E-values for anytime-valid inference with exponential families

Hao, Y.

Bibliography

[1] Reuben Adams. Safe hypothesis tests for the 2× 2 contingency table. Master's thesis, Delft University of Technology, 2020.

[2] Akshay Agrawal. Lecture notes on Loewner order, 2018.

[3] Theodore W Anderson and Donald A Darling. A test of goodness of fit. *Journal of the American statistical association*, 49(268):765–769, 1954.

[4] Francis J. Anscombe. Fixed-sample size analysis of sequential observations. *Biometrics*, 10(1):89–100, 1954.

[5] Yaser Awad, Shaul K Bar-Lev, and Udi Makov. A new class of counting distributions embedded in the Lee–Carter model for mortality projections: A Bayesian approach. *Risks*, 10(6):111, 2022.

[6] Akshay Balsubramani and Aaditya Ramdas. Sequential nonparametric testing with the law of the iterated logarithm. *Uncertainty in Artificial Intelligence*, 2016.

[7] Shaul K Bar-Lev. Independent, though identical results: the class of Tweedie on power variance functions and the class of Bar-Lev and Enis on reproducible natural exponential families. *International Journal of Statistics and Probability*, 9(1):30–35, 2020.

[8] Shaul K Bar-Lev. The exponential dispersion model generated by the Landau distribution —A comprehensive review and further developments. *Mathematics*, 11(20):4343, 2023.

[9] Shaul K Bar-Lev, Gérard Letac, and Ad Ridder. A delineation of new classes of exponential dispersion models supported on the set of nonnegative integers. *Annals of the Institute of Statistical Mathematics*, 2024.

[10] Shaul K Bar-Lev and Ad Ridder. New exponential dispersion models for count data – the ABM and LM classes. *ESAIM: Probability and Statistics*, 25:31–52, 2021.

[11] Shaul K Bar-Lev and Ad Ridder. Exponential dispersion models for overdispersed zero-inflated count data. *Communications in Statistics-Simulation and Computation*, 52(7):3286–3304, 2023.

[12] George A. Barnard. Review of abraham wald's *Sequential Analysis*. *Journal of the American Statistical Association*, 42(240):658–665, 1947.

[13] Ole E. Barndorff-Nielsen. *Information and exponential families: in statistical theory*. John Wiley & Sons, 2014.

[14] Andrew Barron, Jorma Rissanen, and Bin Yu. The minimum description length principle in coding and modeling. *IEEE transactions on information theory*, 44(6):2743–2760, 1998.

[15] Andrew R Barron and Chyong-Hwa Sheu. Approximation of density functions by sequences of exponential families. *The Annals of Statistics*, 19(3):1347–1369, 1991.

[16] Rabi N. Bhattacharya and R Ranga Rao. *Normal approximation and asymptotic expansions*. SIAM, 2010.

[17] Christopher Bishop. *Pattern recognition and machine learning*. Springer, 2006.

[18] William David Brinda. *Adaptive Estimation with Gaussian Radial Basis Mixtures*. PhD thesis, Yale University, 2018.

[19] Lawrence D. Brown. *Fundamentals of Statistical Exponential Families, with Applications in Statistical Decision Theory*. Institute of Mathematical Statistics, Hayward, CA, 1986.

[20] Lawrence D. Brown, Iain M. Johnstone, and K Brenda MacGibbon. Variation diminishing transformations: a direct approach to total positivity and its statistical applications. *Journal of the American Statistical Association*, 76(376):824–832, 1981.

[21] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[22] Bertrand S Clarke and Andrew R Barron. Information-theoretic asymptotics of bayes methods. *IEEE Transactions on Information Theory*, 36(3):453–471, 1990.

[23] Bertrand S. Clarke and Andrew R. Barron. Jeffreys' prior is asymptotically least favorable under entropy risk. *Journal of Statistical planning and Inference*, 41(1):37–60, 1994.

[24] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, 1991.

[25] Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *Annals of Probability*, 3(1):146–158, 1975.

[26] Imre Csiszár. Sanov property, generalized I-projection and a conditional limit theorem. *Annals of Probability*, 12(3):768–793, 1984.

[27] Imre Csiszár and Frantisek Matus. Information projections revisited. *IEEE Transactions on Information Theory*, 49(6):1474–1490, 2003.

[28] Donald A. Darling and Herbert Robbins. Confidence Sequences for Mean, Variance, and Median. *Proceedings of the National Academy of Sciences*, 58(1):66–68, 1967.

[29] A Philip Dawid. Present position and potential developments: Some personal views statistical theory the prequential approach. *Journal of the Royal Statistical Society: Series A (General)*, 147(2):278–290, 1984.

[30] Martijn de Jong. Tests of significance for linear regression using E-values. Master's thesis, Leiden University, 2021.

[31] Boyan Duan, Aaditya Ramdas, and Larry Wasserman. Interactive rank testing by betting. In *Proceedings of the First Conference on Causal Learning and Reasoning*, volume 177 of *Proceedings of Machine Learning Research*, pages 201–235, 11–13 Apr 2022.

[32] Robin Dunn, Aditya Gangrade, Larry Wasserman, and Aaditya Ramdas. Universal inference meets random projections: a scalable test for log-concavity. *arXiv:2111.09254*, 2021.

[33] Bradley Efron. *Exponential families in theory and practice*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2022.

[34] Willy K Feller. Statistical aspects of ESP. *The Journal of Parapsychology*, 4(2):271, 1940.

[35] Peter Grünwald. *The minimum description length principle*. MIT press, 2007.

[36] Peter Grünwald. The E-posterior. *Philosophical Transactions of the Royal Society, Series A*, 2023.

[37] Peter Grünwald and A Philip Dawid. Game theory, maximum entropy, minimum discrepancy, and robust Bayesian decision theory. *Annals of Statistics*, 32(4):1367–1433, 2004.

[38] Peter Grünwald and Steven De Rooij. Asymptotic log-loss of prequential maximum likelihood codes. In *International Conference on Computational Learning Theory*, pages 652–667. Springer, 2005.

[39] Peter Grünwald, Alexander Henzi, and Tyron Lardy. Anytime-valid tests of conditional independence under model-x. *Journal of the American Statistical Association*, 119(546):1554–1565, 2024.

[40] Peter Grünwald, Tyron Lardy, Yunda Hao, Shaul K. Bar Lev, and Martijn de Jong. Optimal e-values for exponential families: the simple case. *arXiv preprint arXiv:2404.19465*, 2024.

[41] Peter Grünwald and Nishant Mehta. A tight excess risk bound via a unified PAC-Bayesian-Rademacher-Shtarkov-MDL complexity. In *Proceedings of the Thirtieth Conference on Algorithmic Learning Theory (ALT) 2019*, 2019.

[42] Peter Grünwald, Rianne De Heide, and Wouter Koolen. Safe testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, page qkae011, 2024.

[43] Yunda Hao and Peter Grünwald. E-values for exponential families: the general case. *arXiv preprint arXiv:2409.11134*, 2024.

[44] Yunda Hao, Peter Grünwald, Tyron Lardy, Long Long, and Reuben Adams. E-values for k-sample tests with exponential families. *Sankhya A*, 86(1):596–636, 2024.

[45] Allard Hendriksen, Rianne de Heide, and Peter Grünwald. Optional Stopping with Bayes Factors: A Categorization and Extension of Folklore Results, with an Application to Invariant Situations. *Bayesian Analysis*, 16(3):961 – 989, 2021.

[46] Alexander Henzi and Johanna F. Ziegel. Valid sequential inference on probability forecast performance. *Biometrika*, 2022.

[47] Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021.

[48] Jun ichi Takeuchi and Andrew R. Barron. Asymptotic minimax regret for Bayes mixtures. *arXiv preprint arxiv:2406.17929*, 2024.

[49] Kyoungseok Jang, Kwang-Sung Jun, Ilja Kuzborskij, and Francesco Orabona. Tighter PAC-Bayes bounds through coin-betting. In *Proceedings COLT 2023*, 2023.

[50] Bent Jørgensen. *The Theory of Exponential Dispersion Models*, volume 76 of *Monographs on Statistics and Probability*. Chapman and Hall, London, 1997.

[51] Emilie Kaufmann and Wouter M Koolen. Mixture martingales revisited with applications to sequential tests and confidence intervals. *Journal of Machine Learning Research*, 22(246):1–44, 2021.

[52] John L. Kelly. A new interpretation of information rate. *Bell System Technical Journal*, 35:pp. 917–26, 1956.

[53] Wouter Koolen and Peter Grünwald. Log-optimal anytime-valid e-values. *International Journal of Approximate Reasoning*, 2021. Festschrift for G. Shafer's 75th Birthday.

[54] Wojciech Kotłowski and Peter Grünwald. Maximum likelihood vs. sequential normalized maximum likelihood in on-line density estimation. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 457–476, 2011.

[55] Solomon Kullback and Richard A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[56] Tyron Lardy. E-values for hypothesis testing with covariates. Master's thesis, Leiden University, 2021.

[57] Tyron Lardy, Peter Grünwald, and Peter Harremoës. Reverse information projections and optimal e-statistics. *IEEE Transactions on Information Theory*, 2024. To Appear.

[58] Martin Larsson, Aaditya Ramdas, and Johannes Ruf. The numeraire e-variable and reverse information projection. *arXiv preprint arXiv:2402.18810*, 2024.

[59] Alix Lhéritier and Frédéric Cazals. A sequential non-parametric multivariate two-sample test. *IEEE Transactions on Information Theory*, 64(5):3361–3370, 2018.

[60] Qiang Jonathan Li. *Estimation of mixture models*. Yale University, 1999.

[61] Qiang Jonathan Li and Andrew R. Barron. Mixture density estimation. In *Advances in Neural Information Processing Systems*, volume 12, pages 279–285, 2000.

[62] Haoyi Liang. Stratified safe sequential testing for mean effect. Master's thesis, University of Amsterdam, 2023.

[63] Hubert W Lilliefors. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American statistical Association*, 62(318):399–402, 1967.

[64] Michael Lindon, Dae Woong Ham, Martin Tingley, and Iavor Bojinov. *Anytime-valid Inference in Linear Models and Regression-adjusted Inference*. Harvard Business School, 2024.

[65] Alexander Ly, Udo Boehm, Peter Grünwald, Aaditya Ramdas, and Don van Ravenzwaaij. Safe anytime-valid inference: Practical maximally flexible sampling designs for experiments based on e-values, 2024.

[66] Carl N Morris. Natural exponential families with quadratic variance functions. *The Annals of Statistics*, pages 65–80, 1982.

[67] Francesco Orabona and Kwang-Sung Jun. Tight concentrations and confidence sequences from the regret of universal portfolio. *arXiv:2110.14099*, 2021.

[68] Teodora Pandeva, Tim Bakker, Christian A. Naesseth, and Patrick Forré. E-valuating classifier two-sample tests. *Transactions on Machine Learning Research*, 2024.

[69] Muriel Felipe Pérez-Ortiz. *Are we there yet?: Advances in anytime-valid methods for hypothesis testing and prediction.* PhD thesis, Leiden University, 2023.

[70] Muriel Felipe Pérez-Ortiz, Tyron Lardy, Rianne de Heide, and Peter D. Grünwald. E-statistics, group invariance and anytime-valid testing. *The Annals of Statistics*, 52(4):1410 – 1432, 2024.

[71] Aaditya Ramdas, Peter Grünwald, Vladimir Vovk, and Glenn Shafer. Game-theoretic statistics and safe anytime-valid inference. *Statistical Science*, 38(4):576–601, 2023.

[72] Aaditya Ramdas, Johannes Ruf, Martin Larsson, and Wouter Koolen. Admissible anytime-valid sequential inference must rely on nonnegative martingales. *arXiv:2009.03167*, 2020.

[73] Herbert Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.

[74] Richard Royall. *Statistical evidence: a likelihood paradigm.* Chapman and Hall, 1997.

[75] Shalev Shaer, Gal Maman, and Yaniv Romano. Model-Free Sequential Testing for Conditional Independence via Testing by Betting. In *International Conference on Artificial Intelligence and Statistics*, 2023.

[76] Glenn Shafer. Testing by betting: a strategy for statistical and scientific communication (with discussion and response). *Journal of the Royal Statistical Society A*, 184(2):407–478, 2021.

[77] Glenn Shafer, Alexander Shen, Nikolai Vereshchagin, and Vladimir Vovk. Test martingales, Bayes factors and p-values. *Statistical Science*, 26(1):84–101, 2011.

[78] Yu. M. Shtarkov. Universal sequential coding of single messages. *Problems of Information Transmission*, 23(3):3–17, 1987.

[79] Michael A Stephens. EDF statistics for goodness of fit and some comparisons. *Journal of the American statistical Association*, 69(347):730–737, 1974.

[80] Jun-ichi Takeuchi and Andrew R. Barron. Asymptotically minimax regret for exponential and curved exponential families, 1997. Manuscript underlying abstract at the 1998 International Symposium on Information Theory (14 pages). Available at `http://www.stat.yale.edu/ arb4/publications.html`.

[81] Jun-ichi Takeuchi and Andrew R. Barron. Asymptotically minimax regret for exponential families. In *Proceedings SITA '97*, pages 665–668, 1997.

[82] Jun-ichi Takeuchi and Andrew R. Barron. Asymptotically minimax regret by Bayes mixtures for non-exponential families. In *2013 IEEE Information Theory Workshop (ITW)*, pages 1–5. IEEE, 2013.

[83] Judith Ter Schure, Muriel Felipe Pérez-Ortiz, Alexander Ly, and Peter Grünwald. The anytime-valid logrank test: Error control under continuous monitoring with unlimited horizon. *The New England Journal of Statistics in Data Science*, 2(2):190–214, 2024.

[84] Flemming Topsøe. Information-theoretical optimization techniques. *Kybernetika*, 15(1):8–27, 1979.

[85] Timmy Tse and Anthony C Davison. A note on universal inference. *Stat*, 11(1):e501, 2022.

[86] Rosanne Turner and Peter Grünwald. Anytime-valid confidence intervals for contingency tables and beyond. *Statistics and Probability Letters*, 2023.

[87] Rosanne Turner and Peter Grünwald. Safe sequential testing and effect estimation in stratified count data. In *Annual AI and Statistics Conference*, 2023.

[88] Rosanne Turner, Alexander Ly, and Peter Grünwald. Generic e-variables for exact sequential k-sample tests that allow for optional stopping. *Journal of Statistical Planning and Inference*, 230:106116, 2024.

[89] Tim van Erven. Blog post on large deviations: Cramér vs. Sanov, 2012.

[90] Jean Ville. *Etude critique de la notion de collectif*. Gauthier-Villars Paris, 1939.

[91] Vladimir Vovk and Ruodu Wang. E-values: Calibration, combination and applications. *The Annals of Statistics*, 49(3):1736–1754, 2021.

[92] Vladimir Vovk and Ruodu Wang. Nonparametric e-tests of symmetry. *The New England Journal of Statistics in Data Science*, 2(2):261–270, 2024.

[93] Abraham Wald. *Sequential Analysis*. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London, 1947.

[94] Qiuqi Wang, Ruodu Wang, and Johanna Ziegel. E-backtesting. *arXiv preprint arXiv:2209.00991*, 2022.

[95] Larry Wasserman, Aaditya Ramdas, and Sivaraman Balakrishnan. Universal inference. *Proceedings of the National Academy of Sciences*, 117(29):16880–16890, 2020.

[96] Ian Waudby-Smith, David Arbour, Ritwik Sinha, Edward H Kennedy, and Aaditya Ramdas. Time-uniform central limit theory and asymptotic confidence sequences. *arXiv:2103.06476*, 2021.

[97] Ulla-Britt Wennerholm, Sissel Saltvedt, Anna Wessberg, Mårten Alkmark, Christina Bergh, Sophia Brismar Wendel, Helena Fadl, Maria Jonsson, Lars Ladfors, Verena Sengpiel, et al. Induction of labour at 41 weeks versus expectant management and induction of labour at 42 weeks (Swedish post-term induction study, swepis): multicentre, open label, randomised, superiority trial. *British Medical Journal*, 367, 2019.

[98] David Williams. *Probability with martingales*. Cambridge university press, 1991.

[99] William Henry Young. On classes of summable functions and their fourier series. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 87(594):225–229, 1912.

[100] Zhenyuan Zhang, Aaditya Ramdas, and Ruodu Wang. On the existence of powerful p-values and e-values for composite hypotheses. *arXiv preprint arXiv:2304.16539*, 2023.

List of publications

This dissertation is based on the following papers.

- Chapter 2 is based on
  Peter Grünwald, Tyron Lardy, Yunda Hao, Shaul K Bar-Lev, and Martijn Jong. Optimal E-Values for Exponential Families: the Simple Case. arXiv preprint arXiv:2404.19465, 2024. It has been submitted as a contribution to the *Festschrift for Andrew Barron on the Occasion of his 65th Birthday*.

- Chapter 3 is based on
  Yunda Hao, and Peter Grünwald. E-Values for Exponential Families: the General Case. arXiv preprint arXiv:2409.11134, 2024, under submission.

- Chapter 4 is based on
  Yunda Hao, Peter Grünwald, Tyron Lardy, Long Long, and Reuben Adams. E-values for k-Sample Tests with Exponential Families. Sankhya A, 86(1):596–636, 2024.

- Chapter 5 is based on
  Peter Grünwald, Yunda Hao, and Akshay Balsubramani. Growth-Optimal E-Variables and an extension to the multivariate Csiszár-Sanov-Chernoff Theorem. arXiv preprint arXiv:2412.17554, 2024.

The following paper was completed during the PhD period but is not included in this dissertation.

- Chengli Tan, Jiangshe Zhang, Junmin Liu, Yicheng Wang, and Yunda Hao. Stabilizing Sharpness-aware Minimization Through A Simple Renormalization Strategy. arXiv preprint arXiv:2401.07250, 2024. Journal of Machine Learning Research, accept after minor revision.