1 PRS derivation

Site (S), genetic cluster (A), PRS score (G), and phenotype (Y). For subject $i \in \{1..n\}$. All distributions are conditionally indepednent unless otherwise stated

$$\begin{aligned} p(S_i = 1) &\sim Ber(p) \\ G_i|S_i = s &\sim p_s N(0,1) + (1-p_s)N(\mu, 1) \\ Y_i|G_i &\sim N(G_i\beta, 1) \end{aligned}$$

Then, the covariance

$$Cov \begin{bmatrix} Y_i \\ G_i \\ S_i \end{bmatrix} = \begin{bmatrix} Var(Y_i) & Cov(Y_i, G_i) & Cov(Y_i, S_i) \\ Cov(Y_i, G_i) & Var(G_i) & Cov(G_i, S_i) \\ Cov(Y_i, S_i) & Cov(G_i, S_i) & Var(S_i) \end{bmatrix}$$

Now we solve for each entry saving intermediate results along the way because they are likely to help out later

$$\begin{aligned} & EVar(Y_i|G_i) = E1 = 1 \\ & EVar(G_i|S_i) = E1 = 1 \\ & VarE(G_i|S_i) = Var(1-p)\mu = \mu p(1-p) \\ & VarE(Y_i|G_i) = Var(G_i\beta) = \beta^2 Var(G_i) \\ & Var(G_i) = EVar(G_i|S_i) + Var(E(G_i|S_i)) \\ & = 1 + \mu p(1-p) \\ Var(Y_i) &= EVar(Y_i|G_i) + VarE(Y_i|G_i) \\ &= 1 + \beta^2 Var(G_i) \\ Var(G_i) &= EVar(G_i|S_i) + Var(E(G_i|S_i)) \\ &= 1 + \mu p(1-p) \\ Var(S_i) &= p(1-p) \end{aligned}$$

Now for the off diagonal elements

$$
\begin{aligned}
&ES_i = p \\
&EG_i = EEG_i|S_i = p\mu \\
&EY_i = E(EY_i|G_i) = E(G_i\beta) = \beta p\mu \\
&E(G_i|S_i = 0) = (1 - p_0)\mu \\
&E(G_i|S_i = 1) = (1 - p_1)\mu \\
&E(G_i|S_i) = E(G_i|S_i = 1)p + E(G_i|S_i = 0)(1 - p) \\
&= (1 - p_1)\mu p + (1 - p_0)\mu(1 - p) \\
&EY_iG_i = E(EY_iG_i|G_i) = \beta E(G_i^2) \\
&EG_i^2 = Var(G_i) + (EG_i)^2 = 1 + \mu p(1 - p) + (p\mu)^2 \\
&= 1 + \mu p - \mu p^2) + p^2\mu^2 \\
&EY_iS_i = E(EY_iS_i|S_i) = E(Y_i|S_i = 1)p = p\beta(1 - p_1)\mu
\end{aligned}
$$

$$
\begin{aligned}
EG_iS_i &= E(G_iS_i|S_i = 1)p + E(G_iS_i|S_i = 0)(1 - p) \\
&= (1 - p_1)\mu p + (1 - p_0)\mu(1 - p) \\
&= \mu(p - pp_1 + 1 - p_0 - p + pp_0) \\
&= \mu(1 - pp_1 - p_0 + pp_0) \\
Cov(S_i, G_i) &= ES_iG_i - ES_iEG_i \\
&= \mu(1 - pp_1 - p_0 + pp_0 - p^2) \\
Cov(Y_i, G_i) &= EY_iG_i - EY_iEG_i \\
&= \beta E(G_i^2) - p\mu(\beta p\mu) \\
&= \beta(E(G_i^2) - p^2\mu^2) \\
&= \beta(1 + \mu p - \mu p^2 - 2p^2\mu^2) \\
Cov(Y_i, S_i) &= EY_iS_i - EY_iES_i \\
&= p\beta(1 - p_1)\mu - p^2\beta\mu \\
&= p\beta\mu(1 - p_1 - p)
\end{aligned}
$$

2 Conditioning

Conditioning on S_i we know $\Sigma = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

$$
\Sigma = \begin{bmatrix} Var(Y_i) & Cov(Y_i, G_i) \\ Cov(Y_i, G_i) & Var(G_i) \end{bmatrix} - \begin{bmatrix} Cov(Y_i, S_i) \\ Cov(G_i, S_i) \end{bmatrix} \begin{bmatrix} Var(S_i)^{-1} \end{bmatrix} \begin{bmatrix} Cov(Y_i, S_i) & Cov(G_i, S_i) \end{bmatrix}
$$

Computing the second term

$$
\begin{aligned}
&\begin{bmatrix} Cov(Y_i,S_i) \\ Cov(G_i,S_i) \end{bmatrix} \begin{bmatrix} Var(S_i)^{-1} \end{bmatrix} \begin{bmatrix} Cov(Y_i,S_i) & Cov(G_i,S_i) \end{bmatrix} \\
&= \frac{1}{p(1-p)} \begin{bmatrix} p\beta\mu(1-p_1-p) \\ \mu(1-pp_1-p_0+pp_0-p^2) \end{bmatrix} \begin{bmatrix} p\beta\mu(1-p_1-p) & \mu(1-pp_1-p_0+pp_0-p^2) \end{bmatrix} \\
&= \frac{1}{p(1-p)} \begin{bmatrix} p^2\beta^2\mu^2(1-p_1-p)^2 & p\beta\mu^2(1-p_1-p)(1-pp_1-p_0+pp_0-p^2) \\ & \mu^2(1-pp_1-p_0+pp_0-p^2)^2 \end{bmatrix} \\
&= \frac{\mu^2}{p(1-p)} \begin{bmatrix} p^2\beta^2(1-p_1-p)^2 & p\beta(1-p_1-p)(1-pp_1-p_0+pp_0-p^2) \\ & (1-pp_1-p_0+pp_0-p^2)^2 \end{bmatrix}
\end{aligned}
$$

3 Delta method

$$
\begin{aligned}
h^2 &= \frac{\sigma_g^2}{\sigma_g^2+\sigma_e^2} \\
\hat{V}ar(\hat{h}^2) &= \frac{\delta h^2}{\delta\sigma} Cov(\sigma) \frac{\delta h^2}{\delta\sigma} \\
\frac{\delta h^2}{\delta\sigma} &= (\sigma_g^2+\sigma_e^2)^{-2}(\sigma_e^2,-\sigma_g^2) \\
\hat{V}ar(\hat{h}^2) &= (\sigma_g^2+\sigma_e^2)^{-4}(\sigma_e^2,-\sigma_g^2)Cov(\sigma)(\sigma_e^2,-\sigma_g^2)'
\end{aligned}
$$

The first multiplication

$$
(\sigma_e^2,-\sigma_g^2)Cov(\sigma) = \begin{bmatrix} \sigma_e^2p^2\beta^2(1-p_1-p)^2 - \sigma_g^2p\beta(1-p_1-p)(1-pp_1-p_0+pp_0-p^2) \\ \sigma_e^2p\beta(1-p_1-p)(1-pp_1-p_0+pp_0-p^2) - \sigma_g^2(1-pp_1-p_0+pp_0-p^2)^2 \end{bmatrix}
$$

Second multiplication

$$
= (\sigma_e^2)^2p^2\beta^2(1-p_1-p)^2 - 2\sigma_e^2\sigma_g^2p\beta(1-p_1-p)(1-pp_1-p_0+pp_0-p^2)) + (\sigma_g^2)^2(1-pp_1-p_0+pp_0-p^2)^2
$$

This is a quadratic form. Let $a = \sigma_e^2p\beta(1-p_1-p), b = \sigma_g^2(1-pp_1-p_0+pp_0-p^2)$

$$
\begin{aligned}
&= a^2 - 2ab + b^2 \\
&= (a-b)^2 \\
&= (\sigma_e^2p\beta(1-p_1-p) - \sigma_g^2(1-pp_1-p_0+pp_0-p^2))^2
\end{aligned}
$$

Note that $\sigma_e^2 = 1$ for this consideration and that σ_g^2

$$
\begin{aligned}
\sigma_g^2 &= Cov(Y,G)/Var(Y) \\
&= \frac{\beta(1+\mu p-\mu p^2-2p^2\mu^2)}{1+\beta^2+\beta^2\mu p(1-p))}
\end{aligned}
$$

To get an idea of how this looks, we note $\sigma_e^2 = 1$ and we'll fix $\beta = \mu = 1, p = p_0 = 0.5$ and let $p_1 \in (0, 0.5)$. Note that this makes $\sigma_g^2 = \frac{0.75}{2.25} = 1/3$. Then, plugging into the variance formula

$$Var(h^2) = \frac{1}{9}(p_1 - 1/4)^2$$