# Non-Parametric Instrumental Variable Models for Categorical Data

Thomas Richardson
University of Washington

JSM 2016
Chicago

Joint work with James Robins (Harvard)

# Outline

- Randomized experiments
- Observational studies
- Combining observational studies
- Instrumental variables
- Testing instrumental variable models
- Bayesian inference

# Potential outcomes with binary treatment and outcome

For binary treatment $X$, we define two potential outcome variables:

- $Y(x = 0)$: the value of $Y$ that *would* be observed for a given unit *if* assigned $X = 0$;
- $Y(x = 1)$: the value of $Y$ that *would* be observed for a given unit *if* assigned $X = 1$;

*Notation*: We will use $Y(x_i)$ as an abbreviation for $Y(x = i)$

# Drug Response Types:

In the simplest case where $Y$ is a binary outcome we can think of patients as belonging to one of 4 'types':

| $Y(x_0)$ | $Y(x_1)$ | Name |
|---|---|---|
| 0 | 0 | *Never Recover* (NR) |
| 0 | 1 | *Helped* (HE) |
| 1 | 0 | *Hurt* (HU) |
| 1 | 1 | *Always Recover* (AR) |

# Identification Problem

Want: $P(Y(x_0), Y(x_1))$;     Given: $P(Y \mid X=0), P(Y \mid X=1)$

Under randomization: $X \perp\!\!\!\perp Y(x_i)$ implies:

$$P(Y(x_i) = 1) = P(Y(x_i) = 1 \mid X = i) = P(Y = 1 \mid X = i).$$

# Identification Problem

Want: $P(Y(x_0), Y(x_1))$; Given: $P(Y \mid X=0), P(Y \mid X=1)$

Under randomization: $X \perp\!\!\!\perp Y(x_i)$ implies:

$$P(Y(x_i) = 1) = P(Y(x_i) = 1 \mid X = i) = P(Y = 1 \mid X = i).$$

Thus the observed joint $P(Y|X)$ puts two restrictions on $P(Y(x_0), Y(x_1))$:

$$\begin{aligned} P(Y=1 \mid X=0) &= P(Y(x_0)=1, Y(x_1)=0) + P(Y(x_0)=1, Y(x_1)=1) \\ P(Y=1 \mid X=1) &= P(Y(x_0)=0, Y(x_1)=1) + P(Y(x_0)=1, Y(x_1)=1). \end{aligned}$$

# Identification Problem

Want: $P(Y(x_0), Y(x_1))$;  Given: $P(Y \mid X=0), P(Y \mid X=1)$

Under randomization: $X \perp\!\!\!\perp Y(x_i)$ implies:

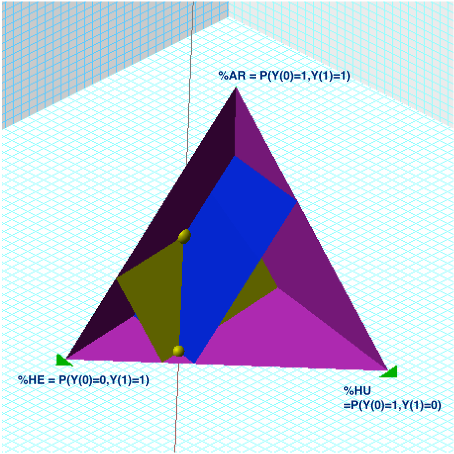
$$P(Y(x_i) = 1) = P(Y(x_i) = 1 \mid X = i) = P(Y = 1 \mid X = i).$$

Thus the observed joint $P(Y|X)$ puts two restrictions on $P(Y(x_0), Y(x_1))$:

$$\begin{aligned} P(Y=1 \mid X=0) &= P(Y(x_0)=1, Y(x_1)=0) + P(Y(x_0)=1, Y(x_1)=1) \\ P(Y=1 \mid X=1) &= P(Y(x_0)=0, Y(x_1)=1) + P(Y(x_0)=1, Y(x_1)=1). \end{aligned}$$

Each restriction implies a 2-d subset in $\Delta_3$.
Intersection forms a 1-d subset on which ACE is constant.

# 3-d Plot




In this plot:

$P(Y=1 \mid X=0) = P(Y(x_0) = 1) = \%HU + \%AR = 0.3$, (yellow)

$P(Y=1 \mid X=1) = P(Y(x_1) = 1) = \%HE + \%AR = 0.6$, (blue)

# Two-way Table

Under randomization, the relationship between the counterfactual distribution $P(Y(x_0), Y(x_1))$ and the observed distributions $\{P(Y \mid x_0), P(Y \mid x_1)\}$ is:

| | | col sums | |
|---|---|---|---|
| | | $P(Y=0 \mid X=0)$ | $P(Y=1 \mid X=0)$ |
| row sums | $P(Y=0 \mid X=1)$ | $P(Y(x_0)=0, Y(x_1)=0)$ | $P(Y(x_0)=1, Y(x_1)=0)$ |
| | $P(Y=1 \mid X=1)$ | $P(Y(x_0)=0, Y(x_1)=1)$ | $P(Y(x_0)=1, Y(x_1)=1)$ |

Here $P(Y=i \mid X=j) = P(Y(x_j)=i)$ due to randomization.

# Two-way Table

Under randomization, the relationship between the counterfactual distribution $P(Y(x_0), Y(x_1))$ and the observed distributions $\{P(Y \mid x_0), P(Y \mid x_1)\}$ is:

| | | col sums | |
|---|---|---|---|
| | | $P(Y=0 \mid X=0)$ | $P(Y=1 \mid X=0)$ |
| row sums | $P(Y=0 \mid X=1)$ | $P(Y(x_0)=0, Y(x_1)=0)$ | $P(Y(x_0)=1, Y(x_1)=0)$ |
| | $P(Y=1 \mid X=1)$ | $P(Y(x_0)=0, Y(x_1)=1)$ | $P(Y(x_0)=1, Y(x_1)=1)$ |

Here $P(Y=i \mid X=j) = P(Y(x_j)=i)$ due to randomization.

Equivalently we may write this in terms of types

| | $P(Y=0 \mid X=0)$ | $P(Y=1 \mid X=0)$ |
|---|---|---|
| $P(Y=0 \mid X=1)$ | $P(\text{NR})$ | $P(\text{HU})$ |
| $P(Y=1 \mid X=1)$ | $P(\text{HE})$ | $P(\text{AR})$ |

# Observational study: one-way table!

| Observed | Counterfactual |
|---|---|
| $p(X=0, Y=0)$ | $p(X=0, \text{NR}) + p(X=0, \text{HE})$ |
| $p(X=0, Y=1)$ | $p(X=0, \text{HU}) + p(X=0, \text{AR})$ |
| $p(X=1, Y=0)$ | $p(X=1, \text{NR}) + p(X=1, \text{HU})$ |
| $p(X=1, Y=1)$ | $p(X=1, \text{HE}) + p(X=1, \text{AR})$ |

# Bounds on joints $P(Y(x_0), Y(x_1))$

| Observed | Counterfactual |
|---|---|
| $p(X=0, Y=0)$ | $p(X=0, \text{NR}) + p(X=0, \text{HE})$ |
| $p(X=0, Y=1)$ | $p(X=0, \text{HU}) + p(X=0, \text{AR})$ |
| $p(X=1, Y=0)$ | $p(X=1, \text{NR}) + p(X=1, \text{HU})$ |
| $p(X=1, Y=1)$ | $p(X=1, \text{HE}) + p(X=1, \text{AR})$ |

$$0 \leqslant \quad \%HE \quad \leqslant P(X = 0, Y = 0) + P(X = 1, Y = 1)$$

# Bounds on joints $P(Y(x_0), Y(x_1))$

| Observed | Counterfactual |
|---|---|
| $p(X=0, Y=0)$ | $p(X=0, \text{NR}) + p(X=0, \text{HE})$ |
| $p(X=0, Y=1)$ | $p(X=0, \text{HU}) + p(X=0, \text{AR})$ |
| $p(X=1, Y=0)$ | $p(X=1, \text{NR}) + p(X=1, \text{HU})$ |
| $p(X=1, Y=1)$ | $p(X=1, \text{HE}) + p(X=1, \text{AR})$ |

$$
\begin{aligned}
0 \leqslant \quad \%HE \quad &\leqslant P(X=0, Y=0) + P(X=1, Y=1) \\
0 \leqslant \quad \%HU \quad &\leqslant P(X=0, Y=1) + P(X=1, Y=0) \\
0 \leqslant \quad \%NR \quad &\leqslant P(X=0, Y=0) + P(X=1, Y=0) \; = \; P(Y=0) \\
0 \leqslant \quad \%AR \quad &\leqslant P(X=0, Y=1) + P(X=1, Y=1) \; = \; P(Y=1)
\end{aligned}
$$

# Bounds on margins $P(Y(x_i))$

| Observed | Counterfactual |
|---|---|
| $p(X=0, Y=0)$ | $p(X=0, \text{NR}) + p(X=0, \text{HE})$ |
| $p(X=0, Y=1)$ | $p(X=0, \text{HU}) + p(X=0, \text{AR})$ |
| $p(X=1, Y=0)$ | $p(X=1, \text{NR}) + p(X=1, \text{HU})$ |
| $p(X=1, Y=1)$ | $p(X=1, \text{HE}) + p(X=1, \text{AR})$ |

We also have the following inequalities on the marginals:

$$\begin{cases} P(Y(x_0) = 1) = P(\text{HU}) + P(\text{AR}) \\ P(Y(x_1) = 1) = P(\text{HE}) + P(\text{AR}) \end{cases}$$

$= P(Y(X=0) = 1)$

$$P(X = 0, Y = 1) \leqslant P(Y(x_0) = 1) \leqslant 1 - P(X = 0, Y = 0)$$

$$P(X = 1, Y = 1) \leqslant P(Y(x_1) = 1) \leqslant 1 - P(X = 1, Y = 0)$$

# Bounds on margins $P(Y(x_i))$

| Observed | Counterfactual |
|---|---|
| $p(X=0, Y=0)$ | $p(X=0, \text{NR}) + p(X=0, \text{HE})$ |
| $p(X=0, Y=1)$ | $p(X=0, \text{HU}) + p(X=0, \text{AR})$ |
| $p(X=1, Y=0)$ | $p(X=1, \text{NR}) + p(X=1, \text{HU})$ |
| $p(X=1, Y=1)$ | $p(X=1, \text{HE}) + p(X=1, \text{AR})$ |

We also have the following inequalities on the marginals:

$$P(Y(x_0)=1) = P(\text{HU}) + P(\text{AR})$$

$$P(Y(x_1)=1) = P(\text{HE}) + P(\text{AR})$$

$P(Y(x_0)=0) = P(HU) + P(NR)$

$P(Y(x_1)=0) = P(NR) + P(HU)$

$P(HU) + P(AR)$

$$P(X=0, Y=1) \leqslant P(Y(x_0)=1) \leqslant 1 - P(X=0, Y=0)$$

$$P(X=1, Y=1) \leqslant P(Y(x_1)=1) \leqslant 1 - P(X=1, Y=0)$$

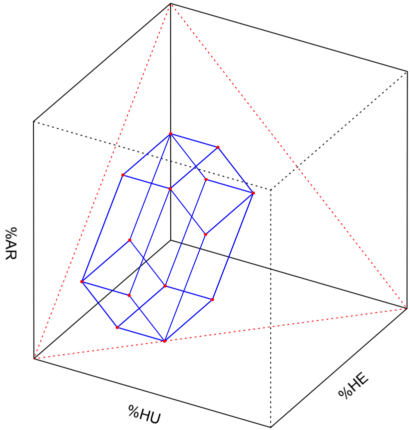
$P(X=1, Y=0) \leq P(Y(x_0)=0) \leq 1 - P(X=1, Y=1)$

Thus we have 6 pairs of parallel planes.

$\therefore P(Y(x_1)=0) = 1 - P(Y(x_0)=1) - P(Y(x_1)=0) - P(Y(x_0)=0)$

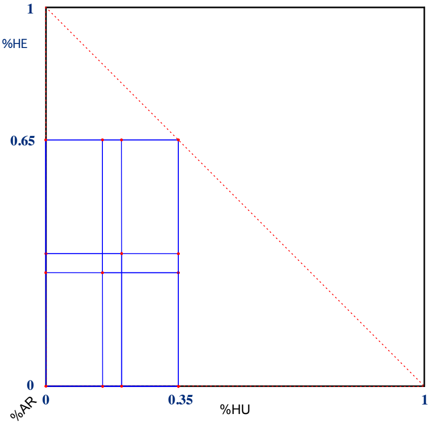
# Polytope for observational study

Set of margins $P(Y(x_0), Y(x_1))$ compatible with the Obs. Study.



# ACE bounds

ACE = % HE − % HU




(But why is this helpful!?)

# Inference for the ACE without randomization

## (Robins-Manski bounds)

Suppose that we do not know that $X \perp\!\!\!\perp Y(x_0)$ and $X \perp\!\!\!\perp Y(x_1)$.

What can be inferred from the observed distribution $P(X, Y)$?

General case:

$$
\begin{aligned}
-(P(X=0, Y=1) &+ P(X=1, Y=0)) \\
&\leqslant \; ACE(X \to Y) \\
&\qquad \leqslant P(X=0, Y=0) + P(X=1, Y=1)
\end{aligned}
$$

# Inference for the ACE without randomization

## (Robins-Manski bounds)

Suppose that we do not know that $X \perp\!\!\!\perp Y(x_0)$ and $X \perp\!\!\!\perp Y(x_1)$.

What can be inferred from the observed distribution $P(X, Y)$?

General case:

$$
\begin{aligned}
-(P(X{=}0, Y{=}1) + P(X{=}1, Y{=}0))& \\
\leqslant \quad ACE(X \to Y)& \\
\leqslant P(X{=}0, Y{=}0) + P(X{=}1, Y{=}1)&
\end{aligned}
$$

$\Rightarrow$ Bounds are of length 1 and will always include zero.

# Combining two Obs Studies: Cholestyramine data

$Z$: assignment to treatment or control arm (randomized);
$X$: whether patient takes (more than certain amount of) drug;
$Y$: patient's health outcome.

# Combining two Obs Studies: Cholestyramine data

$Z$: assignment to treatment or control arm (randomized);
$X$: whether patient takes (more than certain amount of) drug;
$Y$: patient's health outcome.

| Z | X | Y | count |
|---|---|---|---|
| 0 | 0 | 0 | 158 |
| 0 | 0 | 1 | 14 |
| 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 52 |
| 1 | 0 | 1 | 12 |
| 1 | 1 | 0 | 23 |
| 1 | 1 | 1 | 78 |

(Data originally considered by Efron and Feldman (1991); dichotomized by Pearl.)

We wish to find ACE$(X \rightarrow Y)$. Note $Z = 0 \;\Rightarrow\; X = 0$

# Combining two Obs Studies: Cholestyramine data

$Z$: assignment to treatment or control arm (randomized);
$X$: whether patient takes (more than certain amount of) drug;
$Y$: patient's health outcome.

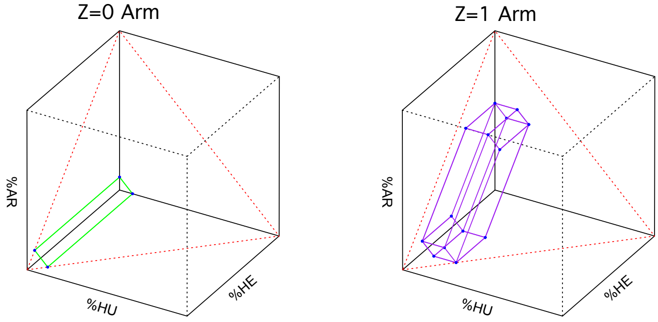
| Z | X | Y | count |
|---|---|---|---|
| 0 | 0 | 0 | 158 |
| 0 | 0 | 1 | 14 |
| 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 1 | 0 | 0 | 52 |
| 1 | 0 | 1 | 12 |
| 1 | 1 | 0 | 23 |
| 1 | 1 | 1 | 78 |

(Data originally considered by Efron and Feldman (1991); dichotomized by Pearl.)

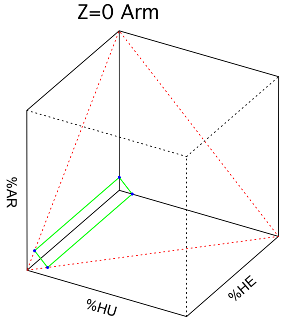
We wish to find ACE$(X \to Y)$. Note $Z = 0 \;\Rightarrow\; X = 0$

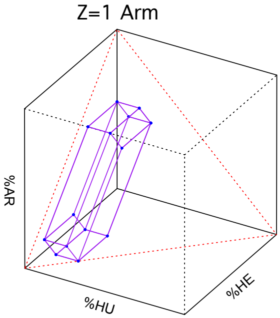
Idea: Analyze each $Z$ arm as an observational study.

# Each Z Arm



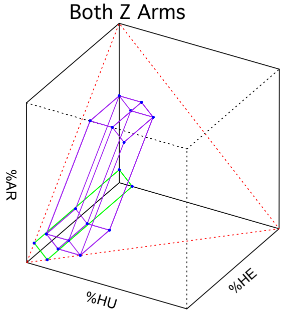
# Each Z Arm

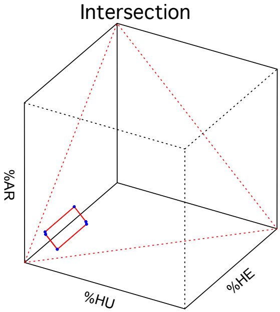


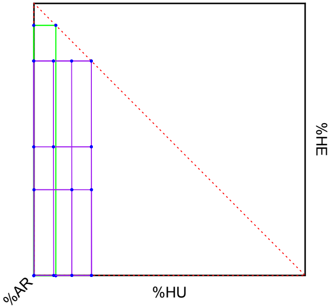

$Z = 0$ arm polytope is 2-d since $Z = 0 \Rightarrow X = 0$

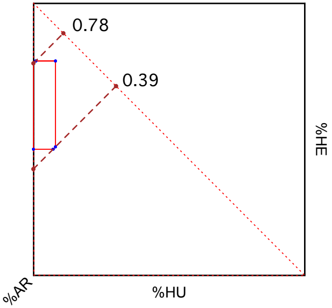
# Combining the Arms



# Obtaining ACE bounds

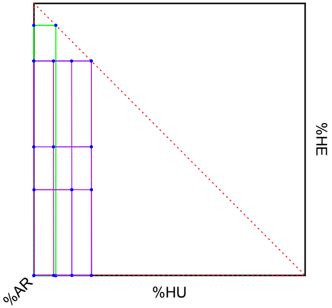


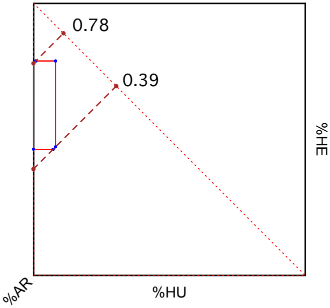




Upper bound is: 0.78; lower bound is 0.39

Note: ACE bounds for each arm contain 0, but not when combined. *Why?*

# Obtaining ACE bounds



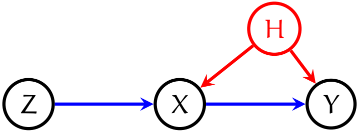

Upper bound is: 0.78; lower bound is 0.39

Note: ACE bounds for each arm contain 0, but not when combined. *Why?*

These are the Balke and Pearl (1993) bounds, obtained via prinicipal strata and computational algebra;

# Instrumental Variable Model

We assume that assignment $Z$ has no direct effect on the outcome $Y$ except through $X$, so $Y(x, z_0) = Y(x, z_1) \equiv Y(x)$;




$Z$ is said to be an 'instrument' for the effect of $X$ on $Y$.

Potential outcomes:

$$X(z_0), X(z_1), Y(x_0), Y(x_1)$$

Randomization becomes: for all $x, z$, $\quad Z \perp\!\!\!\perp \{X(z), Y(x)\}$

The assumptions: (i) No direct effect of $Z$ on $Y$, and (ii) $Z$ randomized are sufficient to justify the intersecting polytope analysis.

# Bounds on ACE: Previous results

- Robins (1989) and Manski (1990) derived the 'natural' bounds:

$$
\begin{aligned}
& p(y_1 \mid z_1) - p(y_1 \mid z_0) - p(y_1, x_0 \mid z_1) - p(y_0, x_1 \mid z_0) \\
& \quad \leqslant \quad \mathrm{ACE}(X \to Y) \quad \leqslant \\
& \qquad p(y_1 \mid z_1) - p(y_1 \mid z_0) + p(y_0, x_0 \mid z_1) + p(y_1, x_1 \mid z_0)
\end{aligned}
$$

  These are not sharp if there are Defiers;

  they follow from (and are sharp) under $Z \perp\!\!\!\perp Y(x_0)$, $Z \perp\!\!\!\perp Y(x_1)$.

# Bounds on ACE: Previous results

- Robins (1989) and Manski (1990) derived the 'natural' bounds:

$$p(y_1 \mid z_1) - p(y_1 \mid z_0) - p(y_1, x_0 \mid z_1) - p(y_0, x_1 \mid z_0)$$

$$\leqslant \quad ACE(X \to Y) \quad \leqslant$$

$$p(y_1 \mid z_1) - p(y_1 \mid z_0) + p(y_0, x_0 \mid z_1) + p(y_1, x_1 \mid z_0)$$

  These are not sharp if there are Defiers;

  they follow from (and are sharp) under $Z \perp\!\!\!\perp Y(x_0)$, $Z \perp\!\!\!\perp Y(x_1)$.

- Balke and Pearl (1993) derived closed form expression for ACE bounds via computational algebra;

  Resulting expressions are maxima and minima over 8 different expressions.

# Bounds on ACE: Previous results

- Robins (1989) and Manski (1990) derived the 'natural' bounds:

$$p(y_1 \mid z_1) - p(y_1 \mid z_0) - p(y_1, x_0 \mid z_1) - p(y_0, x_1 \mid z_0)$$
$$\leqslant \quad ACE(X \to Y) \quad \leqslant$$
$$p(y_1 \mid z_1) - p(y_1 \mid z_0) + p(y_0, x_0 \mid z_1) + p(y_1, x_1 \mid z_0)$$

  These are not sharp if there are Defiers;

  they follow from (and are sharp) under $Z \perp\!\!\!\perp Y(x_0)$, $Z \perp\!\!\!\perp Y(x_1)$.

- Balke and Pearl (1993) derived closed form expression for ACE bounds via computational algebra;

  Resulting expressions are maxima and minima over 8 different expressions.

- Dawid (2002) re-derived these bounds without (explicitly!) using potential outcomes, again using computational algebra.

An advantage of the approach taken here is that it extends to Z with finitely many levels.

# Extending to $p$ levels of $Z$: Identification

Solving the identification problem requires no further work:

Intersect $p$ polytopes.

# Extending to $p$ levels of $Z$: Identification

Solving the identification problem requires no further work:
Intersect $p$ polytopes.

The set of distributions $P(Y(x_0), Y(x_1))$ that are compatible with $P(X, Y|Z)$ is still given by the following six pairs of parallel planes:

$$
\begin{array}{rcl}
0 \leqslant & \%NR & \leqslant \min_i p(y_0 \mid z_i) \\
0 \leqslant & \%AR & \leqslant \min_i p(y_1 \mid z_i) \\
0 \leqslant & \%HE & \leqslant \min_i \{p(x_1, y_1 \mid z_i) + p(x_0, y_0 \mid z_i)\} \\
0 \leqslant & \%HU & \leqslant \min_i \{p(x_0, y_1 \mid z_i) + p(x_1, y_0 \mid z_i)\} \\
\max_i p(x_0, y_1 \mid z_i) \leqslant & P(Y(x_0)=1) & \leqslant \min_i \{1 - p(x_0, y_0 \mid z_i)\} \\
\max_i p(x_1, y_1 \mid z_i) \leqslant & P(Y(x_1)=1) & \leqslant \min_i \{1 - p(x_1, y_0 \mid z_i)\}
\end{array}
$$

# Generalizing to more levels of Z

**Theorem (R+Robins 2014)**

If $X, Y$ binary and $Z$ with $p$ levels then under the IV model:

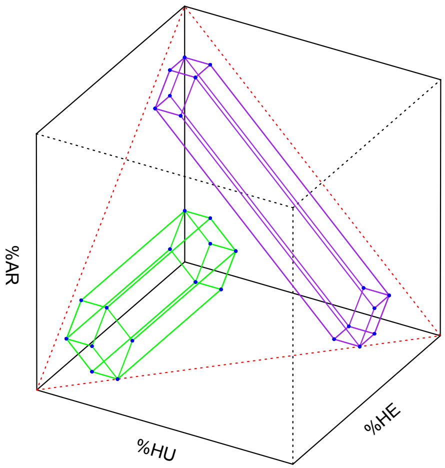
$$1 - g(1,0) - g(0,1) \leqslant ACE(X \to Y) \leqslant g(0,0) + g(1,1) - 1,$$

where

$$\begin{aligned} g(i,j) \equiv \min \Big\{ & \min_{z} \left[ P(X{=}i, Y{=}j \mid Z{=}z) + P(X{=}1-i \mid Z{=}z) \right], \\ & \min_{z,\tilde{z}:\, z \neq \tilde{z}} \big[ P(X{=}i, Y{=}j \mid Z{=}z) + P(X{=}1-i, Y{=}0 \mid Z{=}z) \\ & \quad + \; P(X{=}i, Y{=}j \mid Z{=}\tilde{z}) + P(X{=}1-i, Y{=}1 \mid Z{=}\tilde{z}) \big] \Big\}. \end{aligned}$$

This exploits the fact that $P(Y(x_0) = 1)$ and $P(Y(x_1) = 1)$ are variation independent.

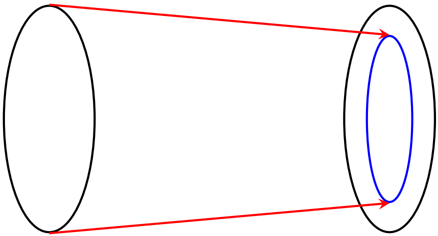
# Polytopes may not intersect




$\Rightarrow$ Model places testable constraints on $P(X, Y \mid Z)$.

# Model for $P(X, Y \mid Z)$ (Z binary)

*Counterfactual* *Observed*



$$\Delta_{15} \quad \to \quad \Delta_3 \times \Delta_3$$

$$P(\{X(z), Y(x)\}) \quad \mapsto \quad (P(X, Y \mid Z=0), P(X, Y \mid Z=1))$$

# Model for observables

For $Z$ binary requiring that the polytopes intersect leads to the following:
If $p(X, Y \mid Z)$ is compatible with the binary IV model iff

$$\begin{aligned}
p(Y{=}0, X{=}0 \mid Z{=}0) + p(Y{=}1, X{=}0 \mid Z{=}1) &\leqslant 1,\\
p(Y{=}0, X{=}0 \mid Z{=}1) + p(Y{=}1, X{=}0 \mid Z{=}0) &\leqslant 1,\\
p(Y{=}0, X{=}1 \mid Z{=}0) + p(Y{=}1, X{=}1 \mid Z{=}1) &\leqslant 1,\\
p(Y{=}0, X{=}1 \mid Z{=}1) + p(Y{=}1, X{=}1 \mid Z{=}0) &\leqslant 1,
\end{aligned}$$

This describes a subset of $\Delta_3 \times \Delta_3$.

These are the IV inequalities of Pearl (1995) and Bonet (2001);
they provide a test of the binary IV model.

# Model for observables

For $Z$ binary requiring that the polytopes intersect leads to the following:
If $p(X, Y \mid Z)$ is compatible with the binary IV model iff

$$\begin{aligned}
p(Y=0, X=0 \mid Z=0) + p(Y=1, X=0 \mid Z=1) &\leqslant 1,\\
p(Y=0, X=0 \mid Z=1) + p(Y=1, X=0 \mid Z=0) &\leqslant 1,\\
p(Y=0, X=1 \mid Z=0) + p(Y=1, X=1 \mid Z=1) &\leqslant 1,\\
p(Y=0, X=1 \mid Z=1) + p(Y=1, X=1 \mid Z=0) &\leqslant 1,
\end{aligned}$$

This describes a subset of $\Delta_3 \times \Delta_3$.

These are the IV inequalities of Pearl (1995) and Bonet (2001); they provide a test of the binary IV model.

Can be interpreted as bounding away from zero $E[Y(z_1, x) - Y(z_0, x)]$, the average direct effect of $Z$ on $Y$, holding $X$ fixed at $x$.

# Model for observables

For $Z$ binary requiring that the polytopes intersect leads to the following:
If $p(X, Y \mid Z)$ is compatible with the binary IV model iff

$$\begin{aligned}
p(Y{=}0, X{=}0 \mid Z{=}0) + p(Y{=}1, X{=}0 \mid Z{=}1) &\leqslant 1,\\
p(Y{=}0, X{=}0 \mid Z{=}1) + p(Y{=}1, X{=}0 \mid Z{=}0) &\leqslant 1,\\
p(Y{=}0, X{=}1 \mid Z{=}0) + p(Y{=}1, X{=}1 \mid Z{=}1) &\leqslant 1,\\
p(Y{=}0, X{=}1 \mid Z{=}1) + p(Y{=}1, X{=}1 \mid Z{=}0) &\leqslant 1,
\end{aligned}$$

This describes a subset of $\Delta_3 \times \Delta_3$.

These are the IV inequalities of Pearl (1995) and Bonet (2001); they provide a test of the binary IV model.

Can be interpreted as bounding away from zero $E[Y(z_1, x) - Y(z_0, x)]$, the average direct effect of $Z$ on $Y$, holding $X$ fixed at $x$. Leads to smooth variation independent parametrization of model (R, Evans & Robins , 2011).

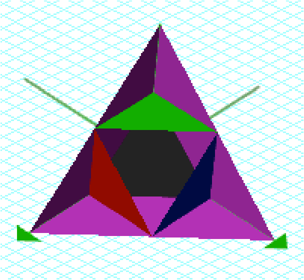
# Model for observables

Hard to visualize this subset of $\Delta_3 \times \Delta_3$.
Can project onto the co-ordinates given by the LHS of the inequalities:

$$u_1 \equiv \sum_j p(y_j, x_0 \mid z_j) \qquad u_2 \equiv \sum_k p(y_k, x_0 \mid z_{1-k})$$

$$u_3 \equiv \sum_j p(y_j, x_1 \mid z_j) \qquad u_4 \equiv \sum_k p(y_k, x_1 \mid z_{1-k})$$

then since $\sum_i u_i = 2$, $(u_1, u_2, u_3, u_4)$ lies in a 3-d simplex with sum 2.
The IV inequalities are $u_i \leqslant 1$ for $i = 1, \ldots, 4$. This gives an octahedron:



# Model for observables

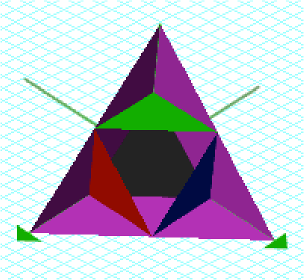
Hard to visualize this subset of $\Delta_3 \times \Delta_3$.
Can project onto the co-ordinates given by the LHS of the inequalities:

$$u_1 \equiv \sum_j p(y_j, x_0 \mid z_j) \qquad u_2 \equiv \sum_k p(y_k, x_0 \mid z_{1-k})$$

$$u_3 \equiv \sum_j p(y_j, x_1 \mid z_j) \qquad u_4 \equiv \sum_k p(y_k, x_1 \mid z_{1-k})$$

then since $\sum_i u_i = 2$, $(u_1, u_2, u_3, u_4)$ lies in a 3-d simplex with sum 2.

The IV inequalities are $u_i \leqslant 1$ for $i = 1, \ldots, 4$. This gives an octahedron:



**Corollary:** At most one inequality is ever violated. (see also Cai, Kuroki, Pearl, Tian, 2008)

# Extending to $p$ levels of $Z$: Observable model

Explicitly characterizing the set of observed distributions $P(X, Y|Z)$ does require more work. Three known types of constraint:

- $2p(p-1)$ IV inequalities, these involve 2 levels of $Z$;
- $4p(p-1)(p-2)$ inequalities found by Bonet (2001), these involve 3 levels of $Z$
- $p!/(p-4)!$ inequalities found by R, which involve 4 levels of $Z$.

**Conjecture**: These are the only constraints.
Confirmed by direct computation for $p \leqslant 7$.

(Note: Implicit solution given by seeing if the set of distributions $P(Y(x_0), Y(x_1))$ is non-empty.)

# Bayesian Inference

What about sampling variability?

The true population distribution $p(x, y \mid z)$ is not equal to the empirical distribution observed in the sample.

# Bayesian Inference

What about sampling variability?

The true population distribution $p(x, y \mid z)$ is not equal to the empirical distribution observed in the sample.

$\Rightarrow$ Simple Bayesian approach (Pearl, 2000):
put a (Dirichlet) prior on the set of distributions:

$$\{P(X(z_0), X(z_1), Y(x_0), Y(x_1))\}.$$

# Bayesian Inference

What about sampling variability?

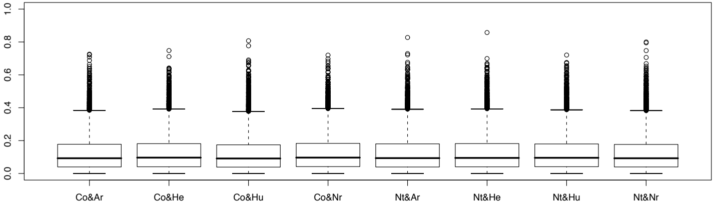
The true population distribution $p(x, y \mid z)$ is not equal to the empirical distribution observed in the sample.

$\Rightarrow$ Simple Bayesian approach (Pearl, 2000):
put a (Dirichlet) prior on the set of distributions:

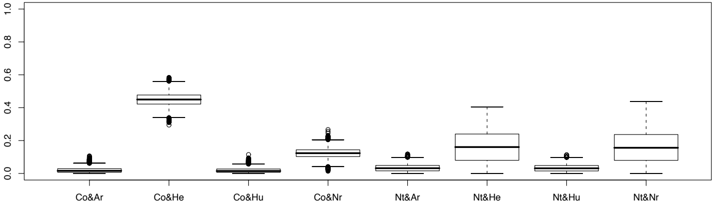
$$\{P(X(z_0), X(z_1), Y(x_0), Y(x_1))\}.$$

$\Rightarrow$ Use MCMC to sample from posterior distribution.

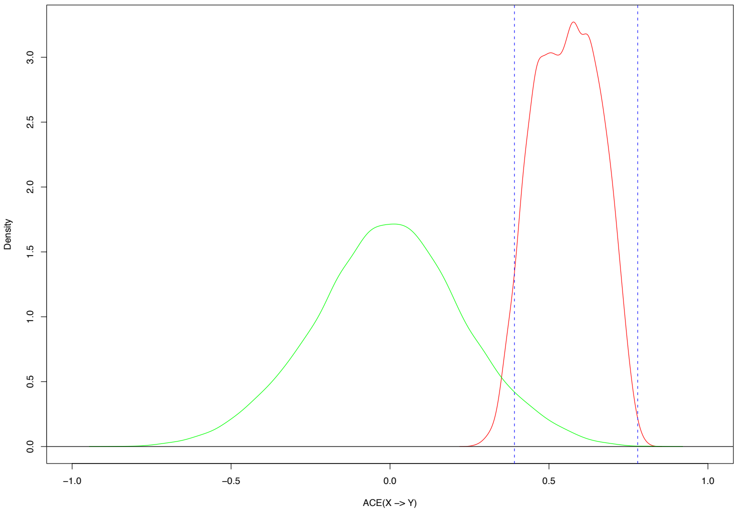
**Marginal Prior distributions; Dir(1,...,1)**

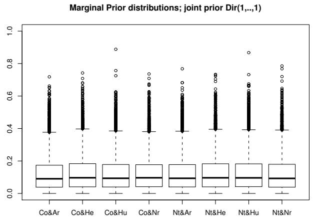


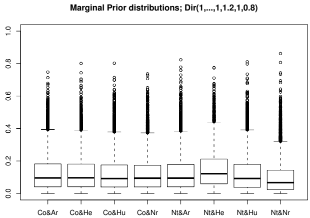

**Marginal Posterior distributions from Lipid Data; Prior Dir(1,...,1)**

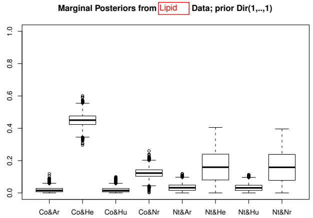


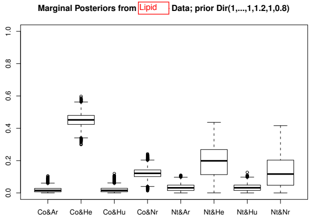
Prior Dir(1,...,1) and Posterior for ACE(X–>Y) from Lipid Data



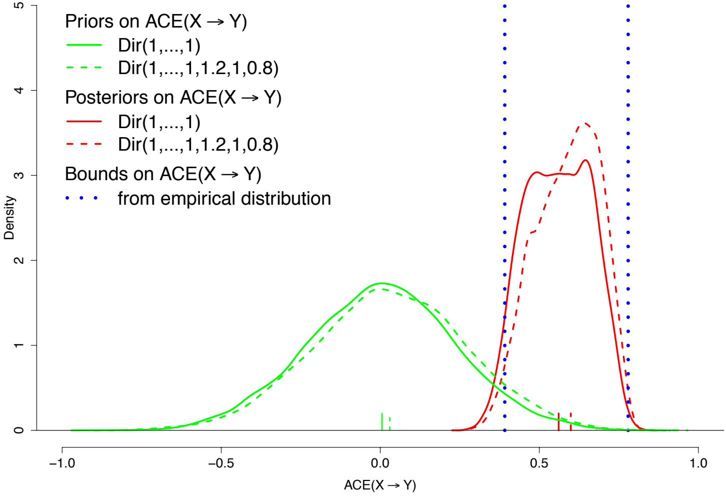

Marginal Prior distributions; joint prior Dir(1,..,1)
0.0
0.2
0.4
0.6
0.8
1.0
Co&Ar
Co&He
Co&Hu
Co&Nr
Nt&Ar
Nt&He
Nt&Hu
Nt&Nr


Marginal Prior distributions; Dir(1,...,1,1.2,1,0.8)
0.0
0.2
0.4
0.6
0.8
1.0
Co&Ar
Co&He
Co&Hu
Co&Nr
Nt&Ar
Nt&He
Nt&Hu
Nt&Nr


Marginal Posteriors from Lipid Data; prior Dir(1,..,1)
0.0
0.2
0.4
0.6
0.8
1.0
Co&Ar
Co&He
Co&Hu
Co&Nr
Nt&Ar
Nt&He
Nt&Hu
Nt&Nr


Marginal Posteriors from Lipid Data; prior Dir(1,...,1,1.2,1,0.8)
0.0
0.2
0.4
0.6
0.8
1.0
Co&Ar
Co&He
Co&Hu
Co&Nr
Nt&Ar
Nt&He
Nt&Hu
Nt&Nr

# Prior and posterior on ACE(X → Y) for Lipid data
## Uniform & perturbed uniform priors on potential outcomes



# Is the problem caused by the priors?

Try a ‘unit’ information prior:

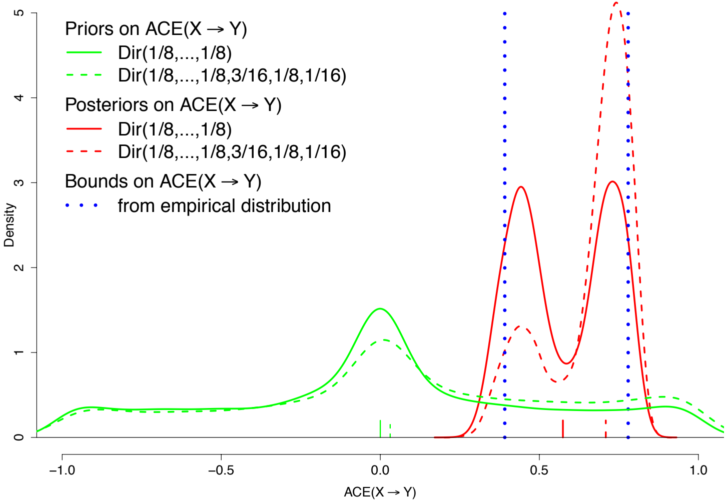
$$p(\{X(z), Y(x)\}) \sim \text{Dir}(1/8, \ldots, 1/8)$$

vs.

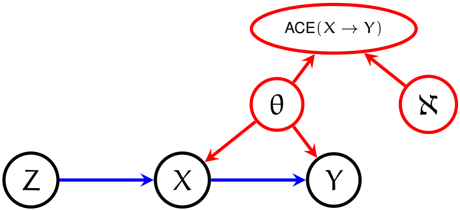
$$p(\{X(z), Y(x)\}) \sim \text{Dir}(1/8, \ldots, 1/8, 3/16, 1/8, 1/16)$$

(?Though ‘unit’ of whose information?)

## Prior and posterior on ACE(X → Y) for Lipid data
### Unit & perturbed unit priors on potential outcomes



# Big picture: How to formulate priors transparently




We will re-parameterize:

$$\{P(X(z_0), X(z_1), Y(x_0), Y(x_1))\} \quad \leftrightarrow \quad (\theta, \aleph)$$

$\theta$ is a 6 dim. parameter, (completely!) identifiable from $P(X, Y \mid Z)$.
$\aleph$ is a 9 dim. parameter, (completely!) non-identifiable.

$$P(\theta, \aleph) = P(\theta)P(\aleph)$$
$$P(\theta, \aleph \mid Z, X, Y) = P(\theta|Z, X, Y)p(\aleph)$$

Note that $\aleph \perp\!\!\!\perp Z, X, Y$

# Simple implementation (I)

Recall that the binary IV model is defined by the inequalities:

$$
\begin{aligned}
P(Y=0, X=0 \mid Z=0) + P(Y=1, X=0 \mid Z=1) &\leqslant 1,\\
P(Y=0, X=0 \mid Z=1) + P(Y=1, X=0 \mid Z=0) &\leqslant 1,\\
P(Y=0, X=1 \mid Z=0) + P(Y=1, X=1 \mid Z=1) &\leqslant 1,\\
P(Y=0, X=1 \mid Z=1) + P(Y=1, X=1 \mid Z=0) &\leqslant 1.
\end{aligned} \tag{1}
$$

*Prior*: Dirichlet on $P(X, Y \mid Z)$ restricted (and re-normalized) to those distributions obeying (1).

*Posterior*: The usual Dirichlet posterior, again restricted to those distributions obeying (1).

Inference may be performed by 'straight' Monte-Carlo.

# Simple Implementation (II)

To obtain the posterior distribution on the ACE bounds, perform the following steps:

1. Specify Dirichlet $(\alpha_{00z}, \alpha_{01z}, \alpha_{10z}, \alpha_{11z})$ priors on $p(x, y|z)$ for $z = 0, 1$.
2. Compute the posteriors in the usual way:
   Dirichlet $(\alpha_{00z} + n_{00z}, \alpha_{01z} + n_{01z}, \alpha_{10z} + n_{10z}, \alpha_{11z} + n_{11z})$
   where $n_{ijz}$ is the number of observations with $X = i$, $Y = j$, $Z = z$.
3. Simulate $p^{(1)}(x, y|z), \ldots, p^{(N)}(x, y|z)$ from this posterior.
4. Throw out any $p^{(i)}(x, y|z)$ violating the inequalities (1).
5. Compute upper and lower bounds on the ACE from each distribution $p^{(i)}(x, y|z)$ remaining after step 4.

# Back to Lipid Data

Parametrize IV model as:

$$p(X, Y \mid Z) \text{ obeying IV inequalities}$$

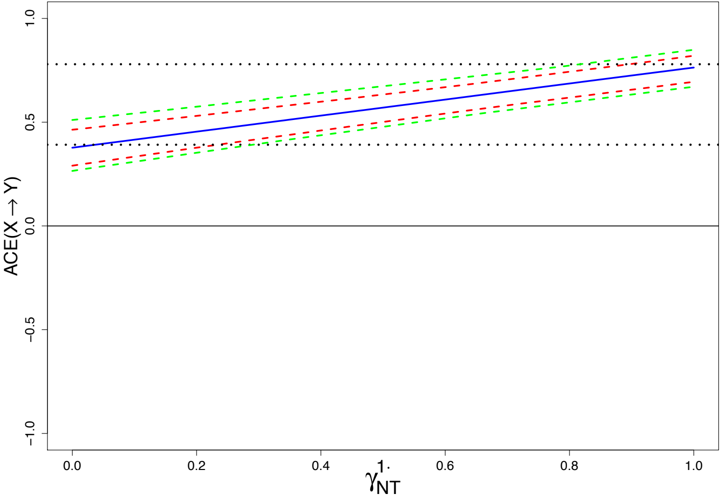
Use $\text{Dir}(1,1)$ for $p(x = 0, y \mid z = 0)$ (since $p(x = 1, y \mid z = 0) = 0$)

and $\text{Dir}(1,1,1,1)$ for $p(x, y \mid z = 1)$.

(Posterior probability that IV model holds: 0.368; prior probability was 0.5.)

We then truncate and renormalize using the IV inequalities.

ACE(X → Y) as a function of $\gamma_{NT}^{1\cdot}$




Here $\gamma_{NT}^{1\cdot} \equiv P(Y(x_1) = 1 \mid NT)$, the probability of a good outcome for Never Takers, if they were to get the drug; this is completely unidentified.

# Summary

- Potential outcome models provide a way to formulate causal models in terms of missing data
- Geometric analysis of randomized and observational studies
- Instrumental variable design may be seen as combining observational studies
- Non-parametric IV models impose testable constraints
- These constraints lead to a *transparent* method of Bayesian analysis.

**Thank you!**

# What if I don't believe (deterministic) potential outcomes exist?

- Dawid (2003), Cai *et al.* (2008) show that IV inequalities may be derived without assuming existence of deterministic potential outcomes, likewise for ACE bounds

  $\Rightarrow$ multivariate link functions for GLMs are also useful even under probabilistic potential outcome models.

Similar 'coupling' arguments may also be given for the generalizations to more states for $X$ and $Z$.

# Extending to $p$ levels of $Z$, $q$ levels of $X$: identification

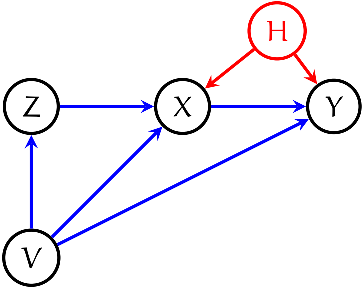
**Conjecture**: The set of distributions $P(Y(x_1), \ldots, Y(x_q))$ compatible with $P(X, Y|Z)$ is given by the following constraints: For every non-empty subset $A \subseteq \{1, \ldots, q\}$, every assignment $(y^i, i \in A)$, with $y^i \in \{0, 1\}^{|A|}$ and every $z \in \{1, \ldots, p\}$:

$$0 \leqslant P(\{Y(x_i) = y^i, i \in A\}) \leqslant \sum_{j \in A} p(Y = y^j, X = j \mid z) + \sum_{k \notin A,} p(X = k \mid z) \tag{2}$$

*Necessity is established. Sufficiency has been proved up to $p \leqslant 7$; sketch of proof in general.*

# Incorporating baseline covariates

May wish to model ACE$(X \rightarrow Y)$ as a function of possibly continuous baseline covariates.

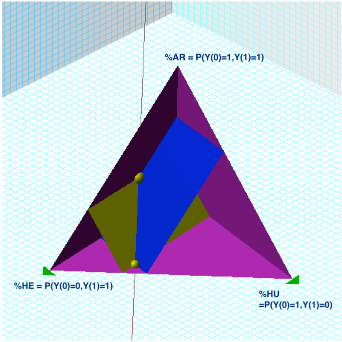


# Generalized Linear Models

Need: a multivariate link function (aka diffeomorphism) from $\mathbb{R}^6$ into subset of $\Delta_3 \times \Delta_3$ given by the inequalities defining the IV model

$\Rightarrow$ incorporate covariates as in a multivariate GLM.

# Fréchet inequalities




Equation for line segment in simplex:

$$\left\{\begin{array}{lcl} P(1,1) & = & {\color{red}t} \\ P(1,0) & = & c_0 - {\color{red}t} \\ P(0,1) & = & c_1 - {\color{red}t} \\ P(0,0) & = & 1 - c_0 - c_1 + {\color{red}t} \end{array} \quad \begin{array}{r} {\color{red}t} \in \left[\max\{0, (c_0 + c_1) - 1\}, \min\{c_0, c_1\}\right] \\ c_0 \equiv P(Y=1 \mid x_0) \\ c_1 \equiv P(Y=1 \mid x_1) \end{array}\right\}$$

Extreme points are given by ‘Fréchet inequalities’.