

1. Page 2, (see note0101.pdf at <https://github.com/SorensenD/SLGDS>) for more details on the subject)

CHANGE: "This requires equating the expected proportion of genotypes shared identical by descent, given a pedigree, with the observed phenotypic correlation between relatives."

TO: "This requires equating the expected proportion of the genome shared identical by descent, given a pedigree, with the observed phenotypic correlation between relatives."

2. Page 258, eqn. (6.1),

CHANGE second line to (brackets missing):

$$\begin{aligned} E_{y_0x} [\text{MSE}(f(x))] &= E_{y_0x} [(y_0 - f(x))^2] \\ &= E_x [E_{y_0|x} [(y_0 - f(x))^2 | x]] \\ &= E_x [\text{Var}(y_0|x) + (E(y_0|x) - f(x))^2], \end{aligned} \quad (6.1)$$

3. Page 259, below eqn. (6.7),

CHANGE: "that is in the best linear approximation ..."

TO: "that is the best linear approximation ..."

4. Page 270, eqn. (6.31),

CHANGE: the left hand side of the first line:

$$\text{Var}(y_0 - \hat{y}_0 | x_0)$$

TO:

$$\text{Var}[(y_0 - \hat{y}_0) | x_0]$$

(brackets missing)

5. Page 277,

CHANGE:

$$E_{y_v y} (\text{MSE}_v) = \text{Var}(y_{v,i} | x_i) + \text{bias}^2 + \frac{1}{N} \text{Var}(\hat{f}(x_i) | x_i)$$

and

$$E_{y_v y} (\text{MSE}_t) = E_{y_v y} (\text{MSE}_v) - \frac{2}{N} \sum_i \text{Cov}(y_i, \hat{f}(x_i))$$

TO:

$$E_{y_v y} (\text{MSE}_v) = \frac{1}{N} \sum_i \text{Var}(y_{v,i} | x_i) + \frac{1}{N} \sum_i \text{bias}(i)^2 + \frac{1}{N} \sum_i \text{Var}(\hat{f}(x_i) | x_i)$$

and when covariates in training and validating data take the same values,

$$E_{y_v y} (\text{MSE}_t) = E_{y_v y} (\text{MSE}_v) - \frac{2}{N} \sum_i \text{Cov}(y_i, \hat{f}(x_i))$$

6. Page 286, below eqn. (6.66),
CHANGE: "and from (6.65)"
TO: "and from (6.65) when covariates in training and validating data take the same values,"
(see note0603.pdf for elaborations on the subject).

7. Page 325, eqn. (7.41),
CHANGE:
The conditional distribution of the data takes the form

$$y|\mu, \alpha, \delta, \sigma^2 \sim N\left(1\mu + \sum_{i=1}^m X_i(\alpha_i \delta_i), I\sigma^2\right), \quad i = 1, \dots, m \quad (7.41)$$

where X_i is the i th column of X .

TO: The conditional distribution of the data takes the form

$$y|\mu, \alpha, \delta, \sigma^2 \sim N\left(1\mu + \sum_{i=1}^m X_i(\alpha_i \delta_i), I\sigma^2\right), \quad i = 1, \dots, m \quad (7.41)$$

where X_i is the i th column of X .

(remove extra ", " before $i = 1, \dots, m$).

8. Top of page 393, eqn. (9.30),
CHANGE:
An unbiased estimator of incidence (given n , sen , spe and bounded between 0 and 1) is

$$\frac{\Pr(\hat{Y} = 1) - (1 - spe)}{sen + spe - 1} = \frac{T - n(1 - spe)}{n(sen + spe - 1)}. \quad (9.30)$$

TO: An unbiased estimator of prevalence (given n , sen , spe and bounded between 0 and 1) is

$$\frac{\widehat{\Pr}(\hat{Y} = 1) - (1 - spe)}{sen + spe - 1} = \frac{T - n(1 - spe)}{n(sen + spe - 1)}. \quad (9.30)$$

(change "incidence" for "prevalence" and "hat" missing on the biased estimator of prevalence; see note0901.pdf for further details).

9. Page 675,
CHANGE: "Boyle EA, Li YI, Pritchard JK (2017) An expanded view of complex **trits**: from polygenic to omnigenic. Cell 169:1177–1186"
TO: "Boyle EA, Li YI, Pritchard JK (2017) An expanded view of complex traits: from polygenic to omnigenic. Cell 169:1177–1186"