Statistics and Probability Letters 229 (2026) 110577

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A generalisation of Ville's inequality to monotonic lower bounds and thresholds[☆]

Wouter M. Koolen[a,b,*], Muriel F. Pérez-Ortiz[c], Tyron Lardy[a,d]

[a] Centrum Wiskunde & Informatica, Amstedam, The Netherlands
[b] University of Twente, Enschede, The Netherlands
[c] Technical University Eindhoven, Eindhoven, The Netherlands
[d] Leiden University, Leiden, The Netherlands

ARTICLE INFO

Keywords:
Ville's inequality
Supermartingales
Anytime-valid
Optional stopping
Law of iterated logarithm

ABSTRACT

Essentially all anytime-valid methods hinge on Ville's inequality to gain validity across time without incurring a union bound. Ville's inequality is a proper generalisation of Markov's inequality. It states that a non-negative supermartingale will only ever reach a multiple of its initial value with small probability. In the classic rendering both the lower bound (of zero) and the threshold are constant in time. We generalise both to monotonic curves. That is, we bound the probability that a supermartingale which remains above a given decreasing curve exceeds a given increasing threshold curve. We show our bound is tight by exhibiting a supermartingale for which the bound is an equality. Using our generalisation, we derive a cleaner finite-time version of the law of the iterated logarithm.

1. Introduction

We revisit Ville's inequality. To set the stage, consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$. Recall that an adapted process $(M_n)_{n\geq 0}$ is a *supermartingale* if for all time steps $n \geq 0$

$$\mathbb{E}[M_{n+1} | \mathcal{F}_n] \leq M_n \qquad \text{and} \qquad \mathbb{E}[|M_n|] < \infty. \tag{1}$$

The classic inequality by Ville (1939) states that a supermartingale which is bounded below will only ever reach a large threshold with small probability. In the customary presentation with zero as the lower bound, it reads

Theorem 1.1 (*Ville's Inequality*). *Let $(M_n)_{n\geq 0}$ be a non-negative supermartingale. Then for every threshold $C > 0$, $\mathbb{P}\left\{\exists n \geq 0 : M_n \geq C\right\} \leq \frac{\mathbb{E}[M_0]}{C}$.*

This strengthens Markov's inequality by including a search over time steps, while keeping the threshold C and probability $\frac{1}{C}$ inversely proportional. This reveals that, as test statistics, supermartingales are immune to the multiple testing problem that would otherwise require a Bonferroni-type union bound. The statistical use of Ville's inequality stems from the fact that many events of interest can be encoded as a fixed supermartingale not growing. As such Ville's inequality is a workhorse in the fields of sequential analysis (Wald, 1952), game-theoretic probability (Shafer and Vovk, 2019), e-values (Ramdas et al., 2023) and many more.

☆ This article is part of a Special issue entitled: 'E-values and Multiple Testing' published in Statistics and Probability Letters.
* Corresponding author at: Centrum Wiskunde & Informatica, Amstedam, The Netherlands.
E-mail address: wmkoolen@cwi.nl (W.M. Koolen).

https://doi.org/10.1016/j.spl.2025.110577

Available online 18 October 2025

Generalisations. Multiple extensions of Ville's inequality have been proposed in discrete and continuous time. Ruf et al. (2022) study a composite null hypothesis, and they characterise the events that can be expressed as a uniform (in $\mathbb{P}$) supermartingale getting big. Wang and Ramdas (2023) study the non-integrable case, dropping the right hand side condition from (1), and obtain a correction for the threshold/probability relationship. Ramdas and Manole (2023) develop a randomised sharpening of Markov's inequality, which also extends to a Ville-type inequality under external stopping.

Our question. We investigate an orthogonal direction: what if the lower bound and threshold are not constant in time? That is, suppose supermartingale $(M_n)_{n\geq 0}$ is bounded below by $M_n \geq -f(n)$ for some increasing function f, and fix some increasing target threshold $g(n)$. What can we say about $\mathbb{P}\left\{\exists n \geq 0 : M_n \geq g(n)\right\}$? And what applications could this tackle?

Ville & LIL. We focus on the paradigmatic application of Ville's inequality for deriving anytime-valid concentration inequalities of iterated-logarithm type (for the history, see Shafer and Vovk, 2001, § 5.5). These state that for $X_1, X_2, \ldots$ conditionally sub-Gaussian with mean μ,[1] with empirical mean abbreviated by $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^n X_i$, for any confidence $\delta \in (0,1]$, we have

$$\mathbb{P}\left\{\exists n : \hat{\mu}_n - \mu \geq (1+o(1))\sqrt{2\frac{\ln\frac{1}{\delta} + \ln\ln n}{n}}\right\} \leq \delta. \tag{2}$$

Darling and Robbins (1967), Robbins (1970), De la Peña (1999), Balsubramani and Ramdas (2016), Kaufmann and Koolen (2021) prove versions of (2) by applying Ville's inequality to a mixture martingale, crafted to implement a "slicing" weighted union bound over near-exponentially spaced time intervals. Could a more powerful or versatile version of Ville's inequality result in a cleaner path to form (2) with an explicit and interpretable $o(1)$ term?

Partial proof of concept. Our prime example of a load-bearing supermartingale with a diverging lower bound comes from Koolen and Van Erven (2015, § 3.3). For bounded outcomes $X_1, X_2, \ldots \in [-1,+1]$ with conditional mean at most $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq \mu$, they build supermartingales of the form $n \mapsto$

$$\int_0^b \frac{\prod_{i=1}^n (1+\eta(X_i-\mu)) - 1}{\eta}\,\mathrm{d}\eta \quad \text{or} \quad \int_0^b \frac{e^{\eta\sum_{i=1}^n (X_i-\mu) - \frac{n}{2}\eta^2} - 1}{\eta}\,\mathrm{d}\eta. \tag{3}$$

These processes are used in hypothesis testing: they quantify evidence against the stated hypothesis of conditional means in $[-1,\mu]$, and they grow exponentially (in n) when data are in fact better explained by any alternative mean in $(\mu,1]$. The details of how (3) transcend testing canon are intriguing. It is well known (see e.g. Shafer and Vovk, 2001) that for $\eta \geq 0$ sufficiently small, the factors $1+\eta(X_i-\mu)$ and $e^{\eta(X_i-\mu)-\frac{1}{2}\eta^2}$ are non-negative and have expectation bounded by one. Hence taking products over time-steps $i = 1,\ldots,n$, and mixing over $\eta \in [0,b]$ according to some finite (typically probability) measure results in a supermartingale. The innovation in (3) is to mix the *increments* (over the starting value 1) with an *improper* prior $\frac{1}{\eta}\,\mathrm{d}\eta$. The resulting processes (3) are *bona fide* supermartingales, satisfying both parts of (1), but they are not bounded below by any constant, but instead by $-1-\ln n$ at time n. These processes connect to applications as follows. If we assume that supermartingale (3) is not high, a concentration inequality of iterated-logarithm type follows deterministically. But why would it not be high? In the context of individual sequence prediction where (3) arose, a supermartingale can be prevented from growing by the method of defensive forecasting. In that setting, the key advantage of improper vs standard mixtures is computational: the integral required for defensive forecasting admits a closed form. Instead, in the context of stochastic concentration inequalities the growth of a non-negative supermartingale is controlled by Ville's inequality. But how to control (3) and friends? And what are the trade-offs?

Our contributions. In this article we extend Ville's inequality to handle supermartingales with arbitrary lower bounds and thresholds. We arrive at a tight characterisation of the threshold-crossing probability for any monotonic lower bound and threshold curves. Comparing to Koolen and Van Erven (2015), we find a way to handle sub-Gaussian distributions with unbounded outcomes (for which (3) would be unbounded below at every n). Using our extensions, we prove a cleaner iterated logarithm concentration inequality.

2. Main result

We show the following generalisation of Ville's inequality Theorem 1.1.

Theorem 2.1. *Consider two non-decreasing functions f and g defined on $\{0,1,\ldots\}$ such that $-f(0) < g(0)$.*

(a) *For any supermartingale $(M_n)_{n\geq 0}$ bounded below by $M_n \geq -f(n)$ for all $n \geq 0$ and with initial expectation $\mathbb{E}[M_0] \in [-f(0), g(0)]$, we have*

$$\mathbb{P}\left\{\exists n \geq 0 : M_n \geq g(n)\right\} \leq 1 - \frac{g(0)-\mathbb{E}[M_0]}{g(0)+f(0)}\prod_{n=1}^{\infty}\frac{g(n)+f(n-1)}{g(n)+f(n)}. \tag{4}$$

[1] That is, $\mathbb{E}[e^{\eta(X_{n+1}-\mu)}|\mathcal{F}_n] \leq e^{\frac{1}{2}\eta^2}$ for all $\eta \in \mathbb{R}$ and $n \geq 0$.

(b) *For any* $m \in [-f(n), g(n)]$ *there is a martingale* $(M_n)_{n\geq 0}$ *bounded below by* $M_n \geq -f(n)$ *and with* $\mathbb{E}[M_0] = m$ *such that* (4) *holds with equality.*

We recover Theorem 1.1 for constant $f(n) = 0$ and $g(n) = C$ in (4).[2] In Section 2.1 we discuss the martingale that proves tightness (b), and then we prove the upper bound (a) in Section 2.2. A more interpretable continuous-time bound and examples of the use of Theorem 2.1 are obtained in Section 3.

2.1. *Proof of Tightness (Theorem 2.1 item b) by Floor Hugging*

We will first consider the case $M_0 = -f(0)$ deterministically. Then to create a witness martingale for $\mathbb{E}[M_0] \in (-f(0), g(0)]$, we simply randomise (with weights chosen to control the mean) between starting at either $M_0 = -f(0)$ or $M_0 = g(0)$.

We now define the *Floor-Hugger* martingale. It starts with $M_0 = -f(0)$. If it finds itself at time n at value $M_n = -f(n)$, then it either jumps up to $M_{n+1} = g(n+1)$ or drops down to $M_{n+1} = -f(n+1)$—it hugs the floor. The martingale property forces the jump probability p_n to satisfy

$$-f(n) = p_n g(n+1) - (1-p_n) f(n+1) \quad \text{i.e.} \quad p_n = \frac{f(n+1)-f(n)}{g(n+1)+f(n+1)}. \tag{5}$$

Let $s(n) := \mathbb{P}(\forall t \geq n : M_t = -f(t) | M_n = -f(n))$ denote the Floor-Hugger probability to never reach g, starting from time n at value $-f(n)$. Then

$$s(n) = \prod_{t=n}^{\infty} \left(1 - \frac{f(t+1)-f(t)}{g(t+1)+f(t+1)}\right). \tag{6}$$

We observe that $s(n) \in [0,1]$. Taking $n = 0$, we see that the Floor-Hugger martingale witnesses (4) with equality upon plugging in $\mathbb{E}[M_0] = -f(0)$. By randomising between $M_0 = g(0)$ and the Floor Hugger, we obtain a supermartingale witnessing (4) with equality for any starting value $\mathbb{E}[M_0]$.

2.2. *Proof of Upper Bound (Theorem 2.1 item a) using Classic Ville*

Fix any supermartingale $(M_n)_{n\geq 0}$ bounded below by $M_n \geq -f(n)$. We construct the auxiliary[3] non-negative supermartingale $(K_n)_{n\geq 0}$ by

$$K_n := 1 - \frac{g(n) - M_n}{g(n)+f(n)} s(n), \tag{7}$$

where $s(n)$ is the probability, defined in (6), that the Floor-Hugger martingale never reaches g when starting from value $-f(n)$ at time n. We prove that $(K_n)_{n\geq 0}$ is a non-negative supermartingale, and then derive our upper bound using Ville's inequality.

Lemma 2.2. *The process* $(K_n)_{n\geq 0}$ *in* (7) *is a non-negative supermartingale.*

Proof. First, K_n is non-decreasing in M_n, and hence non-negativity follows by plugging in the lowest possible value $M_n = -f(n)$, at which $K_n \geq 1 - s(n) \geq 0$. For the supermartingale claim, we have

$$\begin{aligned}\mathbb{E}[K_{n+1} | \mathcal{F}_n] &\leq 1 - \frac{g(n+1) - M_n}{g(n+1)+f(n+1)} s(n+1) = 1 - \frac{g(n+1) - M_n}{g(n+1)+f(n)} s(n) \\ &\leq 1 - \frac{g(n) - M_n}{g(n)+f(n)} s(n) = K_n,\end{aligned}$$

where we used the definition of K_n, the fact that M_n is a supermartingale, the recurrence $s(n) = \left(1 - \frac{f(n+1)-f(n)}{g(n+1)+f(n+1)}\right) s(n+1)$ by (6), and $M_n \geq -f(n)$. □

Corollary 2.3. *The first claim of Theorem 2.1, i.e.* (4) *holds.*

Proof. As $s(n) \in [0,1]$, K_n is non-decreasing in M_n and hence $M_n \geq g(n)$ implies $K_n \geq 1$. Hence by Ville's inequality (Theorem 1.1), we find

$$\mathbb{P}\left\{\exists n \geq 0 : M_n \geq g(n)\right\} \leq \mathbb{P}\left\{\exists n \geq 0 : K_n \geq 1\right\}$$

---

[2] The result also implies that for constant f, an increasing g (no matter how fast) is no different from the constant $g(0)$; we only get traction when f increases.

[3] In fact K_n is the *upper probability* of the event $\{\exists t \geq n : M_t > g(t)\}$, i.e. the largest probability of this event occurring among all supermartingale futures $(M_t)_{t\geq n}$ that are bounded below by $M_t \geq -f(t)$. In dual fashion, K_n is also the *upper price* of this event, i.e. the smallest capital from which the indicator $\mathbf{1}\{\exists t \geq n : M_t > g(t)\}$ can be super-replicated by engaging sequentially in one-round bets (aka e-variables) that are fair under the composite null that $(M_t)_{t\geq n}$ is a supermartingale bounded below by $M_t \geq -f(t)$.

$$\leq \mathbb{E}[K_0] = 1 - \frac{g(0) - \mathbb{E}[M_0]}{g(0) + f(0)} s(0),$$

which, upon rewriting, yields (4). □

The two key properties used in the proof are that the auxiliary martingale $(K_n)_{n\geq 0}$ is non-negative and that $M_n \geq g(n) \Rightarrow K_n \geq 1$. However, the definition of K_n might seem unintuitive because it involves $s(n)$, which consists of future values of f and g. We were only able to imagine K_n after first finding the Floor-Hugger martingale in Section 2.1. Still, could one possibly design a non-trivial non-negative martingale $(L_n)_{n\geq 0}$ based merely on the history, $L_n = L_n(\{M_i, g(i), f(i)\}_{i\leq n})$, such that $M_n \geq g(n) \Rightarrow L_n \geq 1$? For instance, one could consider defining L_n as a scaled and/or translated version of M_n based on past and present values of f and g (e.g. (7) without the factor $s(n)$). Such a construction turns out to be impossible. This follows from the fact that we must have $\mathbb{E}[L_{n+1} | \mathcal{F}_n] \leq L_n$ regardless of what the original martingale $(M_n)_{n\geq 0}$ does after time n; in particular, this must hold if it does a "Floor-Hugger" type jump (see Section 2.1) from M_n to $g(n+1)$. If this jump is successful, then $M_{n+1} = g(n+1)$, which would imply that $L_{n+1} \geq 1$. However, the probability of success (5) can be made arbitrarily close to one by taking $f(n+1) \to \infty$, which gives $\mathbb{E}[L_{n+1} | \mathcal{F}_n] = 1 \leq L_n$. It follows that $L_n \geq 1$ for all n, so the implication $M_n \geq g(n) \Rightarrow L_n \geq 1$ is powerless. The only way to avoid this is for K_n to account for future values of f and g.

3. A continuous bound and examples

We present a more interpretable relaxation of (4) by considering a lower bound f and threshold g defined on real-valued inputs.

Corollary 3.1. *Let $f, g : \mathbb{R}_+ \to \mathbb{R}_+$ be increasing with f differentiable. Under the setup of Theorem 2.1(a),*

$$\mathbb{P}\left\{\exists n \geq 0 : M_n \geq g(n)\right\} \leq 1 - \frac{g(0) - \mathbb{E}[M_0]}{g(0) + f(0)} \exp\left(-\int_0^\infty \frac{f'(t)}{g(t) + f(t)} \mathrm{d}t\right).$$

Proof. For every $n \geq 1$, as g is increasing,

$$-\int_{n-1}^{n} \frac{f'(t)}{g(t) + f(t)} \mathrm{d}t \leq -\int_{n-1}^{n} \frac{f'(t)}{g(n) + f(t)} \mathrm{d}t = \ln\left(1 - \frac{f(n) - f(n-1)}{g(n) + f(n)}\right). \quad \square$$

This bound is convenient in discrete time, and tight in continuous time (by an analogue of the Floor-Hugger martingale from Section 2.1). Furthermore, its right-hand side is pleasantly invariant under reparametrisations of time by any increasing bijection. Next we explore convenient choices for f and g.

Example 3.2 (*Quadratic*)**.** Fix increasing differentiable f with $f(0) = 0$ and $\lim_{n\to\infty} f(n) = \infty$. We propose to study $g(t) = (af(t) + b)^2 - f(t)$, which is increasing when $2ab \geq 1$. The exponent in Corollary 3.1 is

$$-\int_0^\infty \frac{f'(t)}{g(t) + f(t)} \mathrm{d}t = -\int_0^\infty \frac{f'(t)}{(af(t) + b)^2} \mathrm{d}t = \left.\frac{1}{a(af(t) + b)}\right|_0^\infty = -\frac{1}{ab}.$$

Equating the threshold crossing probability bound $1 - e^{-\frac{1}{ab}}$ to δ requires picking $ab = \frac{1}{-\ln(1-\delta)}$.

The quadratic is in fact a member of a more general construction, which we present in Appendix A. On the other hand, if g vs f grows too slowly, (4) trivialises to 1:

Example 3.3 (*Counterexample*)**.** For diverging $f(t)$, when we take $g = O(f)$, the upper bound in Corollary 3.1 is the trivial 1: let $t_0 \geq 0$ and $c > 0$ be such that $g(t) \leq cf(t)$ for all $t \geq t_0$. Then

$$\int_0^\infty \frac{f'(t)}{g(t) + f(t)} \mathrm{d}t \geq \int_{t_0}^\infty \frac{f'(t)}{(1+c)f(t)} \mathrm{d}t = \ln f(n)|_{t_0}^\infty = \infty.$$

4. Finite-time law of the iterated logarithm (LIL)

We now apply our main result to obtain a non-asymptotic deviation inequality of LIL type. Multiple such inequalities exist in the literature. We will prove a particularly clean new one, by applying Theorem 2.1 to a new, possibly negative, supermartingale. To set the stage, let $X_1, X_2, \ldots$ be a sequence of independent sub-Gaussian random variables with common conditional mean μ, and let $S_n = \sum_{i\leq n}(X_i - \mu)$ be their centred sum. In this sub-Gaussian case we can derive an iterated-logarithm bound of form (2) using the supermartingale

$$M_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left\{\exp\left(\eta S_n - \frac{1}{2}\eta^2 n\right) - 1\right\} \mathrm{e}^{-\eta^2/2} \frac{\mathrm{d}\eta}{|\eta|}. \tag{8}$$

This choice of M_n is inspired by (3), but now instead of using the improper "prior" $\frac{\mathbf{1}\{|\eta|\leq b\}}{|\eta|}\mathrm{d}\eta$, (8) uses $\frac{\mathrm{e}^{-\eta^2/2}}{|\eta|}\mathrm{d}\eta$, which eases some Gaussian-integral computations. Lemma Appendix B.2 shows that $M_n \geq -\frac{1}{\sqrt{2\pi}} \ln(1+n)$. Thus, we can apply Theorem 2.1 with $f(n) = \frac{1}{\sqrt{2\pi}} \ln(1+n)$. The following is shown using a specific choice of threshold g in Theorem 2.1 and a linearisation of the ensuing implicit bound. A slightly tighter version is in Appendix B.

Lemma 4.1. *Let $X_1, X_2, \ldots$ be a sequence of independent sub-Gaussian random variables with common conditional mean μ, and let $S_n = \sum_{i \le n}(X_i - \mu)$. As long as $\delta \le 3/5$, with probability larger than $1-\delta$ and $\delta' := -\ln(1-\delta)$,*

$$\forall n : \left|\frac{S_n}{\sqrt{n+1}}\right| \le \sqrt{2 \ln \left\{ 1 + \frac{\ln(n+1)/\sqrt{\pi} + \mathrm{e}\sqrt{2}}{\delta'} \ln^2\left(\frac{\ln(n+1)}{\sqrt{2\pi}} + \mathrm{e}\right)\right\}} + \frac{1}{\sqrt{2}}.$$

5. Discussion

We extended Ville's inequality to supermartingales that are bounded below by a time-decreasing lower bound, and to time-increasing thresholds. We then put the result to use in deriving concentration inequalities of iterated-logarithm type. Let us discuss what we learned on the way.

Is our extended Ville's inequality strictly more general? The answer is, surprisingly, *no*. This becomes apparent in the proof of the upper bound in Section 2.2, which operates by applying Ville's inequality to an auxiliary non-negative supermartingale. Indeed if one is able to encode an event of interest directly in the form of that auxiliary supermartingale, classic Ville suffices.

Is our extension more user friendly? We strongly believe that the auxiliary supermartingale from Section 2.2 is not intuitive or natural, and we are not aware of any result that can be rendered as having guessed it. We certainly discovered it last, when trying to prove that the Floor-Hugger martingale from Section 2.1 is the worst case. This suggests that possible consequences of our extension, even though accessible in principle, were not practically available. In that sense we claim our extension is more empowering.

Is our finite-time LIL better? To put our result to use, we develop a sub-Gaussian concentration inequality of iterated logarithm type. If we juxtapose the consequence of applying Ville to proper and improper mixture approaches, and ignore constants, we need to unpack respectively

$$\int_{-b}^{b} \frac{e^{\eta S_n - \frac{1}{2} n \eta^2}}{|\eta| \, (\ln|\eta|)^{1+c}} \,\mathrm{d}\eta \;\le\; \frac{1}{\delta} \quad \text{or} \quad \int_{-b}^{b} \frac{e^{\eta S_n - \frac{1}{2} n\eta^2} - 1}{|\eta|} \,\mathrm{d}\eta \;\le\; \frac{\ln n}{\delta}(\ln\ln n)^{1+c}.$$

Any lower bound of either integral gives a LIL bound. The comparison then boils down to which form is easier to bound tightly. We show in Lemma Appendix B.2 that the right-hand inequality rewrites to a fixed function of the deviation $\frac{S_n}{\sqrt{n+1}}$ being below the threshold $\frac{\ln n}{\delta}(\ln\ln n)^{1+c}$. Due to this separation, obtaining a LIL bound reduces to inverting that function. A similar modularity is not present in the left-hand inequality, and its tight analysis is considerably more involved (Robbins, 1970; Koolen and Van Erven, 2015).

One-sided LIL. The proper mixture technique easily adapts to delivering one-sided LIL bounds, simply by mixing only over $\eta \ge 0$. A one-sided improper mixture analogue of (8) would be well-defined. But it would not be bounded below, as can be seen by taking $S_n \to -\infty$, which takes the exponential to zero. This is a curious downside that invites further investigation.

Multiple testing. Theorem 2.1 reveals that the worst-case classic nonnegative martingale only makes a single all-or-nothing attempt to cross the threshold. This is how Ville generalises Markov's inequality without any overhead. The story changes fundamentally with a receding lower bound $-f$. Now, every time margin is created by growing f, the worst-case martingale makes a new independent attempt to hit the then-current threshold g. These attempts all contribute to the overall probability of ever reaching g. In that light, non-constant lower bounds correspond to multiple testing scenarios. Our relation between f, g and the probability of ever reaching g quantifies the exact multiple testing correction required.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2025.110577.

Data availability

No data was used for the research described in the article.

References

Balsubramani, A., Ramdas, A., 2016. Sequential nonparametric testing with the law of the iterated logarithm. In: Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence.

Darling, D.A., Robbins, H., 1967. Iterated logarithm inequalities. Proc. Natl. Acad. Sci. 57 (5), 1188–1192.

Kaufmann, E., Koolen, W.M., 2021. Mixture martingales revisited with applications to sequential tests and confidence intervals. J. Mach. Learn. Res. 22, 1–44.

Koolen, W.M., Van Erven, T., 2015. Second-order quantile methods for experts and combinatorial games. In: Proceedings of the 28th Annual Conference on Learning Theory. COLT.

De la Peña, V.H., 1999. A general class of exponential inequalities for martingales and ratios. Ann. Probab. 27 (1), 537–564.

Ramdas, A., Grünwald, P., Vovk, V., Shafer, G., 2023. Game-theoretic statistics and safe anytime-valid inference. Statist. Sci. 38 (4), 576–601.

Ramdas, A., Manole, T., 2023. Randomized and exchangeable improvements of Markov's, Chebyshev's and Chernoff's inequalities. arXiv:2304.02611.

Robbins, H., 1970. Statistical methods related to the law of the iterated logarithm. Ann. Math. Stat. 41 (5), 1397–1409.

Ruf, J., Larsson, M., Koolen, W.M., Ramdas, A., 2022. A composite generalization of Ville's martingale theorem. arXiv:2203.04485.
Shafer, G., Vovk, V., 2001. Probability and Finance – It's Only a Game!. Wiley, New York.
Shafer, G., Vovk, V., 2019. Game-Theoretic Probability: Theory and Applications to Prediction, Science and Finance. Wiley.
Ville, J., 1939. Étude critique de la notion de collectif. Gauthier-Villars.
Wald, A., 1952. Sequential Analysis. Wiley, New York.
Wang, H., Ramdas, A., 2023. The extended Ville's inequality for nonintegrable nonnegative supermartingales. arXiv:2304.01163.