

the polyTest model

polyTest is a statistical method for identifying genomic annotations that are enriched with GWAS signals. The **polyTest** model is detailed in the supplement to [this paper](#). It makes the assumption that, at each SNP i , the GWAS effect β_i is a draw from a normal distribution centered at 0 with some variance σ_i^2 :

$$\beta_i \sim N(0, \sigma_i^2)$$

It further assumes that σ_i^2 is determined by the functional annotations at SNP i , so that:

$$\sigma_i^2 = s_i + \exp\left(\sum_j x_{ij}\gamma_j\right)$$

Where s_i is a baseline value depending on properties of the SNP (allele frequency, imputation quality, etc), x_{ij} is 0 or 1¹ for absence or presence of annotation j at SNP i , and γ_j is the model parameter for annotation j .

polyTest determines a maximum-likelihood fit for the parameters $(\gamma_1, \gamma_2, \gamma_3, \dots)$ to the observed GWAS data. If annotation j is uninformative about the GWAS signal, γ_j will be close to 0. However, if annotation j is a good model for the GWAS signal, it will be significantly greater than 0, because SNPs with annotation j will have large effects and require large γ_j (leading to large σ^2) for the maximum-likelihood fit.

¹ x_{ij} could also be a continuous value