

In order to fully specify the likelihood function in equation 3, we need to calculate $p(A_i|Y_{1i}, Y_{2i})$ and $p(Y_{1i}, Y_{2i}|\bar{X}_i, \bar{Z}_i)$. According the null model, the primary and secondary traits in the general population follow a multivariate normal distribution. Therefore, $p(Y_{1i}, Y_{2i}|\bar{X}_i, \bar{Z}_i)$ satisfies

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix} \sim \text{MVN} \left(\begin{pmatrix} \beta_{10} + \sum_s \beta_{1s} X_i^s + \sum_{j_1} \alpha_{1j_1} Z_{ij_1} \\ \beta_{20} + \sum_{j_2} \alpha_{2j_2} Z_{ij_2} \end{pmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$$

When selected samples are used, it is assumed that the sampling proportions from each extreme tail are known. For a selective sampling study that collects N^H individuals from the upper extreme (i.e. with primary trait $Y_{1i} \geq y^H$) and N^L individuals from the lower extreme (i.e. with primary trait $Y_{1i} \leq y^L$), the probability $p(A_i|Y_{1i}, Y_{2i})$ in the likelihood satisfies

$$p(A_i|Y_{1i}, Y_{2i}) \propto \begin{cases} N^H / \Pr(Y_{1i} \geq y^H) & \text{if } Y_{1i} \geq y^H \\ N^L / \Pr(Y_{1i} \leq y^L) & \text{if } Y_{1i} \leq y^L \end{cases}$$

The model is also applicable to random population based studies, where each individual has equal probability of being sampled. In this case, the sampling probability is proportional to a constant, regardless of the primary and secondary trait values, i.e. $p(A_i|Y_{1i}, Y_{2i}) \propto C$.

In order to obtain interchangeable residuals under the null hypothesis, the likelihood model needs to be re-parameterized. Specifically, we use $\tau_1 = \sigma_1, \tau_2 = (1 - \rho^2)\sigma_2, \beta_{2\tau} = \sigma_2\rho/\sigma_1$. The Jacobian for the transformation is given by

$$\partial(\tau_1, \tau_2, \beta_{2\tau_1}) / \partial(\sigma_1, \sigma_2, \rho) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 - \rho^2 & -2\rho\sigma_2 \\ -\rho\sigma_2/\sigma_1^2 & \rho/\sigma_1 & \sigma_2/\sigma_1 \end{pmatrix}. \text{ The determinant of}$$

$\partial(\tau_1, \tau_2, \beta_{2\tau_1}) / \partial(\sigma_1, \sigma_2, \rho)$ is always positive and non-zero, i.e.

$\det(\partial(\tau_1, \tau_2, \beta_{2\tau_1}) / \partial(\sigma_1, \sigma_2, \rho)) = 1 + \rho^2 \sigma_2 / \sigma_1 > 0$ Therefore, the transformation is one-to-one

and non-degenerate. After the re-parameterization, an equivalent mean model can be fitted,

$$\begin{cases} \mu(Y_{1i}) = \beta_{10} + \sum_s \beta_{1s} X_s + \sum_{j_1} \alpha_{1j_1} Z_{j_1} \\ \mu(Y_{2i}) = \beta_{20} + \beta_{2\tau_1} (Y_{1i} - \mu(Y_{1i})) + \sum_{j_2} \alpha_{2j_2} Z_{j_2} \end{cases}.$$

The residuals $e_{1i} = Y_{1i} - \mu(Y_{1i})$ and $e_{2i} = Y_{2i} - \mu(Y_{2i})$ from the model are interchangeable under

the null hypothesis.