x 27 matrix. $Z=(\boldsymbol{\delta}_1:\boldsymbol{\delta}_2: \ \ldots \ :\boldsymbol{\delta}_{27})$
Where $\boldsymbol{\delta}_i$ is the indicator vector for the ith subject.

Notice that ZZ' is a so-called block diagonal matrix. Each 'block' is a 4 x 4 sub-matrix of 1's.
The upper left 8 x 8 part of ZZ' would look like:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

so the upper left 8 x 8 part of

$$VAR(\boldsymbol{y}) \text{ is } \begin{pmatrix} \sigma_e^2+\sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & 0 & 0 & 0 & 0 \\ \sigma_u^2 & \sigma_e^2+\sigma_u^2 & \sigma_u^2 & \sigma_u^2 & 0 & 0 & 0 & 0 \\ \sigma_u^2 & \sigma_u^2 & \sigma_e^2+\sigma_u^2 & \sigma_u^2 & 0 & 0 & 0 & 0 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_e^2+\sigma_u^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_e^2+\sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 \\ 0 & 0 & 0 & 0 & \sigma_u^2 & \sigma_e^2+\sigma_u^2 & \sigma_u^2 & \sigma_u^2 \\ 0 & 0 & 0 & 0 & \sigma_u^2 & \sigma_u^2 & \sigma_e^2+\sigma_u^2 & \sigma_u^2 \\ 0 & 0 & 0 & 0 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_e^2+\sigma_u^2 \end{pmatrix}$$

Notice that measurements from 2 different children are independent but 2 measurements from the same child have covariance σ_u^2 . [Compound Symmetry]

Notice that the addition of $Z\boldsymbol{u}$ to $X\boldsymbol{\beta}$, here, shifts β_0 to β_0+u_i for the ith subject.

The Marginal Approach

Begin with the marginal mean $E(\boldsymbol{y})=X\boldsymbol{\beta}$ and

directly model the marginal Variance matrix $VAR\,\boldsymbol{y}$.

If this matrix is block diagonal with diagonal elements σ^2 and off diagonal elements $\rho\sigma^2$, then we obtain a correlation structure very close to the structure implied by the conditional approach.