Empirical Bayes for Compound Adaptive Experiments

Jiaying Gu

University of Toronto

Indiana University
April 2, 2025

Joint work with Karun Adusumilli (U Penn)and Junfan Tao (Kyoto U)
Preliminary, Comments are welcome!

Sequential experiments

- Sequential experiments are widely used in a number of fields
  - Online advertising, clinical trials, economic interventions...
  - Allow one to target and achieve optimal balance of welfare, ethical, and economic criteria
  - E.g., since 2006, the FDA has actively recommended using sequential experiments in clinical trials
- Now commonplace to run multiple sequential experiments
- Questions:
  - How do we perform estimation and make decisions following these experiments?
  - Can we use information across experiments to improve decision-making?

Estimation following sequential experiments

- Consider question of estimation of treatment effects following sequential experiments
- Classical methods: MLE/using sample means does not work
  - Sample size is random and dependent on data
  - MLE can be badly biased due to selection
- Solution 1: De-bias MLE, e.g, through inverse probability weighting (Hadad et al, 2021)
  - Restores asymptotic normality and unbiasedness of MLE.
  - This only works for restrictive classes of algorithms (e.g., deterministic algorithms are excluded)
  - Moreover, we need to know the algorithm.

- Solution 2: Use Bayesian methods
  - Probably the most common approach in practice
  - Start with a prior, then estimate treatment effects through posterior mean
  - By likelihood principle (proper stopping rule), posterior does not depend on how data is obtained
  - But how should we choose the prior?

Empirical Bayes (EB) methods

- EB methods
  - Aim to improve decision making across a collective by 'learning from the experience of others'
  - Specifically, they let us learn the 'prior' from data
  - Two main approaches to EB modeling (Efron, 2014): g-modeling and f-modeling
- Our contribution: extend EB methodology to sequential experiments
  - Short summary: g-modeling works but f-modeling fails
- In fact, we can simply employ g-modeling by pretending data is exogenously generated!

What is g-modeling?

- Suppose $Y_i|\theta_i \sim N(\theta_i, 1)$, $\theta_i \in \Theta = \mathbb{R}$ and $\theta_i \sim G$, $i = 1, \ldots, n$.
- If we know G, the optimal Bayes estimator minimizes

$$\mathbb{E}[(\delta(Y_i) - \theta_i)^2] = \int \int (\delta(y) - \theta)^2 \varphi(y - \theta) dy dG(\theta)$$

- and takes the form

$$\hat{\theta}_i^* = \mathbb{E}[\theta | Y_i] = \frac{\int \theta \varphi(Y_i - \theta) dG(\theta)}{\int \varphi(Y_i - \theta) dG(\theta)}$$

- g-modeling: estimate G via deconvolution, then plug in.

What is f-modeling?

- For any G, given the base density belongs to exponential family, we have the Tweedie formula:
$$\mathbb{E}[\theta | Y = y] = y + \frac{f'(y)}{f(y)}$$
with
$$f(y) = \int \varphi(y - \theta) dG(\theta)$$
- f-modeling: estimate f and f' from the sample $Y_1, \ldots, Y_n$, then plug in
- Under iid sampling, both f and g-modeling can lead to good EB estimator.

EB for adaptive sampling

- We show that g-modeling works for adaptive experiments, but not f-modeling.
- g-modeling has some remarkable properties under adaptive experiments:
  - Does not require knowledge of algorithms used to generate the data.
  - Algorithms could even vary across experiments
- We analyze parametric g-modeling method, e.g. linear shrinkage, Simple GMM
- We also analyze non-parametric g-modeling methods, specifically, NPMLE (non-parametric maximum likelihood) to estimate G.
- We provide finite sample regret analysis on both.

Related literature

- Adaptive experiments: Bandit experiments, optimal stopping, p-hacking,...
  - See Lattimore & Szepesvari (2020), Wassmer & Brannath (2016)
  - Our contribution: New way to analyze multiple adaptive experiments
- Empirical Bayes: Large literature
  - See Efron (2016), Walters (2024), Koenker and Gu (2024) for surveys
  - Our contribution: Extension to adaptive experiments, new interpretation of g-modeling
- Statistical gaurantee for NP-g-modeling: Jiang and Zhang (2009), Polyanskiy & Wu (2020), Jiang (2020), Chen (2024)
  - Our contribution: Extension of regret analysis to adaptive setting

Motivating example; the likelihood principle

OCEs

- Online controlled experiment (OCEs):
  - Web-based randomized controlled trials for evaluating digital products and services
  - Users are randomly assigned to a control group or one of the K treatment groups
  - When $K = 1$, they are called A/B tests
- OCEs use adaptive stopping algorithms to determine when to stop experimentation
  - E.g., Wald's stopping rule: stop if average treatment differences multiplied by time exceeds a threshold

The ASOS digital experiments dataset

- Between 2019-20 fashion retailer ASOS conducted $n = 61$ A/B tests (web designs).
- In each dataset, arms were sampled in exactly equal proportions
- However: algorithms used to stop are proprietary and unknown
  - Could even have differed across experiments!
- We are interested in estimating treatment effects for all experiments

Data generating model

- Experiments indexed by i
- Data in each experiment collected in stages $j = 1, 2, \ldots$
- Each stage: observe difference

$$Y_{j,i} = Y_{j,i}^{(1)} - Y_{j,i}^{(0)}$$

  between a single treatment and single control observation
- Outcomes are Gaussian with known variance

$$Y_{j,i} \sim \mathcal{N}(\theta_i, \omega_i^2)$$

  - θ_i is unknown treatment effect we want to estimate

Likelihood for Each Experiment

- Let $A_{j,i} \in \{0,1\}$ indicate whether we stop or continue sampling in stage j
- Information set until period j denoted $\mathcal{I}_{j,i} \equiv \{Y_{1,i}, A_{1,i}, \ldots, Y_{j-1,i}, A_{j-1,i}\}$
- Actions determined by algorithm (which could be experiment specific)

$$\pi_{j,i} : \mathcal{I}_{j,i} \to [0,1]$$

- At the end of experiment, define
  - N_i: number of observations sampled
  - $Z_i = N_i^{-1} \sum_j Y_{j,i}$: sample mean of observations

Likelihood for Each Experiment

- Likelihood of data $\mathcal{D}_i$ is:

$$
\begin{aligned}
p(\mathcal{D}_i|\theta_i) &= \prod_{j=1}^{N_i} p(A_{j,i}, Y_{j,i}|\theta_i, \mathcal{I}_{j,i}) \\
&= \prod_{j=1}^{N_i - 1} p(A_{j,i} = 1|\mathcal{I}_{j,i}, \theta_i) p(A_{N_i,i} = 0|\mathcal{I}_{N_i,i}, \theta_i) \cdot \prod_{j=1}^{N_i} p(Y_{j,i}|A_{j-1,i} = 1, \mathcal{I}_{j,i}, \theta_i) \\
&\overset{*}{=} \left[\prod_{j=1}^{N_i} \pi_{j,i}(A_{j,i}|\mathcal{I}_{j,i})\right] \cdot \left[\prod_{j=1}^{N_i} p(Y_{j,i}|\theta_i)\right]
\end{aligned}
$$

- (*): whether to draw one more sample only depends on $\mathcal{I}_{j,i}$, not θ_i.
- (*): Conditional on drawing $(A_{j-1,i} = 1)$, the distribution of $Y_{j,i}$ only depends on θ_i, not $\mathcal{I}_{j,i}$.

Likelihood for Each Experiment

- Likelihood is proportional to $\prod_{j=1}^{N_i} \varphi(y_{j,i} - \theta_i)$
- That is, N_i is exogenously fixed, and $Y_{j,i}$ are iid draws.
- By normality,

$$p(\mathcal{D}_i|\theta_i) = c(\mathcal{D}_i) \cdot \frac{1}{\sigma_i}\varphi\left(\frac{Z_i - \theta_i}{\sigma_i}\right)$$

with $\sigma_i^2 = \omega_i^2 / N_i$.

Working likelihood and likelihood principle

- If parallel experiments are independent, then full likelihood

$$p(\mathcal{D}|\theta_1, \ldots, \theta_n) = c(\mathcal{D}) \cdot \prod_{i=1}^{n} \frac{1}{\sigma_i} \varphi\left(\frac{Z_i - \theta_i}{\sigma_i}\right)$$

- Consider a scenario where N_i are determined exogenously for all i.

- And $Z_i|\theta_i \sim \mathcal{N}(\theta_i, \sigma_i^2)$ and likelihood of observations $(Z_1, \ldots, Z_n)$ would be given by the working likelihood

$$\prod_i \frac{1}{\sigma_i} \varphi\left(\frac{Z_i - \theta_i}{\sigma_i}\right)$$

- The Likelihood principle: Given a prior over $\theta_1, \ldots, \theta_n$, posterior is same whether we use true likelihood or working one!

Working likelihood and likelihood principle

- This equivalence goes beyond the ASOS example; it applies to, e.g.,
  - Multi-armed experiments: each arm is treated as its own experiment (as long as arm label doesn't contain information about true effect).
  - Panel data with attrition and missing observations (missing at random given past outcomes)
  - Multiple 'p-hacked' experiments

Empirical Bayes methdology

The Empirical Bayes strategy

- Key EB assumption: θ_i are iid draws from some unknown prior G_0
- Aim of EB methods: Estimate G_0, and mimic oracle Bayes performance
- Marginal distribution of data for a given prior G is

$$p_G(\mathcal{D}) = \int p(\mathcal{D}|\theta_1, \ldots, \theta_n) dG^{(n)} = c(\mathcal{D}) \cdot \prod_i f_{G,i}(Z_i)$$

where

$$f_{G,i}(Z_i) = \int \frac{1}{\sigma_i} \varphi \left(\frac{Z_i - \theta_i}{\sigma_i} \right) dG(\theta_i)$$

- Note: true marginal likelihood is proportional to 'working marginal likelihood', but not equal!

g-modeling

- Estimate G_0 by maximizing marginal likelihood over family of priors $\mathcal{G}$:

$$\hat{G} = \text{argmax}_{G \in \mathcal{G}} \ln p_G(\mathcal{D}) = \text{argmax}_{G \in \mathcal{G}} \sum_i \ln f_{G,i}(Z_i)$$

- Examples of candidate classes $\mathcal{G}$:
  - Parametric: $\mathcal{G}$ is some exponential family, e.g., $\mathcal{G} \equiv \{\mathcal{N}(0, \gamma^{-1}) : \gamma > 0\}$
  - Non-parametric: $\mathcal{G}$ is unrestricted leading to NPMLE

- Clearly g-modeling is numerically invariant to how data is generated

- But what 'information' is being used to obtain these estimates?
  - And why is it still statistically valid under adaptive sampling?

The variational interpretation of Bayesian updating

- The Donsker-Varadhan variational formula: for a given G, and fix (Z_i, σ_i)

$$\ln f_{G,i}(Z_i) = \max_{q_i} \left\{ \mathbb{E}_{q_i} \left[\ln \frac{1}{\sigma_i} \varphi \left(\frac{Z_i - \theta_i}{\sigma_i} \right) \right] - \mathrm{KL}\left(q_i \mid\mid G\right) \right\}$$

  - $\mathrm{KL}()$ is KL-divergence and q_i is any distribution over θ_i
  - Optimum occurs at $q^*_{i,G}(\cdot)$, the posterior (given Z_i, σ_i) corresponding to G

- Hence,

$$\max_{G \in \mathcal{G}} \sum_i \ln f_{G,i}(Z_i) = \max_{G \in \mathcal{G}} \max_{\{q_i\}_i} \sum_i \left\{ \mathbb{E}_{q_i} \left[\ln \frac{1}{\sigma_i} \varphi \left(\frac{Z_i - \theta_i}{\sigma_i} \right) \right] - \mathrm{KL}\left(q_i \mid\mid G\right) \right\}$$

  - The two max operations are just the E-step and M-step in EM (see also Neal & Hinton, 1998).

Interpretation of g-modeling

- At optimum $\hat{G}$, EM reaches a fixed point, and we can show

$$\hat{G} = \operatorname{argmin}_{G \in \mathcal{G}} \mathrm{KL}\left(\bar{q}_G \mid\mid G\right), \text{ where } \bar{q}_G := \frac{1}{n} \sum_i q^*_{i,G}$$

  - Note that $\bar{q}_G$ is average posterior

- Consider exponential family $\mathcal{G}$ with sufficient statistic $u(\cdot)$:
  - Minimizing KL divergence is equivalent to matching moments of sufficient statistic:

$$\mathbb{E}_{\hat{G}}[u(\theta)] = \mathbb{E}_{\bar{q}_{\hat{G}}}[u(\theta) | \mathcal{D}]$$

- This is just sample analogue of law of iterated expectations (LIE):

$$\mathbb{E}_{G_0}[u(\theta)] = \mathbb{E}_{G_0}\left[\mathbb{E}_{G_0}\left[u(\theta) | \mathcal{D}\right]\right]$$

Parametric Example

- Say $\mathcal{G}$ is Gaussian family $\mathcal{N}(0, \frac{1}{\gamma})$, with true γ being γ_0
- Posterior $\theta_i|\mathcal{D}_i \sim N\left(\frac{Z_i}{1+\gamma_0\sigma_i^2}, \frac{\sigma_i^2}{1+\gamma_0\sigma_i^2}\right)$
  - Sufficient statistic is $u(\theta) = \theta^2$
  - Moment-matching: estimated γ solves

$$\frac{1}{n}\sum_i m(Z_i, \sigma_i; \hat{\gamma}) = 0, \text{ where}$$

$$m(Z_i, \sigma_i; \gamma) = \left(\frac{Z_i}{1+\gamma\sigma_i^2}\right)^2 + \frac{\sigma_i^2}{1+\gamma\sigma_i^2} - \frac{1}{\gamma}$$

- Note common EB practise (e.g. James-Stein) of estimating γ via $Var_n(Z_i) - E_n(\sigma_i^2)$ is inconsistent for γ^{-1}.

Interpretation of NPMLE

- For NPMLE, moment-matching requires:

$$\hat{G} = \bar{q}_{i,\hat{G}}$$

  - Self Consistency Property: i.e., $\hat{G}$ is prior s.t it is the same as the average posterior
  - Again a consequence of LIE: prior must equal expected posterior
  - And the validity of LIE is algorithm independent.

Tweedie's formula and failure of f-modeling

- Tweedie's formula applies to the working likelihood:

$$\sigma_i^2 \nabla_z \ln f_{G,\sigma_i}(z)\Big|_{z=Z_i} = \mathbb{E}_G[\theta_i|\mathcal{D}] - Z_i$$

- In classical settings $f_{G_0,\sigma_i}(z)$ equals marginal density $p_{\sigma_i}(z)$ of Z_i

- But f-modeling fails in adaptive settings
    - Under adaptive sampling, $f_{G_0,\sigma_i}(z)$ does not equal $p_{\sigma_i}(z)$
    - The true conditional distribution $Z_i|\theta_i$ is not really $\mathcal{N}(\theta_i, \sigma_i^2)$
    - The actual joint density of (Z_i, σ_i) is very complicated and algorithm dependent.

Theoretical properties

Compound Bayes risk and regret

- Aim: provide statistical guarantees for EB estimates of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$
  - Recall EB assumption: $\theta_i \sim_{\text{i.i.d}} G_0$
- Let $\delta_i(\mathcal{D})$ denote some proposed estimator of θ_i and denote $\boldsymbol{\delta}(\mathcal{D}) = (\delta_1(\mathcal{D}), \ldots, \delta_n(\mathcal{D}))$
- Compound Bayes risk under MSE loss:

$$R(\boldsymbol{\delta}, G_0) = \mathbb{E}_{G_0^{(n)}} \left[\frac{1}{n} \sum_i |\delta_i(\mathcal{D}) - \theta_i|^2 \right]$$

- An Oracle who knows G_0 would choose $\delta_i^*(\mathcal{D}) = \mathbb{E}_{G_0}[\theta_i | \mathcal{D}]$

Compound Bayes risk and regret

- Compound Bayes regret: difference in Bayes risk between candidate $\boldsymbol{\delta}$ and oracle

$$\mathcal{R}(\boldsymbol{\delta}, G_0) = R(\boldsymbol{\delta}, G_0) - R(\boldsymbol{\delta}^*, G_0)$$

  - This will be our evaluation criterion

- Straightforward to show

$$\mathcal{R}(\boldsymbol{\delta}, G_0) = \mathbb{E}_{G_0^{(n)}}\left[\frac{1}{n}\sum_i |\delta_i(\mathcal{D}) - \delta_i^*(\mathcal{D})|^2\right]$$

- EB strategy: replace G_0 with $\hat{G}$ to obtain $\hat{\delta}_i^{\mathrm{EB}} = \mathbb{E}_{\hat{G}}[\theta_i|\mathcal{D}]$

- Regret consistency: we say $\boldsymbol{\delta}$ is regret consistent if its regret goes to 0 as $n \to \infty$

Regret consistency under Gaussian priors

- Suppose G_0 lies in Gaussian prior family $\mathcal{G} \equiv \{\mathcal{N}(0, \gamma^{-1}) : \gamma > 0\}$
  - We saw true γ_0 can be estimate using method of moment procedure
- We have a simple proof of regret consistency under leave-one-out-estimation
- Leave-one-out EB:
  - Assume experiments i are independent of each other
  - Estimate γ_0 using data excluding experiment i; term estimate $\hat{\gamma}^{-i}$
  - Compute EB estimate of θ_i using $\hat{\gamma}^{-i}$:

$$\tilde{\delta}_i^{\mathrm{EB}} = \frac{Z_i}{1 + \hat{\gamma}^{-i}\sigma_i^2}$$

- Compare to MLE estimate: $\delta_i^{\mathrm{MLE}} = Z_i$

- A non-asymptotic bound on regret ratio:

$$\frac{\mathcal{R}(\tilde{\boldsymbol{\delta}}^{\mathrm{EB}}, G_0)}{\mathcal{R}(\hat{\boldsymbol{\delta}}^{\mathrm{MLE}}, G_0)} \leq \sup_i \mathbb{E}_{G_0,i} \left[\left(\frac{\hat{\gamma}^{-i} - \gamma_0}{\gamma_0} \right)^2 \right]$$

- Standard GMM arguments: RHS is $O(n^{-1})$
    - Subject to regularity conditions: compact support of $\hat{\gamma}^{-i}$ etc.
    - Denominator in LHS is finite as long as $\mathbb{E}_{G_0}[Z_i^2] < \infty$
    - So $\mathcal{R}(\tilde{\boldsymbol{\delta}}^{\mathrm{EB}}, G_0) = O(n^{-1})$ under mild conditions

- Leave one-out-estimation not really needed, nor is independence of algorithms; more generally,

$$\mathcal{R}(\tilde{\boldsymbol{\delta}}^{\mathrm{EB}}, G_0) \leq \mathbb{E}_{G_0^{(n)}} \left[\left(\frac{1}{n} \sum_i \frac{Z_i^2}{\sigma_i^2} \right) \left(\frac{1}{\hat{\gamma}} - \frac{1}{\gamma_0} \right)^2 \right]$$

    - $\mathcal{R}(\tilde{\boldsymbol{\delta}}^{\mathrm{EB}}, G_0) = O(n^{-1})$ under some stronger regularity conditions

Non-parametric EB

- The non-parametric EB (NPEB) estimator $\hat{\delta}_i^{\mathrm{NPEB}} = \mathbb{E}_{\hat{G}}[\theta_i|\mathcal{D}]$
  - Uses NPMLE estimate $\hat{G}$ in place of G_0
- No parametric requirements on G_0 but regularity conditions more stringent:
  - Experiments are all independent of each other
  - G_0 is compactly supported
  - σ_i's are uniformly bounded away from 0 and ∞
  - There exists $\bar{c} < \infty$ such that for all z_i, σ_i

$$p(z_i, \sigma_i|\theta_i = 0) \leq \frac{\bar{c}}{\sqrt{2\pi}\sigma_i} e^{-z_i^2/2\sigma_i^2}$$

Main result

Theorem: Regret consistency of NPEB

Under the regularity conditions described previously,

$$\mathcal{R}(\hat{\delta}^{\mathrm{NPEB}}, G_0) \lesssim \frac{(\ln n)^5}{n}$$

Remarks:

- Regret rate same as in classical setting (Soloff et al (2023), Chen (2023))
- We pay only a small price $(\ln n)^5$ for estimate G non-parametrically.

Simulations

Data generating process

- Simulate multiple one-armed bandit experiments
  - Algorithms used: Thompson sampling and Upper Confidence Bound algorithm
  - These are commonly used algorithm that balances exploration and exploitation, to maximize expected reward.
- Outcomes drawn from Gaussian $\mathcal{N}(\theta_i, 1)$
- Maximum rounds $N_i \leq 50$.
- Consider two priors over θ_i:
  - Gaussian: $G_0 \equiv \mathcal{N}(0, 1/4)$
  - Two-point prior: $G_0 \equiv \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_3$

Gaussian prior

	Oracle	NPMLE	James-Stein	L-loo	L-posterior	MLE
			Thompson Sampling			
$n = 100$	0.0560	0.0686	0.0646	0.0574	0.0572	0.1744
$n = 500$	0.0561	0.0605	0.0641	0.0564	0.0564	0.1790
$n = 1000$	0.0562	0.0586	0.0637	0.0564	0.0564	0.1770
$n = 5000$	0.0562	0.0570	0.0635	0.0563	0.0563	0.1775
			UCB algorithm			
$n = 100$	0.0605	0.0717	0.0720	0.0623	0.0622	0.1959
$n = 500$	0.0605	0.0647	0.0715	0.0608	0.0608	0.1976
$n = 1000$	0.0605	0.0628	0.0716	0.0606	0.0606	0.1986
$n = 5000$	0.0607	0.0613	0.0716	0.0607	0.0607	0.1983

Notes: Gaussin prior $\theta_i \sim N(0, \frac{1}{4})$. Results are based on 500 simulations.

- L-loo: Leave-one-out Gaussian g-modeling
- L-posterior: Gaussian g-modeling

Two-point prior

	Oracle	NPMLE	James-Stein	L-loo	L-posterior	MLE
			Thompson Sampling			
$n = 100$	0.0000	0.0139	0.1267	0.1227	0.1243	0.2096
$n = 500$	0.0000	0.0029	0.1222	0.1196	0.1199	0.2037
$n = 1000$	0.0000	0.0016	0.1227	0.1202	0.1204	0.2047
$n = 5000$	0.0000	0.0004	0.1228	0.1205	0.1205	0.2047
			UCB algorithm			
$n = 100$	0.0000	0.0104	0.1039	0.1013	0.1023	0.1654
$n = 500$	0.0000	0.0032	0.1049	0.1030	0.1033	0.1683
$n = 1000$	0.0000	0.0018	0.1050	0.1032	0.1033	0.1684
$n = 5000$	0.0000	0.0004	0.1046	0.1029	0.1029	0.1682

Notes: Two point prior: $\theta_i \sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_3$. 500 simulation repetitions.

- L-loo: Leave-one-out Gaussian g-modeling
- L-posterior: Gaussian g-modeling

Performance of Hadad et al

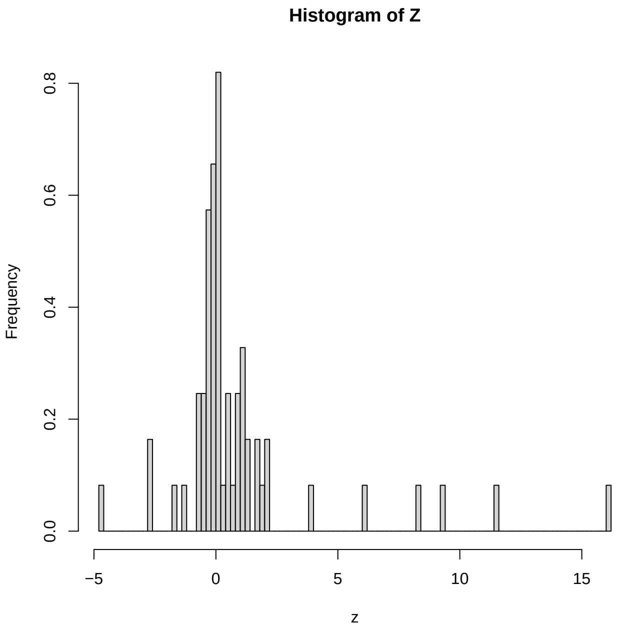
n=	Oracle	NPMLE	J-S	L-pos	MLE	MLE^H	J-S^H	$NPMLE^H$
100	0.0470	0.0553	0.0564	0.0481	0.1566	0.1443	0.0594	0.0833
500	0.0474	0.0498	0.0555	0.0476	0.1547	0.1430	0.0578	0.0764
$1K$	0.0475	0.0488	0.0556	0.0476	0.1557	0.1443	0.0580	0.0758
$5K$	0.0475	0.0479	0.0556	0.0476	0.1554	0.1438	0.0577	0.0743

- Thompson sampling, with lower bound on sampling probability (Hadad et al condition)
- Maximum Round $N = 150$.
- $G_0 = N(0, 1/8)$, $Y|\theta \sim N(\theta, 4)$.
- Takeaway: Hadad et al corrects bias, but sacrifice on variances.

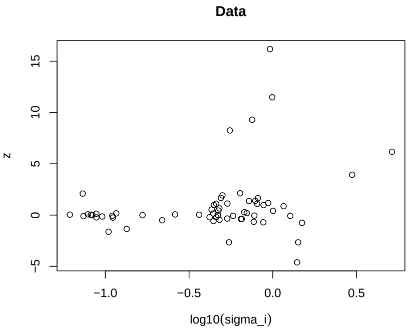
Empirical Illustration

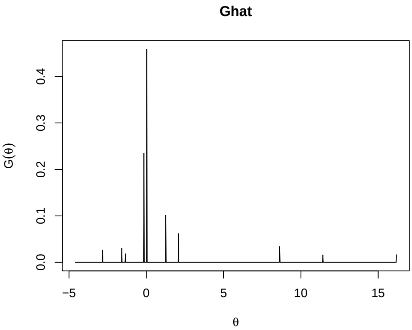
ASOS digital experiments dataset

- Dataset was described earlier:
    - $n = 61$ A/B tests
    - Sampled in equal proportions, but stopping time adaptive and unknown
- Outcomes are actually binary instead of Gaussian
    - Specifically, $Y_{j,i}^{(a)} \sim \mathrm{Bernoulli}(\tilde{\theta}_i^{(a)})$
    - We are interested in scaled treatment effects $\theta_i = \sqrt{N}\left(\tilde{\theta}_i^{(1)} - \tilde{\theta}_i^{(0)}\right)$
- We employ local asymptotics
    - Reasonable since $N \approx 10^5$
    - Take Z_i to be (scaled) difference in sample means

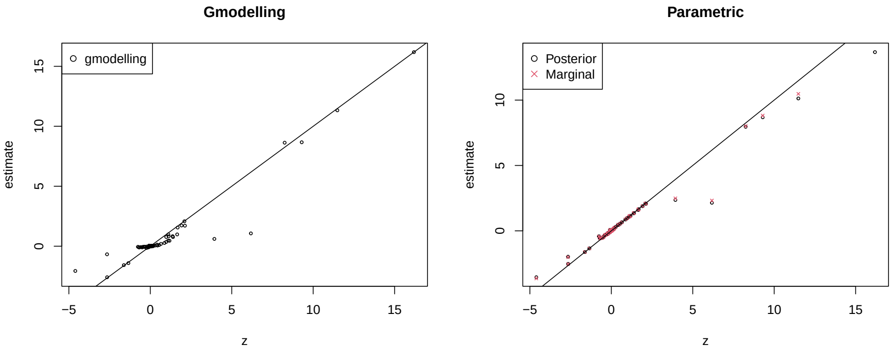

Histogram of Z
Frequency
0.0
0.2
0.4
0.6
0.8
−5
0
5
10
15
z

Estimated prior from NPMLE

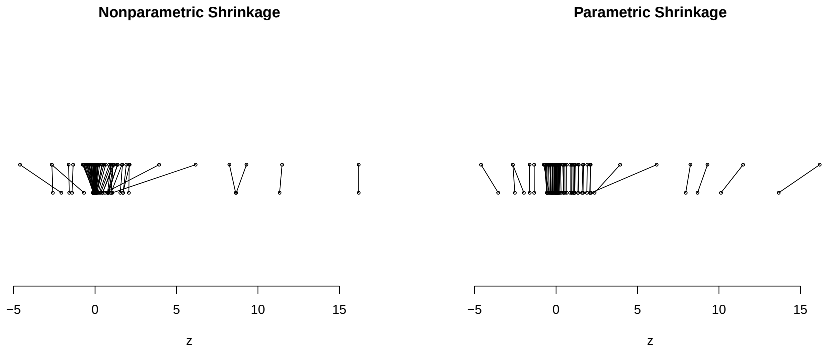




EB estimates



EB estimates



Conclusion

- We extend EB methodology to adaptive experiments
- g-modeling remains valid: simply pretend data was generated exogenously
  - We do not need any knowledge of algorithms used to sample data
- In contrast, f-modeling fails
- Further work and open questions:
  - What is efficient way to estimate prior? (does knowing algorithm help?)
  - How can we use estimated prior to design future experiments?

Thank you!

Why can we learn the prior?

- Let's revisit the very first EB paper: Robbins (1951).
- Consider a random sample Y from $N(\theta, 1)$ with $\theta \in \Theta = \{-1, 1\}$.
- Goal is to estimate θ.
- R.A. Fisher: $\hat{\theta}^{\text{MLE}} = \underset{\theta \in \Theta}{\text{argmax}} \ln \varphi(y - \theta) = sgn(y)$. This is also the minimax estimator.
- Bayesian: Given prior $p = \mathbb{P}(\theta = 1)$, $\hat{\theta}^{\text{Bayes}} = sgn(y - \frac{1}{2} \log \frac{1-p}{p})$.
- MLE, being the minimax estimator, uses the least favorable prior $p = 1/2$.

Why can we learn the prior?

- Now imagine we have in parallel n such experiments, $Y_i \sim N(\theta_i, 1)$ and $\theta_1, \ldots, \theta_n$ are iid draws from a distribution supported on Θ.

- If we see most of the outcome $Y_1, \ldots, Y_n$ to be positive, then more likely $p > 1/2$.

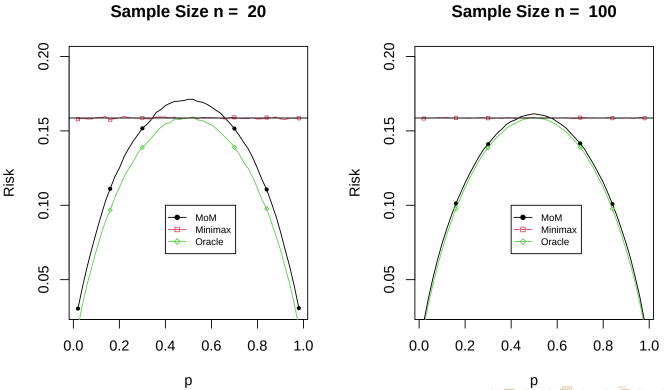
- The n collective experiment provides an opportunity to learn about the properties of the bulk of parameters $\{\theta_1, \ldots, \theta_n\}$.

- Robbins proposed to estimate $\hat{p} = \frac{\bar{Y}+1}{2}$.

Empirical Bayes (EB)

- The EB estimator is $\hat{\theta}_i^{\text{EB}} = sgn(y_i - \frac{1}{2}\log\frac{1-\hat{p}}{\hat{p}})$.
- Performs better, for $p \neq 1/2$, than the MLE when evaluated based on

$$\frac{1}{n}\mathbb{E}\sum_i |\hat{\theta}_i - \theta_i|$$



Proof sketch

- Basic strategy similar to Jiang and Zhang (2009), Zhang (2020), Soloff et al (2023):
- By Tweedie's formula,

$$\mathcal{R}(\hat{\boldsymbol{\delta}}^{\text{NPEB}}, G_0) = \mathbb{E}_{G_0^{(n)}}\left[\frac{1}{n}\sum_i \left(\sigma_i^2 \frac{f'_{\hat{G},i}(Z_i)}{f_{\hat{G},i}(Z_i)} - \sigma_i^2 \frac{f'_{G_0,i}(Z_i)}{f_{G_0,i}(Z_i)}\right)^2\right]$$

  - Recall that $f_{G_0,i}$ and $f_{\hat{G},i}$ are marginals under working likelihood
- Step 1: Show that $f_{G_0,i}$ and $f_{\hat{G},i}$ are close to each other in some 'average' Hellinger-distance sense
- Step 2: Convert Hellinger distance bound to bound between $\nabla_z \ln f_{\hat{G},i}$ and $\nabla_z \ln f_{G_0,i}$
- Novelty vis-a-vis classical setting: true marginal of Z_i is not $f_{G_0,i}(Z_i)$

Group Sequential Clinical Trials

- We have fixed number of patients N, split equally into K rounds.

$$\underbrace{Y_{11}, \dots, Y_{1n_1}}_{n_1 \text{ observations}}, \underbrace{Y_{21}, \dots, Y_{2n_2}}_{n_2 \text{ observations}}, \dots, \underbrace{Y_{K1}, \dots, Y_{Kn_K}}_{n_K \text{ observations}}$$

  Draws from $N(\mu, \sigma^2)$.

- Cumulative mean up to round k

$$\bar{Y}^{(k)} = \frac{1}{\sum_{\tilde{k}=1}^{k} n_{\tilde{k}}} \sum_{\tilde{k}=1}^{k} n_{\tilde{k}} \bar{Y}_{\tilde{k}}$$

  with $\bar{Y}_k$ the mean at round k.

- At each round, we test using

$$Z_k^* = \frac{\bar{X}^{(k)} - \mu_0}{\sigma} \sqrt{\sum_{\tilde{k}=1}^{k} n_{\tilde{k}}}$$

- Stop at round k if $|Z_k^*| \geq c_k$.

One arm Thompson Sampling

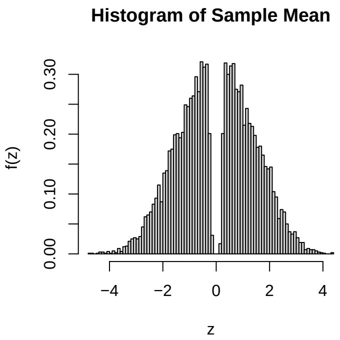
- Let cumulative rewards be X.
- Arm distribution $N(\theta, \sigma^2)$.
- Specify a prior $\theta \sim N(\mu, \sigma^2) = N(0, 1/\tau)$
- Let q_j be the number of samples collected up to round j.
- In round j, update

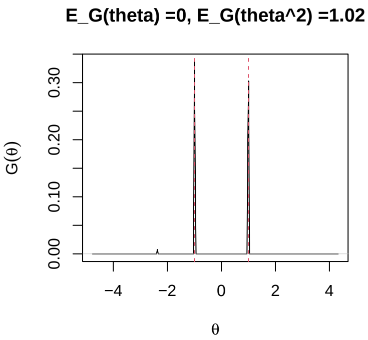
$$\mu_j^{post} = \frac{x_{j-1}}{q_{j-1} + \tau}$$
$$se_j^{post} = \frac{\sigma}{q_{j-1} + \tau}$$

- Calculate sampling probabiity $\pi_j = P(\theta > 0 | \mathcal{D}) = \Phi(\frac{\mu^{post}}{se^{post}})$.
- Sample decision $a_j \sim Bernounlli(\pi_j)$.
- Sample if $a_j = 1$ and see reward.
- $q_j = q_{j-1} + a_j$.

g-modelling: Non-informative stopping rule

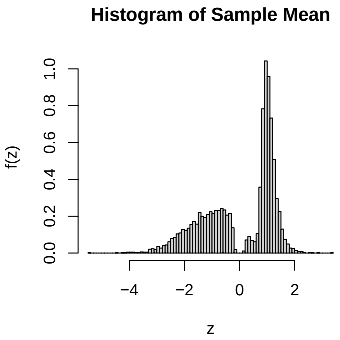
- $\theta_i \in \{-1, 1\}$ with equal probability: $G_0 = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$. $(E_{G_0}(\theta) = 0, E_{G_0}(\theta^2) = 1)$.
- Sampling distribution: $Y_{j,i} \sim N(\theta_i, 1)$.
- Noninformative stopping rule: stop when $|\frac{1}{N}\sum_{j=1}^{N} Y_{j,i}| \geq k/\sqrt{N}$, $k = 0.25$.

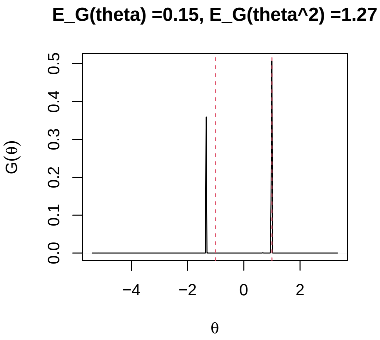




g-modelling: Informative stopping rule

- $\theta_i \in \{-1, 1\}$ with equal probability: $G_0 = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$. $(E_{G_0}(\theta) = 0, E_{G_0}(\theta^2) = 1)$.
- Sampling distribution: $Y_{j,i} \sim N(\theta_i, 1)$.
- Informative stopping rule: stop when $|\frac{1}{N}\sum_{j=1}^{N} Y_{j,i}| \geq k_i/\sqrt{N}$.
- $k_i = \begin{cases} 0.25 & \theta_i = -1 \\ 5 & \theta_i = 1 \end{cases}$





g-modelling: Heterogenous non-informative stopping rule

- $\theta_i \in \{-1, 1\}$ with equal probability: $G_0 = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$. $(E_{G_0}(\theta) = 0, E_{G_0}(\theta^2) = 1)$.
- Sampling distribution: $Y_{j,i} \sim N(\theta_i, 1)$.
- Heterogeneous non-informative stopping rule: stop when $|\frac{1}{N}\sum_{j=1}^{N} Y_{j,i}| \geq k_i/\sqrt{N}$.
- $k_i \sim Unif[0.25, 5]$.