Novel refinement of Pinsker's inequality via medical trial design

Charlie Cowen-Breen

August 26, 2023

We will prove a refinement of Pinsker's inequality by way of a thought experiment: we show that a hypothetical medical trial would extract and exploit a relatively large amount of information about the treatments it is testing, but an amount which is necessarily constrained by the classical upper-bound of Lai and Robbins. We propose the manifest inequality as a Pinsker-like inequality which may be of independent interest.

Theorem 1. *Let $0 < p_1, p_2 < 1$, and $d(p_1, p_2)$ denote the relative entropy between the distributions* $\mathrm{Bern}(p_1)$ *and* $\mathrm{Bern}(p_2)$*. Then*

$$\max\left(2, \frac{1}{24\overline{p}(1-\overline{p})}\right)(p_1 - p_2)^2 \leq d(p_1, p_2)$$

where $\overline{p}$ is the arithmetic mean of p_1 and p_2.

This is stronger than Pinsker's inequality for $|\overline{p} - 1/2| > \sqrt{11/48}$, as well as its refinements proven in [1] and [2], for $|\overline{p} - 1/2| \gtrsim 0.485$. While Pinsker's inequality states

$$2(p_1 - p_2)^2 \leq d(p_1, p_2)$$

in [1], the authors compute an optimal prefactor $\phi(p_1)$ such that

$$\phi(p_1)(p_1 - p_2)^2 \leq d(p_1, p_2)$$

One can view our result here as the analogous statement except for a prefactor $\psi(\overline{p}) = 1/24\overline{p}(1-\overline{p})$ such that

$$\psi(\overline{p})(p_1 - p_2)^2 \leq d(p_1, p_2)$$

Another way of interpreting Theorem 1 is the following: Pinsker's inequality compares $d(p_1, p_2)$ to the absolute difference between p_1 and p_2. Theorem 1 provides a tighter bound by comparing $d(p_1, p_2)$ not only to the *difference* $|p_1 - p_2|$ between p_1 and p_2, but also to their the *sum*, $p_1 + p_2 = 2\overline{p}$. In particular, it bounds $d(p_1, p_2)$ in terms of a rational function of the sum and difference of p_1 and p_2.

1 Setup

1.1 Clinical trials

In what follows, we will give an informal overview of clinical trial design and its connection with the Lai-Robbins theorem; for a more thorough explanation, see [3]. Given repeated choices between two options, say with expected utility θ_0 and θ_1, respectively, the classical theorem of Lai and Robbins provides an upper bound on the amount of *effective information* which may be extracted and exploited in real-time, made precise in terms of the relative entropy $d(\theta_0, \theta_1)$. One case in which the Lai-Robbins theorem applies is that of trials of experimental medical treatments. Here, we propose a system of extracting and exploiting information about experimental medical treatments by first performing a *randomized controlled trial* to determine which treatment, if any, is statistical superior to the others, and subsequently applying that treatment to the remainder of the population.

The Lai-Robbins upper bound, in terms of the relative entropy between the treatment effects, applies to this situation as well. In fact, the expected utility achieved by this situation is close enough to the Lai-Robbins upper bound to demonstrate Theorem 1.

1.2 Bandit formulation

In the case of particular interest here, we will consider the following scenario: Suppose that $0 < p_A, p_B < 1$ are fixed and known. Consider a game in which you repeatedly draw prizes from two boxes, one with expected utility $\theta_1 = (p_A + p_B)/2$, and the other with expected utility θ_0, where θ_0 is either p_A or p_B (but it is not known which is equal to θ_0). In each round of the game, you choose which box to draw a prize from based on your previous choices and the utilities you subsequently observed.

We will use the strategy for solving this setup proposed in [3], together with the extension of the Lai-Robbins theorem proven by Burnetas and Katehakis, to prove Theorem 1.

1.3 Utility

In this section, we briefly review the medical trial design proposed in [3]. Suppose that to each patient we may administer two treatments with expected utilities θ_0 and θ_1, respectively. Given a total number of patients to be treated N, The proposal of [3] is to select a subset of $n \leq N$ for a randomized controlled trial, and then apply the statistically superior treatment to the remaining $N - n$ patients. A randomized controlled trial consists of applying the treatment 0 (with expected utility θ_0) to half of the n trial patients, and treatment 1 (with expected utility θ_1) to the other half of the trial patients.

Based on the reactions of these first n patients to the treatments, a Z-test may be performed; it will indicate the treatment found to be statistically superior. In [3], the probability of incorrectly concluding that the *inferior* treatment is in fact *superior* is calculated to be

$$\Phi\left(\frac{-|A|\sqrt{n}}{\sqrt{B}}\right)$$

This is a standard calculation. If X_i represents the (random) utility of the i^{th} patient, the total expected utility, as a proportion of the maximum achievable offline utility $N \max \theta$, is

$$\tilde{U}(\theta) = \frac{\mathbb{E}^{\pi}[\sum_{i=1}^{N} X_i]}{\max_{\pi'} \mathbb{E}^{\pi'}[\sum_{i=1}^{N} X_i]} = \frac{\mathbb{E}^{\pi}[\sum_{i=1}^{N} X_i]}{N \max \theta}$$

As in [3], one can then derive that with this choice of utility function, if $\tilde{U}_n(\theta)$ represents the expected utility of the policy in which $n \leq N$ patients are chosen for trial in the procedure described above, we have

$$\tilde{U}_n(\theta) = \frac{n}{2N}\left(1 + \frac{\min \theta}{\max \theta}\right) + \frac{N-n}{N}\left(1 - \left(1 - \frac{\min \theta}{\max \theta}\right)\Phi\left(\frac{-|A|\sqrt{n}}{\sqrt{B}}\right)\right)$$

2 Proof of Theorem 1

Choosing $n^* = c \log N$ samples, we see that in the large sample limit

$$U = 1 - \frac{c}{2}\left(1 - \frac{\min\theta}{\max\theta}\right)\frac{\log N}{N}(1 + o(1))$$

which corresponds to regret equal to

$$\mathcal{R} = \frac{c}{2}|\delta|(1 + o(1))\log N$$

for all θ such that $c \geq \frac{4(\sigma_0^2+\sigma_1^2)}{\delta^2}$ (i.e. $F \geq 1$). Since $\sigma_0^2 + \sigma_1^2 = \theta_0(1-\theta_0) + \theta_1(1-\theta_1) \leq 1/2$, it suffices to take c such that $c \geq \frac{2}{\delta^2} \geq \frac{4(\sigma_0^2+\sigma_1^2)}{\delta^2}$.

To this end, we will do the following. Say that we are given $p_A < p_B \in (0,1)$. Let $\overline{p} = (p_A + p_B)/2 \in (p_A, p_B)$ and $\Delta = |p_B - p_A|/2 > 0$. Put $\Theta_0 = \{p_A, p_B\}$ and $\Theta_1 = \{\overline{p}\}$ so that, conceptually, we imagine a one-parameter hypothesis test

$$\begin{aligned} H_0 &: \theta = p_A \text{ or } \theta = p_B \\ H_1 &: \theta = \overline{p} \end{aligned}$$

For all $(\theta_0, \theta_1) \in \Theta_0 \times \Theta_1$, we then immediately have that

$$|\delta| \geq \Delta \tag{1}$$

where $\delta = \theta_1 - \theta_0$. Define $\sigma^2_{\max} = \max\{\sigma^2(p_A), \sigma^2(p_B)\}$, where where $\sigma^2(\lambda) = \lambda(1-\lambda)$. Trivially, this means that

$$\sigma^2(\theta_0) \leq \sigma^2_{\max} \quad \forall \theta_0 \in \Theta_0 \tag{2}$$

Finally, let us take

$$c = \frac{4(\sigma^2_{\max} + \sigma^2(\overline{p}))}{\Delta^2}$$

Now, it follows from Equation 1 and Equation 2 that for all $(\theta_0, \theta_1) \in \Theta_0 \times \Theta_1$ we have

$$\frac{4(\sigma^2(\theta_0) + \sigma^2(\theta_1))}{\delta^2} \leq \frac{4(\sigma^2_{\max} + \sigma^2(\overline{p}))}{\Delta^2}$$

and thus $F \geq 1$ for all $\theta \in \Theta_0 \times \Theta_1$. This means that we have regret bounded by

$$\mathcal{R} \leq \frac{2(\sigma^2_{\max} + \sigma^2(\overline{p}))}{\Delta^2}|\theta_1 - \theta_0| \log N(1 + o(1)) \tag{3}$$

for all $\theta \in \Theta_0 \times \Theta_1$.

We now show that this setup satisfies the conditions for the forward direction of the Burnetas & Katehakis theorem. In the notation of the paper, we have two populations, Θ_0 and Θ_1, each one dimensional so that $\underline{\theta}_0 = \theta_0$ and $\underline{\theta}_1 = \theta_1$. Then with $\Theta = \Theta_0 \times \Theta_1$, $\underline{\underline{\theta}} = (\theta_0, \theta_1) \in \Theta$. We also have $\mu_a(\underline{\theta}_a) = \theta_a$ since the expectation of a Bernoulli distribution is equal to its parameter.

Consider $\underline{\underline{\theta}} = (\theta_0, \theta_1) = (p_A, \overline{p}) \in \Theta$. Then $\mu^*(\underline{\underline{\theta}}) := \max_{a=0,1}\{\mu_a(\underline{\theta}_a)\} = \theta_1 = \overline{p}$ since by construction, $\overline{p} > p_A$. Then $O(\underline{\underline{\theta}}) := \{a : \mu_a = \mu^*(\underline{\underline{\theta}})\} = \{a : \theta_a = \overline{p}\} = \{1\}$. We also compute

$$\Delta\Theta_0(\underline{\theta}_0) = \{\underline{\theta}_0' \in \Theta_0 : \theta_0' > \overline{p}\} = \{p_B\} \neq \emptyset$$

Now, recall that $B(\underline{\underline{\theta}}) := \{a : a \notin O(\underline{\underline{\theta}}) \text{ and } \Delta\Theta_a(\underline{\underline{\theta}}_a) \neq \emptyset\}$. Since $O(\underline{\underline{\theta}}) = \{1\}$, this forces $1 \notin B(\underline{\underline{\theta}})$. However, $a = 0$ satisfies the conditions required to be in $B(\underline{\underline{\theta}})$ as computed above, so we have $B(\underline{\underline{\theta}}) = \{0\}$.

We have one last quantity to compute. $K_0(\underline{\underline{\theta}}) = \inf\{I(\underline{\underline{\theta}}_0, \underline{\underline{\theta}}'_0) : \underline{\underline{\theta}}'_0 \in \Delta\Theta_0(\underline{\underline{\theta}}_0\} = \inf\{d(\theta_0, \theta'_0) : \theta'_0 \in \{p_B\}\}$. where $d(x, y)$ denotes the Bernoulli KL divergence between x and y. Thus,

$$K_0(\underline{\underline{\theta}}) = d(\theta_0, p_B)$$

This quantity is nonzero because $\theta_0 = p_A < p_B$ by construction and $d(x, y)$ vanishes only for $x = y$.

We can finally state and use the main theorem of Burnetas and Katehakis. Indeed, $\underline{\underline{\theta}} \in \Theta$ is such that $0 \in B(\underline{\underline{\theta}})$ and $K_0(\underline{\underline{\theta}}) \neq 0$. It is also true that the policy π which dictates the RCT mechanism we are proposing lies in the class of *uniformly fast convergent* policies $\mathcal{C}_{UF}$, as defined in the paper: $\pi \in \mathcal{C}_{UF}$ if $\forall \underline{\underline{\theta}} \in \Theta$, $\mathcal{R}^\pi_N(\underline{\underline{\theta}}) = o_{N\to\infty}(N^\alpha)$, $\forall \alpha > 0$. This criterion follows by our regret bound, Equation 3, that pathwise for each fixed $\underline{\underline{\theta}} \in \Theta$,

$$\mathcal{R}^\pi_N = O(\log N) = o(N^\alpha) \quad \forall \alpha > 0.$$

Given that the conditions of the theorem are satisfied, we have the conclusion

$$\varliminf_{N\to\infty} \frac{\mathbb{E}^\pi_{\underline{\underline{\theta}}} T_N(0)}{\log N} \geq \frac{1}{K_0(\underline{\underline{\theta}})} = \frac{1}{d(p_A, p_B)}$$

Using the relation $\mathcal{R}^\pi_N = (\theta_1 - \theta_0)\mathbb{E}^\pi_{\underline{\underline{\theta}}} T_N(0)$ for $\underline{\underline{\theta}}$ such that $\theta_0 < \theta_1$, it follows that

$$\varliminf_{N\to\infty} \frac{\mathcal{R}^\pi_N}{\log N} \geq \frac{(\theta_1 - \theta_0)}{d(p_A, p_B)}$$

In view of Equation 3, we have then that

$$\varliminf_{N\to\infty} \frac{2(\sigma^2_{\max} + \sigma^2(\overline{p}))}{\Delta^2}(1 + h(N)) \geq \frac{1}{d(p_A, p_B)}$$

where $h(x)$ is a suitable function having the property that $\lim_{x\to\infty} h(x) = 0$. Evaluating the limit for our fixed choices of p_A, p_B, we conclude that

$$\frac{2(\sigma^2_{\max} + \sigma^2(\overline{p}))}{\left(\frac{p_B - p_A}{2}\right)^2} \geq \frac{1}{d(p_A, p_B)}$$

Rearranging,

$$d(p_A, p_B) \geq \frac{\left(\frac{p_B - p_A}{2}\right)^2}{2(\sigma^2_{\max} + \sigma^2(\overline{p}))} = \frac{(p_B - p_A)^2}{8(\sigma^2_{\max} + \sigma^2(\overline{p}))}$$

With a final bound $\sigma^2(p_A), \sigma^2(p_B) \leq 2\sigma^2(\overline{p})$ by Jensen's inequality, we have

$$d(p_A, p_B) \geq \frac{(p_B - p_A)^2}{24\sigma^2(\overline{p})}$$

which is stronger than Pinsker's inequality for sufficiently small $\overline{p}$.

References

[1] E. Ordentlich and M.J. Weinberger. A distribution dependent refinement of Pinsker's inequality. IEEE Transactions on Information Theory, 51(5):1836–1840, 2005.

[2] Sebastien Gerchinovitz, Pierre Ménard, Gilles Stoltz. Fano's inequality for random variables. 2017. hal-01470862v1

[3] Charlie Cowen-Breen, Sofia Villar. Guidelines for Sample Sizes: Randomised Controlled Trials are nearly Optimal for Overall Patient Health, if the Right Number of Patients is Chosen. 2023.