Errata for:

Applied Statistical Inference: Likelihood and Bayes

Leonhard Held and Daniel Sabanés Bové

July 2, 2019

The required changes are in marked in red.

- Section 1.1.4 on page 4 on "Hardy-Weinberg Equilibrium" should start with

  "The *Hardy-Weinberg Equilibrium* (...) plays a central role in population genetics."

- Definition 2.8 on page 41 should read

  "Let $x_{1:n}$ denote the realization of a random sample $X_{1:n}$ from a distribution with probability mass or density function $f(x;\theta)$. Any function $T = h(X_{1:n})$ of $X_{1:n}$ with realization $t = h(x_{1:n})$ is called a *statistic*."

- Example 3.14 on page 67 should read:

  "Besides the mean transformation factor μ (see Example 3.13) we may also be interested in a bootstrap confidence interval for the *coefficient of variation* $\phi = \sigma/\mu$. The corresponding *sample coefficient of variation* $\hat{\phi} = s/\bar{x}$ can be computed for every bootstrap sample, the corresponding histogram is shown in Figure 3.1b). Based on the quantiles of the bootstrap distribution we obtain the 95 % bootstrap confidence interval from 0.086 to 0.108, with point estimate 0.097. In the same manner, a 95 % bootstrap confidence interval can be constructed *e. g.* for the median transformation factor, with the result $[2411, 2479]$."

- Example 3.17 on page 71 on Fisher's exact test: The first paragraph of this example should read:

  "Let θ denote the odds ratio and suppose we want to test the null hypothesis $H_0 : \theta = 1$ against the alternative $H_1 : \theta > 1$. Let $\hat{\theta}_{\text{ML}} = \hat{\theta}_{\text{ML}}(x)$ denote the observed odds ratio, assumed to be larger than 1. The one-sided P-value for the alternative $H_1 : \theta > 1$ is then

  $$P\text{-value} = \mathsf{Pr}\{\hat{\theta}_{\text{ML}}(X) \geq \hat{\theta}_{\text{ML}} \mid H_0 : \theta = 1\},$$

  where $\hat{\theta}_{\text{ML}}(X)$ denotes the MLE of θ viewed as a function of the data X."

- Property 2 on top of page 98 should read:

  "The variance of the ML estimator asymptotically attains the Cramér-Rao lower bound, so the ML estimator is efficient."

- Section 4.4.1 on page 106 on "The Likelihood Ratio Test": The sentence in the fifth line from bottom should read

  "If negative, it is less conservative due to $T_4 \leq T_1$."

- In Definition 5.2 on page 125, "0" should be in bold face:

$$\boldsymbol{S}(\boldsymbol{\theta}) = \mathbf{0}$$

- Exercise 3 and 5 a) in Chapter 5 (page 156/157): "Example 5.2" should be replaced by "Result 5.2".

- The penultimate sentence on page 168 should read:

  "Thus, we can calculate $\mathsf{Pr}(T+)$ if we know the sensitivity $\mathsf{Pr}(T+ \mid D+)$, the prevalence $\mathsf{Pr}(D+)$, and $\mathsf{Pr}(T+ \mid D-) = 1 - \mathsf{Pr}(T- \mid D-)$, *i. e.* 1 minus the specificity."

- Page 172, Definition 6.3: The phrase "credible level" should be "credibility level". The same applies to Definition 6.4 on page 176, the sentence just before equation (6.15) on page 177, the penultimate sentence on page 207 in Section 6.6.2, Example 8.9 on page 264, and the corresponding index entry on page 381.

- Caption of Figure 6.1 on page 173: The last sentence should read

  "The equi-tailed 95% credible interval $[0.492, 0.916]$ is also shown"

- Example 6.4 on page 174-175 should be better presented as follows:

  "We now revisit the diagnostic testing problem discussed in Example 6.1 under the more realistic scenario that the disease prevalence $\pi = \mathsf{Pr}(D+)$ is not known, but only estimated from a *prevalence study*. We will describe how the Bayesian approach can be used to assess the uncertainty of the positive and negative predictive values in this case.

  For example, suppose there was $x = 1$ diseased individual in a prevalence study with $n = 100$ participants. A $\mathrm{Be}(0.5, 5)$ prior distribution for π expresses our initial prior beliefs that the prevalence is below 33% with approximately 95% probability and below 5% with approximately 50% probability. The posterior distribution of π is then $\mathrm{Be}(\tilde{\alpha}, \tilde{\beta})$ with updated parameters $\tilde{\alpha} = 0.5 + x = 1.5$ and $\tilde{\beta} = 5 + n - x = 104$. The posterior mean, median and mode are 0.014, 0.011 and 0.005, respectively. The difference between the point estimates indicates that the posterior distribution is quite skewed. The equi-tailed 95% credible interval is [0.001,0.044].

In the posterior odds (6.4), the prevalence π enters:

$$\omega = \frac{\mathsf{Pr}(D+\,|\,T+)}{\mathsf{Pr}(D-\,|\,T+)} = \frac{\mathsf{Pr}(T+\,|\,D+)}{1-\mathsf{Pr}(T-\,|\,D-)} \cdot \frac{\pi}{1-\pi}.$$

Therefore, the posterior odds ω are easily obtained from π and we can transform ω into the corresponding probability $\mathsf{Pr}(D+\,|\,T+)$ as follows:

$$\theta = \mathsf{Pr}(D+\,|\,T+) = \frac{\omega}{1+\omega} = (1+\omega^{-1})^{-1}. \tag{1}$$

Suppose we want to replace the fixed prevalence $\pi = 1/100$ with the posterior distribution $\pi \sim \mathrm{Be}(\tilde{\alpha}, \tilde{\beta})$ to appreciate the uncertainty involved in the estimation of π. Note that (6.12) is a monotone function of π. Therefore, any quantile of the posterior distribution of π can be transformed to the θ-scale. Based on the corresponding quantiles of π stated above, we easily obtain the posterior median 0.09 and the 95% equi-tailed credible interval [0.01, 0.29] for θ.

It is also possible to analytically compute the implied distribution of $\theta = \mathsf{Pr}(D+\,|\,T+)$. In Exercise 2 it is shown that the density of θ is

$$f(\theta) = c \cdot \theta^{-2} \cdot f_{\mathrm{F}}\left(c(1/\theta - 1)\,;\, 2\tilde{\beta}, 2\tilde{\alpha}\right), \tag{2}$$

where

$$c = \frac{\tilde{\alpha}\,\mathsf{Pr}(T+\,|\,D+)}{\tilde{\beta}\{1-\mathsf{Pr}(T-\,|\,D-)\}}$$

and $f_{\mathrm{F}}(x\,;\,2\tilde{\beta}, 2\tilde{\alpha})$ is the density of the F distribution with parameters $2\tilde{\beta}$ and $2\tilde{\alpha}$, *cf*. Appendix A.5.2. Note that c does not depend on θ. One can analogously proceed for the negative predictive value $\mathsf{Pr}(D-\,|\,T-)$, see Exercise 2. The posterior mean and the posterior mode of θ can now be computed using numerical integration and optimization, respectively. The posterior mean turns out to be 0.109 while the posterior mode is 0.049. The positive predictive value 8.3% obtained in Example 6.1 with fixed prevalence $\pi = 0.01$ lies in between."

- Example 6.8 (Normal model): The first two equations on page 182 should start with "$\mu\,|\,x_{1:n}$", not with "$\mu\,|\,x$".

- Example 6.8 (Normal model): The last sentence should read:

  "Furthermore, we can interpret $n_0 = \delta/\kappa$ as a prior sample size to obtain the relative prior sample size $n_0/(n_0+n)$, cf. Example 6.3."

- Example 6.15 (Inference for a proportion) on page 190: The reference should be to Figure 4.9e, not to Figure 6.9e:

  "They behave similarly to the likelihood confidence interval in Figure 4.9e with slightly lower coverage of the HPD interval as a result of the smaller interval width."

- Example 6.24 (Normal model) on page 200: It should say "for $\boldsymbol{\theta} = (\mu, \sigma^2)^\top$." instead of "for $\boldsymbol{\theta} = (\mu, \kappa)^\top$." at the end of the first sentence.

- Section 6.6.2 (Continuous Asymptotics) on page 206: The first sentence of the second paragraph should be:

  "Let $X_{1:n}$ denote a random sample from a distribution with probability mass or density function $f(x \mid \boldsymbol{\theta})$."

- In Example 6.31 (Comparison of proportions), the penultimate equation on page 212 should read

$$\hat{\psi}_i \mid \psi_i \quad \sim \quad \mathrm{N}(\psi_i, \sigma_i^2).$$

- To make Exercise 6 a) in Chapter 6 consistent with the LDL cholesterol level application given in that exercise, we should have $c > d$ e.g. $c = 3$ and $d = 1$, so that $|a - \theta|$ is penalised more when $a \leq \theta$.

- Example 7.7 (Blood alcohol concentration), page 235: The left-hand side of Formula (7.18) should be changed to $f(x_{1:n} \mid M_1)$, so that Formula (7.18) reads:

$$f(x_{1:n} \mid M_1) = \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\delta}{n\kappa + \delta}\right)^{\frac{1}{2}} \exp\left[-\frac{\kappa}{2}\left\{\sum_{i=1}^n (x_i - \bar{x})^2 + \frac{n\delta}{n\kappa + \delta}(\bar{x} - \nu)^2\right\}\right].$$

- Section 7.2.4 on page 241 on "Model Averaging": In the third paragraph, it should say:

  "Especially in situations where the posterior model probabilities are quite similar in size, so that no clear "winner" model arises, choosing the MAP-model and discarding the other models appears as an inappropriate approach."

- Exercise 2 b) in Chapter 7: The formula should read:

$$C_p = \frac{\mathrm{SS}}{\hat{\sigma}^2_{\mathrm{ML}}} + 2p - n.$$

- Exercise 3 b) in Chapter 7 should say:

  "Next, calculate explicitly the posterior probabilities of the two (*a priori* equally probable) models M_1 and M_2 using a $\mathrm{NG}(2000, 5, 1, 50\,000)$ distribution as prior for κ and μ."

- Exercise 4 b) in Chapter 7 should be changed to:

  "As an example, calculate the Bayes factor for the centred alcohol concentration data using $\mu_0 = 0$, $\nu = 0$ and $\delta = 1/100$."

- Section 8.3.1 on page 264 on "Monte Carlo integration": In the definition of the function `outerdens`, the last comment on page 264 should relate to outside, not inside the interval:

```
## compute the probability outside of the interval (lower, upper):
prob <- pbeta(lower, alpha, beta) +
     pbeta(upper, alpha, beta, lower.tail = FALSE)}
```

- Section 8.3.3, page 268: The equation after (8.17) should read:

$$\mathsf{Pr}\{a \cdot U \cdot f_Z(Z) \leq f_X(Z) \,|\, Z = z\} = \mathsf{Pr}\left\{U \leq \frac{f_X(z)}{a \cdot f_Z(z)}\right\} = \frac{f_X(z)}{a \cdot f_Z(z)}$$

- Exercise 1 d) in Chapter 8: The second part should be changed to:

  "Second, compute the Laplace approximation of the posterior expectation of $\lambda = \exp(\theta)$ and compare again with the exact value which you have obtained by numerical integration using the `R`-function `integrate`."

- Section 9.3.2 on "Computation of the Posterior Predictive Distribution": The first line on page 304 should read:

  "where the last equation follows from conditional independence of X and Y given θ."

- Update for Table A.1 on page 335: The syntax of the functions related to the beta-binomial distribution in the `R`-package `VGAM` has changed as follows:

  $\{\mathtt{r}, \mathtt{d}, \mathtt{p}\}\mathtt{betabinom.ab}(\ldots, \mathtt{size} = n, \mathtt{shape1} = \alpha, \mathtt{shape2} = \beta)$.

- Updates for Table A.2, pages 336-338: The syntax of some `R` functions has changed as follows

  - for the normal distribution

    $\_\mathtt{norm}(\ldots, \mathtt{mean} = \mu, \mathtt{sd} = \sigma)$
  - for the Student t distribution in the package `sn`:

    $\_\mathtt{st}(\ldots, \mathtt{xi} = \mu, \mathtt{omega} = \sigma, \mathtt{nu} = \alpha)$.
  - for the folded normal distribution in the package `VGAM`:

    $\_\mathtt{foldnorm}(\ldots, \mathtt{mean} = \mu, \mathtt{sd} = \sigma)$.

- Addition to Table A.3: For some multivariate distributions, the modes are missing:

  - Dirichlet: $\mathrm{D}_k(\boldsymbol{\alpha})$

$$\mathrm{Mod}(\boldsymbol{X}) = \left(\frac{\alpha_1 - 1}{\sum_{i=1}^k \alpha_i - k}, \ldots, \frac{\alpha_k - 1}{\sum_{i=1}^k \alpha_i - k}\right)^\top$$

    if $\alpha_i > 1$ for all $i \in \{1, \ldots, k\}$

  - Multivariate normal: $\mathrm{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$\mathrm{Mod}(\boldsymbol{X}) = \boldsymbol{\mu}$$

– Wishart: $\mathrm{Wi}_k(\alpha, \boldsymbol{\Sigma})$

$$\mathrm{Mod}(\boldsymbol{X}) = (\alpha - k - 1)\boldsymbol{\Sigma} \quad \text{if } \alpha \geq k+1$$

– Inverse Wishart: $\mathrm{IWi}_k(\alpha, \boldsymbol{\Psi})$

$$\mathrm{Mod}(\boldsymbol{X}) = \frac{\boldsymbol{\Psi}}{\alpha + k + 1}.$$

- Appendix B.2.5, page 349: In the first equation, the function f should be replaced by g, *i. e.* the equation should read

$$\frac{\partial}{\partial \boldsymbol{x}} g(\boldsymbol{x}_0) = \left(\frac{\partial}{\partial x_1} g(\boldsymbol{x}_0), \ldots, \frac{\partial}{\partial x_m} g(\boldsymbol{x}_0) \right)^\top \neq \mathbf{0}.$$