Jason Abaluck

optimal Representation

Ephraim Sutherland

Setup

1. Suppose a physician can only see see ATE and some measure of representativeness. They have prior $\bar{\beta}$ and $\beta_{ATE} = (1/N)\sum \beta_i$.

2. need model for betas related to each other based on x's. WLOG, suppose

$$\beta(x_i) = x_i\gamma$$

   Where x_i is a vector of characteristics and γ is a vector of coefficients.
   If you know γ, then you know β for any given patient.

3. However, you don't observe γ, you instead observe: $\beta_{ATE} = \bar{x}\gamma$ where $\bar{x} = (\frac{1}{N})\sum x_i$

4. We know β_i for patients with characteristics $\bar{x}$ (it is β_{ATE}).

5. For other patients, need to solve

$$\beta_{i,post} = \mathrm{E}(x_i\gamma|\bar{x}\gamma = \beta_{ATE})$$

6. to solve

   (a)

$$\begin{aligned}\beta_{i,post} &= \mathrm{E}(x_i\gamma|\bar{x}\gamma = \beta_{ATE})\\ &= \mathrm{E}((x_i - c_i\bar{x})\gamma|\bar{x}\gamma = \beta_{ATE}) + c_i\mathrm{E}(\bar{x}\gamma|\bar{x}\gamma = \beta_{ATE})\\ &= \mathrm{E}((x_i - c_i\bar{x})\gamma|\bar{x}\gamma = \beta_{ATE}) + c_i\beta_{ATE}\end{aligned}$$

   For any constant c_i.
   Choose c_i so that

$$\mathrm{Cov}((x_i - c_i\bar{x})\gamma, \bar{x}\gamma) = 0$$

   maybe assume normality so that this guarantees independence. Then,

$$\mathrm{E}((x_i - c_i\bar{x})\gamma|\bar{x}\gamma = \beta_{ATE}) = (x_i - c_i\bar{x})\mathrm{E}(\gamma)$$

   So then

$$(x_i - c_i\bar{x})\mathrm{E}(\gamma) + c_i\beta_{ATE} = x_i\mathrm{E}(\gamma) + c_i(\beta_{ATE} - \bar{x}\mathrm{E}(\gamma))$$

   (c_i depends on x_i)

In other words, your belief is your prior, adjusted based on the difference between the observed ATE and your prior about the ATE. The key question is how much adjustment you do which depends on "c_i". We choose c_i to solve:

$$\text{Cov}((x_i - c_i\bar{x})\gamma, \bar{x}\gamma) = 0$$

$$\iff \text{Cov}(x_i\gamma, \bar{x}\gamma) - c_i\text{Cov}(\bar{x}\gamma, \bar{x}\gamma) = 0$$

$$\iff \text{Cov}(x_i\gamma, \bar{x}\gamma) = c_i\text{Var}(\bar{x}\gamma)$$

$$\iff c_i = \frac{\text{Cov}(\beta_i, \beta_{ATE})}{\text{Var}(\beta_{ATE})}$$

The random variable in this context is γ (the coefficients on the x's) in this case $\text{Var}(\beta_{ATE})$ is a measure of how uncertain one was about what β_{ATE} would be before doing the trial.

c_i is the equation for a regression of β_i on β_{ATE}. In other words, we take a bunch of patients with characteristics x_i and we keep redrawing the gammas from our prior distribution the we ask how correlated β_i and β_{ATE} are. If they are more correlated (as they would be for patients where the x_i are closer to $\bar{x}$ we update more.

To compute c_i, we just need to know x_i $\bar{x}$, and the distribution of γ.

Suppose we want to design the trial to minimize:

$$\min \text{E}[(\beta_i - \beta_{i,post})^2]$$

Simple Cases

1. There is just one x and it is binary (old v young). Can it be solved analytically?
2. Can you solve a 2-dimensional case?

Solving

we want to solve

$$\min \text{E}[(\beta_i - \beta_{i,post})^2]$$

let's start by considering

$$\begin{aligned}
\text{E}[(\beta_i - \beta_{i,post})^2] &= \text{E}([x'\gamma - ((x - c\bar{x})\text{E}(\gamma) + c\beta_{ATE})]^2) \\
&= \text{E}([x'\gamma - x\text{E}(\gamma) + c\bar{x}\text{E}(\gamma) - c\bar{x}\gamma]^2) \\
&= \text{E}([x'(\gamma - \text{E}(\gamma)) - c\bar{x}(\gamma - \text{E}(\gamma))]^2) \\
&= \text{E}([(x' - c\bar{x})(\gamma - \text{E}(\gamma))]^2)
\end{aligned}$$

let $A = (x' - c\bar{x})$ and $B = (\gamma - \mathrm{E}(\gamma))$. Now if A and B are $1x1$ (i.e. there is no intercept, then:

$$= \mathrm{E}[(x' - c\bar{x})^2(\gamma - \mathrm{E}(\gamma))^2]$$

$$= \mathrm{E}[(x'x - 2c\bar{x}x + (c\bar{x})^2)(\gamma - \mathrm{E}(\gamma))^2]$$

$$= \mathrm{E}[(x^2(\gamma - \mathrm{E}(\gamma))^2] - \mathrm{E}[(2c\bar{x}x(\gamma - \mathrm{E}(\gamma))^2] + \mathrm{E}[((c\bar{x})^2)\bar{x}x(\gamma - \mathrm{E}(\gamma))^2]$$

but if A and B are not scalars but we instead allow for another arbitrary characteristic (including an intercept)

$$= \mathrm{E}[(A)^2(B)^2] + 2 * \Pi_i(A_i B_i)$$

$$= \mathrm{E}[(A)^2(\gamma - \mathrm{E}(\gamma))^2]$$

$$= \mathrm{E}[(x'x - 2c\bar{x}x + (c\bar{x})^2)(\gamma - \mathrm{E}(\gamma))^2]$$

$$= \mathrm{E}[(x^2(\gamma - \mathrm{E}(\gamma))^2] - \mathrm{E}[(2c\bar{x}x(\gamma - \mathrm{E}(\gamma))^2] + \mathrm{E}[((c\bar{x})^2)\bar{x}x(\gamma - \mathrm{E}(\gamma))^2]$$

We must also solve c and make a claim about the distribution of x.
We have that

$$c = \frac{\mathrm{Cov}(x\gamma, \bar{x}\gamma)}{\mathrm{Var}(\bar{x}\gamma)}$$

so if we have a single characteristic and no intercept, then for individual i, c reduces to

$$c = \frac{\mathrm{Cov}(x\gamma, \bar{x}\gamma)}{\mathrm{Var}(\bar{x}\gamma)} = \frac{x\bar{x}\mathrm{Var}(\gamma)}{\bar{x}^2\mathrm{Var}(\gamma)} = \frac{x}{\bar{x}}$$

note that the above formula also captures if we added an intercept. In this case, if we assume $\beta_i = \gamma_1, \gamma_2 x$ and both gammas are drawn from the same distribution and are independent, then the equation would then simply become

$$c = \frac{(1^2 + x\bar{x})\mathrm{Var}(\gamma)}{(1^2 + \bar{x}^2)\mathrm{Var}(\gamma)} = \frac{1 + x\bar{x}}{1 + \bar{x}^2} = \frac{x'\bar{x}}{\bar{x}'\bar{x}}$$

note, this is for individual i so x is given. Thus the only random variable in this context is γ

going back to the objective function, then now we are taking the expectation over all individuals, so we can look at the expected population characteristic as a bernoulli r.v. so $\mathrm{E}(x^2) = \mathrm{E}(x) = p$ and we get

$$\mathrm{E}[(x^2(\gamma - \mathrm{E}(\gamma))^2] - \mathrm{E}[(2c\bar{x}x(\gamma - \mathrm{E}(\gamma))^2] + \mathrm{E}[((c\bar{x})^2)\bar{x}x(\gamma - \mathrm{E}(\gamma))^2]$$

$$= \mathrm{Var}(\gamma)[\mathrm{E}(x) - \mathrm{E}(2c\bar{x}x) + \mathrm{E}((c\bar{x})^2)]$$

In our case, we're assuming $x = [x_0, x_1]$ and $\gamma = [\gamma_0, \gamma_1]$. Ignore dimensions for now. We can take care of proper transposes etc later to make this formally correct. But this implies we get $\mathrm{Var}(\gamma) = (\mathrm{Var}(\gamma_0), \mathrm{Var}(\gamma_1))$ (by assumption of being uncorrelated).

$$\mathrm{Var}(\gamma)[\mathrm{E}(x) - \mathrm{E}(2c\bar{x}x) + \mathrm{E}((c\bar{x})^2)]$$

$$= (\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1)) \left[1 + \mathrm{E}(x) - \frac{2\bar{x}}{1+\bar{x}^2}\mathrm{E}(x + x^2\bar{x}) + \frac{\bar{x}^2}{1+\bar{x}^2}\mathrm{E}(1 + x\bar{x})\right]$$

$$= (\mathrm{Var}(\gamma_0) + \mathrm{Var}(\gamma_1)) \left[1 + \mathrm{E}(x) - \frac{2\bar{x}}{1+\bar{x}^2}\left(\mathrm{E}(x) + \bar{x}\mathrm{E}(x)\right) + \frac{\bar{x}^2}{1+\bar{x}^2} + \bar{x}\frac{\bar{x}^2}{1+\bar{x}^2}\mathrm{E}(x)\right)\right]$$

notes

integrate over the γ 's (taking expecation over possible gammas).

FOC

Let

$$J = \frac{1}{(1+\bar{x}^2)^2}\left[\mathrm{E}(\alpha^2)(\bar{x}^4 - 2\bar{x}^3 + p\bar{x}^2) + \mathrm{E}(\beta^2)(p - 2\bar{x}p + \bar{x}^2) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(2\bar{x}^2p - p\bar{x} - \bar{x}^3)\right]$$

Then taking derivative wrt $\bar{x}$ we get

$$\frac{\partial J}{\partial \bar{x}} = \frac{-4\bar{x}_1}{(1+\bar{x}_1{}^2)^3}\left[\mathrm{E}(\alpha^2)(\bar{x}_1{}^4 - 2\bar{x}_1{}^3 + p\bar{x}_1{}^2) + \mathrm{E}(\beta^2)(p - 2\bar{x}_1p + \bar{x}_1{}^2) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(2\bar{x}_1{}^2p - p\bar{x}_1 - \bar{x}_1{}^3)\right]$$

$$+ \frac{1}{(1+\bar{x}_1{}^2)^2}\left[\mathrm{E}(\alpha^2)(4\bar{x}_1{}^3 - 6\bar{x}_1{}^2 + 2p\bar{x}_1) + \mathrm{E}(\beta^2)(2p + 2\bar{x}_1) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(4\bar{x}_1p - p - 3\bar{x}_1{}^2)\right]$$

$$\iff$$

$$\frac{4\bar{x}_1}{(1+\bar{x}_1{}^2)^3}\left[\mathrm{E}(\alpha^2)(\bar{x}_1{}^4 - 2\bar{x}_1{}^3 + p\bar{x}_1{}^2) + \mathrm{E}(\beta^2)(p - 2\bar{x}_1p + \bar{x}_1{}^2) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(2\bar{x}_1{}^2p - p\bar{x}_1 - \bar{x}_1{}^3)\right]$$

$$= \frac{1}{(1+\bar{x}_1{}^2)^2}\left[\mathrm{E}(\alpha^2)(4\bar{x}_1{}^3 - 6\bar{x}_1{}^2 + 2p\bar{x}_1) + \mathrm{E}(\beta^2)(2p + 2\bar{x}_1) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(4\bar{x}_1p - p - 3\bar{x}_1{}^2)\right]$$

canceling

$$\frac{4\bar{x}_1}{(1+\bar{x}_1{}^2)}\left[\mathrm{E}(\alpha^2)(\bar{x}_1{}^4 - 2\bar{x}_1{}^3 + p\bar{x}_1{}^2) + \mathrm{E}(\beta^2)(p - 2\bar{x}_1p + \bar{x}_1{}^2) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(2\bar{x}_1{}^2p - p\bar{x}_1 - \bar{x}_1{}^3)\right]$$

$$= \left[\mathrm{E}(\alpha^2)(4\bar{x}_1{}^3 - 6\bar{x}_1{}^2 + 2p\bar{x}_1) + \mathrm{E}(\beta^2)(2p + 2\bar{x}_1) + 2\mathrm{E}(\alpha)\mathrm{E}(\beta)(4\bar{x}_1p - p - 3\bar{x}_1{}^2)\right]$$

multiplying out

$$4\alpha^2 x_1^5 - 8x_1^4\alpha^2 + 4x_1^3\alpha^2 p + 4x_1 p\beta^2 - 8x_1^2 p\beta^2 + 4x_1^3\beta^2 + 16\alpha\beta p x_1^3 - 8\alpha\beta p x_1^2 - 8\alpha\beta x_1^4$$

$$= 4\alpha^2 x_1^3 + 4\alpha^2 x_1^5 - 6\alpha^2 x_1^2 - 6x_1^4\alpha^2 + 2p\alpha^2\bar{x_1} + 2p\alpha^2\bar{x_1}^3 + 2\beta^2 p + 2\beta^2 p\bar{x_1}^2 + 2\beta^2\bar{x_1} + 2\beta^2\bar{x_1}^3$$

$$+8\alpha\beta p\bar{x_1} + 8\alpha\beta p\bar{x_1}^3 - 2\alpha\beta p - 2\alpha\beta p\bar{x_1}^2 - 6\alpha\beta\bar{x_1}^2 - 6\alpha\beta\bar{x_1}^4$$

Cancelling

$$-2x_1^4\alpha^2 + 4x_1^3\alpha^2 p + 4x_1 p\beta^2 - 8x_1^2 p\beta^2 + 4x_1^3\beta^2 + 16\alpha\beta p x_1^3 - 6\alpha\beta p x_1^2 - 2\alpha\beta x_1^4$$

$$= 4\alpha^2 x_1^3 - 6\alpha^2 x_1^2 - 6x_1^4\alpha^2 + 2p\alpha^2\bar{x_1} + 2p\alpha^2\bar{x_1}^3 + 2\beta^2 p + 2\beta^2 p\bar{x_1}^2 + 2\beta^2\bar{x_1} + 2\beta^2\bar{x_1}^3$$

$$+8\alpha\beta p\bar{x_1} + 8\alpha\beta p\bar{x_1}^3 - 2\alpha\beta p - 6\alpha\beta\bar{x_1}^2$$

more cancelling

$$-2x_1^4\alpha^2 + 4x_1^3\alpha^2 p + 4x_1 p\beta^2 - 8x_1^2 p\beta^2 + 4x_1^3\beta^2 + 16\alpha\beta p x_1^3 - 6\alpha\beta p x_1^2 - 2\alpha\beta x_1^4$$

$$= 4\alpha^2 x_1^3 - 6\alpha^2 x_1^2 - 6x_1^4\alpha^2 + 2p\alpha^2\bar{x_1} + 2p\alpha^2\bar{x_1}^3 + 2\beta^2 p + 2\beta^2 p\bar{x_1}^2 + 2\beta^2\bar{x_1} + 2\beta^2\bar{x_1}^3$$

$$+8\alpha\beta p\bar{x_1} + 8\alpha\beta p\bar{x_1}^3 - 2\alpha\beta p - 6\alpha\beta\bar{x_1}^2$$

reduces to

$$(x_1^4)(a_2 p + a_1 b_1) + (x_1^3)(2a_2 - a_2 p - b_2 - 8a_1 b_1 p) +$$

$$(x_1^2)(3b_2 p + 7a_1 b_1 p - 3a_2 p - 3a_1 b_1) + x_1(pa_2 - 2b_2 p + b_2 + 4a_1 b_1 p) - (b_2 p + a_1 b_1 p)$$

Now $a_2 = \mathrm{Var}(\gamma_0) = 1 = B_2 = \mathrm{Var}(\gamma_1)$ and $a_1 = \mathrm{E}(\gamma_0 - E(\gamma_0)) = 0 = b_1 = (\gamma_1 - E(\gamma_1))$ so we further reduce to

$$(x_1^4)(\mathrm{Var}(\gamma_0)p) + (x_1^3)(2\mathrm{Var}(\gamma_0) - \mathrm{Var}(\gamma_0)p - \mathrm{Var}(\gamma_1)) +$$

$$(x_1^2)(3p\mathrm{Var}(\gamma_1) - 3\mathrm{Var}(\gamma_0)p) + x_1(p\mathrm{Var}(\gamma_0) - 2\mathrm{Var}(\gamma_1)p + \mathrm{Var}(\gamma_1)) - (\mathrm{Var}(\gamma_1)p)$$

So we finally get

$$(x_1^4)(p) + (x_1^3)(1-p) +$$

$$+ x_1(1-p) - p$$

solving we get

$$x = \pm\frac{\sqrt{5p^2 - 2p + 1} + p - 1}{2p}$$

1 Symmetric Version

Now we impose symmetric variation on both groups so now we have $\beta_i = \gamma_0 * (1-x) + \gamma_1 * x$ and $\beta_{ate} = \gamma_0 * (1-\bar{x}) + \gamma_1 * \bar{x}$

C still has the same analytical solution.

To solve this, let's make this slightly more general where we have $x = [x_0, x_1]$ and $\bar{x} = [\bar{x_0}, \bar{x_1}]$ as above. In the end we can reduce by letting $x_0 = 1 - x_1$ and $\bar{x_0} = 1 - \bar{x_1}$

our objective function then becomes:

$$
\begin{aligned}
\mathrm{E}_\gamma\left(\mathrm{E}_x[(\beta_i - \beta_{i,post})^2]\right) &= \mathrm{E}_\gamma\left(\mathrm{E}_x([x'\gamma - ((x - c\bar{x})\mathrm{E}_x(\gamma) + c\beta_{ATE})]^2)\right) \\
&= \mathrm{E}_\gamma\left(\mathrm{E}_x[(x' - c\bar{x}')^2(\gamma - \mathrm{E}_x(\gamma))^2] + 2\mathrm{E}_x(\Pi(x' - c\bar{x}')\Pi(\gamma - \mathrm{E}_x(\gamma)))\right) \\
&= \mathrm{E}_x[(x' - c\bar{x}')^2]\mathrm{E}_\gamma[(\gamma - \mathrm{E}_x(\gamma))^2] + 2\mathrm{E}_x(\Pi(x' - c\bar{x}'))\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) \\
&= \mathrm{E}_x[(x' - c\bar{x}')^2]\mathrm{Var}(\gamma) + 2\mathrm{E}_x(\Pi(x' - c\bar{x}'))\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1) \\
&= \left(\frac{1}{\bar{X}'\bar{X}}\right)^2 \mathrm{E}_x\left([X(\bar{X}\bar{X}') - \bar{X}(X\bar{X}'))][X(\bar{X}\bar{X}') - \bar{X}(X\bar{X}'))]'\right)\mathrm{Var}(\gamma) \\
&+ 2\mathrm{E}_x(\Pi(X(\bar{X}\bar{X}') - \bar{X}(X\bar{X}')))\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)
\end{aligned}
$$

Now taking it element wise

$$
\begin{aligned}
&\frac{1}{\bar{x}_1^4 + 2\bar{x}_1^2\bar{x}_2^2 + \bar{x}_2^4}\mathrm{E}_x\Big(\left(x_1^2\bar{x}_2^4 - 2x_1\bar{x}_2^3x_2\bar{x}_1 + x_2^2\bar{x}_1^2\bar{x}_2^2\right)\mathrm{Var}(\gamma_0) \\
&+ \left(x_2^2\bar{x}_1^4 - 2x_2\bar{x}_1^3x_1\bar{x}_2 + x_1^2\bar{x}_1^2\bar{x}_2^2\right)\mathrm{Var}(\gamma_1) \\
&+ 2\mathrm{Var}(\gamma_0)\mathrm{Var}(\gamma_1)(x_1\bar{x}_2^2 - x_2\bar{x}_1\bar{x}_2)(x_2\bar{x}_1^2 - x_1\bar{x}_1\bar{x}_2)\Big)
\end{aligned}
$$