Asymptotically Optimal Search Policy for Odd Arm Identification[1]

Srikrishna Bhashyam

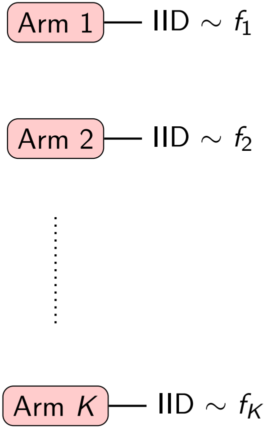
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai 600036

BITS 2018

Joint work with Gayathri R. Prabhu, Aditya Gopalan and Rajesh Sundaresan

[1] G. R. Prabhu, S. Bhashyam, A. Gopalan, R. Sundaresan, "Optimal Odd Arm Identification with Fixed Confidence," http://arxiv.org/abs/1712.03682, Dec. 2017.

Odd Arm Identification: Model



- K arms
- $K - 1$ arms have identical distribution
- Odd arm has a different distribution

- Choose the arm to observe at each stage
- Identify the odd arm
- Metrics
  - Probability of false detection
  - Delay in arriving at the decision
  - Switching cost

Applications: Anomaly Detection, Search tasks, Controlled sensing

Odd Arm Identification: Problem

Objective

Find the policy that minimizes expected cost for a given probability of false detection constraint

Setting

- Distribution of arms belong to the exponential family

$$f(x|\eta) = h(x)\exp\left(\eta^T T(x) - A(\eta)\right) \quad \forall x \in \mathbb{R}^d,$$

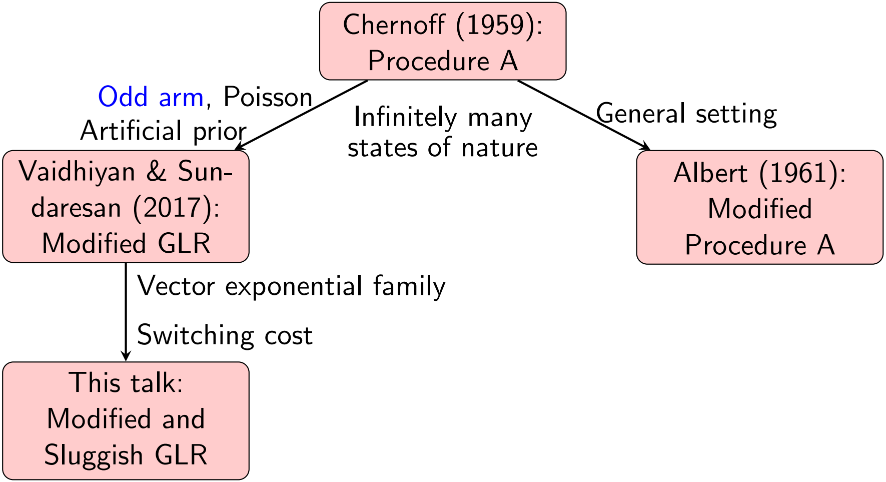
  where η is the vector parameter,
  η_1: Parameter of the odd arm (unknown),
  η_2: Parameter of the other arms (unknown)
- Probability of false detection, $P_F \leq \alpha$
- Cost(C) = Delay(τ) + Switching cost(g)

Related Work

- Sequential hypothesis with control
  - Chernoff (1959): Sequential design of experiments
  - Albert (1961): Composite hypothesis with infinitely many states of nature
- Odd arm identification
  - Vaidhiyan & Sundaresan (2017): Poisson observations, artificial prior
- Best arm identification
  - Garivier & Kaufmann (2016): One-parameter exponential family

- This work: Odd arm identification
  - Builds on Vaidhiyan & Sundaresan (2017)
  - General vector exponential family
  - Switching costs
  - Same asymptotic optimality

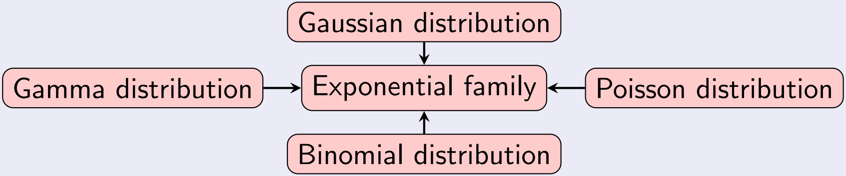
Related Work



Exponential family: Useful facts

$$f\left(x|\eta\right) = h\left(x\right) \exp\left(\eta^T T\left(x\right) - A\left(\eta\right)\right) \quad \forall x \in \mathbb{R}^d,$$

Unified model



Useful parameters and expressions

Dual parameters: $\eta \rightleftarrows \kappa = E[T(X)]$, Conjugate functions: $A(\eta) \rightleftarrows F(\kappa)$

$$\begin{aligned} D\left(\eta_1 || \eta_2\right) &:= D\left(f(\cdot|\eta_1) || f(\cdot|\eta_2)\right) \\ &= \left(\eta_1 - \eta_2\right)^T \kappa_1 - A\left(\eta_1\right) + A\left(\eta_2\right) \\ &= \left(\kappa_2 - \kappa_1\right)^T \eta_2 + F\left(\kappa_1\right) - F\left(\kappa_2\right). \end{aligned}$$

Lower Bound and Some Observations

Lower Bound and Interpretation

$$E[C(\pi)|\psi] \geq E[\tau|\psi] \geq \frac{-\log\alpha}{D^*(i,\eta_1,\eta_2)} \text{ as } \alpha \to 0$$

Config. $\psi = (i, \eta_1, \eta_2)$

Arm 1 — $\sim f(.|\eta_2)$

⋮

Arm i — $\sim f(.|\eta_1)$

⋮

Arm K — $\sim f(.|\eta_2)$

Config. $\psi' = (j, \eta_1', \eta_2')$

Arm 1 — $\sim f(.|\eta_2')$

⋮

Arm j — $\sim f(.|\eta_1')$

⋮

Arm K — $\sim f(.|\eta_2')$

Extract $\log\left(\frac{1}{\alpha}\right)$ units of information

Lower Bound and Interpretation

How much information can we get in each slot, on average?

$$D^*(i,\eta_1,\eta_2) = \max_{\lambda\in\mathcal{P}(K)} \min_{\eta_1',\eta_2',j\neq i} \left[\lambda(i)\, D\left(\eta_1||\eta_2'\right) + \lambda(j)\, D\left(\eta_2||\eta_1'\right) + (1-\lambda(i)-\lambda(j))\, D\left(\eta_2||\eta_2'\right)\right]$$

Max-min-drift of log-likelihood ratio process between configurations (i,η_1,η_2) and (j,η_1',η_2')

- Minimum over all possible error configurations
- Maximum over all IID sampling policies

$$\text{Expected delay} \geq \frac{\log(1/\alpha)}{D^*}$$

Simplifications of the lower bound: Exponential family

One-dimensional optimization

$$D^*(i,\eta_1,\eta_2) = \max_{0\le\lambda(i)\le 1}\left[\lambda(i)\,D(\eta_1||\tilde{\eta}) + (1-\lambda(i))\,\frac{K-2}{K-1}D(\eta_2||\tilde{\eta})\right]$$

where $\tilde{\eta} = f(\tilde{\kappa})$ with

$$\tilde{\kappa} = \hat{\lambda}(i)\,\kappa_1 + (1-\hat{\lambda}(i))\kappa_2, \quad \hat{\lambda}(i) = \frac{\lambda(i)}{\lambda(i) + (1-\lambda(i))\,\frac{K-2}{K-1}}.$$

Also, we have $\lambda^*(i,\eta_1,\eta_2)(j)$ of the form

$$\left\{\frac{1-\lambda^*(i)}{K-1},\ldots,\frac{1-\lambda^*(i)}{K-1},\lambda^*(i),\frac{1-\lambda^*(i)}{K-1},\ldots,\frac{1-\lambda^*(i)}{K-1}\right\}$$

1 $\quad$ i $\quad$ K

Nontrivial sampling of all actions

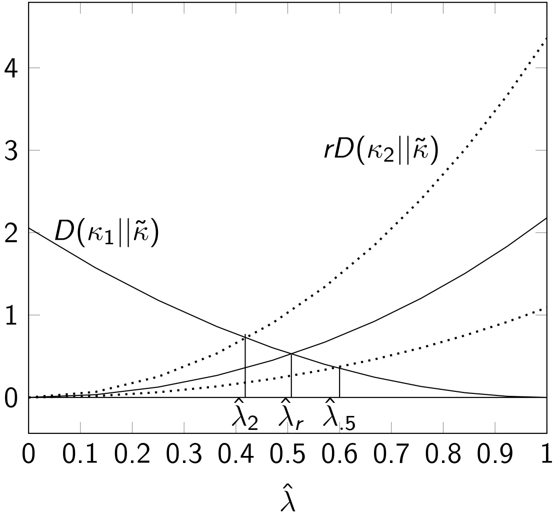
Nontrivial sampling strategy

$$\lambda^* (k, \eta_1, \eta_2) (j) \geq c_K > 0$$

for all $j \in 1, 2, \ldots, K$ and for all (k, η_1, η_2) such that $\eta_1 \neq \eta_2$

- Each arm sampled atleast c_K fraction of time independent of true configuration
- Useful to show convergence of parameter estimates
- Proof for Poisson case in Vaidhiyan & Sundaresan (2017)
- Need to show λ (or $\hat{\lambda}$) bounded away from 0 and 1

Nontrivial sampling of all actions: Exponential family



- Optimal $\hat{\lambda}$ satisfies

$$D(\kappa_1||\tilde{\kappa}) = \frac{K-2}{K-1} D(\kappa_2||\tilde{\kappa})$$

- Sufficient to show $\hat{\lambda}_{0.5} < 1$ and $\hat{\lambda}_2 > 0$
- Further simplification using Taylor's theorem

Nontrivial sampling of all actions: Exponential family

Conjugate functions: $A(\eta) \rightleftarrows F(\kappa)$

Sufficient condition

$\exists \;\; \hat{\lambda}^* < 1$ such that

$$\int_{\hat{\lambda}^*}^{1} (1-u)\Delta\kappa^T \, \text{Hess}\,(F)\Delta\kappa du - \frac{1}{2}\int_{0}^{\hat{\lambda}^*} u\Delta\kappa^T \, \text{Hess}\,(F)\Delta\kappa du < 0.$$

- Condition proved for Poisson, single parameter Gaussian
- Numerically checked for Bernoulli, two-parameter Gaussian

Proposed Policy: Modified and Sluggish GLR

Proposed Policy: Modified Generalized Likelihood Ratio

$$Z_{ij}(n) := \log \frac{\tilde{f}(X^n, A^n | H=i)}{\hat{f}(X^n, A^n | H=j)}$$

$\hat{f}(X^n, A^n | H = j)$: Maximum likelihood of observations and actions till time n under $H = j$

$\tilde{f}(X^n, A^n | H = i)$: Averaged likelihood (according to the conjugate prior)

$$Z_i(n) := \min_{j \neq i} Z_{ij}(n)$$

Proposed Policy: Estimates of the Expectation parameter

Estimate of odd and non-odd expectation parameters under $H = j$

$$\hat{\kappa}_1^n(j) = \frac{Y_j^n}{N_j^n} \quad \text{and} \quad \hat{\kappa}_2^n(j) = \frac{Y^n - Y_j^n}{n - N_j^n},$$

where $N_j^n = \sum_{t=1}^{n} 1_{\{A_t = j\}}$, Y_j^n is the sum of sufficient statistic of arm j up to time n, i.e.,

$$Y_j^n = \sum_{t=1}^{n} T(X_t) 1_{\{A_t = j\}},$$

and $Y^n = \sum_{j=1}^{K} Y_j^n$.

Proposed Policy: Modified and Sluggish GLR ($\pi_{SM}(\alpha, \gamma)$)

Estimate parameters under each hypothesis

↓

Modified GLR: Compute $Z_i(n)$ for each i

↓

Arm with largest modified GLR: $i^*(n) = \arg\max_i Z_i(n)$

↓

Stopping Rule: If $Z_{i^*(n)}(n) \geq \log\left(\frac{K-1}{\alpha}\right)$ stop and declare $i^*(n)$ as the odd arm location

↓

Sampling Rule: Else sample according to $\lambda^*\left(i^*(n), \hat{\eta}_1^n(i^*(n)), \hat{\eta}_2^n(i^*(n))\right)$ w.p. γ

Performance of Proposed Policy

Performance of Proposed Policy: Stops in finite time

Policy stops in finite time with probability 1

- Parameter estimates converge to true values, almost surely
- When $H = i^*$, Test statistic $Z_{i^*}(n)$ has a positive drift
- And crosses threshold $\log(\frac{K-1}{\alpha})$ in finite time, almost surely

Performance of Proposed Policy: Satisfies false detection constraint

Policy satisfies the constraint on the probability of false detection α

- Threshold $= \log(\frac{K-1}{\alpha})$
- Proof relies on conjugate prior on parameters

Performance of Proposed Policy: Asymptotically optimal in total cost

$$\limsup_{L\to\infty} \frac{E[C(\pi_{SM}(\alpha,\gamma))|\psi]}{\log(L)} \le \frac{1}{D^*(i,\eta_1,\eta_2)} + \frac{g_{max}\gamma}{D^*(i,\eta_1,\eta_2)},$$

where $L = 1/\alpha$. Proof uses

- Convergence of parameter estimates to the actual parameters, almost surely
- Convergence of positive drift of the test statistic to $D^*(i,\eta_1,\eta_2)$ as $\alpha \to 0$
- Exponential bound on $P[Z_i(n) < \log((K-1)L)]$ for large n and L

Choose γ to be arbitrarily close to 0 to approach the lower bound

Summary

- Proposed modified and sluggish GLR for odd arm identification
- Asymptotically optimal cost as false detection constraint $\alpha \to 0$
- Generalization of result in Vaidhiyan & Sundaresan (2017)
  - Vector exponential family for observations
  - Include switching costs
- Growth rate of the cost, as both α and γ are driven to 0, is the same as that without switching costs.

Current Work

- Other structures: Best arm identification

Thank you

http://www.ee.iitm.ac.in/∼skrishna/

Supported by DST