SI Appendix for Bayesian Posteriors For Arbitrarily Rare Events

Drew Fudenberg, Kevin He, and Lorens A. Imhof

This appendix contains omitted proofs. Before we prove Theorem 1 we prove Proposition 1 and then state and prove a simple consequence of Chernoff's inequality. Both results are needed in the proof of Theorem 1.

Proof of Proposition 1. The assumption that $\pi(p)/\prod_{i=1}^{K} p_i^{\alpha_i-1}$ is uniformly continuous on int Δ implies that the function has a continuous extension $\tilde{\pi} : \Delta \to \mathbb{R}$, see [1], Theorem 5.2, page 302. Let $\tilde{\pi}_0 = \min\{\tilde{\pi}(p) : p \in \Delta\}$. Then $\tilde{\pi}_0 > 0$. Given $\epsilon > 0$, choose $\delta \in (0, \tilde{\pi}_0)$ so small that

$$\frac{1+\frac{\delta}{\tilde{\pi}_0}}{1-\frac{\delta}{\tilde{\pi}_0}} \leq 1+\epsilon. \tag{12}$$

To approximate the integrals in the assertion by sums of Dirichlet integrals we use the fact that the continuous function $\tilde{\pi}$ can be uniformly approximated by Bernstein polynomials, see [2], pages 6 and 51. Thus, there is a polynomial

$$h(p) = \sum_{\substack{\nu_1,\dots,\nu_K \geq 0 \\ \nu_1+\cdots+\nu_K=m}} c_\nu \prod_{i=1}^{K} p_i^{\nu_i}, \qquad c_\nu = \tilde{\pi}\left(\frac{\nu_1}{m}, \dots, \frac{\nu_K}{m}\right) \frac{m!}{\nu_1! \cdots \nu_K!},$$

so that

$$|\tilde{\pi}(p) - h(p)| \leq \delta, \qquad p \in \Delta.$$

Using the formula

$$\int \prod_{i=1}^{K} p_i^{s_i-1}\, d\lambda(p) = \frac{\prod_{i=1}^{K} \Gamma(s_i)}{\Gamma(\sum_{i=1}^{K} s_i)}, \qquad s_1, \dots, s_K > 0,$$

and the relation $\Gamma(s+1) = s\Gamma(s)$, we get

$$\frac{\int p_k \left(\prod_{i=1}^{K} p_i^{s_i-1}\right) h(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^{K} p_i^{s_i-1}\right) h(p)\, d\lambda(p)} = \frac{1}{m+\sum_{i=1}^{K} s_i} \frac{\sum_\nu c_\nu (\nu_k + s_k) \prod_{i=1}^{K} \Gamma(\nu_i + s_i)}{\sum_\nu c_\nu \prod_{i=1}^{K} \Gamma(\nu_i + s_i)}.$$

Since $c_\nu > 0$ for every ν, it follows that

$$\frac{s_k}{m+\sum_{i=1}^{K} s_i} \leq \frac{\int p_k \left(\prod_{i=1}^{K} p_i^{s_i-1}\right) h(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^{K} p_i^{s_i-1}\right) h(p)\, d\lambda(p)} \leq \frac{m+s_k}{m+\sum_{i=1}^{K} s_i}. \tag{13}$$

For all $p \in \Delta$, $h(p) \geq \tilde{\pi}_0$, and so $|\tilde{\pi}(p) - h(p)| \leq \delta \leq \frac{\delta}{\tilde{\pi}_0} h(p)$. Thus,

$$\left(1 - \frac{\delta}{\tilde{\pi}_0}\right) h(p) \leq \tilde{\pi}(p) \leq \left(1 + \frac{\delta}{\tilde{\pi}_0}\right) h(p).$$

It follows from these inequalities together with (12) and (13) that for $n, n_1, \ldots, n_K \in \mathbb{N}_0$ with $\sum_{i=1}^K n_i = n$,

$$\begin{aligned}\frac{\int p_k \left(\prod_{i=1}^K p_i^{n_i}\right) \pi(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^K p_i^{n_i}\right) \pi(p)\, d\lambda(p)} &= \frac{\int p_k \left(\prod_{i=1}^K p_i^{n_i+\alpha_i-1}\right) \tilde{\pi}(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^K p_i^{n_i+\alpha_i-1}\right) \tilde{\pi}(p)\, d\lambda(p)} \\ &\leq \frac{1+\frac{\delta}{\tilde{\pi}_0}}{1-\frac{\delta}{\tilde{\pi}_0}} \frac{\int p_k \left(\prod_{i=1}^K p_i^{n_i+\alpha_i-1}\right) h(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^K p_i^{n_i+\alpha_i-1}\right) h(p)\, d\lambda(p)} \\ &\leq (1+\epsilon) \frac{m + n_k + \alpha_k}{m + n + \sum_{i=1}^K \alpha_i}.\end{aligned}$$

Similarly, using the inequality $1/(1+\epsilon) > 1 - \epsilon$, we obtain

$$\begin{aligned}\frac{\int p_k \left(\prod_{i=1}^K p_i^{n_i}\right) \pi(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^K p_i^{n_i}\right) \pi(p)\, d\lambda(p)} &\geq \frac{1-\frac{\delta}{\tilde{\pi}_0}}{1+\frac{\delta}{\tilde{\pi}_0}} \frac{\int p_k \left(\prod_{i=1}^K p_i^{n_i+\alpha_i-1}\right) h(p)\, d\lambda(p)}{\int \left(\prod_{i=1}^K p_i^{n_i+\alpha_i-1}\right) h(p)\, d\lambda(p)} \\ &\geq (1-\epsilon) \frac{n_k + \alpha_k}{m + n + \sum_{i=1}^K \alpha_i}.\end{aligned}$$

The assertion follows with $\gamma = m + \sum_{i=1}^K \alpha_i$. □

Remark 3′. Using results on the degree of approximation by Bernstein polynomials, one may compute explicit values for the constant γ in Proposition 1. If, for example, $K = 2$ and $\phi(p_1) = \tilde{\pi}(p_1, 1-p_1)$ has a continuous derivative on $[0, 1]$, one can apply Theorem 1.6.1 in [2] to show that (3) holds with

$$\gamma = \alpha_1 + \alpha_2 + \left[\frac{5}{4}\left(1 + \frac{2}{\epsilon}\right) \frac{\max\{|\phi'(p_1)| : 0 \leq p_1 \leq 1\}}{\min\{\phi(p_1) : 0 \leq p_1 \leq 1\}}\right]^2.$$

If $K \geq 2$ and $\tilde{\pi}$ coincides with a polynomial on Δ, then, by a result of [3], π can be written as a finite mixture of densities of Dirichlet distributions and Theorem 3 of [4] gives a computable upper bound on the support of the mixing distribution. Thus, the inequalities in (4) hold with computable constants a and A.

Lemma 3. *Let S_n be a binomial random variable with parameters n and p. Let $1 < c < 2$ and $d > 0$. Then*

$$\mathbb{P}\left(\frac{S_n}{n} \geq cp + \frac{d}{n}\right) \leq e^{(1-c)d}, \qquad \mathbb{P}\left(\frac{S_n}{n} \leq \frac{p}{c} - \frac{d}{n}\right) \leq e^{(1-c)d}.$$

Proof. By Chernoff's inequality,

$$\mathbb{P}\left(\frac{S_n}{n} \geq cp + \frac{d}{n}\right) \leq \inf_{t>0} \left[e^{-t(cp+\frac{d}{n})}(1 - p + pe^t)\right]^n \leq e^{(1-c)d}[\psi(p)]^n,$$

where $\psi(s) = e^{(1-c)cs}(1 - s + se^{c-1})$. For $0 \le s \le 1$,

$$\frac{\psi'(s)}{e^{(1-c)cs}} = e^{c-1} - 1 - (c-1)c - s(c-1)c(e^{c-1} - 1) \le e^{c-1} - 1 - (c-1)c.$$

Set $\phi(u) = e^{u-1} - 1 - (u-1)u$. The function ϕ' is convex, $\phi'(1) = 0$ and $\phi'(2) < 0$. Thus, ϕ' is negative on $(1, 2)$, so that $\phi(c) < \phi(1) = 0$. It now follows that ψ is decreasing on $[0, 1]$, so that $\psi(p) \le \psi(0) = 1$. This proves the first claim. The proof of the second claim is similar. □

Proof of Theorem 1. Let $0 < \epsilon < 1$. Choose $c \in (1, 2)$ and $\delta > 0$ so that

$$\frac{1-\delta}{c} > 1 - \frac{\epsilon}{2}, \qquad (1+\delta)c < 1 + \frac{\epsilon}{2}.$$

Let $d > 0$ be so that the bound in Lemma 3 satisfies $e^{(1-c)d} < \frac{\epsilon}{2}$. By Proposition 1, there exists $\gamma > 0$ so that for every $n \in \mathbb{N}$,

$$(1-\delta)\frac{X_k^n}{n+\gamma} \le \hat{p}_k(X^n) \le (1+\delta)\frac{X_k^n + \gamma}{n}, \qquad k = 1, \ldots, K.$$

Let N be so large that

$$(1-\delta)\left(\frac{1}{c} - \frac{d}{N}\right)\frac{1}{1+\gamma/N} > 1 - \epsilon, \quad (1+\delta)\frac{d+\gamma}{N} < \frac{\epsilon}{2}.$$

Fix k, p_k and n with $np_k \ge N$. Set $A = \{\frac{1}{n}X_k^n < cp_k + \frac{d}{n}\}$ and $B = \{\frac{1}{n}X_k^n > \frac{p_k}{c} - \frac{d}{n}\}$. On A,

$$\frac{\hat{p}_k(X^n)}{p_k} \le (1+\delta)\frac{X_k^n + \gamma}{np_k} \le (1+\delta)\left(c + \frac{d+\gamma}{np_k}\right) \le (1+\delta)\left(c + \frac{d+\gamma}{N}\right) < 1 + \epsilon$$

and on B,

$$\frac{\hat{p}_k(X^n)}{p_k} \ge (1-\delta)\frac{X_k^n}{np_k}\frac{n}{n+\gamma} \ge (1-\delta)\left(\frac{1}{c} - \frac{d}{np_k}\right)\frac{1}{1+\gamma/n}$$
$$\ge (1-\delta)\left(\frac{1}{c} - \frac{d}{N}\right)\frac{1}{1+\gamma/N} > 1 - \epsilon.$$

By Lemma 3, $\mathbb{P}_p(A \cap B) \ge 1 - \mathbb{P}_p(A^c) - \mathbb{P}_p(B^c) \ge 1 - \epsilon$. □

Remark 3″. In the proof of Theorem 1 one can choose $c = 1 + \frac{\epsilon}{4}$, $\delta = \frac{\epsilon}{5}$, and $d = 3\epsilon^{-2}$. If the prior-dependent constant $\gamma > 0$ is so chosen that the inequalities in (3) hold with ϵ replaced by $\frac{\epsilon}{5}$, then it follows by a small variation of the above proof that the conclusion of Theorem 1 holds for $N = 8\epsilon^{-3} + 3\gamma\epsilon^{-1}$.

The proof of Example 1 uses the following lower bound for the Bayes estimates of p_1.

Lemma 4. *Let* $\pi(p) = e^{-1/p}$, $0 < p \le 1$. *Then*

$$\frac{\int_0^1 p^{\nu+1}(1-p)^{n-\nu}\pi(p)\,dp}{\int_0^1 p^{\nu}(1-p)^{n-\nu}\pi(p)\,dp} \ge \frac{1}{8\sqrt{1 \vee n}}$$

for every $n \in \mathbb{N}_0$ *and* $\nu = 0, \dots, n$.

Proof. Let U be a random variable with density proportional to $p^{\nu}(1-p)^{n-\nu}\pi(p)$ and let V be a random variable with density proportional to $(1-p)^n\pi(p)$, $0 < p < 1$. Then U is larger than V in the likelihood ratio order since $p^{\nu}(1-p)^{n-\nu}\pi(p)/[(1-p)^n\pi(p)] = (p/(1-p))^{\nu}$ is increasing in p. This implies that $\mathbb{E}(U) \ge \mathbb{E}(V)$, that is,

$$\frac{\int_0^1 p^{\nu+1}(1-p)^{n-\nu}\pi(p)\,dp}{\int_0^1 p^{\nu}(1-p)^{n-\nu}\pi(p)\,dp} \ge \frac{\int_0^1 p(1-p)^{n}\pi(p)\,dp}{\int_0^1 (1-p)^{n}\pi(p)\,dp},$$

see [5], page 70. It is therefore enough to prove the claim for $\nu = 0$.

Let $f_n(p) = c_n(1-p)^n\pi(p)$, where $c_n = [\int_0^1 (1-p)^n\pi(p)\,dp]^{-1}$. We have

$$f_n'(p) = c_n \frac{e^{-1/p}(1-p)^{n-1}}{p^2}(1 - p - np^2),$$

showing that f_n is increasing on $[0, 2a_n]$, where $a_n = 1/(4\sqrt{1 \vee n})$. Hence

$$\frac{\int_{a_n}^1 f_n(p)\,dp}{1 - \int_{a_n}^1 f_n(p)\,dp} = \frac{\int_{a_n}^1 f_n(p)\,dp}{\int_0^{a_n} f_n(p)\,dp} \ge \frac{\int_{a_n}^{2a_n} f_n(p)\,dp}{a_n f_n(a_n)} \ge \frac{(2a_n - a_n)f(a_n)}{a_n f_n(a_n)} = 1.$$

Thus $\int_{a_n}^1 f_n(p)\,dp \ge \frac{1}{2}$, and therefore

$$\int_0^1 p f_n(p)\,dp \ge \int_{a_n}^1 p f_n(p)\,dp \ge a_n \int_{a_n}^1 f_n(p)\,dp \ge \frac{1}{2}a_n = \frac{1}{8\sqrt{1 \vee n}}. \qquad \square$$

Proof of Example 1. Let $N \in \mathbb{N}$. For $n > N^2$ define $p(n) \in \Delta$ by $p_1(n) = Nn^{-\frac{1}{2}-\delta}$. By Lemma 4, $\hat{p}_1(X^n) - 2p_1(n) \ge n^{-\frac{1}{2}}(\frac{1}{8} - 2Nn^{-\delta})$, and so, for n sufficiently large, $\mathbb{P}_{p(n)}(|\hat{p}_1(X^n) - p_1(n)| > p_1(n)) = 1$. $\square$

Proof of Example 2. Suppose π satisfies Condition $\mathcal{P}(\alpha)$, $\alpha \in (0,\infty)^K$. By Proposition 1, there exists $\gamma > 0$ so that $\hat{p}_1(X^n) \ge \alpha_1/[2(n+\gamma)]$. For every $n > \alpha_1/8$ pick $p(n) \in \Delta$ with $p_1(n) = \alpha_1/(8n)$. Let $n_0 = \max(\alpha_1/8, \gamma)$. If $n > n_0$, then $\alpha_1/[2(n+\gamma)] > 2p_1(n)$, and so $\mathbb{P}_{p(n)}(|\hat{p}_1(X^n) - p_1(n)| > p_1(n)) = 1$. Since $\limsup_{n\to\infty} \zeta(n)/n = \infty$, there exists for every $N \in \mathbb{N}$ an $n > n_0$ with $\zeta(n)p_1(n) \ge N$. $\square$

The following result was used in Remark 4.

Proposition 2. *Let* $K > 2$ *and* $\bar{k} \in \{1, \dots, K\}$. *Suppose the density* π *of the prior distribution on* Δ *satisfies Condition* $\mathcal{P}(\alpha_1, \dots, \alpha_K)$ *with* $\alpha_1, \dots, \alpha_K > 0$. *Then the image measure induced by the mapping* $(p_1, \dots, p_K) \mapsto (p_{\bar{k}}, \sum_{k \ne \bar{k}} p_k)$ *has a density that satisfies Condition* $\mathcal{P}(\alpha_{\bar{k}}, \sum_{k \ne \bar{k}} \alpha_k)$.

Proof. Suppose without loss of generality that $\bar{k} = 1$. Then the image measure has a density π_1 with respect to the normalized Lebesgue measure on $\Delta_1 = \{q \in [0,1]^2 : q_1 + q_2 = 1\}$ which is given by

$$\pi_1(q) = \int_{A(q_2)} \pi\left(q_1, p_2, \dots, p_{K-1}, q_2 - \sum_{k=2}^{K-1} p_k\right) d(p_2, \dots, p_{K-1}),$$

where

$$A(q_2) = \{(p_2, \dots, p_{K-1}) \in (0,1)^{K-2} : p_2 + \dots + p_{K-1} < q_2\}.$$

Making the change of variable $t = (t_2, \dots, t_{K-1}) = q_2^{-1}(p_2, \dots, p_{K-1})$ we get

$$\pi_1(q) = q_2^{K-2} \int_{A(1)} \pi\left(q_1, q_2 t, q_2\left(1 - \sum_{k=2}^{K-1} t_k\right)\right) dt$$

for $q \in \Delta_1$ with $q_2 > 0$. Since π satisfies Condition $\mathcal{P}(\alpha_1, \dots, \alpha_K)$, there exists a continuous positive function $\tilde{\pi}$ on Δ such that $\tilde{\pi}(p) = \pi(p)/\prod_{k=1}^K p_k^{\alpha_k - 1}$ for all $p \in \operatorname{int} \Delta$. Hence, for $q \in \operatorname{int} \Delta_1$,

$$\frac{\pi_1(q)}{q_1^{\alpha_1 - 1} q_2^{(\sum_{k=2}^K \alpha_k) - 1}} = \int_{A(1)} \tilde{\pi}\left(q_1, q_2 t, q_2\left(1 - \sum_{k=2}^{K-1} t_k\right)\right) \prod_{k=2}^{K-1} t_k^{\alpha_k - 1} \left(1 - \sum_{k=2}^{K-1} t_k\right)^{\alpha_K - 1} dt.$$

The integral is positive for every $q \in \Delta_1$ and, by dominated convergence, depends continuously on $q \in \Delta_1$. Thus, π_1 satisfies Condition $\mathcal{P}(\alpha_1, \alpha_2 + \dots + \alpha_K)$. $\square$

Proof of Example 3. Let $N \in \mathbb{N}$. For every $n \geq \max(N, \frac{N}{c})$ let $p(n) = (\frac{N}{n}, 1 - \frac{N}{n})$, $q(n) = (\frac{N}{cn}, 1 - \frac{N}{cn})$, $\vartheta_n = (p(n), q(n))$, and

$$A_n = \left\{\hat{p}_1(X^n) \geq \frac{c}{2}\hat{q}_1(X^n)\right\}.$$

We will prove more than is stated, namely that $\mathbb{P}_{\vartheta_n}(A_n) \to 0$ as $n \to \infty$. Let Y_n denote the number of times the blue die lands on side 1 in the first n periods. By Proposition 1, there exists $\gamma > 0$ so that $\hat{p}_1(X^n) \leq \frac{3}{2}(Y_n + \gamma)/(B_n + \gamma)$. For every $n \geq \max(N, \frac{N}{c})$ and $b \in \{0, 1, \dots, n\}$, by Lemma 4,

$$\mathbb{P}_{\vartheta_n}(A_n | B_n = b) \leq \mathbb{P}_{\vartheta_n}\left(\frac{3}{2}\frac{Y_n + \gamma}{b + \gamma} \geq \frac{c}{16\sqrt{1 \vee (n-b)}} \middle| B_n = b\right).$$

If $b > \frac{n}{2}\mu_B$, then $c(b+\gamma)/(24\sqrt{1 \vee (n-b)}) \geq d\sqrt{n}$ with $d := c\mu_B/(48\sqrt{1 - \mu_B/2})$, and it follows that

$$\mathbb{P}_{\vartheta_n}(A_n | B_n = b) \leq \mathbb{P}_{\vartheta_n}(Y_n \geq -\gamma + d\sqrt{n} | B_n = b).$$

To bound the probability on the right-hand side we use a Poisson approximation to the conditional distribution of Y_n. Let W_ν be a Poisson random variable with mean ν. Then, by [6], (43) on page 89,

$$\begin{aligned}\mathbb{P}_{\vartheta_n}\left(Y_n \geq -\gamma + d\sqrt{n} | B_n = b\right) &\leq \mathbb{P}\left(W_{bp_1(n)} \geq -\gamma + d\sqrt{n}\right) + p_1(n) \\ &\leq \mathbb{P}\left(W_N \geq -\gamma + d\sqrt{n}\right) + \frac{N}{n}.\end{aligned}$$

In the second line we used the fact that W_N is stochastically larger than $W_{bp_1(n)}$ because $N \geq bp_1(n)$, see [5], pages 67-70. Hence

$$\begin{aligned}\mathbb{P}_{\vartheta_n}(A_n) &\leq \mathbb{P}_{\vartheta_n}\left(B_n \leq \frac{n}{2}\mu_B\right) + \sum_{b:b>\frac{n}{2}\mu_B} \mathbb{P}_{\vartheta_n}(A_n|B_n = b)\mathbb{P}_{\vartheta_n}(B_n = b) \\ &\leq \mathbb{P}_{\vartheta_n}\left(\frac{1}{n}B_n \leq \frac{1}{2}\mu_B\right) + \mathbb{P}\left(W_N \geq -\gamma + d\sqrt{n}\right) + \frac{N}{n}.\end{aligned}$$

As $n \to \infty$, $\mathbb{P}(W_N \geq -\gamma + d\sqrt{n}) \to 0$ and, by the weak law of large numbers, $\mathbb{P}_{\vartheta_n}(\frac{1}{n}B_n \leq \frac{1}{2}\mu_B) \to 0$. Thus, $\mathbb{P}_{\vartheta_n}(A_n) \to 0$ as $n \to \infty$. $\square$

Proof of Example 4. Let Y_n and Z_n be the respective number of times the blue and the red die land on side 1 in the first n periods. By Proposition 1, there exists $\gamma > 0$ so that

$$\begin{aligned}\mathbb{P}_\vartheta\left(\hat{p}_1(X^n) < \frac{c}{2}\hat{q}_1(X^n)\right) &\geq \mathbb{P}_\vartheta\left(\frac{3}{2}\frac{Y_n + \gamma}{B_n + \gamma} < \frac{c}{4}\frac{Z_n}{n + \gamma}\right) \\ &\geq \mathbb{P}_\vartheta\left(Y_n = 0, \frac{6\gamma}{c} < \frac{B_n}{n}Z_n\right).\end{aligned}$$

For every $n \in \mathbb{N}$ with $n \geq c$ pick $\vartheta_n = (p(n), q(n)) \in \Delta^2$ with $p_1(n) = \frac{c}{n}$ and $q_1(n) = \frac{1}{n}$. Let $\mu_0 \in (0, \mu_B)$ and $\mu_1 \in (\mu_B, 1)$. Then, for $b = \lceil \mu_0 n \rceil, \ldots, \lfloor \mu_1 n \rfloor$,

$$\mathbb{P}_{\vartheta_n}\left(Y_n = 0, \frac{6\gamma}{c} < \frac{B_n}{n}Z_n \middle| B_n = b\right) \geq [1 - p_1(n)]^n \mathbb{P}_{\vartheta_n}\left(\frac{6\gamma}{c\mu_0} < Z_n \middle| B_n = \lfloor \mu_1 n \rfloor\right).$$

Now $[1 - p_1(n)]^n \to e^{-c} > 0$ and, by [6], (43) on page 89,

$$\mathbb{P}_{\vartheta_n}\left(\frac{6\gamma}{c\mu_0} < Z_n \middle| B_n = \lfloor \mu_1 n \rfloor\right) \geq \mathbb{P}\left(W > \frac{6\gamma}{c\mu_0}\right) - \frac{1}{n},$$

where W is a Poisson random variable with mean $1 - \mu_1$. Hence

$$\liminf_{n\to\infty} \mathbb{P}_{\vartheta_n}\left(Y_n = 0, \frac{6\gamma}{c} < \frac{B_n}{n}Z_n \middle| \mu_0 n \leq B_n \leq \mu_1 n\right) > 0.$$

Since $\mathbb{P}(\mu_0 n \leq B_n \leq \mu_1 n) \to 1$, it follows that there exists $\epsilon_0 > 0$ and $n_0 \in \mathbb{N}$ so that

$$\mathbb{P}_{\vartheta_n}\left(\hat{p}_1(X^n) < \frac{c}{2}\hat{q}_1(X^n)\right) > \epsilon_0$$

for all $n \geq n_0$. Since $\zeta(p_1(n))/p_1(n) \to \infty$ as $n \to \infty$, there exists for every $N \in \mathbb{N}$ an $n \geq n_0$ with $n\zeta(p_1(n)) \geq N$ and ϑ_n has the required properties. $\square$

Proof of Lemma 1. Set $\ell = d/(n \wedge m)$. By Markov's inequality, for every $t > 0$,

$$\mathbb{P}\left(\frac{T_m}{m} \geq \frac{1}{c'}\frac{S_n}{n} + \ell\right) = \mathbb{P}\left(e^{t(c'T_m - \frac{m}{n}S_n)} \geq e^{tc'\ell m}\right) \leq \frac{\mathbb{E}[e^{t(c'T_m - \frac{m}{n}S_n)}]}{e^{tc'\ell m}}. \tag{14}$$

We will determine a suitable value for t so that the expectation is at most 1. Let ξ and τ be Bernoulli variables with $\mathbb{P}(\xi = 1) = p$ and $\mathbb{P}(\tau = 1) = q$. Then

$$\mathbb{E}[e^{t(c'T_m - \frac{m}{n}S_n)}] = \mathbb{E}(e^{tc'T_m})\mathbb{E}(e^{-t\frac{m}{n}S_n}) = [\mathbb{E}(e^{tc'\tau})]^m[\mathbb{E}(e^{-t\frac{m}{n}\xi})]^n. \tag{15}$$

For $t > 0$ and $s \in \mathbb{R}$ let $\psi_t(s) = (1 - s + se^{c't})(1 - cs + cse^{-t})$. Since $p \geq cq$,

$$\mathbb{E}(e^{tc'\tau})\mathbb{E}(e^{-t\xi}) = (1 - q + qe^{c't})(1 - p + pe^{-t}) \leq \psi_t(q).$$

We have $\psi_t(0) = 1$, and $\psi_t''(s) = 2c(e^{c't} - 1)(e^{-t} - 1) < 0$, so that ψ_t is concave. For $t_0 := (c' + 1)^{-1}\log(c/c')$,

$$\psi_{t_0}'(0) = e^{c't_0} - 1 + c(e^{-t_0} - 1) = \int_0^{t_0} e^{-u}[c'e^{(c'+1)u} - c]\,du < 0,$$

so that $\psi_{t_0}(s) \leq 1$ for $s \geq 0$. Hence,

$$\mathbb{E}(e^{c't_0\tau})\mathbb{E}(e^{-t_0\xi}) \leq 1. \tag{16}$$

If $m \leq n$, then by Lyapunov's inequality, $[\mathbb{E}(e^{-t_0\frac{m}{n}\xi})]^n \leq [\mathbb{E}(e^{-t_0\xi})]^m$. Combining this inequality with (15) and (16) yields

$$\mathbb{E}[e^{t_0(c'T_m - \frac{m}{n}S_n)}] \leq [\mathbb{E}(e^{t_0c'\tau})]^m[\mathbb{E}(e^{-t_0\xi})]^m \leq 1,$$

and so, by (14),

$$\mathbb{P}\left(\frac{T_m}{m} \geq \frac{1}{c'}\frac{S_n}{n} + \ell\right) \leq e^{-t_0c'\ell m} = \left(\frac{c'}{c}\right)^{c'd/(c'+1)}.$$

If $m > n$, then Lyapunov's inequality gives $[\mathbb{E}(e^{tc'\tau})]^m \leq [\mathbb{E}(e^{tc'\frac{m}{n}\tau})]^n$. Setting $t_1 = \frac{n}{m}t_0$, we get in this case

$$\mathbb{E}[e^{t_1(c'T_m - \frac{m}{n}S_n)}] \leq [\mathbb{E}(e^{t_1c'\frac{m}{n}\tau})]^n[\mathbb{E}(e^{-t_1\frac{m}{n}\xi})]^n \leq 1,$$

and so

$$\mathbb{P}\left(\frac{T_m}{m} \geq \frac{1}{c'}\frac{S_n}{n} + \ell\right) \leq e^{-t_1c'\ell m} = \left(\frac{c'}{c}\right)^{c'd/(c'+1)}. \qquad \square$$

Proof of Lemma 2. We will use a Poisson approximation to the binomial distribution. If W_ν is a Poisson random variable with mean $\nu > 0$, then $\mathbb{P}(W_\nu \leq M) \to 0$ as $\nu \to \infty$. Thus there exists $N_0 \in \mathbb{N}$ so that $\mathbb{P}(W_\nu \leq M) < \frac{1}{2}\epsilon$ for $\nu > N_0$. By [6], (43) on page 89, $|\mathbb{P}_p(S_n \leq M) - \mathbb{P}(W_{np} \leq M)| \leq p$. Thus if $np \geq N_0$ and $p \leq \frac{1}{2}\epsilon$, then $\mathbb{P}_p(S_n \leq M) \leq \epsilon$. In particular, for $p = \frac{1}{2}\epsilon$ and $n = \lceil 2N_0/\epsilon \rceil$, we have $\mathbb{P}_{\epsilon/2}(S_{\lceil 2N_0/\epsilon \rceil} \leq M) \leq \epsilon$.

On the other hand, if $p > \frac{1}{2}\epsilon$ and $n \geq 2N_0/\epsilon$, then

$$\mathbb{P}_p(S_n \leq M) \leq \mathbb{P}_{\epsilon/2}(S_n \leq M) \leq \mathbb{P}_{\epsilon/2}(S_{\lceil 2N_0/\epsilon \rceil} \leq M) \leq \epsilon,$$

where we used the fact that the family of binomial distributions is stochastically increasing in both parameters, see e.g. [5], pages 67-70. The claim follows with $N = 2N_0/\epsilon$. $\square$

References

[1] Dugundji J (1966) *Topology.* (Allyn and Bacon, Boston).
[2] Lorentz GG (1986) *Bernstein Polynomials.* (Chelsea Publishing Company, New York).
[3] Handelman D (1988) Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific Journal of Mathematics* 132(1):35–62.
[4] Powers V, Reznick B (2001) A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *Journal of Pure and Applied Algebra* 164(1–2):221–229.
[5] Lehmann EL, Romano JP (2005) *Testing Statistical Hypotheses.* (Springer Science & Business Media, New York).
[6] Stein C (1986) *Approximate Computation of Expectations*, Lecture Notes-Monograph Series. (Institute of Mathematical Statistics, Hayward, CA) Vol. 7.