Solving a coin-flip l2 mimimax estimation problem

John Mount*

August 8, 2014

Abstract

We set up and solve a minimax under l2 loss problem derived from a coin flipping problem. The solution is interesting as it involves use of cancellation to solve the mimimax problem.

1 Introduction

Wald [Wald, 1949] set up statistical estimation as a game played against nature where the researcher picks a (possibly probabilistic) decision function and nature picks an adversarial distribution. Nature's distribution plays the role of Bayesian priors, but is not considered to be the an objective true distribution or a subjective estimate. It is instead a worst-possible distribution so that any inference bounds proven in this formalism hold in general. This game theoretic form of probability is fascinating and leads quickly to interesting questions and procedures.

2 The Problem

Take as our problem the task of estimating the unknown win-rate (or heads-rate) p of a random process or coin. We assume the process is memory-less and stationary (p is not changing and does not depend on earlier flips). We observe a sequence of n flips showing h wins/heads, and then are asked to return an estimate $\phi_n(h)$ for p. This problem was discussed and given context in [Mount, 2014a], [Mount, 2014c], [Mount, 2014b], and [Bauer et al., 2014].

Fix $n \in \mathbb{N}$, $n \geq 1$. Let $p \in [0,1]$ and $\phi = (\phi(0), \ldots, \phi(n))$ be a $(n+1)$-dimensional real vector in $[0,1]^{n+1}$, and define:

$$L_n(p, \phi) := \sum_{h=0}^{n} \binom{n}{h} p^h (1-p)^{n-h} (\phi(h) - p)^2. \tag{1}$$

$L_n(p, \phi)$ represents the expected square-error of encountered when using ϕ to estimate the win-rate of a coin with (unknown) true win-rate p by observing n flips/outcomes. The estimate is: use $\phi(h)$ when you see h wins/heads. This is related to Wald's game-theoretic formalism, but we are insisting on pure strategies for both the estimate (a single deterministic ϕ) and a single unknown true probability p. We are going to assume that nature picks p in an adversarial manner with full knowledge of ϕ.

Define:

$$f_n(\phi) = \max_{p \in [0,1]} L_n(p, \phi). \tag{2}$$

What we are looking for is $\mathrm{argmin}_{\phi \in [0,1]^{n+1}} f_n(\phi)$. The issue is: the definition of $f_n()$ has two quantifiers so it seems like it will be difficult to derive or even check solutions.

*email: mailto:jmount@win-vector.com web: http://www.win-vector.com/

3 A Solution

Lemma 1. *Suppose ϕ is in the interior of $[0,1]^{n+1}$ and is such that $L_n(p,\phi)-\phi(0)^2 = 0$ simultaneously for all p. Then: ϕ is the unique global minimizer of $f_n()$.*

Proof. Suppose ϕ is as stated. We will confirm ϕ is an isolated local minimum by checking partial derivatives. Look at $\frac{\partial}{\partial \phi(h)} f_n(\phi)$ and $\frac{\partial}{\partial - \phi(h)} f_n(\phi)$. If we can show these are always both positive for all h we are done.

Because $L_n(p,\phi)$ is a constant independent of p we know $\frac{\partial}{\partial \phi(h)} f_n(\phi) \geq \frac{\partial}{\partial \phi(h)} L_n(p,\phi)$ for any $p \in [0,1]$ (i.e. all p curves are active or on the boundary boundary). So

$$\begin{aligned}
\frac{\partial}{\partial \phi(h)} f_n(\phi) &\geq \max_p \frac{\partial}{\partial \phi(h)} L_n(p,\phi) \\
&= \max_p \binom{n}{h} p^h (1-p)^{n-h} 2(\phi(h) - p) \\
&\geq \binom{n}{h} p^h (1-p)^{n-h} 2(\phi(h) - p)\Big|_{p=\phi(h)/2} \\
&> 0
\end{aligned}$$

Similarly we know $\frac{\partial}{\partial - \phi(h)} f_n(\phi) \geq \frac{\partial}{\partial - \phi(h)} L_n(p,\phi)$ for any $p \in [0,1]$. So

$$\begin{aligned}
\frac{\partial}{\partial - \phi(h)} f_n(\phi) &\geq \max_p \frac{\partial}{\partial - \phi(h)} L_n(p,\phi) \\
&= \max_p \binom{n}{h} p^h (1-p)^{n-h} 2(p - \phi(h)) \\
&\geq \binom{n}{h} p^h (1-p)^{n-h} 2(p - \phi(h))\Big|_{p=(1+\phi(h))/2} \\
&> 0
\end{aligned}$$

So we know ϕ is an isolated local minimum of $f_n()$ But $L_n(p,\phi)$ is convex in ϕ for any fixed n, p ($n \geq 1, p \in \mathbb{R}$), so $f_n(\phi)$ is also convex in ϕ. So an isolated local minimum ϕ is also the unique global minimum. □

Define: ϕ_n as the vector in $\mathbb{R}^{n+1}$ such that

$$\phi_n(h) := (\frac{1}{2}\sqrt{n} + h)/(\frac{1}{2}\sqrt{n} + n). \tag{3}$$

Lemma 2. *ϕ_n from equation 3 is in the interior of $[0,1]^{n+1}$ has $L_n(p,\phi) - \phi(0)^2 = 0$ simultaneously for all p.*

Proof. It is obvious is in the interior of $[0,1]^{n+1}$. So it is just a matter of checking $L_n(p,\phi_n) - \phi_n(0)^2 = 0$ using arguments from [Bauer et al., 2014] or by checking ϕ_n obeys the recurrences in [Mount, 2014b]. □

Theorem 1. *ϕ_n from equation 3 is the unique minimizer of $f_n(\phi)$ and the only ϕ in the interior of $[0,1]^{n+1}$ such that $L_n(p,\phi) - \phi(0)^2 = 0$.*

Proof. By lemma 2 we know ϕ_n meets the conditions of lemma 1. Therefore ϕ_n is the unique global minimizer of $f_n()$. This follows by combining lemma 1 and lemma 2. The uniqueness of the minimizer of $f_n()$ means there can be no other solutions of $L_n(p,\phi) - \phi(0)^2 = 0$ that meet the pre-conditions of lemma 1, so ϕ_n must be the only solution $L_n(p,\phi) - \phi(0)^2 = 0$ in the interior of $[0,1]^{n+1}$. Note $L_n(p,\phi) - \phi(0)^2 = 0$ can have solutions outside of $[0,1]^{n+1}$ (for example it is known to have non-real solutions). □

It is kind of neat we get that these is no more than one solution of $L_n(p, \phi) - \phi(0)^2 = 0$ in the interior of $[0, 1]^{n+1}$ from the convexity of the related optimization problem.

References

[Bauer et al., 2014] Bauer, A., Mount, J., and Dotsenko, V. (2014). Mathoverflow: Existence of solutions of a polynomial system. http://mathoverflow.net/questions/177574/existence-of-solutions-of-a-polynomial-system. Accessed: 2014-08-08.

[Mount, 2014a] Mount, J. (2014a). Frequentist inference only seems easy. http://www.win-vector.com/blog/2014/07/frequenstist-inference-only-seems-easy/. Accessed: 2014-08-08.

[Mount, 2014b] Mount, J. (2014b). Winvector: explicit solution. https://github.com/WinVector/Examples/blob/master/freq/python/explicitSolution.rst. Accessed: 2014-08-08.

[Mount, 2014c] Mount, J. (2014c). Winvector: frequentist soln. https://github.com/WinVector/Examples/blob/master/freq/python/freqMin.rst. Accessed: 2014-08-08.

[Wald, 1949] Wald, A. (1949). Statistical Decision Functions. *Ann. Math. Statist.*, 20(2):165–205.