Errata etc. for *Algorithmic Learning in a Random World* (second edition, 2022)

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer

January 28, 2024

1 Errata

- **Page 142, the line following (4.38):** (X, Y) is not used and can be removed.
- **Page 302, 2nd line up from Lemma 9.14:** Remove "and any $\epsilon > 0$". When we say "p-variable" at the end of this sentence, we, of course, mean a p-variable with respect to P.
- **Page 359, 3rd line up from bottom:** "Asarin [8]" should be "Asarin [9]".
- **Page 375, 3rd line from bottom:** "cliques" should be "maximal cliques".
- **Page 449, reference 2:** remove "Johnson, W. E." from the author list and add "W. E. Johnson and" in front of the title. (This was a very bold edit made by the Springer copy editor, which the authors missed.)
- **Page 449, reference 8:** replace "δ-random" by "Δ-random".

Some of these deficiencies (such as those on pp. 142 and 302) might not be errors from the point of view of formal logic, but they are superfluous and confusing.

2 Complements

- **Page 138, first line after the proof of Proposition 4.8:** We refer to "[385, proof of Theorem 1]". That proof is unnecessarily complicated; namely, Lemma 1 in [385] (in the book's bibliography) is easily obtained by the method of coupling [Dubhashi and Panconesi, 2009, Sect. 7.4].
- **Page 358, Sect. 11.6:** Here we should have mentioned that generalized conformal prediction was introduced in Vovk [2003].

- **Page 362, Sect. 11.6.4:** Gibbs and Maxwell worked in the 3D Euclidean space, of course. In Bourbaki's [Bourbaki, 1969, Chap. 24] notation, we consider a gas made up of N molecules of mass m at (absolute) temperature T, and $v_1, \dots, v_N$ are their velocities. The kinetic energy of the system is
$$\frac{m}{2}\left(\|v_1\|^2 + \cdots + \|v_N\|^2\right) = 3NkT, \tag{1}$$
where k is Boltzmann's constant. If T is given (this is our summary of the system), the components of $v_1, \dots, v_N$ lie on the sphere (1). Gibbs's model is that the actual velocities are chosen from the uniform distribution on that sphere. Maxwell's law of velocities is that the velocities have a Gaussian distribution (with given variances, $2kT/m$ for each component, which are independent).

- **Page 405, line 2:** We mention that Fisher regarded his verification protocol as somewhat unnatural but do not mention that the protocol given in [405, Sect. 9 of the arXiv report] is completely natural (since it involves testing predictions based on all the available data).

- **Page 473:** We should have included

  Probability integral transformation, 307, 327

  in the index.

3 Comments

3.1 Lemma 9.6 on p. 269

Lemma 9.6 is a special case of a result on the optimality of the likelihood ratio. Let P and Q be probability measures on the same sample space, playing the role of the null and alternative hypotheses, respectively. The *e-power* of an e-variable E w.r.t. to P is then defined as $\int \log E \, \mathrm{d}Q$.

Lemma A. *For given null and alternative hypotheses P and Q, respectively, such that $Q \ll P$, the largest e-power is attained by the likelihood ratio $\mathrm{d}Q/\mathrm{d}P$: for any e-variable E,*
$$\int \log E \, \mathrm{d}Q \le \int \log \frac{\mathrm{d}Q}{\mathrm{d}P} \, \mathrm{d}Q.$$
And if $Q \ll P$ is violated, the largest e-power is ∞.

The likelihood ratio $\mathrm{d}Q/\mathrm{d}P$ in Lemma A is understood to be the Radon–Nikodym derivative of Q w.r.t. to P. For a simple proof, see, e.g., [Shafer, 2021, Sect. 2.2.1] or Vovk and Wang [2022].

Lemma 9.6 in the book is a special case of Lemma A in which P is the uniform probability measure on $[0, 1]$ and $\rho = \mathrm{d}Q/\mathrm{d}P$.

3.2 Section 9.3 on pp. 294–298

The behaviour of the test martingales and related processes in this section appears anomalous; e.g., the mean in the majority of rows in Table 9.1 lies outside the interquartile range. In particular, the final values of our test martingales have heavy-tailed distributions. The reason is that we consider multiplicative processes, in the terminology of [Nair et al., 2022, Chap. 6].

References

Nicolas Bourbaki. *Éléments de mathématique, Livre VI, Intégration, Chapitre IX, Intégration sur les espaces topologiques séparé, Note historique*, volume 1343 of *Actualités scientifiques et industrielles*, pages 113–123. Hermann, Paris, 1969.

Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms.* Cambridge University Press, Cambridge, 2009.

Jayakrishnan Nair, Adam Wierman, and Bert Zwart. *The Fundamentals of Heavy Tails: Properties, Emergence, and Estimation.* Cambridge, Cambridge University Press, 2022.

Glenn Shafer. The language of betting as a strategy for statistical and scientific communication (with discussion). *Journal of the Royal Statistical Society A*, 184:407–478, 2021.

Vladimir Vovk. Well-calibrated predictions from on-line compression models. In Ricard Gavaldà, Klaus P. Jantke, and Eiji Takimoto, editors, *Proceedings of the Fourteenth International Conference on Algorithmic Learning Theory*, volume 2842 of *Lecture Notes in Artificial Intelligence*, pages 268–282, Berlin, 2003. Springer. Extended version published in *Theoretical Computer Science* (Special Issue devoted to ALT 2003) **364**, 10–26 (2006).

Vladimir Vovk and Ruodu Wang. Efficiency of nonparametric e-tests. Technical Report arXiv:2208.08925 [math.ST], arXiv.org e-Print archive, August 2022. Version 2, to appear.