# A linear optimization approach to decision support with CANS data

David Offner [1] Dan Warner [2] Cara Yi [1] Wani Zhang [1]

[1] Carnegie Mellon University

[2] Community Data Roundtable

CDR Data Talk
August 18, 2022

# Research question

Given a CANS assessment for a new patient, suggest whether IBHS or FBMHS would be more likely to lead to improvement.

# Research question

Given a CANS assessment for a new patient, suggest whether IBHS or FBMHS would be more likely to lead to improvement.

**Goal:** Use historical data to rate each CANS item.

- INPUT: Historical data for 5,894 patients
  - 53 (top-level) items from CANS assessment
  - Initial and final assessments
- OUTPUT: A score between -1 (IBHS) and 1 (FBMHS) for each item

# Research question

Given a CANS assessment for a new patient, suggest whether IBHS or FBMHS would be more likely to lead to improvement.

**Goal:** Use historical data to rate each CANS item.

- INPUT: Historical data for 5,894 patients
  - 53 (top-level) items from CANS assessment
  - Initial and final assessments
- OUTPUT: A score between -1 (IBHS) and 1 (FBMHS) for each item

Example:

- Sensory/Motor Functioning -0.70
- Intentional Misbehavior 0.06
- Attachment 0.86

# Use in decision support

**Toy example:** CANS Assessment

| Patient | Sensory/Motor Funct.(-.70) | Int. Misbehavior(.06) | Attachment(.86) |
|---|---|---|---|
| Patient 1 | 2 | 3 | 0 |
| Patient 2 | 0 | 1 | 2 |

# Use in decision support

**Toy example:** CANS Assessment

| Patient | Sensory/Motor Funct.(-.70) | Int. Misbehavior(.06) | Attachment(.86) |
|---|---|---|---|
| Patient 1 | 2 | 3 | 0 |
| Patient 2 | 0 | 1 | 2 |

Recommendation:

- Patient 1: $2(-.70) + 3(0.06) + 0(0.86) = -1.22 \to$ IBHS
- Patient 2: $0(-.70) + 1(0.06) + 2(0.86) = 1.78 \to$ FBMHS

# Use in decision support

**Toy example:** CANS Assessment

| Patient | Sensory/Motor Funct.(-.70) | Int. Misbehavior(.06) | Attachment(.86) |
|---|---|---|---|
| Patient 1 | 2 | 3 | 0 |
| Patient 2 | 0 | 1 | 2 |

Recommendation:

- Patient 1: $2(-.70) + 3(0.06) + 0(0.86) = -1.22 \rightarrow$ IBHS
- Patient 2: $0(-.70) + 1(0.06) + 2(0.86) = 1.78 \rightarrow$ FBMHS

**Importance/interest**

- Novel way to do decision support
- Emphasis on specific items

# Linear optimization

We use linear optimization to produce the item scores.

# Linear optimization

We use linear optimization to produce the item scores.

Examples of previous work:

- HS Leff, S. Graves, J. Natkins, J. Bryan. A system for allocating mental health resources, *Administration in Mental Health* Vol. 13, No. 1, Fall 1985
- A. Miniguano-Trujillo, F. Salazar, R. Torres, P. Arias, K. Sotomayor. An integer programming model to assign patients based on mental health impact for tele-psychotherapy intervention during the Covid-19 emergency, *Health Care Manag Sci.* 2021 Jun;24(2):286–304.

# Linear optimization

Given linear **constraints**, we want to maximize/minimize a linear **objective function**.

# Linear optimization

Given linear **constraints**, we want to maximize/minimize a linear **objective function**.

E.g.

$$\begin{array}{rrrrl} \max & 4x_1 & -5x_2 & +2x_3 & \\ \text{s.t.} & 2x_1 & -x_2 & +x_3 & = 4 \\ & 4x_1 & -2x_2 & +4x_3 & \leq 9 \end{array}$$

$$x_1, x_2, x_3 \geq 0$$

# Linear optimization

Given linear **constraints**, we want to maximize/minimize a linear **objective function**.

E.g.

$$\begin{array}{rrrrl} \max & 4x_1 & -5x_2 & +2x_3 & \\ \text{s.t.} & 2x_1 & -x_2 & +x_3 & = 4 \\ & 4x_1 & -2x_2 & +4x_3 & \leq 9 \end{array}$$

$$x_1, x_2, x_3 \geq 0$$

We have powerful tools for finding optimal solutions to LOPs (even with 1000s of variables and constraints!).

# Notation for assessments and improvements

Given historical data

- A set of **patients** $P$ and
- A set of **items** $N$
- **Initial** and **final** assessments for every patient $p \in P$.

# Notation for assessments and improvements

Given historical data

- A set of **patients** $P$ and
- A set of **items** $N$
- **Initial** and **final** assessments for every patient $p \in P$.

Define

- $A_0(n, p) =$ **initial** assessment of item $n$ for patient $p$.

# Notation for assessments and improvements

Given historical data

- A set of **patients** $P$ and
- A set of **items** $N$
- **Initial** and **final** assessments for every patient $p \in P$.

Define

- $A_0(n, p) =$ **initial** assessment of item $n$ for patient $p$.
- $A_1(n, p) =$ **final** assessment of item $n$ for patient $p$.
  - When we have $> 2$ assessments, we use first and last.

# Notation for assessments and improvements

Given historical data

- A set of **patients** $P$ and
- A set of **items** $N$
- **Initial** and **final** assessments for every patient $p \in P$.

Define

- $A_0(n, p)$ = **initial** assessment of item $n$ for patient $p$.
- $A_1(n, p)$ = **final** assessment of item $n$ for patient $p$.
  - When we have $> 2$ assessments, we use first and last.
- $I(p)$ = **cumulative improvement** for patient $p$

$$I(p) = \sum_{n \in N} A_0(n, p) - A_1(n, p)$$

# Example of cumulative improvement

Say we have a patient $p$ with the following items:

| Item | Initial Assessment($A_0$) | Final Assessment($A_1$) | Improvement |
|---|---|---|---|
| 1 | 3 | 2 | $3-2=1$ |
| 2 | 1 | 2 | $1-2=-1$ |
| 3 | 2 | 0 | $2-0=2$ |
| 4 | 1 | 1 | $1-1=0$ |

The cumulative improvement, $I(p)$, is: $1-1+2+0=2$.

# Notation for item scores

Define variables

- $x_n$ for each item $n \in N$

# Notation for item scores

Define variables

- $x_n$ for each item $n \in N$
  - Constraint: $-1 \leq x_n \leq 1$

# Notation for item scores

Define variables

- $x_n$ for each item $n \in N$
  - Constraint: $-1 \leq x_n \leq 1$
- $x_p$ for each patient $p \in P$

# Notation for item scores

Define variables

- $x_n$ for each item $n \in N$
  - Constraint: $-1 \leq x_n \leq 1$
- $x_p$ for each patient $p \in P$
  - Sum of the cumulative improvements, weighted by (variable) item score
  - $x_p = \sum_{n \in N} x_n \cdot (A_0(n,p) - A_1(n,p))$

# Objective function

Define

- $T(p)$: the **actual treatment** received by patient $p$
  - If $p$ received FBMHS, $T(p) = 1$. If $p$ received IBHS, $T(p) = -1$.
- $|IB|$: total number of patients who went to IBHS
- $|FB|$: total number of patients who went to FBMHS

# Objective function

Define

- $T(p)$: the **actual treatment** received by patient $p$
  - If $p$ received FBMHS, $T(p) = 1$. If $p$ received IBHS, $T(p) = -1$.
- $|IB|$: total number of patients who went to IBHS
- $|FB|$: total number of patients who went to FBMHS

The objective is to maximize

$$\frac{1}{|FB|} \sum_{p \in FB} x_p T(p) I(p) + \frac{1}{|IB|} \sum_{p \in IB} x_p T(p) I(p)$$

# Objective function

Define

- $T(p)$: the **actual treatment** received by patient $p$
  - If $p$ received FBMHS, $T(p) = 1$. If $p$ received IBHS, $T(p) = -1$.
- $|IB|$: total number of patients who went to IBHS
- $|FB|$: total number of patients who went to FBMHS

The objective is to maximize

$$\frac{1}{|FB|}\sum_{p\in FB} x_p T(p) I(p) + \frac{1}{|IB|}\sum_{p\in IB} x_p T(p) I(p)$$

Because $T(p) = 1$ for FBMHS, and $T(p) = -1$ for IBHS, this simplifies to

$$\frac{1}{|FB|}\sum_{p\in FB} x_p I(p) - \frac{1}{|IB|}\sum_{p\in IB} x_p I(p)$$

# Objective function example

Say we have 4 patients $p_1, p_2, p_3, p_4$ with 3 items, so we have 3 decision variables $x_1, x_2, x_3$

| Patient | T(p) | I(item$_1$) | I(item$_2$) | I(item$_3$) | $I(p)$ |
|---|---|---|---|---|---|
| $p_1$ | 1 | 3 | 2 | 1 | $3+2+1=6$ |
| $p_2$ | 1 | 1 | 2 | 0 | $1+2+2=3$ |
| $p_3$ | $-1$ | $-1$ | 3 | 0 | $-1+3+0=2$ |
| $p_4$ | $-1$ | 1 | $-1$ | 3 | $1-1+3=3$ |

The objective is to maximize

$$\frac{1}{2}[\mathbf{(6)}(3x_1 + 2x_2 + x_3) + \mathbf{(3)}(x_1 + 2x_2 + 0)] -$$

$$\frac{1}{2}[\mathbf{(2)}(-2x_1 + x_2 + 0) + \mathbf{(3)}(x_1 - x_2 - 3x_3)]$$

# Full statement of the linear optimization problem

$$\max \quad \frac{1}{|FB|}\sum_{p\in FB} x_p I(p) - \frac{1}{|IB|}\sum_{p\in IB} x_p I(p)$$

$$\text{s.t.} \quad x_p = \sum_{n\in N} x_n \cdot (A_0(n,p) - A_1(n,p)) \qquad \text{for all } p \in P$$

$$-1 \le x_n \le 1 \qquad \text{for all } n \in N$$

# Subdividing data and averaging item scores

When we ran the previous model once on the full data set, the item scores are all 1 or -1.

# Subdividing data and averaging item scores

When we ran the previous model once on the full data set, the item scores are all 1 or -1.

To get more nuanced scores we

- Randomly divided the historical data into 100 groups
- Ran the model on each group
- Averages each item score

# Subdividing data and averaging item scores

When we ran the previous model once on the full data set, the item scores are all 1 or -1.

To get more nuanced scores we

- Randomly divided the historical data into 100 groups
- Ran the model on each group
- Averages each item score

We then repeated this 60 times and then averaged those scores.

# Validation

We validated the model using "dummy data"

- 200 patients
- Items 1-5 always improve with IBHS, never with FBMHS
- Items 6-10 always improve with FBMHS, never with IBHS
- Items 11-100 change randomly

# Validation

We validated the model using "dummy data"

- 200 patients
- Items 1-5 always improve with IBHS, never with FBMHS
- Items 6-10 always improve with FBMHS, never with IBHS
- Items 11-100 change randomly

| | Expected/Ideal $x_n$ Values | Model's Values |
|---|---|---|
| Items 1-5 | All $-1$ | All $-1$ |
| Items 6-10 | All 1 | All 1 |
| Items 11-90 | Between $-1$ and 1, close to 0 | All between $-1$ and 1 |

67.5% of items 11-90 are between -0.2 and 0.2.

# Results: Item scores for CANS data

| Problem Presentation | |
|---|---|
| **Item** | **Score** |
| Psychosis | 0.329 |
| Attention Deficit/Impulse | 0.339 |
| Autism Spectrum | -0.394 |
| Depression/Anxiety | 0.744 |
| Oppositional Behavior | 0.388 |
| Antisocial Behavior | 0.340 |
| Anger Control | 0.369 |
| Substance Abuse | 0.270 |
| Adjustment to Trauma | 0.745 |
| Attachment | 0.862 |

| Risk | |
|---|---|
| **Item** | **Score** |
| Suicide Risk | 0.415 |
| Danger to Others | -0.006 |
| Other Self Harm | 0.366 |
| Runaway | -0.136 |
| Exploitation | 0.487 |
| Sexual Aggression | 0.123 |
| Intentional Misbehavior | 0.063 |
| Delinquency | 0.370 |
| Firearms Risk | 0.148 |
| Fire setting | 0.002 |

| Functioning | |
|---|---|
| **Item** | **Score** |
| Intellectual Disability | 0.019 |
| Physical/Medical | -0.117 |
| Sleep | 0.010 |
| Family Functioning | 0.585 |
| Living Situation | 0.298 |
| Social Functioning - Peer | 0.061 |
| School Achievement | 0.420 |
| School Behavior | -0.055 |
| School Attendance | 0.513 |
| Sexual Development | 0.449 |
| Sensory/Motor Functioning | -0.703 |
| Communication | -0.688 |
| Maladaptive Behaviors | -0.706 |

| Caregiver | |
|---|---|
| **Item** | **Score** |
| Safety | 0.719 |
| Physical/Behavioral Health | 0.459 |
| Supervision | 0.647 |
| Involvement | 0.617 |
| Knowledge | 0.691 |
| Organization | 0.502 |

| Strengths | |
|---|---|
| **Item** | **Score** |
| Resources | 0.508 |
| Residential Stability | 0.368 |
| Family | 0.592 |
| Interpersonal | 0.045 |
| Relationship Permanence | 0.537 |
| Educational | 0.037 |
| Vocational | 0.028 |
| Well-Being | 0.153 |
| Optimism | 0.104 |
| Spiritual/Religious | -0.396 |
| Talents/Interests | -0.004 |
| Inclusion | 0.145 |
| Resiliency | 0.157 |
| Resourcefulness | 0.136 |

# Notable items

**Most positive items (improve with FBMHS):**

- Attachment 0.86
- Adjustment to Trauma 0.75
- Depression/Anxiety 0.74
- Safety 0.72
- Knowledge 0.69

**Most negative items (improve with IBHS):**

- Maladaptive Behaviors -0.71
- Sensory/Motor Functioning -0.70
- Communication -0.69
- Spiritual/Religious -0.40
- Autism Spectrum -0.39

# Using scores for decision support

**Decision Output**

For a treatment recommendation for a new patient, each item assessment (0 to 3) of the patient is multiplied with the item score (-1 to 1) computed by the model.

If the sum across all items is

- larger than a threshold $\rightarrow$ FBMHS
- smaller than the threshold $\rightarrow$ IBHS

# Using scores for decision support

**Decision Output**

For a treatment recommendation for a new patient, each item assessment (0 to 3) of the patient is multiplied with the item score (-1 to 1) computed by the model.

If the sum across all items is

- larger than a threshold$\rightarrow$ FBMHS
- smaller than the threshold$\rightarrow$ IBHS

New patient data: 14,510 patients who have been sent to IBHS, and 3,527 patients who have been sent to FBMHS.

# Scores for IBHS patients

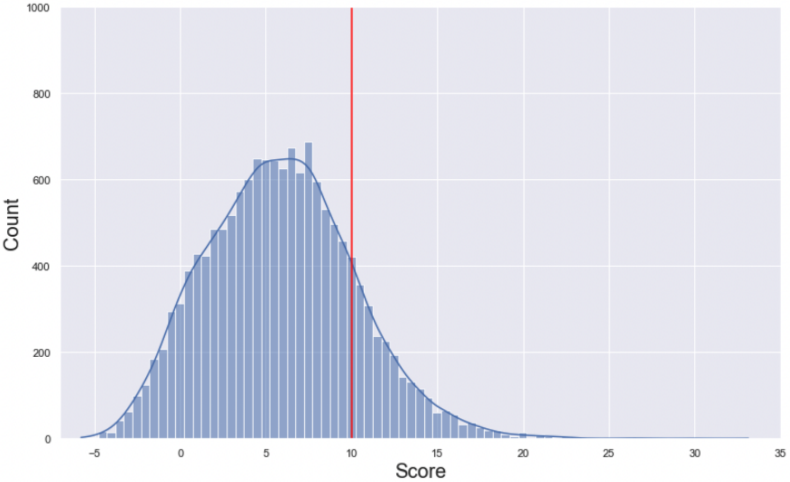



Figure: Frequency of total scores for patients who have received IBHS

# Scores for FBMHS patients

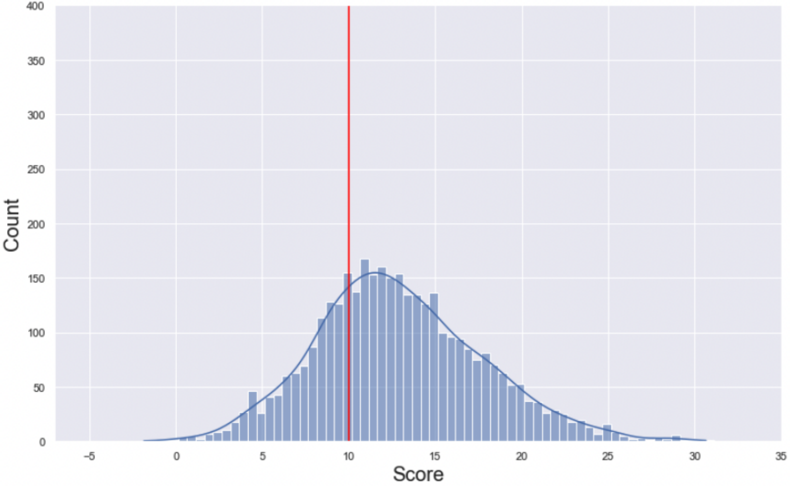
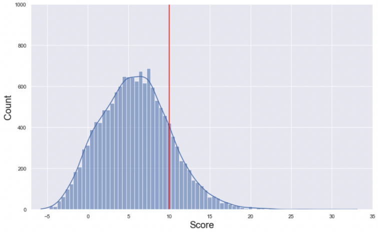


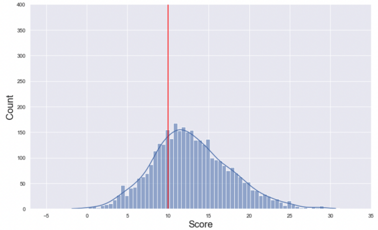

Figure: Frequency of total scores for patients who have received FBMHS

# IBHS FBMHS comparison

(a) Patient Score for IBHS patients




(b) Patient Score for FBMHS patients



# Technical details

Tools used for cleaning and organizing data and building and running the model.

- Python
- Jupyter
- Pandas
- Numpy
- Gurobi
- Seaborn
- Excel

# Thank you!


- David Offner
  CMU Mathematical Sciences
  doffner@andrew.cmu.edu
- Cara Yi
  CMU Mathematical Sciences '22 → CMU Heinz School
  fyi@andrew.cmu.edu
- Wani Zhang
  CMU Mathematical Sciences '22 → London Business School
  wzhang@london.edu