STATISTICAL DECISION THEORY FOR TREATMENT CHOICE

Charles F. Manski

Department of Economics and Institute for Policy Research

Northwestern University

Background

An important objective of empirical research is to analyze treatment response (policy impacts) in order to inform treatment choice.

Identification problems combine with the necessity of inference from sample data to limit the informativeness of studies.

Researchers commonly use methods of statistical inference whose foundations are distant from the decision problem.

It is common to use local asymptotic theory to motivate inferential methods, even though the available data often derive from samples too small to make asymptotic approximations credible.

It is common to perform hypothesis tests and to judge estimates by statistical significance, even though these concepts are remote from decision making.

The Wald (*Statistical Decision Functions*, 1950) development of statistical decision theory provides a coherent framework for the use of sample data to make decisions.

Wald posed the task as choice of a *statistical decision function*, which maps potentially available data into a choice among the feasible actions.

He recommended ex ante evaluation of statistical decision functions as procedures, chosen prior to realization of the data, specifying how a decision maker would use whatever data may be realized.

Thus, the theory is frequentist.

Expressing the objective as minimization of a loss function, Wald proposed that the decision maker evaluate a statistical decision function by its mean performance across realizations of the sampling process (*risk*).

In the presence of uncertainty about the loss function and the sampling process yielding the data, he prescribed a three-step decision process.

1. Specify the state space (parameter space), which indexes the loss functions and sampling distributions that the decision maker deems possible.

2. Eliminate inadmissible statistical decision functions.

A decision function is inadmissible (weakly dominated) if there exists another that yields at least as good mean sampling performance in every state of nature and strictly better mean performance in some state.

3. Use some criterion to choose an admissible statistical decision function.

Wald focused on the *minimax* criterion and on minimization of a subjective mean of the risk function (*Bayes risk*). Savage proposed *minimax regret*.

Statistical decision theory makes no direct reference to statistical inference.

However, there exists a prominent finding connecting minimization of Bayes risk to Bayesian inference.

Minimization of Bayes risk yields the same outcome as would occur if one

1. performs Bayesian inference, combining the prior distribution on the state space with the data to form a posterior subjective distribution.
2. chooses an action to minimize the posterior mean of expected loss.

This holds as a consequence of Fubini's Theorem.

There is no optimal way to choose an admissible decision function.

There at most are reasonable ways.

Wald (1950) motivated his focus on the minimax criterion in part by stating:

> "a minimax solution seems, in general, to be a reasonable solution of the decision problem when an a priori distribution in Ω does not exist or is unknown to the experimenter."

Ferguson (*Mathematical Statistics*,1967) wrote: "A *reasonable* rule is one that is better than just guessing." He continued:

> "It is a natural reaction to search for a 'best' decision rule, a rule that has the smallest risk no matter what the true state of nature. Unfortunately, *situations in which a best decision rule exists are rare and uninteresting*. For each fixed state of nature there may be a best action for the statistician to take. However, this best action will differ, in general, for different states of nature, so that no one action can be presumed best over all."

Savage (*JASA*, 1951) argued against the minimax criterion, writing:

> "Application of the minimax rule . . . is indeed ultra-pessimistic; no serious justification for it has ever been suggested, and it can lead to the absurd conclusion in some cases that no amount of relevant experimentation should deter the actor from behaving as though he were in complete ignorance."

He emphasized that minimax regret is not "ultra-pessimistic."

Chernoff (*ECMA*, 1954) argued against minimax regret on the basis that it sometimes violates the choice axiom called *independence of irrelevant alternatives*.

Savage (*Foundations of Statistics*,1954) developed choice axioms that imply maximization of subjective expected utility and argued that adherence to his axioms constitutes *rationality*. Yet Allais (*ECMA*, 1953) and Ellsberg (*QJE*, 1961) questioned the positive accuracy and normative appeal of the Savage axioms.

The Wald framework has breathtaking generality.

In principle, it enables comparison of all statistical decision functions whose risk functions exist. (Moreover, it may be extended beyond this class.)

It enables comparison of alternative sampling processes as well as decision rules.

It uses no asymptotic approximations.

It applies whatever information the decision maker may have.

> The state space may be finite dimensional or larger (nonparametric).
>
> The true state of nature may be point or partially identified.

After publication of Wald (1950), a surge of important extensions and applications followed in the 1950s.

Much research focused on best point prediction with sample data, analysis of which began with Hodges and Lehmann (*AMS*, 1950).

However, this period of rapid development came to a close by the 1960s with the exception of Bayesian statistical decision theory.

Conditional Bayesian analysis has continued to developed as a self-contained field of study, disconnected from the Wald frequentist framework.

Why did statistical decision theory lose momentum long ago?

One reason may have been the technical difficulty of the subject.

Wald's ideas are easy to describe in the abstract, but applying them can be analytically and computationally demanding. Determination of admissible decision functions and minimax/minimax-regret rules is often difficult.

Another reason may have been diminishing interest in decision making as the motivation for analysis of sample data.

Modern statisticians and econometricians tend to view their objectives as estimation and hypothesis testing rather than decision making.

I cannot be sure what role these or other reasons played in the vanishing of statistical decision theory from statistics and econometrics in the latter part of the twentieth century.

However, the near absence of the subject in mainstream journals and textbooks of the period is indisputable.

I think this unfortunate.

Application of Statistical Decision Theory to Treatment Choice

My interest in treatment choice with sample data grew out of my research in the 1990s on partial identification of treatment response.

Consider choice between treatments a and b. The sampling process and maintained assumptions may not identify the sign of the average treatment effect.

Then optimal treatment choice is infeasible even with exact knowledge of all observable features of the population.

The decision problem is further complicated if one only observes sample data.

Finding the Wald theory appealing, I suggested in Manski (*ECMA*, 2004) that it be used to study how a planner wanting to maximize social welfare might use sample data on treatment response to choose treatments for a heterogeneous population.

In this setting, a statistical decision function uses the data to choose a treatment allocation.

I called such a function a *statistical treatment rule* (*STR*).

I called the mean sampling performance of an STR its *expected welfare.*

I used the minimax-regret criterion to evaluate STRs in Manski (*ECMA*, 2004; *Social Choice with Partial knowledge of Treatment Response*, 2005; *JoE*, 2007; and *Identification for Prediction and Decision*, 2007).

Manski and Tetenov (*JSPI*, 2007) studied admissibility in a particular setting, building on a classical result of Karlin and Rubin (*AMS*, 1956).

Schlag (2006), Hirano and Porter (*ECMA*, 2009), Stoye (*JoE*, 2009; *JoE*, 2012), and Tetenov (*JoE*, 2012) continued study of the minimax regret and related criteria.

Manski and Tetenov (2014, 2015) study quantile performance of STRs and near-optimality of experimental designs. Kitagawa and Tetenov (2016) propose empirical welfare maximization.

A Leading Case: Allocating a Population of Observationally Identical Persons to Two Treatments

A planner must assign treatment a or b to each member of population J.

$j \in J$ has response function $u_j(\cdot): T \rightarrow Y$ mapping treatments $t \in T$ into individual welfare outcomes $u_j(t) \in R$.

The population is a probability space (J, Ω, P). $P(j) = 0, j \in J$.

The members of the population are observationally identical to the planner.

The planner can randomly allocate persons to the two treatments with specified allocation probabilities $\delta(a)$ and $\delta(b)$ such that $\delta(a) + \delta(b) = 1$.

The planner wants to choose δ to maximize an additive welfare function

$$U(\delta, P) = E[u(a)]\cdot\delta(a) + E[u(b)]\cdot\delta(b) = \alpha\cdot\delta(a) + \beta\cdot\delta(b).$$

It is optimal to set $\delta(a) = 1$ if $\alpha \geq \beta$ and $\delta(b) = 1$ if $\alpha \leq \beta$.

The problem of interest is treatment choice when incomplete knowledge of P makes it impossible to determine the ordering of α and β.

Suppose that sample data are available.

Let Q be the sampling distribution and Ψ be the sample space. For example, the data may be treatment response observed in a randomized experiment.

A statistical treatment rule (STR) is a function δ: $T \times \Psi \rightarrow [0, 1]$ that maps sample data into a treatment allocation. Thus, $\delta(a, \psi) + \delta(b, \psi) = 1, \forall \psi \in \Psi$.

Henceforth, let $\delta(\psi) = \delta(b, \psi)$ and $1 - \delta(\psi) = \delta(a, \psi)$.

The welfare realized with δ and data ψ is the random variable

$$U(\delta, P, \psi) = \alpha \cdot [1 - \delta(\psi)] + \beta \cdot \delta(\psi).$$

The state space $[(P_s, Q_s), s \in S]$ is the set of pairs that the planner deems possible.

Suppose S contains at least one state such that $\alpha_s > \beta_s$ and another such that $\alpha_s < \beta_s$.

The mean sampling performance of rule δ in state s is

$$W(\delta, P_s, Q_s) = \alpha_s \cdot \{1 - E_s[\delta(\psi)]\} + \beta_s \cdot E_s[\delta(\psi)].$$

$E_s[\delta(\psi)] \equiv \int_\Psi \delta(\psi)dQ_s(\psi)$ is the mean fraction of persons allocated to b, across potential samples.

Rule δ is admissible if there exists no rule δ' such that $W(\delta', P_s, Q_s) \geq W(\delta, P_s, Q_s)$ for all $s \in S$ and $W(\delta', P_s, Q_s) > W(\delta, P_s, Q_s)$ for some s.

Let Δ_a be the admissible STRs. Three decision criteria are

Maximin (MM): $\max_{\delta \in \Delta_a} \min_{s \in S} W(\delta, P_s, Q_s)$.

Minimax-regret (MMR): $\min_{\delta \in \Delta_a} \max_{s \in S} [\max(\alpha_s, \beta_s) - W(\delta, P_s, Q_s)]$.

Minimization of Bayes Risk (Bayes): $\max_{\delta \in \Delta_a} \int_S W(\delta, P_s, Q_s)d\pi(s)$,

where π is a subjective distribution on the state space.

Optimization over Δ

The set Δ_a is knowable in principle, but it may be difficult to determine in practice.

Lacking knowledge of Δ_a, it is common to modify the criteria to choose over the full set Δ of feasible rules.

If the optimal (MM, MMR, Bayes) rule is unique, the optimal rule is admissible. No other rule dominates a unique optimal rule.

When optimization over Δ yields multiple optimal rules, some may be inadmissible.

The MM and MMR criteria tolerate inadmissibility when the states where a rule is inferior are not those that yield worst performance across the state space.

Bayes criteria tolerate it when the states where a rule is inferior collectively have subjective probability zero.

Using a Randomized Experiment to Evaluate an Innovation

(Source: Manski, *Identification for Prediction and Decision*, 2007, Ch. 12)

To illustrate, consider a simple case of treatment choice with sample data.

$t = a$ is the status quo treatment and $t = b$ is an innovation. Outcomes are binary.

The planner knows the success probability $\alpha \equiv P[y(a) = 1]$ of the status quo treatment, but not the success probability $\beta \equiv P[y(b) = 1]$ of the innovation.

The planner wants to choose a treatment to maximize the success probability.

An experiment is performed to learn about outcomes under the innovation, with N subjects randomly drawn from the population and assigned to treatment b. There is full compliance.

n subjects realize outcome $y = 1$ and $N - n$ realize $y = 0$. Outcomes are observed. Thus, the experiment point-identifies β. The planner only faces a problem of statistical inference.

N indexes the sampling process and the number n of experimental successes is a sufficient statistic for the sample data.

The feasible statistical treatment rules are functions $\delta(\cdot)$: $[0, \ldots, N] \rightarrow [0, 1]$ that map the number of experimental successes into a treatment allocation.

The expected welfare of rule δ is

$$W(\delta, P, N) = \alpha \cdot E[1 - \delta(n)] + \beta \cdot E[\delta(n)] = \alpha + (\beta - \alpha) \cdot E[\delta(n)].$$

The number n of experimental successes is distributed binomial $\mathbf{B}[\beta, N]$.

$$E[\delta(n)] = \sum_{i=0}^{N} \delta(i) \cdot f(n = i; \beta, N).$$

$f(n = i; \beta, N) \equiv N![i! \cdot (N - i)!]^{-1} \beta^i (1 - \beta)^{N-i}$ is the probability of i successes.

The only unknown determinant of expected welfare is β. Hence, S indexes the feasible values of β. $\beta_s \equiv P_s[y(b) = 1]$.

The Admissible Treatment Rules

It is reasonable to conjecture that admissible treatment rules should be ones in which the fraction of the population allocated to treatment b increases with n.

The admissible treatment rules are a simple subclass of these rules.

A theorem of Karlin and Rubin (*AMS*, 1956) shows that the admissible rules are the *monotone treatment rules*.

Monotone rules assign all persons to the status quo if the experimental success rate is below some threshold and all to the innovation if the success rate is above the threshold. Thus, δ is admissible if and only if

$$\delta(n) = 0 \quad \text{for } n < n_0,$$
$$\delta(n) = \lambda \quad \text{for } n = n_0,$$
$$\delta(n) = 1 \quad \text{for } n > n_0,$$

for some $0 \leq n_0 \leq N$ and $0 \leq \lambda \leq 1$.

Monotone rules have simple expressions for the expected treatment allocation:

$$E_s[\delta(n)] = f(n > n_0;\, \beta_s, N) + \lambda \cdot f(n = n_0;\, \beta_s, N).$$

The collection of monotone treatment rules is a mathematically "small" subset of the space of all feasible treatment rules.

Nevertheless, it still contains a broad range of rules. These include

Data-Invariant Rules: These are the rules $\delta(\cdot) = 0$ and $\delta(\cdot) = 1$, which assign all persons to treatment a or b respectively, whatever n may be.

Empirical Success Rules: An optimal treatment rule allocates all persons to treatment a if $\beta < \alpha$ and all to b if $\beta > \alpha$. An empirical success rule emulates the optimal rule by replacing β with its sample analog, the empirical success rate n/N.

Bayes Rules: The form of the Bayes rule depends on the prior subjective distribution placed on β. Consider the class of Beta priors, which form the conjugate family for a Binomial likelihood. Let $(\beta_s, s \in S) = (0, 1)$ and let the prior be Beta with parameters (c, d). Then the posterior mean for β is $(c + n)/(c + d + N)$. The resulting Bayes rule is

$$\delta(n) = 0 \quad \text{for } (c + n)/(c + d + N) < \alpha,$$

$$\delta(n) = \lambda \quad \text{for } (c + n)/(c + d + N) = \alpha, \quad \text{where } 0 \le \lambda \le 1,$$

$$\delta(n) = 1 \quad \text{for } (c + n)/(c + d + N) > \alpha.$$

As (c, d) tend to zero, the Bayes rule approaches an empirical success rule. The class of Bayes rules includes the data-invariant rules $\delta(\cdot) = 0$ and $\delta(b, \cdot) = 1$.

Statistical Significance Rules: These use a one-sided hypothesis test to choose between the status quo and the innovation. The null hypothesis is $[\beta = \alpha]$. The alternative is $[\beta > \alpha]$. Treatment b is chosen if the null is rejected, and treatment a otherwise. Thus, the rule is

$$\delta(n) = 0 \quad \text{for} \quad n \leq d(\alpha, \theta, N),$$

$$\delta(n) = 1 \quad \text{for} \quad n > d(\alpha, \theta, N),$$

θ is the specified size of the test and $d(\alpha, \theta, N)$ is the associated critical value.

The use of one-sided hypothesis tests to make treatment choices is institutionalized in the U. S. Food and Drug Administration drug approval process, which calls for comparison of a new drug under study (t = b) with an approved treatment if one exists, or a placebo otherwise (t = a).

Although statistical significance rules are monotone treatment rules, the conventional practice of hypothesis testing is remote from the problem of treatment choice with sample data.

Maximin Rule: Minimum expected welfare for rule δ is

$$\min_{s \in S} W(\delta, P_s, N) = \alpha + \min_{s \in S} (\beta_s - \alpha)E_s[\delta(n)].$$

$E_s[\delta(n)] > 0$ for all $\beta_s > 0$ and for all monotone rules except $\delta(\cdot) = 0$.

Let S contain states with $\beta_s < \alpha$. Then the maximin rule is $\delta(\cdot) = 0$.

Minimax-Regret Rule: The regret of rule δ in state s is

$$U^*(P_s) - W(\delta, P_s, N) = \max(\alpha, \beta_s) - \{\alpha + (\beta_s - \alpha)\cdot E_s[\delta(n)]\}$$
$$= (\beta_s - \alpha)\{1 - E_s[\delta(n)]\}\cdot 1[\beta_s \geq \alpha] + (\alpha - \beta_s)E_s[\delta(n)]\cdot 1[\alpha \geq \beta_s].$$

Thus, regret is the mean welfare loss when a member of the population is assigned the inferior treatment, multiplied by the expected fraction of the population assigned this treatment.

The minimax-regret rule does not have an analytical solution but it can be determined numerically. When all values of β are feasible, the minimax-regret rule is well approximated by an empirical success rule.

Minimax-Regret Treatment Choice

Setup

There is a finite set T of treatments.

Person j in population J has response function $y_j(\cdot)$: $T \rightarrow Y$ mapping $t \in T$ into outcomes $y_j(t) \in Y$. The population is a "large" probability space (J, Ω, P).

A planner observes covariates $x_j \in X$ for each person. X is finite.

The distribution $P[y(\cdot), x]$ describes treatment response across the population. $P(x)$ is known but not $P[y(\cdot)|x]$.

Let Δ denote all functions $\delta(\cdot, \cdot)$: $T \times X \to [0, 1]$ such that $\sum_{t \in T} \delta(t, \xi) = 1, \forall\, \xi \in X$.

In the absence of data, the feasible treatment rules are the elements of Δ.

The welfare from assigning treatment t to person j is $u[y_j(t), t, x_j]$. Mean welfare under rule δ is

$$U(\delta, P) \quad \equiv \quad \sum_{\xi \in X} P(x = \xi) \sum_{t \in T} \delta(t, \xi) \cdot E\{u[y(t), t, \xi] \,|\, x = \xi\}.$$

The planner wants to solve the problem $\max_{\delta \in \Delta} U(\delta, P)$.

The maximum is achieved by allocating all persons with covariates ξ to a treatment that solves

$$\max_{t \in T} E\{u[y(t), t, \xi] \mid x = \xi\}.$$

The population mean welfare achieved by an optimal rule is

$$U^*(P) \equiv \sum_{\xi \in X} P(x = \xi) \{\max_{t \in T} E\{u[y(t), t, \xi] \mid x = \xi\}\}.$$

The optimum is achievable if the planner knows $E\{u[y(t), t, \xi] \mid x = \xi\}$, all (t, ξ). It may not be achievable with incomplete knowledge.

Statistical treatment rules map covariates and data into treatment allocations.

Let Q denote the sampling process generating the available data and let Ψ denote the sample space.

Let Δ denote the space of functions that map $T \times X \times \Psi$ into the unit interval and that satisfy the adding-up conditions:

$$\delta \in \Delta \;\Rightarrow\; \sum_{t \in T} \delta(t, \xi, \psi) = 1, \;\; \forall\, (\xi, \psi) \in X \times \Psi.$$

Each function $\delta \in \Delta$ defines a statistical treatment rule.

The expected welfare yielded by rule δ in repeated samples is

$$W(\delta, P, Q) \equiv \int \{ \sum_{\xi \in X} P(x = \xi) \sum_{t \in T} \delta(t, \xi, \psi) \cdot E\{u[y(t), t, \xi] \mid x = \xi\} dQ(\psi)$$

$$= \sum_{\xi \in X} P(x = \xi) \sum_{t \in T} E[\delta(t, \xi, \psi)] \cdot E\{u[y(t), t, \xi] \mid x = \xi\}.$$

$E[z(t, \xi, \psi)] \equiv \int \delta(t, \xi, \psi) dQ(\psi)$ is the mean (across potential samples) fraction of persons with covariates ξ who are assigned to treatment t.

The state space gives $[(P_s, Q_s), s \in S]$. The minimax regret criterion is

$$\min_{\delta \in \Delta} \max_{s \in S} U^*(P_s) - W(\delta, P_s, Q_s),$$

where $U^*(P_s)$ is the optimal population mean welfare achievable if $P = P_s$.

Why Focus on Minimax Regret?

* Interpretable as maximization of uniform closeness to optimal treatment choice.

* Applicable in absence of subjective distribution on state space.

* Balanced consideration of states, in contrast to conservatism of maximin.

* Yields "Reasonable" decision rules in many settings.

Arguments Against Minimax Regret

* Sometimes violates axiom of menu independence (Chernoff, *ECMA*, 1954).

* Yields a "No Data" decision rule with continuous covariates and no cross-covariate restrictions on the state space (Stoye, *JoE*, 2009).

Approaches to Analysis of Maximum Regret

* Exact analytical evaluation feasible only in some very simple cases.

* Analytical evaluation of bounds feasible in some cases.

* Numerical computation feasible in some cases.

* Application of a game theoretic approach feasible in some cases.

Bounds on the Maximum Regret of Empirical Success Rules with Data from a Randomized Experiment

(Source, Manski, *ECMA*, 2004)

Let T = {a, b}. The outcome has bounded range [0, 1]. Welfare is $y_j(t)$.

A randomized experiment is performed and all outcomes are observed. I consider stratified random sampling and simple random sampling designs.

Stratified Random Sampling

For $(t, \xi) \in T \times X$, the experimenter draws $N_{t\xi}$ subjects at random from the population and assigns them to treatment t. $N_{TX} \equiv [N_{t\xi}, (t, \xi) \in T \times X]$ indexes the sampling process. $N(t, \xi)$ is the sub-sample assigned to t.

The data are the experimental outcomes $Y_{TX} \equiv [y_j, j \in N(t, \xi); (t, \xi) \in T \times X]$.

Rule δ allocates fraction $\delta(b, \xi, Y_{TX})$ of persons with covariates ξ to treatment b and $\delta(a, \xi, Y_{TX})$ to treatment a. Its expected welfare is

$$W(\delta, P, N_{TX}) =$$

$$\sum_{\xi \in X} P(x = \xi)\cdot\{E[\delta(a, \xi, Y_{TX})]\cdot E[y(a)|x = \xi] + E[\delta(b, \xi, Y_{TX})]\cdot E[y(b)|x = \xi]\}.$$

Conditional Empirical Success (CES) Rules

The optimal rule solves $\max_{t \in T} E[y(t)|x = \xi]$, $\xi \in X$.

An empirical success rule replaces $E[y(t)|x = \xi]$ by $\bar{y}_{t\xi} \equiv (1/N_{t\xi}) \sum_{j \in N(t, \xi)} y_j$.

Given $v(\cdot): X \rightarrow V$, let δ_{Xv} denote a CES rule conditioning on v. The sample analog of $E[y(t)|v = \nu]$ is $\bar{y}_{tX\nu} \equiv \sum_{\xi \in X} \bar{y}_{t\xi} \cdot P(x = \xi | v = \nu)$.

The expected welfare of δ_{Xv} with tie-breaking in favor of treatment a is

$$W(\delta_{Xv}, P, N_{TX}) = \sum_{\nu \in V} P(v = \nu) \cdot \{P(\bar{y}_{aX\nu} \geq \bar{y}_{bX\nu}) \cdot E[y(a)|v = \nu] + P(\bar{y}_{bX\nu} > \bar{y}_{aX\nu}) \cdot E[y(b)|v = \nu]\}.$$

Direct analysis of CES rules is difficult, but it is possible to develop a useful closed-form bound on expected welfare. The analysis exploits a large deviations theorem of Hoeffding (*JASA*, 1963):

Large Deviations Theorem (Hoeffding, 1963, Theorem 2): Let $w_1, w_2, \ldots, w_K$ be independent real random variables, with bounds $a_k \leq w_k \leq b_k$, $k = 1, 2, \ldots, K$. Let $\bar{w} \equiv (1/K) \sum_{k=1}^{K} w_k$ and $\mu \equiv E(\bar{w})$. Then, for $d > 0$,

$$\Pr(\bar{w} - \mu \geq d) \leq \exp[-2K^2 d^2 \{\sum_{k=1}^{K} (b_k - a_k)^2\}^{-1}]. \qquad \square$$

Proposition 1: Let subjects be drawn by stratified random sampling. Let $v(\cdot): X \rightarrow V$ and consider the CES rule δ_{Xv}.

Let $M_{s\nu} \equiv \max\{E_s[y(a)|v = \nu], E_s[y(b)|v = \nu]\}$ and $\lambda_{s\nu} \equiv |E_s[y(b)|v = \nu] - E_s[y(a)|v = \nu]|$.

Then the expected welfare of rule δ_{Xv} in state s satisfies the inequality

$$\sum_{\nu \in V} P(v = \nu)\, M_{s\nu} - D(\delta_{Xv}, P_s, N_{TX}) \leq W(\delta_{Xv}, P_s, N_{TX}) \leq \sum_{\nu \in V} P(v = \nu)\, M_{s\nu},$$

where

$$D(\delta_{Xv}, P_s, N_{TX}) \equiv \sum_{\nu \in V} P(v = \nu)\cdot\lambda_{s\nu}\cdot\exp[-2\lambda_{s\nu}{}^2\{\sum_{\xi \in X} P(x = \xi|v = \nu)^2(N_{b\xi}{}^{-1} + N_{a\xi}{}^{-1})\}^{-1}]. \;\square$$

The upper bound is the maximum population welfare achievable with v.

The lower bound differs from this ideal by the non-negative *finite-sample penalty* $D(z_{Xv}, P_s, N_{TX})$, which places an upper bound on the loss in welfare that results from estimating mean treatment outcomes rather than knowing them.

The magnitude of the finite-sample penalty decreases with sample size and converges to zero at exponential rate if all elements of N_{TX} grow at the same rate.

Bounding Maximum Regret

The bound on expected welfare in Proposition 1 gives a bound on regret:

$$\sum_{\xi \in X} P(x = \xi)\, M_{s\xi} \;-\; \sum_{\nu \in V} P(v = \nu)\, M_{s\nu} \;\le\; U^*(P_s) - W(z_{Xv}, P_s, N_{TX})$$

$$\le\; \sum_{\xi \in X} P(x = \xi)\, M_{s\xi} \;-\; \sum_{\nu \in V} P(v = \nu)\, M_{s\nu} \;+\; D(z_{Xv}, P_s, N_{TX}).$$

Maximizing this bound over the feasible values of $\{E_s[y(a)|x], E_s[y(b)|x]\}$ yields a bound on maximum regret: $R_{Lv} \le R(z_{Xv}) \le R_U(z_{Xv})$,

where

$$R_{Lv} \;\equiv\; \max_{s \in S} \;\sum_{\xi \in X} P(x = \xi) \max\{E_s[y(a)|x = \xi], E_s[y(b)|x = \xi]\}$$

$$- \sum_{\nu \in V} P(v = \nu) \max\{E_s[y(a) | v = \nu], E_s[y(b) | v = \nu]\},$$

$$R_U(\delta_{Xv}) \equiv \max_{s \in S} \sum_{\xi \in X} P(x = \xi) \max\{E_s[y(a) | x = \xi], E_s[y(b) | x = \xi]\}$$

$$- \sum_{\nu \in V} P(v = \nu) \max\{E_s[y(a) | v = \nu], E_s[y(b) | v = \nu]\}$$

$$+ \sum_{\nu \in V} P(v = \nu)\cdot\lambda_{s\nu}\cdot\exp[-2\lambda_{s\nu}{}^2\{ \sum_{\xi \in X} P(x = \xi | v = \nu)^2(N_{b\xi}{}^{-1} + N_{a\xi}{}^{-1})\}^{-1}],$$

$$\lambda_{s\nu} \equiv | E_s[y(b) | v = \nu] - E_s[y(a) | v = \nu] |,$$

$$E_s[y(t) | v = \nu] = \sum_{\xi \in X} E_s[y(t) | x = \xi]\cdot P(x = \xi | v = \nu).$$

The CES Rule Conditioning on All Observed Covariates

The bound on maximum regret simplifies for rule δ_{Xx}. Then

$$0 \leq R(\delta_{Xx}) \leq \max_{s \in S} \sum_{\xi \in X} P(x = \xi)\cdot\lambda_{s\xi}\cdot\exp[-2\lambda_{s\xi}^{2}(N_{b\xi}^{-1} + N_{a\xi}^{-1})^{-1}].$$

If all distributions of treatment response are feasible, this reduces to

$$0 \leq R(\delta_{Xx}) \leq \tfrac{1}{2} e^{-\frac{1}{2}} \sum_{\xi \in X} P(x = \xi)\,(N_{b\xi}^{-1} + N_{a\xi}^{-1})^{\frac{1}{2}}.$$

If all elements of N_{TX} grow at the same rate, maximum regret converges to zero with the square root of sample size.

Sufficient Sample Sizes for Beneficial Use of Covariate Information

The maximum regret of any rule conditioning on v exceeds that of CES rule z_{Xx} if $R_{Lv} > R_U(z_{Xx})$. That is, conditioning on x is preferable to conditioning on v if

$$R_{Lv} \;>\; \tfrac{1}{2}\, e^{-1/2} \sum_{\xi \in X} P(x = \xi)\, (N_{b\xi}^{-1} + N_{a\xi}^{-1})^{1/2} \,.$$

- When this inequality holds, the maximum regret of δ_{Xx} is smaller than that of *all* rules conditioning treatment choice on v.

- The inequality suffices for superiority of δ_{Xx} to rules that condition on v, but it is not a necessary condition.

- δ_{Xx} need not be the best rule conditioning on x. There may exist a non-CES rule superior to δ_{Xx}.

Bounding Expected Welfare and Regret in Designs with Simple Random Sampling

The experimenter draws N subjects at random from the population and randomly assigns them to treatments a and b with assignment probabilities q and 1 − q. The pair (N, q) indexes the sampling process. The sample data are the stratum sample sizes N_{TX} and the outcomes Y_{TX}. Rule δ allocates fraction $\delta(a, x, N_{TX}, Y_{TX})$ of persons with covariates x to treatment a and $\delta(b, x, N_{TX}, Y_{TX})$ to treatment b. Expected welfare is

$$W(\delta, P, N, q) =$$

$$\sum_{\xi \in X} P(x = \xi)\cdot\{E[y(a)|x = \xi]\cdot E[\delta(a, \xi, N_{TX}, Y_{TX})] + E[y(b)|x = \xi]\cdot E[\delta(b, \xi, N_{TX}, Y_{TX})]\}.$$

With simple random sampling, N_{TX} is random. Proposition 1 applies conditional on N_{TX}. It provides the "inner loop" for analysis of simple random sampling.

Proposition 2: Let subjects be drawn by simple random sampling. Let $v(\cdot): X \rightarrow V$ and consider the CES rule δ_v. For $\nu \in V$, let $B_{N\nu}$ denote the Binomial distribution $\mathbf{B}[P(v = \nu), N]$. For $n = 0, \ldots, N$, let B_{nq} denote the Binomial distribution $\mathbf{B}(q, n)$. Then the expected welfare of rule δ_v in state s satisfies the inequality

$$\sum_{\nu \in V} P(v = \nu) M_{s\nu} - D(\delta_v, P_s, n, q) \leq W(\delta_v, N, P_s, q) \leq \sum_{\nu \in V} P(v = \nu) M_{s\nu},$$

where $D(\delta_v, p, N, q) \equiv$

$$\sum_{\nu \in V} P(v = \nu)\cdot\lambda_{s\nu} \sum_{n=0}^{N} \sum_{m=0}^{n} B_{N\nu}(n)\cdot B_{nq}(m)\cdot \exp\{-2\lambda_{s\nu}{}^2[(n - m)^{-1} + m^{-1}]^{-1}\}. \qquad \square$$

The bound on expected welfare gives this bound on regret:

$$\sum_{\xi \in X} P(x = \xi)M_{s\xi} - \sum_{\nu \in V} P(v = \nu)M_{s\nu} \leq U^*(P_s) - W(\delta_v, P_s, N, q)$$

$$\leq \sum_{\xi \in X} P(x = \xi)\, M_{s\xi} - \sum_{\nu \in V} P(v = \nu)\, M_{s\nu} + D(\delta_v, P_s, N, q).$$

Maximizing over the feasible values of $\{E_s[y(0)|x], E_s[y(1)|x]\}$ gives a bound on maximum regret: $R_{Lv} \leq R(z_v) \leq R_U(z_v)$,

where

$$R_U(\delta_v) \equiv \max_{s \in S} \sum_{\xi \in X} P(x = \xi) \max\{E_s[y(a)|x = \xi], E_s[y(b)|x = \xi]\}$$

$$- \sum_{\nu \in V} P(v = \nu) \max\{E_s[y(a)|v = \nu], E_s[y(b)|v = \nu]\}$$

$$+ \sum_{\nu \in V} P(v = \nu)\cdot\lambda s_\nu \sum_{n=0}^{N} \sum_{m=0}^{n} B_{N\nu}(n)\cdot B_{nq}(m)\cdot\exp\{-2\lambda_{s\nu}{}^2[(n-m)^{-1} + m^{-1}]^{-1}\}.$$

Sufficient sample sizes for productive use of covariate information follow from the bounds on maximum regret. Let $v(\cdot)$: $X \rightarrow V$ and $w(\cdot)$: $X \rightarrow W$ be distinct mappings of X into covariates v and w respectively. Conditioning on w is necessarily preferable to conditioning on v if N is such that $R_{Lv} > R_U(z_w)$.

Minimax Regret Rules with Data from a Randomized Experiment

(Source: Stoye, *JoE*, 2009, 2012)

Background

The minimax theorem of von Neumann (*MA*, 1928) shows that a two-person zero-sum game with finitely many pure strategies $(d, e) \in D \times E$ and payoffs $\pm f(d, e)$ has a mixed-strategy equilibrium that satisfies the minimax equation

$$\max_{p \in P} \min_{q \in Q} \sum_{p \in P} \sum_{q \in Q} p(d) \cdot q(e) \cdot f(d, e) = \min_{q \in Q} \max_{p \in P} \sum_{p \in P} \sum_{q \in Q} p(d) \cdot q(e) \cdot f(d, e).$$

Here P and Q are the sets of possible mixed strategies, p(d) and q(e) being the probabilities that the players choose d and e respectively.

Wald (*AM*, 1945) proves results that relate minimization of Bayes risk to minimax rules.

Let Π denote all distributions on S. For $\pi \in \Pi$ and $\delta \in \Delta$, consider the Bayes risk $\int R(\delta, s)d\pi(s)$. Given regularity conditions, Wald shows:

(i) There exists a decision rule $\delta(\pi)$ that minimizes Bayes risk under π.

Let $R^*(\pi) \equiv \int R[\delta(\pi), s]d\pi(s)$ be minimum Bayes risk under π.

(ii) There exists a *least favorable distribution* $\pi^* = \text{argmax}_{\pi \in \Pi} R^*(\pi)$.

(iii) Rule $\delta(\pi^*)$ is a minimax rule.

(iv) A minimax rule minimizes Bayes risk under π^*.

Wald observes that statistical decision problems may be viewed as fictitious zero-sum games. One player is the decision maker, whose pure strategies are the rules $\delta \in \Delta$. The other is nature, whose pure strategies are the states $s \in S$.

The risk function $R(\delta, s)$ gives the payoffs. The decision maker wants to minimize risk. Nature wants to maximize it.

Wald relates his findings to those of von Neumann. The least favorable distribution is nature's maximin strategy in the fictitious game.

The von Neumann and Wald analyses also apply to minimax-regret rules, with regret rather than risk giving the payoff.

Wald suggested that the equivalence of minimax rules and ones that minimize Bayes risk under a least favorable distribution may be useful in determining minimax rules. He wrote

"An important and unsolved problem is to find a general method which would permit the actual calculation of an optimum statistical decision function. The fact that a decision function which minimizes the average risk relative to a least favorable distribution of θ is an optimum decision function may be very helpful in finding an optimum decision function, since the calculation of a least favorable distribution of θ seems to be a considerably simpler problem."

A common approach is to "guess and verify." One uses heuristic reasoning to conjecture a least favorable distribution, derives the decision rule that minimizes Bayes risk, and then verifies whether the distribution and decision rule solve the minimax equation (are an equilibrium).

Stoye's Application to Treatment Choice with Sample Data

Stoye (2009) applies the "guess and verify" approach. He considers two treatments and data from a randomized experiment.

1. Stoye initially assumes binary outcomes and no covariates. He shows:

(i) With balanced stratified random sampling ($N_a = N_b$), the MMR rule is the empirical success rule that breaks ties by setting $\delta(b) = ½$ when $\bar{y}_a = \bar{y}_b$.

(ii) With balanced random sampling ($q = ½$), the MMR rule sets $\delta(b) = 0$ if $I_N < 0$, $\delta(b) = ½$ if $I_N = 0$, and $\delta(b) = 1$ if $I_N > 0$, where $I_N = N_b(\bar{y}_b - ½) - N_a(\bar{y}_a - ½)$.

(iii) With choice between an innovation and a status quo treatment, the MMR rule solves an implicit equation and is well approximated by the empirical success rule.

2. He then enlarges the state space S so that $P_s[y(a), y(b)]$, $s \in S$ includes all distributions with binary outcomes plus others with bounded outcomes in [0, 1].

He cites previous findings showing that the support of the least favorable distribution on S is concentrated on distributions with only binary outcomes.

The MMR findings proved for the case of binary outcomes continue to hold when experimental outcomes interior to [0, 1]are *coarsened* to make them binary.

Each outcome y_n is replaced by an independent realization of a Bernoulli random variable with parameter y_n. Thus, y_n is replaced by 1 with probability y_n and 0 with probability $1 - y_n$. One then applies the binary MMR rule to the coarsened data.

The MMR rule using coarsened data may not be the unique MMR rule, raising the possibility of inadmissibility.

3. He then considers treatment choice with observable covariates.

Now the state space has the form $\{P_s[y(a), y(b)|x = \xi], \xi \in X, s \in S\}$.

He considers settings where S "is not constrained by cross-covariate restrictions." This means that S is the Cartesian-product $\times_{\xi \in X} \{P_{s\xi}[y(a), y(b)|x = \xi], s_\xi \in S_\xi\}$.

In these settings, the least favorable distribution concentrates on distributions such that $(s_\xi, \xi \in X)$ are statistically independent. Minimization of Bayes risk under the least favorable distribution yields a rule that makes treatment of persons with covariates ξ depend only on the outcomes of subjects with this covariate value.

It follows that the MMR rule separates the outcome data by their covariates. For each $\xi \in X$, it is the MMR rule that would hold if one were to only observe outcomes for subjects with covariates ξ.

Stoye (2009) views his finding with observable covariates as casting doubt on the reasonableness of the minimax-regret criterion. He writes

"Paradoxically, the recommendation based on small sample analysis does not seem desirable in a world of small samples. It requires one to condition on covariates even if this leads to extremely small or empty sample cells. While medical researchers might want to consider the effect of race on treatment outcomes, they will hardly want to altogether ignore experiences made with white subjects when considering treatment for black subjects unless samples are extremely large. And surely, they would not want to consider an arbitrary covariate that happens to be in the data set.

What's more, Proposition 3 implies that as the support of a covariate grows, minimax regret treatment rules approach no data rules, because the proportion of covariate values that have been observed in the sample vanishes. Indeed, one can extend the result to a continuous covariate, and a no-data rule then achieves minimax regret."

In his Conclusion section, he continues as follows:

"A natural question is whether the minimax regret criterion is an attractive alternative to existing approaches. There are two ways to investigate this: One is axiomatic analysis, the other one is to look whether actual minimax regret rules make sense. The former has been done elsewhere . . . The latter was one purpose of this paper, and the results raise some concerns. Consider specifically the findings on covariates. Manski (2004) discovered that for surprisingly small sample sizes, lower bounds on expected regret incurred by pooling the sample exceed upper bounds incurred by conditioning on covariates. This finding was used to criticize prevailing practice, tentatively suggesting that there is too much pooling of observations across covariate values. However, it is now seen that the result was merely an approximation of a much stronger, and pathological, one. The exact finding cannot any more inform a critique of prevailing practice, but rather raises questions about minimax regret."

"Perhaps the problem can be alleviated by properly specifying available prior information. The most obvious way to do this is to restrict S. . . . . Ongoing research indicates that some natural such restrictions lead to reasonable minimax regret treatment rules."

Stoye (*JoE*, 2012) takes up this idea.

The main findings concern restriction of S by placing a uniform upper bound $\kappa \geq 0$ on the variation of mean treatment response across covariate values.

Let outcomes be binary and the design be balanced. Let $N = N_a + N_b$. Assume that, for all $t \in \{a, b\}$, $(\xi, \xi') \in X$, and $s \in S$, $|E_s[y(t)|x = \xi] - E_s[y(t)|x = \xi']| \leq \kappa$.

Let δ_0 = the MMR rule without observation of x,

δ_1 = the MMR rule with observation of x and no cross-covariate restrictions.

Stoye shows that there exist computable $\kappa_0(N) < \kappa_1(N)$ such that δ_0 is a MMR rule if $\kappa \leq \kappa_0(N)$ and δ_1 is a MMR rule if $\kappa \geq \kappa_1(N)$.

$\kappa_0(N)$ and $\kappa_1(N)$ decrease with N at rate $N^{½}$.

Questions for Discussion

What do the Stoye findings suggest about the reasonableness of MMR as a decision criterion?

What do the findings suggest about the reasonableness of decision criteria that pool data across covariates when the state space makes no cross-covariate restrictions?

Treatment Choice with Partial Identification of Treatment Response

(Sources, Manski, *JoE*, 2007; *IPD*, 2007; *IER*, 2009)

Recall that identification problems combine with the necessity of inference from sample data to limit the informativeness of studies.

The preceding analysis of treatment choice with data from a randomized experiment dealt with the polar case in which a planner deals with the imprecision of sample data in a setting with point identification of treatment response.

Now consider the other polar case, in which treatment response is partially identified and there is full knowledge of observable features of the population.

See Stoye (*JoE*, 2012) for some analysis of treatment choice with sample data and partial identification of treatment response.

Setup

There are two treatments, {a, b}.

Each member j of population J has a response function $y_j(\cdot)$: mapping treatments t into outcomes $y_j(t)$.

$P[y(\cdot)]$ is the population distribution of treatment response. The population is large, with $P(j) = 0$ for all $j \in J$.

An allocation $\delta \in [0, 1]$ randomly assigns a fraction δ of the population to treatment b and the remaining $1 - \delta$ to treatment a.

The planner wants to maximize mean welfare.

Let $\alpha \equiv E[u(a)]$ and $\beta \equiv E[u(b)]$.

Mean welfare with allocation δ is

$$W(\delta) \;=\; \alpha(1 - \delta) + \beta\delta \;=\; \alpha + (\beta - \alpha)\delta.$$

$\delta = 1$ is optimal if $\beta \geq \alpha$ and $\delta = 0$ if $\beta \leq \alpha$. The problem is treatment choice when (α, β) is partially known.

Treatment Choice Under Ambiguity

Let S index the feasible states of nature. The planner knows that (α, β) lies in the set $[(\alpha_s, \beta_s), s \in S]$. Let this set be bounded.

Let $\alpha_L \equiv \min_{s \in S} \alpha_s$, $\beta_L \equiv \min_{s \in S} \beta_s$,

$\alpha_U \equiv \max_{s \in S} \alpha_s$, $\beta_U \equiv \max_{s \in S} \beta_s$.

Let $S(a) \equiv \{s \in S: \alpha_s > \beta_s\}$ and $S(b) \equiv \{s \in S: \beta_s > \alpha_s\}$.

The planner faces ambiguity if S(a) and S(b) are both non-empty.

Bayes Rules

A Bayesian planner places a subjective distribution π on S and solves

$$\max_{\delta \in [0, 1]} \quad E_\pi(\alpha) + [E_\pi(\beta) - E_\pi(\alpha)]\delta,$$

where $E_\pi(\alpha) = \int \alpha_s d\pi$ and $E_\pi(\beta) = \int \beta_s d\pi$.

He chooses $\delta = 0$ if $E_\pi(\beta) < E_\pi(\alpha)$ and $\delta = 1$ if $E_\pi(\beta) > E_\pi(\alpha)$.

The Maximin Criterion

A maximin planner solves

$$\max_{\delta \in [0, 1]} \quad \min_{s \in S} \quad \alpha_s + (\beta_s - \alpha_s)\delta.$$

Let (α_L, β_L) be feasible. Then the decision is $\delta = 0$ if $\beta_L < \alpha_L$ and $\delta = 1$ if $\beta_L > \alpha_L$.

The Minimax-Regret Criterion

The minimax-regret criterion is

$$\min_{\delta \in [0, 1]} \quad \max_{s \in S} \quad \max(\alpha_s, \beta_s) - [\alpha_s + (\beta_s - \alpha_s)\delta].$$

$\delta_{MR} = 1$ if S(a) is empty. $\delta_{MR} = 0$ if S(b) is empty.

All δ are MMR if both S(a) and S(b) are empty.

Let S(a) and S(b) both be non-empty.

Let $M(a) \equiv \max_{s \in S(a)} (\alpha_s - \beta_s)$ and $M(b) \equiv \max_{s \in S(b)} (\beta_s - \alpha_s)$. Then

$$\delta_{MR} = \frac{M(b)}{M(a) + M(b)} .$$

If (α_L, β_U) and (α_U, β_L) are feasible, then

$$\delta_{MR} = \frac{\beta_U - \alpha_L}{(\alpha_U - \beta_L) + (\beta_U - \alpha_L)} .$$

Proof:

The maximum regret of δ is max [R(δ, a), R(δ, b)], where

$$R(\delta, a) \equiv \max_{s \in S(a)} \alpha_s - [(1 - \delta)\alpha_s + \delta\beta_s]$$

$$= \max_{s \in S(a)} \delta(\alpha_s - \beta_s) = \delta M(a),$$

$$R(\delta, b) \equiv \max_{s \in S(b)} \beta_s - [(1-\delta)\alpha_s + \delta\beta_s]$$

$$= \max_{s \in S(b)} (1-\delta)(\beta_s - \alpha_s) = (1-\delta)M(b),$$

are maximum regret on S(a) and S(b).

Both treatments are undominated, so $R(1, a) = M(a) > 0$ and $R(0, b) = M(b) > 0$.

As δ increases from 0 to 1, $R(\cdot, a)$ increases linearly from 0 to $M(a)$ and $R(\cdot, b)$ decreases linearly from $M(b)$ to 0.

Hence, the MR rule is the unique $\delta \in (0, 1)$ such that $R(\delta, a) = R(\delta, b)$.

This yields the result.

Planning with Observable Covariates

A planner may differentiate among persons with different covariates $\xi \in X$.

He may segment persons by ξ and treat each group as the population. This works when the objective function is separable in covariates but not otherwise.

The Bayes objective function is always separable.

Maximin is separable if $(\alpha_{\xi L}, \beta_{\xi L})$, $\xi \in X$ is feasible.

Minimax-regret is separable if $(\alpha_{\xi L}, \beta_{\xi U})$, $\xi \in X$ and $(\alpha_{\xi U}, \beta_{\xi L})$, $\xi \in X$ are feasible.

Nonseparability occurs with cross-covariate restrictions that relate $(\alpha_{\xi s}, \beta_{\xi s})$ across $\xi \in X$.

Planning with Multiple Treatments

The MR allocation is not always fractional when a planner allocates the population among more than two treatments.

Stoye (*ET*, 2007) has studied a class of such problems and has found that the MR allocations are subtle to characterize. They often are fractional, but he gives an example in which there exists a unique singleton allocation.

Minimax-Regret Treatment Choice With Missing Outcome Data
(Manski, *JoE*, *IPD*, 2007)

A prominent case of partial identification occurs when outcome data are missing for some members of the study population and one lacks knowledge of the distribution of missing data.

This occurs in randomized experiments when some subjects drop out before their outcomes are measured and one does not know why they dropped out.

It occurs in analysis of observational data when one does not know the process of treatment selection. Outcome data may be missing for other reasons as well.

Let $J_{t\xi}$ denote the sub-population of persons with covariates ξ whose outcome y(t) is observable. By the Law of Iterated Expectations,

$$E[y(t)|x = \xi] =$$
$$E[y(t)|x = \xi, J_{t\xi}]\cdot P(J_{t\xi}|x = \xi) + E[y(t)|x = \xi, \text{not } J_{t\xi}]\cdot P(\text{not } J_{t\xi}|x = \xi).$$

$P(J_{t\xi}|x = \xi)$ and $E[y(t)|x = \xi, J_{t\xi}]$ can be learned empirically. $E[y(t)|x = \xi, \text{not } J_{t\xi}]$ cannot be learned empirically.

Let outcomes take values in [0, 1]. In the absence of knowledge of the process yielding missing data, $E[y(t)|x = \xi, \text{not } J_{t\xi}] \in [0, 1]$. Hence, for $t \in \{a, b\}$,

$$E[y(t)|x = \xi, J_{t\xi}]\cdot P(J_{t\xi}|x = \xi) \leq E[y(t)|x = \xi]$$
$$\leq E[y(t)|x = \xi, J_{t\xi}]\cdot P(J_{t\xi}|x = \xi) + P(\text{not } J_{t\xi}|x = \xi).$$

In terms of the notation used earlier,

$$\alpha_L = E[y(a) \mid x = \xi, J_{a\xi}]\cdot P(J_{a\xi} \mid x = \xi),$$

$$\alpha_U = E[y(a) \mid x = \xi, J_{a\xi}]\cdot P(J_{a\xi} \mid x = \xi) + P(\text{not } J_{a\xi} \mid x = \xi),$$

$$\beta_L = E[y(b) \mid x = \xi, J_{b\xi}]\cdot P(J_{b\xi} \mid x = \xi),$$

$$\beta_U = E[y(b) \mid x = \xi, J_{b\xi}]\cdot P(J_{b\xi} \mid x = \xi) + P(\text{not } J_{b\xi} \mid x = \xi).$$

The state space is $[\alpha_L, \alpha_U] \times [\beta_L, \beta_U]$.

$\delta_{MR} = 1$ if $\alpha_U < \beta_L$. $\delta_{MR} = 0$ if $\alpha_L > \beta_U$. Otherwise,

$$\delta_{MR} = \frac{\beta_U - \alpha_L}{(\alpha_U - \beta_L) + (\beta_U - \alpha_L)} = \frac{E[y(b) \mid x{=}\xi, J_{b\xi}]\cdot P(J_{b\xi} \mid x{=}\xi) + P(\text{not } J_{b\xi} \mid x{=}\xi) - E[y(a) \mid x{=}\xi, J_{a\xi}]\cdot P(J_{a\xi} \mid x{=}\xi)}{P(\text{not } J_{a\xi} \mid x{=}\xi) + P(\text{not } J_{b\xi} \mid x{=}\xi)}.$$

$P(\text{not } J_{a\xi} \mid x{=}\xi) + P(\text{not } J_{b\xi} \mid x{=}\xi) = 1$ given observational data with all outcomes observed.