Errata File

Daniel Sorensen

July 20, 2024

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 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."

5. 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 $f(x) = f(x_1, x_2, \ldots, x_n)$, 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."

6. 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 ..."

7. 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."

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

9. Pages 201-203, REPLACE subsection :
"The acceptance probability for a general Metropolis-Hastings algorithm"
WITH the content of note0401.pdf

10. 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 ..."

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

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

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

13. 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."

14. 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 ..."

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

16. 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)

17. 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,

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

18. 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).

19. 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."

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

21. 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 $i = 1,\ldots,m$).

22. 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, ..."

23. 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."

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

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

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

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

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

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

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

27. 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."

28. 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"