Errata File

Daniel Sorensen

April 17, 2025

This errata file contains a list of errors and modifications to the manuscript.

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 18, CHANGE: "Jeffrey's prior and mathematical form as the likelihood, proportional to $\left((1-\theta)^{27}\right)$ lead to posterior distributions $Be\left(0.5, 27.5\right)$ and $Be\left(1, 28\right)$, respectively;"
   TO: "Jeffrey's prior and the uniform prior, when combined with the likelihood, proportional to $(1-\theta)^{27}$, lead to posterior distributions $Be\left(0.5, 27.5\right)$ and $Be\left(1, 28\right)$, respectively;"

3. Page 32, bottom of paragraph, CHANGE:
   "Spurious results may be obtained by complex interactions between a prediction method and a particular structure in the training data at hand that may not be reproduced when the model is deployed using validating data."
   TO: "Spurious results may be obtained by complex interactions between a prediction method and a particular structure in the data at hand that may not be reproduced when the model is deployed using a new sample of data."

4. Page 39, CHANGE:
   At the population level, using p markers with centred marker genotype codes x_i in vector $x = (x_1, \ldots, x_p)'$, the genomic value (the part of the genetic value explained

by the linear regression on markers) is defined as

$$\begin{aligned} \mathrm{E}\left[G\left(z\right)|x\right] &= \mathrm{E}_{z|x}\left[\mathrm{E}\left(G\left(z\right)|z,x\right)\right] \\ &= \mathrm{E}_{z|x}\left[\mathrm{E}\left(G\left(z\right)|z\right)\right] \\ &= \alpha'\,\mathrm{E}\left(z|x\right) \\ &= \alpha'\Sigma_{zx}\Sigma_x^{-1}x = \alpha'\widehat{z} \\ &= \beta'x. \end{aligned} \tag{1.36}$$

TO:
At the population level, using p markers with centred marker genotype codes x_i in vector $x = (x_1, \ldots, x_p)'$, the genomic value (the part of the additive genetic value explained by the linear regression on markers) is defined as

$$\begin{aligned} \mathrm{E}\left[\alpha'z|x\right] &= \alpha'\,\mathrm{E}\left(z|x\right) \\ &= \alpha'\Sigma_{zx}\Sigma_x^{-1}x = \alpha'\widehat{z} \\ &= \beta'x. \end{aligned} \tag{1.36}$$

5. Page 40, CHANGE:
The genomic variance or variance of genomic values $\beta'x$, is the part of the genetic variance explained by the linear regression on marker genotypes

$$\mathrm{Var}_x\left(\mathrm{E}\left[G\left(z\right)|x\right]\right) = \beta'\Sigma_x\beta = \alpha'\Sigma_{zx}\Sigma_x^{-1}\Sigma_{xz}\alpha. \tag{1.37}$$

TO:
The genomic variance or variance of genomic values $\beta'x$, is the part of the additive genetic variance explained by the linear regression on marker genotypes

$$\mathrm{Var}_x\left(\mathrm{E}\left[\alpha'z|x\right]\right) = \beta'\Sigma_x\beta = \alpha'\Sigma_{zx}\Sigma_x^{-1}\Sigma_{xz}\alpha. \tag{1.37}$$

6. Page 51, CHANGE
"... the subject had been in the air long before Fisher, disguised in the terminology of inverse probability."
TO: "... the subject had been in the air long before Fisher."

7. Page 53, bottom bullet, CHANGE
"When f is a function of several variables $f\left(x\right) = f\left(x_1, x_2, \ldots, x_n\right)$, the *gradient* of f is the column vector of n partial derivatives $\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_{12}}, \ldots, \frac{\partial f}{\partial x_n}\right)'$ and is denoted ∇f."
TO: "When f is a function of several variables $f\left(x\right) = f\left(x_1, x_2, \ldots, x_n\right)$, the *gradient* of f is the column vector of n partial derivatives $\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)'$ and is denoted ∇f."

8. Page 63, CHANGE
"There are m complete bivariate observations $(x_i, y_i), i = 1, 2, \cdots, m$ and $(n - m)$ univariate records ..."
TO: "There are m complete independent bivariate observations $(x_i, y_i), i = 1, 2, \cdots, m$ and $(n - m)$ independent univariate records ..."

9. Page 85, below eqn. (3.18), CHANGE
"For the probit model"
TO:
"For the logistic model"

10. Page 89, CHANGE first line of eqn. (3.31)

$$l'(\beta|y) \sum_{i=1}^{N} \frac{\partial}{\partial \beta} \left\{ (1 - y_i) x_i'\beta - \ln\left[1 + \exp\left(x_i'\beta\right)\right] \right\}$$

TO:

$$l'(\beta|y) \sum_{i=1}^{N} \frac{\partial}{\partial \beta} \left\{ y_i x_i'\beta - \ln\left[1 + \exp\left(x_i'\beta\right)\right] \right\}$$

11. Page 90, below eqn. (3.33) CHANGE:
Now,

$$\begin{aligned} \frac{\partial}{\partial \beta'} \pi\left(x_i'\beta\right) &= \frac{\partial}{\partial \beta'} \left[1 + \exp\left(-x_i'\beta\right)\right]^{-1} \\ &= \left[1 + \exp\left(x_i'\beta\right)\right]^{-2} \exp\left(-x_i'\beta\right) x_i' \\ &= \pi\left(x_i'\beta\right) \left[1 - \pi\left(x_i'\beta\right)\right] x_i'. \end{aligned}$$

TO:
Now,

$$\begin{aligned} \frac{\partial}{\partial \beta'} \pi\left(x_i'\beta\right) &= \left[1 + \exp\left(x_i'\beta\right)\right]^{-2} \exp\left(x_i'\beta\right) x_i' \\ &= \pi\left(x_i'\beta\right) \left[1 - \pi\left(x_i'\beta\right)\right] x_i'. \end{aligned}$$

12. Page 92, top, CHANGE
"Due to the centring, W_{ij} has rank $(n - 1)$, matrix G is singular and $(g|W, \sigma_g^2)$ is singular normally distributed. In practice the expectation and the standard deviation are replaced by their sample estimates."
TO: "In practice, $E(X_{ij})$ and $SD(X_{ij})$ are replaced by the jth column mean and the jth column standard deviation of W. Due to the column centring, when $m > n$, matrix W has rank $(n - 1)$, matrix G is singular and $(g|W, \sigma_g^2)$ is singular normally distributed."

13. Page 121, CHANGE
"... with β ans a viewed as nuisance parameters."
TO:"... with β and a viewed as nuisance parameters."

14. Page 132, CHANGE in (3.135b)
$i = 1, \ldots, n-1,$
TO: $i = 1, \ldots, n,$

15. Page 139,
Replace the two equations and the last line of the chapter by
"

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} y_i &= \hat{\mu}_t, \\ \frac{1}{n}\sum_{i=1}^{n} (y_i - \overline{y})^2 &= \hat{\sigma}_t^2, \end{aligned}$$

where $\mu_t = \mu + i\sigma$ and $\sigma_t^2 = \sigma^2 (1 - i (i - z))$. This yields $\hat{\mu} \approx 9.97$ and $\hat{\sigma}^2 \approx 3.13$."

16. Page 144, below eqn. (4.7)
CHANGE:
"The term S_0^2 can be regarded as a prior variance component, so that $v_0 S_0^2$ is a prior sum of average squared deviations."

TO:
"The term S_0^2 can be regarded as a prior variance component, so that $v_0 S_0^2$ is a prior sum of squares."

17. Page 152, eqn. above (4.32)
CHANGE:

$$p\left(\tilde{y}|\sigma^2, y\right) = \int N\left(\mu, \sigma^2\right) N\left(\mu_1, \sigma_1^2\right) d\mu$$

TO:

$$p\left(\tilde{y}|\sigma^2, y\right) = \int N\left(\tilde{y}|\mu, \sigma^2\right) N\left(\mu|\mu_1, \sigma_1^2\right) d\mu$$

18. Page 172,
CHANGE:
$\log \sigma$
TO:
$\log \sigma^2$

19. Page 174, first paragraph,
CHANGE:
" ..., then σ^2 lognormally distributed".

TO:
" ..., then σ^2 is lognormally distributed".

CHANGE:

$$q\left(\sigma^{2(t)}|Y_{\sigma^2}\right) = (2\pi k_{\sigma^2})^{-\frac{1}{2}} \exp\left[-\frac{\left(\ln\sigma^{2(t)} - \ln Y_{\sigma^2}\right)^2}{2k_{\sigma^2}}\right],$$

TO:

$$q\left(\sigma^{2(t)}|Y_{\sigma^2}\right) = (2\pi k_{\sigma^2})^{-\frac{1}{2}} \exp\left[-\frac{\left(\ln\sigma^{2(t)} - \ln Y_{\sigma^2}\right)^2}{2k_{\sigma^2}}\right] \frac{1}{\sigma^{2(t)}},$$

20. Page 193, bottom line
CHANGE:

$$\begin{aligned}
\widehat{V}_{\texttt{asymp}} &= 3.97 \\
\widehat{\tau} &= 3.03 \\
\widehat{\mathrm{Var}}\left(\hat{\mu}_m\right) &= 0.0004, \\
\widehat{m}_{\texttt{eff}} &= 3294, \\
\widehat{\tau} &= 3.03.
\end{aligned}$$

TO:

$$\begin{aligned}
\widehat{V}_{\texttt{asymp}} &= 3.97, \\
\widehat{\tau} &= 3.03, \\
\widehat{\mathrm{Var}}\left(\hat{\mu}_m\right) &= 0.0004, \\
\widehat{m}_{\texttt{eff}} &= 3294.
\end{aligned}$$

21. Pages 194-206, REPLACE
"4.14 Appendix: A Closer Look at the McMC Machinery"
WITH the content of note0401.pdf

22. Page 213,
CHANGE: "The logistic likelihood shown in (3.30) is "
TO: ""The logistic likelihood is "

23. Page 228, CHANGE
"The last eigenvalue is $\lambda_n = 0$."
TO: "The last eigenvalue is $\lambda_n = 0$ when the rank of W is equal to $(n-1)$."

24. Page 231, CHANGE
"These expressions make it clear that when $\lambda_n = 0$, the mean and variance are zero and the density collapses to a point mass at zero."
TO: "These expressions make it clear that when $\lambda_i = 0$, the mean and variance are zero and the density collapses to a point mass at zero."

25. Page 233, bottom paragraph, CHANGE
" ... where W is the $n \times p$ centred and scaled matrix of marker genotypes, b is the $p \times 1$ vector ..."
TO: " ... where W is the $n \times m$ centred and scaled matrix of marker genotypes, b is the $m \times 1$ vector ..."

26. Page 257
CHANGE: " A distinction is made between the ability to predict (or to fit) those same observations that were used for estimation of parameters, or for prediction of new, yet-to-be observed outcomes."
TO: "A distinction is made between the ability to predict (or to fit) those same observations that were used for estimation of parameters, or to predict new, yet-to-be observed outcomes."

27. Page 258, eqn. (6.1),
CHANGE second line to (brackets missing):

$$\begin{aligned} \mathrm{E}_{y_0 x}\left[\mathrm{MSE}\left(f\left(x\right)\right)\right] &= \mathrm{E}_{y_0 x}\left[\left(y_0 - f\left(x\right)\right)^2\right] \\ &= \mathrm{E}_x\left[\mathrm{E}_{y_0|x}\left[\left(y_0 - f\left(x\right)\right)^2 |x\right]\right] \\ &= \mathrm{E}_x\left[\mathrm{Var}\left(y_0|x\right) + \left(\mathrm{E}\left(y_0|x\right) - f\left(x\right)\right)^2\right], \end{aligned} \tag{6.1}$$

and a little below,
CHANGE: "The error of prediction is"
TO:"Once y_0 is observed, the error of prediction is"

28. Page 259, below eqn. (6.7),
CHANGE: "that is in the best linear approximation ..."
TO: "that is the best linear approximation ..."

29. Page 261, CHANGE from eqn. (6.12) to eqn. (6.15)
TO:

$$y_i = \mu + G\left(z_i\right) + e_i \tag{6.12}$$

where the continuously distributed environmental effect e_i has mean zero and variance σ^2 and marginally, $E(y_i) = \mu$.

The conditional expectation $\mu + G(z_i)$ may not be linear on z_i due to the genetic mechanism operating within and across loci. However, one can always define a linear

relationship of the form

$$m + \alpha' z_i. \tag{6.13}$$

Then m and α are obtained minimising

$$\mathrm{E}\left[\left(y_i - m - \alpha' z_i\right)^2\right]$$

with respect to m and α. The expected squared error is a minimum with

$$\begin{aligned} \alpha &= \left[\mathrm{Var}\left(z_i\right)\right]^{-1} \mathrm{Cov}\left(y_i, z_i\right) \\ &= \left[\mathrm{Var}\left(z_i\right)\right]^{-1} \mathrm{Cov}\left(z_i, G\left(z_i\right)\right) \end{aligned} \tag{6.14}$$

and

$$m = \mathrm{E}\left(y_i\right) - \alpha' \mathrm{E}\left(z_i\right). \tag{6.15}$$

30. Page 262
CHANGE: Substituting (6.14) and (6.15) in (6.13) yields the best linear predictor of the phenotypic value

$$\begin{aligned} \hat{y}_i &= \mathrm{E}\left(y_i\right) + \mathrm{Cov}\left(z_i, G\left(z_i\right)\right)\left[\mathrm{Var}\left(z_i\right)\right]^{-1}\left(z_i - \mathrm{E}\left(z_i\right)\right) \\ &= \mu + \alpha' z_i \end{aligned}$$

TO: Substituting (6.14) and (6.15) in (6.13) yields the best linear predictor of the phenotypic value

$$\begin{aligned} \hat{y}_i &= \mathrm{E}\left(y_i\right) + \mathrm{Cov}\left(z_i', G\left(z_i\right)\right)\left[\mathrm{Var}\left(z_i\right)\right]^{-1}\left(z_i - \mathrm{E}\left(z_i\right)\right) \\ &= \mu + \alpha' z_i \end{aligned}$$

31. Page 262
Delete the bottom four lines of the page:
"Similar algebra yields

$$r^2_{y_0, b\tilde{y}_0} = r^2_{y_0\tilde{y}_0}\left[1 - (1-b)^2\right],$$

indicating that the squared correlation between predictor and predictand is reduced by a factor $\left[1 - (1-b)^2\right]$. The topic is elaborated in de los Campos et al (2013b)."

32. Page 263, first paragraph,
CHANGE: ", ... where $x_i \in \mathbb{R}^p$, and in this section, they are assumed to have mean zero for all i."
TO: ", ... where covariates $x_i \in \mathbb{R}^p$ are centred."

33. Page 264, bottom
CHANGE: "...as the predictor variables have larger sampling variances and are uncorrelated with each other."
TO: "...as the predictor variables have larger sampling variances and are mutually

orthogonal."
CHANGE: "If the covariables are uncorrelated ..."
TO: "If the covariables are orthogonal ..."
CHANGE: "When the covariables are correlated ..."
TO: "When the covariables are not orthogonal ..."

34. Page 265, top
CHANGE: "indicating that when the covariables are correlated ..."
TO: "indicating that when the covariables are not orthogonal ..."

35. Page 270, eqn. (6.31),

CHANGE: the left hand side of the first line:

$$\mathrm{Var}\left(y_0 - \widehat{y}_0 | x_0\right)$$

TO:

$$\mathrm{Var}\left[\left(y_0 - \widehat{y}_0\right) | x_0\right]$$

(brackets missing)

36. Bottom of Page 271, CHANGE:
" On the other hand, if γ in (6.33) is a random variable with zero mean so that $\mathrm{E}(y) = x\beta$ and the fitted model is (6.32) (that ignores the random variable γ), then $\mathrm{E}(\hat{\beta}|x) = (x'x)^{-1}x'\,\mathrm{E}(y) = \beta$, an unbiased estimator."
TO:
"On the other hand, if γ in (6.33) is a random variable with zero mean so that $\mathrm{E}(y|x,z) = x\beta$ and the fitted model is (6.32) (that does not include the random variable γ), then $\mathrm{E}(\hat{\beta}|x,z) = (x'x)^{-1}x'\,\mathrm{E}(y|x,z) = \beta$, an unbiased estimator."

37. Page 273, just before eqn. (6.40), CHANGE:
"As n increases, ..."
TO: "With centred covariates and data, as n increases, ..."

38. Page 277,
CHANGE:

$$\mathrm{E}_{y_v y}\left(\mathrm{MSE}_v\right) = \mathrm{Var}\left(y_{v,i}|x_i\right) + \mathrm{bias}^2 + \frac{1}{N}\mathrm{Var}\left(\hat{f}\left(x_i\right)|x_i\right)$$

and

$$\mathrm{E}_{y_v y}\left(\mathrm{MSE}_t\right) = \mathrm{E}_{y_v y}\left(\mathrm{MSE}_v\right) - \frac{2}{N}\sum_i \mathrm{Cov}\left(y_i, \hat{f}\left(x_i\right)\right)$$

TO:

$$\mathrm{E}_{y_v y}(\mathrm{MSE}_v) = \frac{1}{N}\sum_i \mathrm{Var}\left(y_{v,i}|x_i\right) + \frac{1}{N}\sum_i \mathrm{bias}\left(i\right)^2 + \frac{1}{N}\sum_i \mathrm{Var}\left(\widehat{f}(x_i)|x_i\right)$$

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

$$\mathrm{E}_{y_v y}\left(\mathrm{MSE}_t\right) = \mathrm{E}_{y_v y}\left(\mathrm{MSE}_v\right) - \frac{2}{N}\sum_i \mathrm{Cov}\left(y_i, \hat{f}\left(x_i\right)\right)$$

39. Page 284, first sentence in Subsection 6.6
CHANGE:
" ... leading to $\mathrm{MSE}_t < \mathrm{MSE}_v$."
TO: " ... leading to $\mathrm{E}(\mathrm{MSE}_t) < \mathrm{E}(\mathrm{MSE}_v)$."

40. 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).

41. Page 287, below eqn. (6.69),
CHANGE: "The first term on the right-hand side is a sample estimate of $\mathrm{E}_{yy_v}(\mathrm{MSE}_t)$, and ..."
TO: "The first term on the right-hand side is the sample training mean squared error MSE_t, and ..."

42. Page 288, incorporate the following paragraph at the end of subsection **Correlated data**
"In general, the whole setup of cross-validation builds on the assumption that validating data represent an independent draw from the same distribution as the training data, conditional on the model parameters. When available data with a complicated hierarchical structure (such as a pedigree structure) are divided into training and validating sets, the assumption of conditional independence is often violated and the method is not likely to produce reliable results as a model comparison tool. The problem can be partly mitigated by dividing the data judiciously accounting for the structure in the data and the prediction problem at hand. This may not always be possible and results must therefore be interpreted cautiously, and, as a general rule, supplemented with other approaches of model criticism and comparison."

43. Page 290, left hand side of eqn. (6.74)
CHANGE: $\mathrm{E}_{yy^v}(\mathrm{MSE}_t)$
TO: $\mathrm{E}_{yy_v}(\mathrm{MSE}_t)$
and, left hand side of eqn. (6.75)
CHANGE: $\hat{E}(\mathrm{MSE}_v)$
TO: $\hat{E}_{yy_v}(\mathrm{MSE}_v)$

44. Page 293, first sentence bottom subsection.
CHANGE: " The simulation example compares the expected optimism (6.68) with

its estimate based on (6.75)."
TO: "The simulation example compares the Monte Carlo estimate of the expected optimism (6.68) with its expected value based on the second term of (6.75), for two values of the ratio p/N."

45. Page 316
REPLACE: CODE0701 on page 316
WITH: CODE0701 that can be downloaded from https://github.com/SorensenD/SLGDS in the folder `Codes`.

46. Page 318

CHANGE:

$$p\left(\beta_j\right) = \frac{\tau}{2}\exp\left(-\tau\left|\beta_j\right|\right) \tag{7.26}$$

TO:

$$p\left(\beta_j|\tau\right) = \frac{\tau}{2}\exp\left(-\tau\left|\beta_j\right|\right), \quad \tau > 0, \tag{7.26}$$

47. Page 324, below eqn. (7.37)
CHANGE: " This binary indicator variable with its associated distribution $\Pr\left(\delta_i = 1|\pi\right)$ specifies the a priori probability that a marker effect b_i is non-zero;"
TO: This binary indicator variable specifies the a priori probability that a marker effect b_i is non-zero: $\Pr\left(\delta_i = 1|\pi\right)$."

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

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

where X_i is the ith column of X.
TO: The conditional distribution of the data takes the form

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

where X_i is the ith column of X. (remove extra "," before $i = 1,\ldots,m$).

49. Page 328, bottom line,
CHANGE: "If u is less than or equal to (7.53), set $\delta_{ij}^{[t]} = 1$; otherwise set $\delta_{ij}^{[t]} = 0$."
TO: "If u is less than or equal to (7.53), set $\delta_{ij}^{[t]} = 0$; otherwise set $\delta_{ij}^{[t]} = 1$."

50. Page 332,
CHANGE: "..., while incurring a relatively low proportion of false positives."
TO: "..., while incurring a user-chosen expected proportion of false positives."

51. Page 333,
CHANGE: " If m independent hypotheses H_i are tested, the so-called *family wise error rate* (FWER) is the probability of making one or more Type I errors among the family of hypothesis tests:

$$\begin{aligned}\Pr\left(\text{at least 1 false positive result in } m \text{ tests}|H_1 = 0, \ldots, H_m = 0\right) = \\ 1 - \Pr\left(\text{no false positive results in } m \text{ tests}|H_1 = 0, \ldots, H_m = 0\right) \\ = 1 - (1-\alpha)^m. \end{aligned} \tag{8.1}$$

"

TO: "If m hypotheses H_i are tested, the so-called *family wise error rate* (FWER) is the probability of making one or more Type I errors among the family of hypothesis tests:

$$\begin{aligned}\Pr\left(\text{at least 1 false positive result in } m \text{ tests}|H_1 = 0, \ldots, H_m = 0\right) = \\ 1 - \Pr\left(\text{no false positive results in } m \text{ tests}|H_1 = 0, \ldots, H_m = 0\right) \\ = 1 - (1-\alpha)^m, \end{aligned} \tag{8.1}$$

if the events are independent. "

AND BELOW:
CHANGE: "Then for the m tests, the overall FWER is less than or equal to α. This is a global test that addresses the question: Is there any null hypothesis that is rejected?

The Bonferroni result is readily derived using Boole's inequality and makes no assumptions about the degree of dependence among the tests (see NOTE 1). Therefore, in the example above with $m = 50$ tests, the (global) null hypothesis is rejected if the $p-$value for any particular hypothesis is less than or equal to $(0.05/50) = 0.001$. Controlling FWER is useful ... "

TO:" Then for the whole family of m tests the expected number of false positives is less than or equal to $m(\alpha/m) = \alpha$. This is a global test that addresses the question: is there any null hypothesis that is rejected? In the example above with $m = 50$ tests, each null hypothesis is rejected if its $p-$value is less than or equal to $(0.05/50) = 0.001$.

As shown below (see NOTE 1) the Bonferroni correction controls FWER at level $\leq \alpha$. Controlling FWER is useful ..."

52. Page 335, first equation
CHANGE:

$$\mathrm{E}(V) = \mathrm{E}\left(\sum_{i=1}^{m} V_i\right) = \sum_{i=1}^{m} \mathrm{E}(V_i) \leq m_0 \alpha_B \leq m\alpha_B.$$

TO:

$$\mathrm{E}\left(V\right)=\mathrm{E}\left(\sum_{i=1}^{m}V_i\right)=\sum_{i=1}^{m}\mathrm{E}\left(V_i\right)\leq m\alpha_B.$$

53. Page 339
CHANGE: " Benjamini and Yekutieli (2001) proved that the above theorem holds under certain type of stochastic dependence among the $p-$values, ..."
TO:" Benjamini and Yekutieli (2001) proved that the above theorem holds, with a slight redefinition of the rule (8.14), under certain type of stochastic dependence among the $p-$values, ..."

54. Page 340, second paragraph starting from the bottom
CHANGE: " This rule controls FDR at level $q=0.15$." and
" ... that is expected to include $19\times 0.15\approx 3$ false discoveries."
TO: " This rule controls FDR at level $q\leq 0.15$." and
" ... that is expected to include less than or equal to $19\times 0.15\approx 3$ false discoveries."

55. Page 342,
CHANGE:

$$\left[z_c,\infty\right),\quad\left(-\infty,z_c\right]\cup\left[z_c,\infty\right),\quad\left(-\infty,z_c\right].$$

TO:

$$\left[z_c,\infty\right),\quad\left(-\infty,-z_c\right]\cup\left[z_c,\infty\right),\quad\left(-\infty,-z_c\right].$$

56. Page 345,
CHANGE:
"Then the empirical Byes estimator ..."
TO:
"Then the empirical Bayes estimator ..."

57. Page 346, top
CHANGE: " This may require first mapping the $z-$values to $p-$values using (8.7) with z instead of t. The numerator of (8.27) is

$$\begin{aligned}p_i&=1-F_0(z_i)=\Pr\left(Z\geq z_i|H=0\right)\\&=\int_{z_i}^{\infty}N\left(0,1\right).\end{aligned}$$

Using the ordered $p-$values from smallest to largest, the denominator of (8.27) is"
TO: " The numerator of (8.30) is $\pi_0 p_i$. Using the ordered $p-$values, an estimator of the denominator of (8.27) is"
AND

CHANGE:
" or

$$\widehat{BFDR}\,(p_i) = \widehat{\pi}_0 q \leq q \tag{8.32}$$

as in (8.15). "
TO: " or

$$\widehat{\mathrm{BFDR}}\,(p_i) = \widehat{\pi}_0 \frac{mp_i}{i} \leq \widehat{\pi}_0 q. \tag{8.32}$$

"

58. Page 346, bottom line
REPLACE FROM:
"Under the assumption of a fixed significant threshold p_t ..."
UNTIL THE END OF THE SUBSECTION BELOW eqn. (8.34)
WITH:
Under the same set of assumptions associated with (8.22), there is another way of connecting the positive false discovery rate pBFDR with the Bayesian BFDR and the classical frequentist FDR-BH (Storey, 2003). The expected number of false discoveries V is $E[V(A)] = m\,Pr(Z \in A|H = 0)\,\pi_0$ and the expected number of rejections R is $E[R(A)] = m\,Pr(Z \in A)$. Then

$$\frac{E[V(A)]}{E[R(A)]} = \frac{Pr(Z \in A|H = 0)\,\pi_0}{Pr(Z \in A)} = \mathrm{pBFDR}(A) \tag{8.33}$$

as in (8.19) and (8.23). When m is large, $E[V(A)]/E[R(A)] \approx E\,[V(A)/R\,(A)]$ and the connection with (8.12) is established. Storey(2003) remarks that despite result (8.33), pBFDR captures the joint behaviour of V and R, whereas $E[V(A)]/E[R(A)]$ does not.

59. Page 381, top of the page,
CHANGE:

$$\begin{aligned} \delta_i|\pi \overset{iid}{\sim} Br\,(\pi)\,, \quad ,i &= 1,\ldots,m, \\ \Pr\,(\delta_i = 1|\pi) &= \pi, \quad \Pr\,(\delta_i = 0|\pi) = 1 - \pi. \end{aligned}$$

TO:

$$\begin{aligned} \delta_i|\pi \overset{iid}{\sim} Br\,(\pi)\,, \quad i &= 1,\ldots,m, \\ \Pr\,(\delta_i = 1|\pi) &= \pi, \quad \Pr\,(\delta_i = 0|\pi) = 1 - \pi. \end{aligned}$$

60. Page 392
CHANGE: "Consider a person who feels who has been exposed to ..."
TO: "Consider a person who has been exposed to ..."

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

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

TO: A corrected estimator of prevalence (given n, *sen*, *spe*) is

$$\frac{\widehat{\Pr}\left(\widehat{Y}=1\right)-(1-spe)}{sen+spe-1}, \widehat{\Pr}\left(\widehat{Y}=1\right)>\left(1-spe\right). \tag{9.30}$$

(change "incidence" to "prevalence", "hat" missing on the corrected estimator of prevalence and remove "and bounded between 0 and 1"; see note0901.pdf for further details).

62. Page 413, CHANGE:
The second term in the right hand side is

$$\begin{aligned}\mathrm{E}_{f_i}\left(\mathrm{Var}\left(Y_{ij}|\mu,f_i\right)\right) &\approx E_{f_i}\left[\Phi\left(\mu\right)\left(1-\Phi\left(\mu\right)\right)-\phi\left(\mu\right)f_i-\phi\left(\mu\right)^2f_i^2\right]\\ &= \Phi\left(\mu\right)\left(1-\Phi\left(\mu\right)\right)-\phi\left(\mu\right)^2\sigma_f^2.\end{aligned}$$

TO:

$$\mathrm{E}_{f_i}\left(\mathrm{Var}\left(Y_{ij}|\mu,f_i\right)\right)\approx\Phi\left(\mu\right)\left(1-\Phi\left(\mu\right)\right)-\phi\left(\mu\right)^2\sigma_f^2.$$

63. Page 413, last line
CHANGE:

$$f_i^*|\sigma_f^2\sim N\left(0,\sigma_f^{2*}\right)$$

TO:

$$f_i^*|\sigma_f^{2*}\sim N\left(0,\sigma_f^{2*}\right)$$

64. Page 414,
CHANGE: Assuming improper uniform prior distributions for $\left(\mu,\sigma_f^2\right)$, the approximate posterior density is

$$\begin{aligned}p\left(\mu,\sigma_f^2\right)&\propto p\left(y|\mu,\sigma_f^2\right)p\left(f|\sigma_f^2\right)\\ \propto\left(\sigma_e^{2*}\right)^{-\frac{N}{2}}&\exp\left[-\frac{1}{2\sigma_e^{2*}}\left(y-1\mu^*-Zf^*\right)'\left(y-1\mu^*-Zf^*\right)\right]\\ &\left(\sigma_f^{2*}\right)^{-\frac{n_f}{2}}\exp\left(-\frac{1}{2\sigma_f^{2*}}f^{*'}f^*\right),\quad N=n_fn,\end{aligned} \tag{9.70}$$

TO: Assuming improper uniform prior distributions for $\left(\mu, \sigma_f^2\right)$, the approximate posterior density is

$$
\begin{aligned}
& p\left(\mu, f, \sigma_f^2 | y\right) \propto p\left(y | \mu, f\right) p\left(f | \sigma_f^2\right) \\
& \propto \left(\sigma_e^{2*}\right)^{-\frac{N}{2}} \exp\left[-\frac{1}{2\sigma_e^{2*}}\left(y - 1\mu^* - Zf^*\right)'\left(y - 1\mu^* - Zf^*\right)\right] \\
& \left(\sigma_f^{2*}\right)^{-\frac{n_f}{2}} \exp\left(-\frac{1}{2\sigma_f^{2*}} f^{*'} f^*\right), \quad N = n_f n, \qquad (9.70)
\end{aligned}
$$

65. bottom of Page 414,
CHANGE:" The true model is executed using the `R`-code on page 642."
TO:"The fitted correlated probit model is executed using the `R`-code on page 642; the model is specified on page 220."

66. Page 415
CHANGE:

$$
\begin{aligned}
\widehat{\mathrm{E}}\left(\mu | y\right) &= 0.160, \\
\widehat{\mathrm{E}}\left(\sigma_f^2 | y\right) &= 0.0102, \\
\widehat{\mathrm{E}}\left(\sigma_e^2 | y\right) &= 0.126, \\
\widehat{\mathrm{E}}\left(h^2 | y\right) &= 0.149.
\end{aligned}
$$

TO:

$$
\begin{aligned}
\widehat{\mathrm{E}}\left(\mu^* | y\right) &= 0.160, \\
\widehat{\mathrm{E}}\left(\sigma_f^{2*} | y\right) &= 0.0102, \\
\widehat{\mathrm{E}}\left(\sigma_e^{2*} | y\right) &= 0.126, \\
\widehat{\mathrm{E}}\left(h_o^2 | y\right) &= 0.149.
\end{aligned}
$$

67. Page 415,
CHANGE: "These can be transformed to the estimates on the underlying scale using (9.69):"
TO: "These can be transformed to estimates on the underlying scale using (9.69)"

AND BELOW,

CHANGE:

$$\begin{aligned}\widehat{\mathrm{E}}\left(\mu|y\right) &= -0.995,\\ \widehat{\mathrm{E}}\left(\sigma_f^2|y\right) &= 0.174,\\ \widehat{\mathrm{E}}\left(h^2|y\right) &= 0.291.\end{aligned}$$

TO:

$$\begin{aligned}\widehat{\mathrm{E}}\left(\mu|y\right) &= -0.995,\\ \widehat{\mathrm{E}}\left(\sigma_f^2|y\right) &= 0.174,\\ \widehat{\mathrm{E}}\left(h^2|y\right) &= 0.291,\end{aligned}$$

68. Page 420 and top of Page 421
CHANGE: " ... where it is shown that in large samples,

$$p\left(y_0|y,x_0\right) \approx p\left(y_0|\widehat{\theta},x_0\right)$$

and therefore

$$\begin{aligned}\mathrm{E}\left(y_0|y,x_0\right) &= \int y_0 p\left(y_0|y,x_0\right)\, dy_0\\ &\approx \int y_0 p\left(y_0|\widehat{\theta},x_0\right)\, dy_0\\ &= \mathrm{E}\left[y_0|\widehat{\theta},x_0\right]\\ &= x_0'\widehat{\theta},\end{aligned}$$

where $\widehat{\theta}$ is the maximum likelihood estimator of θ (or the least squares estimator in this linear regression setting)."
TO:
" ... where it is shown that in large samples, for the case of the present normal linear regression,

$$p\left(y_0|y,x_0\right) \approx p\left(y_0|\hat{b},\hat{\sigma}^2,x_0\right)$$

and therefore

$$\begin{aligned}\mathrm{E}\left(y_0|y,x_0\right) &= \int y_0 p\left(y_0|y,x_0\right)\, dy_0\\ &\approx \int y_0 p\left(y_0|\hat{b},\hat{\sigma}^2,x_0\right)\, dy_0\\ &= \mathrm{E}\left[y_0|\hat{b},\hat{\sigma}^2,x_0\right]\\ &= x_0'\hat{b},\end{aligned}$$

where $(\hat{b}, \hat{\sigma}^2)$ is the maximum likelihood estimator of (b, σ^2) ($\hat{b}$ is also the least squares estimator of b in this linear regression setting)."

69. Page 429, in the second lines of the right hand side of equations (10.19) and (10.20),
CHANGE: σ^2
TO: $E_{\sigma^2|y_t}(\sigma^2)$

70. Page 429,
CHANGE: "The step from the first to the second line uses the decomposition of the posterior predictive variance (10.4b)."
TO: "The step from the first to the second line uses the decomposition of the posterior predictive variance (10.4b), here accounting for the unknown variance σ^2."

71. Page 486,
CHANGE: "The three operational models used to analyse the data assumed that the liability can be written in terms of the linear structure:

$$u_i = \mu + x_i'\beta + \varepsilon_i, \quad i = 1, 2, \ldots, 1279,"$$

TO: "The three operational models used to analyse the data assumed that the liability can be written in terms of the linear structure:

$$u_i = \mu + x_i'\beta + \varepsilon_i, \quad i = 1, 2, \ldots, 599,"$$

72. Page 501,
CHANGE in first line of eqn. (11.79)
$a^2 = ...$
TO: $a^{(2)} = ...$

73. Page 525,
CHANGE:
" A matrix of genotypic markers (of order *number of individuals* (sample size)} $\times$ *number of markers* $= 1000$) is generated ..."
TO:"A matrix of genotypic markers (of order *number of individuals* (sample size) $\times$ *number of markers* $= 1000$) is generated ..."

74. Page 661, eqn. (13.82),
CHANGE:

$$\begin{aligned} \mathrm{E}\left(\mathrm{MSE}_v\right) &= \frac{1}{n_v}\left[\left(1\widehat{\mu} - y_v\right)'\left(1\widehat{\mu} - y_v\right) + \sigma^2\left(\mathrm{tr}\left(I_{n_v} + \frac{1}{n_t}\mathrm{tr}\left(11'\right)\right)\right)\right] \\ &= \sigma^2 + \frac{\sigma^2}{n_t} + \frac{1}{n_v}\left(1\widehat{\mu} - y_v\right)'\left(1\widehat{\mu} - y_v\right). \end{aligned} \tag{13.82}$$

TO:

$$\begin{aligned} \mathrm{E}\left(\mathrm{MSE}_v\right) &= \frac{1}{n_v}\left[\left(1\widehat{\mu}-y_v\right)'\left(1\widehat{\mu}-y_v\right)+\sigma^2\left(\mathrm{tr}\left(I_{n_v}+\frac{1}{n_t}\left(11'\right)\right)\right)\right] \\ &= \sigma^2+\frac{\sigma^2}{n_t}+\frac{1}{n_v}\left(1\widehat{\mu}-y_v\right)'\left(1\widehat{\mu}-y_v\right). \end{aligned} \tag{13.82}$$

75. Page 662,
CHANGE:

$$\begin{aligned} z &= \frac{1}{\sigma_e}\left(y_v^*-y_v\right), \\ &\Rightarrow \quad z|\hat{\mu},\hat{b},y_v \sim N\left(\frac{1}{\sigma_e}\left(1_v\widehat{\mu}+W_v\widehat{b}-y_v\right),I\right). \end{aligned}$$

Then

$$z'z \sim \chi^2\left(n_v,\lambda\right),$$

because $A=I\sigma_e^{-2}$, $V=I\sigma_e^2$ and $AV=I$, an idempotent matrix. The non-centrality parameter is

$$\lambda = E\left(z|\hat{\mu},\hat{b},y_v\right)'E\left(z|\hat{\mu},\hat{b},y_v\right)=\frac{1}{\sigma_e^2}\left(1_v\widehat{\mu}+W_v\widehat{b}-y_v\right)'\left(1_v\widehat{\mu}+W_v\widehat{b}-y_v\right).$$

In terms of z, the validating mean squared error is

$$\mathrm{MSE}_v=\frac{\sigma_e^2}{n_v}z'z\sim\frac{\sigma_e^2}{n_v}\chi^2\left(n_v,\lambda\right).$$

TO:

$$\begin{aligned} z &= \left(y_v^*-y_v\right), \\ &\Rightarrow \quad z|\hat{\mu},\hat{b},y_v \sim N\left(1_v\widehat{\mu}+W_v\widehat{b}-y_v,I_{n_v}\sigma_e^2\right). \end{aligned}$$

Let $A=I_{n_v}\sigma_e^{-2}$; since $V=\mathrm{Var}(z)=I_{n_v}\sigma_e^2$, $AV=I_{n_v}$, an idempotent matrix. Then

$$z'Az=\frac{1}{\sigma_e^2}\left(y_v^*-y_v\right)'\left(y_v^*-y_v\right)\sim\chi^2\left(n_v,\lambda\right)$$

where the non-centrality parameter is

$$\lambda=\frac{1}{\sigma_e^2}\left(1_v\widehat{\mu}+W_v\widehat{b}-y_v\right)'\left(1_v\widehat{\mu}+W_v\widehat{b}-y_v\right).$$

In terms of z, the validating mean squared error is

$$\mathrm{MSE}_v=\frac{\sigma_e^2}{n_v}z'Az\sim\frac{\sigma_e^2}{n_v}\chi^2\left(n_v,\lambda\right).$$

76. Page 670,
CHANGE:

$$\mathrm{MSE}_v = \mathrm{MSE}_t + \frac{2}{n} \operatorname{tr}\left(\operatorname{Cov}\left(y, \widehat{y}'\right)\right)$$

TO:

$$\widehat{\mathrm{MSE}}_v = \mathrm{MSE}_t + \frac{2}{n} \operatorname{tr}\left(\operatorname{Cov}\left(y, \widehat{y}'\right)\right)$$

77. 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"

78. Page 680,
CHANGE: "Storey JD, Tibshirani R (2007) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences 100:9440–9445"
TO: "Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences 100:9440–9445"