

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

Vladimir Vovk, Alexander Gammernan, and Glenn Shafer

April 14, 2023

1 Errata

- **Page 302, 2nd line up from Lemma 9.14:** Remove “and any $\epsilon > 0$ ”. (From the point of view of formal logic it might not be an error, but it is superfluous and misleading.) 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 449, reference 8:** replace “ δ -random” by “ Δ -random”.

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 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. to P is then defined as $\int \log E \, dQ$.

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 dQ/dP : for any e -variable E ,*

$$\int \log E \, dQ \leq \int \log \frac{dQ}{dP} \, dQ.$$

And if $Q \ll P$ is violated, the largest e -power is ∞ .

The likelihood ratio dQ/dP in Lemma A is understood to be the Radon–Nikodym derivative of Q w.r. 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 $\rho = dQ/dP$.

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

- 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](https://arxiv.org/abs/2208.08925) [math.ST], [arXiv.org](https://arxiv.org/) e-Print archive, August 2022. Version 2, to appear.