

Supplemental Material I

Corrie Macdonald-Wallis, Debbie A Lawlor, Tom Palmer, Kate Tilling

1 Expressions for regression coefficients adjusting for 2, 3, and 4 covariates

1.1 Two covariates

In a linear model with outcome Y , and covariates X_1, X_2 , in a column matrix $\mathbf{X} = [1, X_1, X_2]$, we estimate parameters $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2]'$. The equation for these parameters is[1]

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y \quad (1)$$

From this it can be shown

$$\beta_1 = \frac{\text{var}(X_2) \text{cov}(X_1, Y) - \text{cov}(X_1, X_2) \text{cov}(X_2, Y)}{\text{var}(X_1) \text{var}(X_2) - (\text{cov}(X_1, X_2))^2}$$
$$\beta_2 = \frac{\text{var}(X_1) \text{cov}(X_2, Y) - \text{cov}(X_1, X_2) \text{cov}(X_1, Y)}{\text{var}(X_1) \text{var}(X_2) - (\text{cov}(X_1, X_2))^2}$$

1.2 Three covariates

With three covariates, X_1, X_2 and X_3 , the vector $\mathbf{X} = [1, X_1, X_2, X_3]$ and we estimate parameters $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2, \beta_3]'$ to solve equation (1). Let us define

$$D_3 = \text{var}(X_1) \text{var}(X_2) \text{var}(X_3) + 2 \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \\ - \text{var}(X_1) (\text{cov}(X_2, X_3))^2 - \text{var}(X_2) (\text{cov}(X_1, X_3))^2 - \text{var}(X_3) (\text{cov}(X_1, X_2))^2$$

Then it can be shown that

$$\beta_1 = \frac{a_{11} \operatorname{cov}(X_1, Y) + a_{12} \operatorname{cov}(X_2, Y) + a_{13} \operatorname{cov}(X_3, Y)}{D_3}$$

$$\beta_2 = \frac{a_{21} \operatorname{cov}(X_1, Y) + a_{22} \operatorname{cov}(X_2, Y) + a_{23} \operatorname{cov}(X_3, Y)}{D_3}$$

$$\beta_3 = \frac{a_{31} \operatorname{cov}(X_1, Y) + a_{32} \operatorname{cov}(X_2, Y) + a_{33} \operatorname{cov}(X_3, Y)}{D_3}$$

where

$$a_{11} = \operatorname{var}(X_2) \operatorname{var}(X_3) - (\operatorname{cov}(X_2, X_3))^2$$

$$a_{12} = \operatorname{cov}(X_1, X_3) \operatorname{cov}(X_2, X_3) - \operatorname{var}(X_3) \operatorname{cov}(X_1, X_2)$$

$$a_{13} = \operatorname{cov}(X_1, X_2) \operatorname{cov}(X_2, X_3) - \operatorname{var}(X_2) \operatorname{cov}(X_1, X_3)$$

$$a_{21} = \operatorname{cov}(X_1, X_3) \operatorname{cov}(X_2, X_3) - \operatorname{var}(X_3) \operatorname{cov}(X_1, X_2)$$

$$a_{22} = \operatorname{var}(X_1) \operatorname{var}(X_3) - (\operatorname{cov}(X_1, X_3))^2$$

$$a_{23} = \operatorname{cov}(X_1, X_2) \operatorname{cov}(X_1, X_3) - \operatorname{var}(X_1) \operatorname{cov}(X_2, X_3)$$

$$a_{31} = \operatorname{cov}(X_1, X_2) \operatorname{cov}(X_2, X_3) - \operatorname{var}(X_2) \operatorname{cov}(X_1, X_3)$$

$$a_{32} = \operatorname{cov}(X_1, X_2) \operatorname{cov}(X_1, X_3) - \operatorname{var}(X_1) \operatorname{cov}(X_2, X_3)$$

$$a_{33} = \operatorname{var}(X_1) \operatorname{var}(X_2) - (\operatorname{cov}(X_1, X_2))^2$$

1.3 Four covariates

With four covariates, X_1 , X_2 , X_3 and X_4 , the vector $\mathbf{X} = [1, X_1, X_2, X_3, X_4]$ and we estimate parameters $\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2, \beta_3, \beta_4]'$ to solve equation (1). Let us define

$$\begin{aligned}
D_4 = & \text{var}(X_1) \text{var}(X_2) \text{var}(X_3) \text{var}(X_4) + 2 \text{var}(X_1) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) \\
& + 2 \text{var}(X_2) \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \text{cov}(X_3, X_4) \\
& + 2 \text{var}(X_3) \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \text{cov}(X_2, X_4) \\
& + 2 \text{var}(X_4) \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \\
& + (\text{cov}(X_1, X_2))^2 (\text{cov}(X_3, X_4))^2 + (\text{cov}(X_1, X_3))^2 (\text{cov}(X_2, X_4))^2 + (\text{cov}(X_1, X_4))^2 (\text{cov}(X_2, X_3))^2 \\
& - \text{var}(X_1) \text{var}(X_2) (\text{cov}(X_3, X_4))^2 - \text{var}(X_1) \text{var}(X_3) (\text{cov}(X_2, X_4))^2 \\
& - \text{var}(X_1) \text{var}(X_4) (\text{cov}(X_2, X_3))^2 - \text{var}(X_2) \text{var}(X_3) (\text{cov}(X_1, X_4))^2 \\
& - \text{var}(X_2) \text{var}(X_4) (\text{cov}(X_1, X_3))^2 - \text{var}(X_3) \text{var}(X_4) (\text{cov}(X_1, X_2))^2 \\
& - 2 \text{cov}(X_1, X_2) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4) \text{cov}(X_1, X_4) \\
& - 2 \text{cov}(X_1, X_2) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) \text{cov}(X_1, X_3) \\
& - 2 \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4) \text{cov}(X_1, X_4)
\end{aligned}$$

Then it can be shown that

$$\beta_1 = \frac{b_{11} \text{cov}(X_1, Y) + b_{12} \text{cov}(X_2, Y) + b_{13} \text{cov}(X_3, Y) + b_{14} \text{cov}(X_4, Y)}{D_4}$$

$$\beta_2 = \frac{b_{21} \text{cov}(X_1, Y) + b_{22} \text{cov}(X_2, Y) + b_{23} \text{cov}(X_3, Y) + b_{24} \text{cov}(X_4, Y)}{D_4}$$

$$\beta_3 = \frac{b_{31} \text{cov}(X_1, Y) + b_{32} \text{cov}(X_2, Y) + b_{33} \text{cov}(X_3, Y) + b_{34} \text{cov}(X_4, Y)}{D_4}$$

$$\beta_4 = \frac{b_{41} \text{cov}(X_1, Y) + b_{42} \text{cov}(X_2, Y) + b_{43} \text{cov}(X_3, Y) + b_{44} \text{cov}(X_4, Y)}{D_4}$$

where

$$b_{11} = \text{var}(X_2) \text{var}(X_3) \text{var}(X_4) + 2 \text{cov}(X_2, X_3) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) \\ - \text{var}(X_2)(\text{cov}(X_3, X_4))^2 - \text{var}(X_3)(\text{cov}(X_2, X_4))^2 - \text{var}(X_4)(\text{cov}(X_2, X_3))^2$$

$$b_{12} = \text{var}(X_3) \text{cov}(X_1, X_4) \text{cov}(X_2, X_4) + \text{var}(X_4) \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \\ + \text{cov}(X_1, X_2)(\text{cov}(X_3, X_4))^2 - \text{var}(X_3) \text{var}(X_4) \text{cov}(X_1, X_2) \\ - \text{cov}(X_1, X_4) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_3) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4)$$

$$b_{13} = \text{var}(X_2) \text{cov}(X_1, X_4) \text{cov}(X_3, X_4) + \text{var}(X_4) \text{cov}(X_1, X_2) \text{cov}(X_2, X_3) \\ + \text{cov}(X_1, X_3)(\text{cov}(X_2, X_4))^2 - \text{var}(X_2) \text{var}(X_4) \text{cov}(X_1, X_3) \\ - \text{cov}(X_1, X_2) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_4) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4)$$

$$b_{14} = \text{var}(X_2) \text{cov}(X_1, X_3) \text{cov}(X_3, X_4) + \text{var}(X_3) \text{cov}(X_1, X_2) \text{cov}(X_2, X_4) \\ + \text{cov}(X_1, X_4)(\text{cov}(X_2, X_3))^2 - \text{var}(X_2) \text{var}(X_3) \text{cov}(X_1, X_4) \\ - \text{cov}(X_1, X_2) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4)$$

$$b_{21} = \text{var}(X_3) \text{cov}(X_1, X_4) \text{cov}(X_2, X_4) + \text{var}(X_4) \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \\ + \text{cov}(X_1, X_2)(\text{cov}(X_3, X_4))^2 - \text{var}(X_3) \text{var}(X_4) \text{cov}(X_1, X_2) \\ - \text{cov}(X_1, X_3) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_4) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4)$$

$$b_{22} = \text{var}(X_1) \text{var}(X_3) \text{var}(X_4) + 2 \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \text{cov}(X_3, X_4) \\ - \text{var}(X_1)(\text{cov}(X_3, X_4))^2 - \text{var}(X_3)(\text{cov}(X_1, X_4))^2 - \text{var}(X_4)(\text{cov}(X_1, X_3))^2$$

$$b_{23} = \text{var}(X_1) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) + \text{var}(X_4) \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \\ + \text{cov}(X_2, X_3)(\text{cov}(X_1, X_4))^2 - \text{var}(X_1) \text{var}(X_4) \text{cov}(X_2, X_3) \\ - \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \text{cov}(X_2, X_4)$$

$$b_{24} = \text{var}(X_1) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4) + \text{var}(X_3) \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \\ + \text{cov}(X_2, X_4)(\text{cov}(X_1, X_3))^2 - \text{var}(X_1) \text{var}(X_3) \text{cov}(X_2, X_4) \\ - \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \text{cov}(X_2, X_3)$$

$$\begin{aligned}
b_{31} = & \text{var}(X_2) \text{cov}(X_1, X_4) \text{cov}(X_3, X_4) + \text{var}(X_4) \text{cov}(X_1, X_2) \text{cov}(X_2, X_3) \\
& + \text{cov}(X_1, X_3) (\text{cov}(X_2, X_4))^2 - \text{var}(X_2) \text{var}(X_4) \text{cov}(X_1, X_3) \\
& - \text{cov}(X_1, X_2) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_4) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4)
\end{aligned}$$

$$\begin{aligned}
b_{32} = & \text{var}(X_1) \text{cov}(X_2, X_4) \text{cov}(X_3, X_4) + \text{var}(X_4) \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \\
& + \text{cov}(X_2, X_3) (\text{cov}(X_1, X_4))^2 - \text{var}(X_1) \text{var}(X_4) \text{cov}(X_2, X_3) \\
& - \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \text{cov}(X_2, X_4) - \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \text{cov}(X_3, X_4)
\end{aligned}$$

$$\begin{aligned}
b_{33} = & \text{var}(X_1) \text{var}(X_2) \text{var}(X_4) + 2 \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \text{cov}(X_2, X_4) \\
& - \text{var}(X_1) (\text{cov}(X_2, X_4))^2 - \text{var}(X_2) (\text{cov}(X_1, X_4))^2 - \text{var}(X_4) (\text{cov}(X_1, X_2))^2
\end{aligned}$$

$$\begin{aligned}
b_{34} = & \text{var}(X_1) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4) + \text{var}(X_2) \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \\
& + \text{cov}(X_3, X_4) (\text{cov}(X_1, X_2))^2 - \text{var}(X_1) \text{var}(X_2) \text{cov}(X_3, X_4) \\
& - \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_2, X_4) - \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \text{cov}(X_2, X_3)
\end{aligned}$$

$$\begin{aligned}
b_{41} = & \text{var}(X_2) \text{cov}(X_1, X_3) \text{cov}(X_3, X_4) + \text{var}(X_3) \text{cov}(X_1, X_2) \text{cov}(X_2, X_4) \\
& + \text{cov}(X_1, X_4) (\text{cov}(X_2, X_3))^2 - \text{var}(X_2) \text{var}(X_3) \text{cov}(X_1, X_4) \\
& - \text{cov}(X_1, X_2) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4) - \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4)
\end{aligned}$$

$$\begin{aligned}
b_{42} = & \text{var}(X_1) \text{cov}(X_2, X_3) \text{cov}(X_3, X_4) + \text{var}(X_3) \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \\
& + \text{cov}(X_2, X_4) (\text{cov}(X_1, X_3))^2 - \text{var}(X_1) \text{var}(X_3) \text{cov}(X_2, X_4) \\
& - \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \text{cov}(X_2, X_3) - \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_3, X_4)
\end{aligned}$$

$$\begin{aligned}
b_{43} = & \text{var}(X_1) \text{cov}(X_2, X_3) \text{cov}(X_2, X_4) + \text{var}(X_2) \text{cov}(X_1, X_3) \text{cov}(X_1, X_4) \\
& + \text{cov}(X_3, X_4) (\text{cov}(X_1, X_2))^2 - \text{var}(X_1) \text{var}(X_2) \text{cov}(X_3, X_4) \\
& - \text{cov}(X_1, X_2) \text{cov}(X_1, X_4) \text{cov}(X_2, X_3) - \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_2, X_4)
\end{aligned}$$

$$\begin{aligned}
b_{44} = & \text{var}(X_1) \text{var}(X_2) \text{var}(X_3) + 2 \text{cov}(X_1, X_2) \text{cov}(X_1, X_3) \text{cov}(X_2, X_3) \\
& - \text{var}(X_1) (\text{cov}(X_2, X_3))^2 - \text{var}(X_2) (\text{cov}(X_1, X_3))^2 - \text{var}(X_3) (\text{cov}(X_1, X_2))^2
\end{aligned}$$

1. Fisher RA. 1925. Chapter 5: Tests of significance of means, differences of means, and regression coefficients, Section 29: Regression with several independent variates. In *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.