Proposer of the vote of thanks to Waudy-Smith and Ramdas and contribution to the Discussion of 'Estimating means of bounded random variables by betting'

Peter Grünwald[1,2]

[1]Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[2]Leiden University, Leiden, The Netherlands

October 13, 2023

The authors derive nonasymptotic anytime-valid confidence sequences for the mean of a sequence $X_1, X_2, \ldots$ of bounded random variables. When put to practice, their new methods beat the best known bounds, sometimes by vast margin — *even for the fixed-sample size, not anytime-valid setting*. It is rare in statistics that one can get such substantial improvements on a decades-old problem, and I congratulate Waudby-Smith and Ramdas on this remarkable achievement. It illustrates once more the relevance of *e-process-based anytime-valid methods* [Grünwald et al., 2024] even when anytime-validity is not required. In particular their results have repercussions for PAC-Bayesian machine learning theory, which relies on concentration bounds for bounded i.i.d. X_t — the main (but not only) setting the authors (WSR from now on) consider and on which I will also focus. So, let $X_1, X_2, \ldots$ be i.i.d. $\sim P$ with P an arbitrary distribution on $[0, 1]$. The null, denoted $\mathcal{P}^{\mu}$, consists of *all* distributions on $[0, 1]$ with some fixed mean μ. We want to test whether the mean is μ, against alternative $\bigcup_{\mu' \neq \mu} \mathcal{P}^{\mu'}$.

1 An Embarrassment of Neyman-Pearson Theory

Assume the X_t arrive sequentially. A company's data science team is instructed to find out whether it can rule out, with high certainty, that $\mu < \mu_0$ for some fixed given value μ_0. They plan to await 5000 outcomes and then check if the lower end of the $1 - \alpha$ confidence interval is above μ_0, for $\alpha = 0.001$.

But now suppose their boss is impatient and, at $t = 1000$, wants to know if there is already sufficiently strong evidence to rule out $\mu < \mu_0$. He thus asks the data science team to peek at the data. They find they already have a significant result, so they stop sampling. As is well-known, this invalidates confidence intervals, and may be viewed as *p-hacking*. What is less known though, is that *even if they had not found a significant result at* $t = 1000$ *and therefore had decided to keep sampling until* $t = 5000$ *after all*, they would already have invalidated the $(1 - \alpha)$-coverage — by the mere act of just checking, even if based on the particular data they saw they did not change course after the check. In this sense, classical methods seem almost like quantum mechanics: you may already destroy the validity of your conclusions merely by looking at the data! Anytime-valid methods like WSR's avoid this issue altogether.

2 An Embarrassment of Bayes Theory

Doesn't Bayesian statistics fare better on this problem? It has often been claimed that 'optional stopping is no problem for Bayesians'. While such claims are problematic anyway [Hendriksen et al., 2021], here I focus on a different issue: the simple problem addressed by WSR is incredibly difficult to solve via a full Bayesian analysis, which requires specifying a prior distribution on some set $\mathcal{P}$ containing P. How to choose $\mathcal{P}$ if, like WSR, we want to make no assumptions at all on P? Even if one adapts a standard nonparametric $\mathcal{P}$ and corresponding prior, one still rules out many possible and reasonable P... While such points have been made since the 1950s, the issue is brought to light particularly clearly in WSR's bounded support setting, since they really need to assume nothing further about P at all and require only *two* parameters to get their results.

Still, their approach does have a *pseudo-Bayesian* flavour. They employ *capital processes* $(\mathcal{K}_t(\mu))_{t=0}^{\infty}$ which are really test martingales relative to the null $\mathcal{P}^{\mu}$, of the form

$$\mathcal{K}_t(\mu) = \prod_{i=1}^{t}(1 + \lambda_i(\mu) \cdot (X_i - \mu)), \tag{1}$$

with $\mathcal{K}_0(\mu) := 1$ and $(\lambda_t(\mu))_{t=1}^{\infty}$ any $\Lambda(\mu)$-valued predictable sequence, with $\Lambda(\mu) = (-1/(1-\mu), 1/\mu)$. Thus, one can let $\lambda_t(\mu)$ depend on X_1^{t-1} and in this way one can learn 'good' values of λ based on past data. In their arguably most sophisticated approach for determining the λ_t's, *GRAPA*, WSR determine a $\hat{\lambda}_t(\mu)$ directly, via a plug-in approach, but, they point out, it can also be done via the *method of mixtures*, by putting a prior density w_μ on $\Lambda(\mu)$ and using in (1) the 'posterior-mean' $\tilde{\lambda}_t(\mu) := \int w_\mu(\lambda \mid X_1^{t-1}) d\lambda$, based on 'pseudo-posterior mixture'

$$w_\mu(\lambda \mid X_1^{t-1}) \propto w_\mu(\lambda) \cdot \prod_{i=1}^{t-1}(1 + \lambda(X_i - \mu)). \tag{2}$$

Orabona and Jun [2021] (OJ) take this approach, with w_μ generalizing Jeffreys' prior for the Bernoulli model.

3 GRAPA vs. REGROW vs. KLinf

What is a good martingale to use in the first place? Grünwald et al. [2024] strongly argued that, if a simple alternative P is given, then the *Kelly criterion* (which they called P-GRO, standing for *growth-rate optimal* relative to P) is the natural anytime-valid replacement for the traditional goal of optimizing power. The P-GRO martingale $(M_t)_{t=0}^{\infty}$ (if it exists) maximizes

$$\mathbf{E}_P[\log M_t] \tag{3}$$

for all t. The natural extension of (3) in case of a large (rather than simple) alternative hypothesis is called REGROW (for *relative growth-optimality in worst-case*) by Grünwald et al. [2024]. WSR's GRAPA can be viewed as approximating the REGROW martingale. This follows from WSR's Proposition 2, Part (d), which shows that *any* test martingale for testing μ must be of the form (1) for some predictable λ_t. Thus, for any P in the alternative $\bigcup_{\mu' \neq \mu} \mathcal{P}^{\mu'}$, there must be some sequence $\{\lambda_t^{(P)}\}_t$ for which the corresponding $\mathcal{K}_t(\mu)$ is GRO. The arguments of Koolen and Grünwald [2021] imply that $\lambda_t^{(P)}$ must be the same for all t; let us denote it as λ_P^*. REGROW then amounts to finding a test martingale $(M_t)_{t=0}^{\infty}$ for testing $\mathcal{H}_0$ for which

$$\max_{P \in \mathcal{H}_1} \mathbf{E}_P \left[\log \mathcal{K}_t^{(\lambda^*_P)}(\mu) - \log M_t\right] \tag{4}$$

is small for each t, where we use $\mathcal{K}_t^{(\lambda)}(\mu)$ to denote the fixed-λ-capital process with $\lambda_t = \lambda$ for all t. Alternatively, one may consider the *expected regret*, given by replacing λ_P^* in (4) by $\lambda^{\text{HS}}(X^t)$ ('optimal fixed λ with hindsight'), the λ for which $\mathcal{K}_t^{(\lambda)}(\mu)$ is maximized at time t.

GRAPA can be thought of as finding an (almost-) REGROW $\mathcal{K}_t(\mu)$ by setting each λ_t to the $\lambda^{\text{HS}}(X^{t-1})$ that would have maximized the empirical counterpart based on the data seen in the past. OJ, in contrast, show that for their prior the regret ((4) with λ_P^* replaced by $\lambda^{\text{HS}}(X^t)$) is within $(1/2)\log t + O(1)$. We suspect that GRAPA will deliver similar REGROWth and regret, taking as our cue the parametric setting, where both REGROW and regret of order $(1/2)\log t + O(1)$ is achieved for both the 'prequential' ML plug-in method (of which GRAPA is a nonparametric analogue) and the Bayesian mixture (for which JO's approach is the nonparametric analogue).

Imposing a regret-minimizing prior on λ's in (1) is also central to the *KLinf method* in the bandit literature [Agrawal et al., 2021], which directly links growth-optimality of (1) to KL divergence, providing a nonparametric analogue of the duality between KL divergence and GRO established by Grünwald et al. [2024] in the parametric case. A further theoretical analysis of the precise relation between GRAPA, KLinf and regret should lead to better understanding and propel potential extensions such as *bounded regression*.

References

Shubhada Agrawal, Wouter M Koolen, and Sandeep Juneja. Optimal best-arm identification methods for tail-risk measures. In *Advances in Neural Information Processing Systems*, volume 34, pages 25578–25590, 2021.

P. Grünwald, Rianne De Heide, and Wouter Koolen. Safe testing. *Journal of the Royal Statistical Society, Series B (to appear, with discussion)*, 2024.

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.

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

F. Orabona and K.S. Jun. Tight concentrations and confidence sequences from the regret of universal portfolio. *arXiv preprint arXiv:2110.14099*, 2021.