第3章_傾向スコア

傾向スコアとは

　傾向スコアは、観察研究において共変量を調整するための統計手法であり、因果効果を推定するために用いられる。無作為化が難しい研究において、交絡因子を制御することが可能。
　傾向スコアは、介入群と非介入群に分けられた状況で、介入群と非介入群の共変量の分布を調整するために使われ、共変量の分布を調整することによって、介入群と非介入群の比較をより正確に行うことができる。

傾向スコアの用途と利点には以下のようなものがある。

- 共変量の影響を調整し、因果関係を推定するために使われる。
- 無作為割付が困難な場合に、介入効果の推定を行うために使われる。
- 研究デザインを改善するための代替方法として使われる。
- 傾向スコアを用いることで、介入群と非介入群のバランスが改善され、効果推定の信頼性が向上することが期待される。

傾向スコアの推定

　介入変数の割り当て確立である傾向スコア$P(X)$を直接観測できる状況ではない。しかし、割り当ての結果はZとして観測されるので、何かしらのモデルを用いることで、手持ちのデータからの傾向スコア$P(X)$を推定できる。

例：ロジスティクス回帰
ロジスティクス回帰は、介入変数Zの値を目的変数とし、以下の回帰式となっている。

$$Z_i = \sigma(\beta X_i + u_i)$$

$$\sigma(x) = 1/(1 + e^{-x})$$

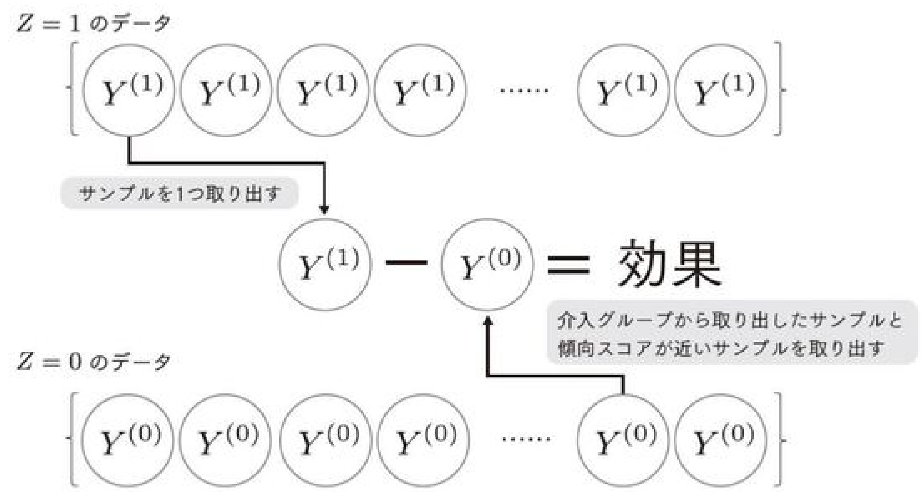
$$\hat{P}(X_i) = \hat{Z}_i = \sigma(\hat{\beta} X_i)$$

このとき。uは誤差項であり、βは推定されるパラメータ。σはシグモイド関数。シグモイド関数はxを入力したとき0から1の間の値を出力するような関数。

傾向スコアマッチング

傾向スコアマッチングのアイディアは、以下のように行う。

- 介入が行われているグループからサンプルを取り出す。
- そのサンプルと近い値の傾向スコアを持つサンプルを、介入を行われていないグループからマッチングしてペアにしていく。
- ペア通しで目的変数の差を計算し、平均を取ったものを効果の推定値とする



■ 図 3.2／傾向スコアマッチング

※わざわざペアを作る理由

　傾向スコアが同じ値を持つサンプルの中で介入変数は、介入をしていないサンプルの目的変数とは無関係に決定されていると考えられるので、この中でグループ間の比較を行っていても、セレクションバイアスの影響を受けないところがある。

　セレクションバイアスは、介入群と非介入群の共変量の分布が異なるため、共変量の影響が介入群と非介入群で異なることによって生じる。傾向スコアマッチングは、傾向スコアに基づいて、共変量の分布を均一化することで、共変量の影響を調整し、セレクションバイアスの影響を軽減することができる。

マッチングは、母集団における以下のような効果を推定しているということになる。

$$\hat{\tau}_{match} = E(E[Y|P(X), Z=1] - E[Y|P(X), Z=0]|Z=1)$$

上記は介入を受けたサンプルにおける介入効果の期待値で、ATT(Average Treatment effect on Treated)と呼ばれる。ATTは介入を受けたサンプルにおける効果で、ATTの推定を行った値は、平均的な効果を推定した値と結果が異なる可能性がある。特に介入群と非介入群で効果料が異なると想定される場合には、この傾向が顕著になる。
マッチングを利用した場合にもATE（Average Treatment effect）を推定することは可能。この場合には、上記のような介入グループのみに対して、マッチングを行われるのではなく、非介入グループに対しても同様にマッチングを行う。

傾向スコアマッチングの例

書籍のように、メールの効果を検証するときに、タイプAとBの２種類のユーザーがそれぞれ300名ずつ存在する状況を考える。

タイプ	Y_0	Y_1	sum(Z)	N
A	1000	1500	100	300
B	1500	2000	200	300

表の内容を整理

- タイプAのユーザ：メールがないときには1000円、メール配信すると1500円の売上が発生
- タイプBのユーザ：メールがないときには1500円、メール配信すると2000円の売上が発生
- メールの配信：タイプAのユーザ100名、タイプBのユーザに200名に配信する

１：単純な集計による効果確認の場合

$Y^{(0)}$の平均と$Y^{(1)}$の平均は、1166と1833になり、差分は666となり、単純な集計ではバイアスセレクションが発生していることになる。（メール配信による介入効果は500円なので）

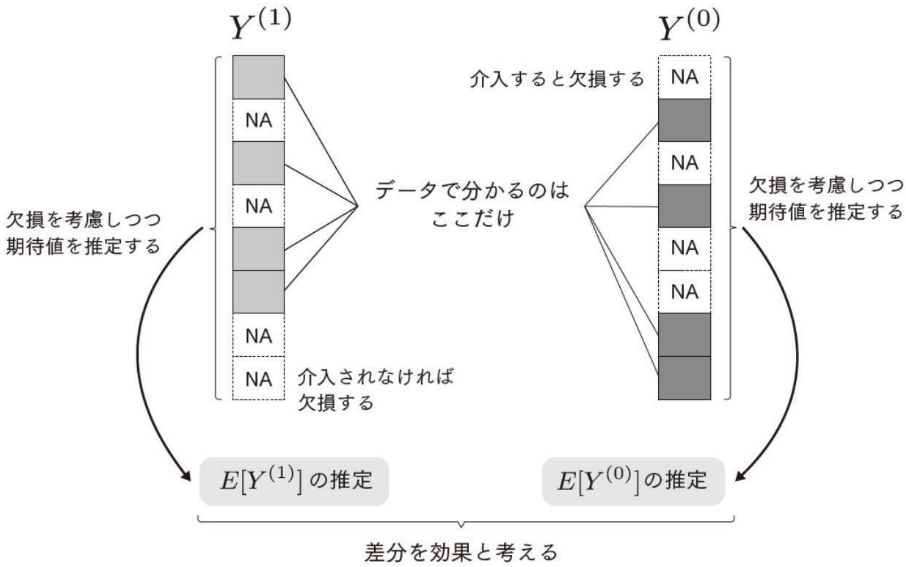
２：傾向スコアマッチングを用いた場合

1：1でマッチさせるので、人数の少ない方に合わせる（今回は100名のタイプAに合わせる）メールが配信された100名のタイプAに対して、メール配信が行われなかった100名のタイプBがペアとして割り当てられる。
このペア通しの売上の差は一律で500円になる。そのため、平均として500円とな

り、メールによる介入効果と等しくなる。（バイアスセレクションの効果を減らすことが出来ている）

逆確率重み付き推定（IWP）

逆確率重み付き推定（Inverse Probability Weighting: IPW）は傾向スコアをサンプルの重みとして利用して、与えられたデータ全体での介入を受けた場合の結果の期待値($E[Y^{(1)}]$)と、介入を受けなかった場合の効果の結果の期待値($E[Y^{(0)}]$)を推定する。その後にそれらの期待値の差分から効果を推定する。



■ 図 3.3／IPWのイメージ

第1章で説明した通り、一つのサンプルでは、$Z=1$か$Z=0$のどちらかしか観測されず、$Y^{(1)}$は$Z=1$となるサンプルのみにおいて観測され、$Y^{(0)}$は$Z=0$となるサンプルのみ観測される。このため、$Y^{(1)}$と$Y^{(0)}$も得られたデータすべてのサンプルでは観測することができない（要は、それぞれのデータで、$Z=1$と$Z=0$のデータを取得することはできない）

このときに単純に手持ちのデータで平均を取ると、期待値($E[Y^{(1)}]$)と、介入を受けなかった場合の効果の結果の期待値($E[Y^{(0)}]$)の推定値となってしまい、セレクションバイアスの影響を受けてしまう。

　IWPでは、傾向スコア$P(X)$を用いて、以下にして重み付きの平均を求めることで対処していく。

$$\bar{Y}^{(1)} = \sum_{i=1}^{N} \frac{Z_i Y_i}{\hat{P}(X_i)} / \sum_{i=1}^{N} \frac{Z_i}{\hat{P}(X_i)}$$

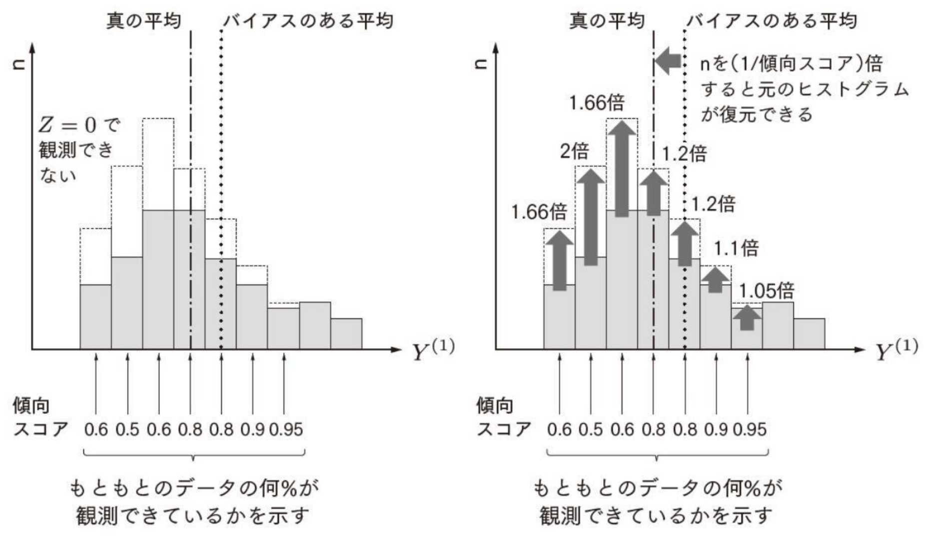
これは、平均を取る際に、確率$\hat{P}(X)$の逆数を重みにしている状態を取っている。確率$\hat{P}(X)$の逆数は確率$\hat{P}(X)$が小さくなるほど大きくなるため、サンプルが含まれない分、重みを増してくれていることになる。

- 傾向スコアが大きい→サンプルの出現確率が高い
- 傾向スコアが小さい→サンプルの出現確率が低い

と考えられるため、傾向スコアの逆数をかけることで、

- 出現確率の高いサンプルは、小さく重み付け
- 出現確率の低いサンプルは、大きく重み付け

することで処置群と対照群の共変量の分布を調整し、ランダム割り付けされた状態に近づけようというイメージ。



■ **図 3.4**／逆数によって重みを付けて本来の期待値に近づける

IPWでは、傾向スコアの値を利用して、観測されたサンプルサイズを水増しする。これによって、Z=0となることが原因で、比較的観測されづらい$Y^{(1)}$の値が小さなサンプルが増やされるため、算出される平均が本来の期待値へと近づく。

$Y^{(0)}$の期待値の推定を行うには、確率$1-\hat{P}(X)$の逆数を重みに利用する。

$$\bar{Y}^{(0)} = \sum_{i=1}^{N} \frac{(1-Z_i)Y_i}{1-\hat{P}(X_i)} / \sum_{i=1}^{N} \frac{(1-Z_i)}{1-\hat{P}(X_i)}$$

それぞれの期待値が推定できたら、あとは差分を算出することで効果の推定値を得ることができる。

$$\hat{\tau}_{IPW} = \bar{Y}^{(1)} - \bar{Y}^{(0)}$$

IPWの簡単な例

タイプ	Y_0	Y_1	sum(Z)	N
A	1000	1500	100	300
B	1500	2000	200	300

タイプAのユーザは300名で、内配信しているユーザは、100名。よってタイプAにメール配信される確率$P(A)$は100/300となる。

同様にタイプBのユーザにメール配信される確率$P(B)$は200/300となる。この確率を用いて、主見つけ平均をZの値ごとに算出し、その差を計算して効果の推定値を得る。

$$\hat{\tau}_{IPW} = \sum_{k=1}^{n} \frac{Y^{(1)}Z}{\hat{P}(A)} / \sum_{k=1}^{n} \frac{Z}{\hat{P}(A)} - \sum_{k=1}^{n} \frac{Y^{(0)}(1-Z)}{1-\hat{P}(A)} / \sum_{k=1}^{n} \frac{1-Z}{1-\hat{P}(A)}$$

$$\bar{Y}^{(1)} = (\frac{1500}{1/3} \times 100 + \frac{2000}{2/3} \times 200)/(\frac{100}{1/3} + \frac{100}{2/3}) = 1750$$

$$\bar{Y}^{(0)} = (\frac{1000}{2/3} \times 100 + \frac{1500}{1/3} \times 200)/(\frac{200}{2/3} + \frac{100}{1/3}) = 1250$$

$$\hat{\tau}_{IPW} = 1750 - 1250 = 500$$

この結果、重みを導入した平均の差は500となり、セレクションバイアスの影響がない正しいメールの効果を推定することができた。

より良い傾向スコアとは

傾向スコアは、回帰分析とどうように、どのように推定してもセレクションバイアスを消し去ってくれるものではない。傾向スコアはデータに対する説明力が一定を超えることが重要であるという解釈がされ、c統計量のような指標が一定の値を上回ることが望ましいとされていた。

しかし、近年では傾向スコアを利用して重み付けかマッチングを行ったあとのデータにおいて、共変量のバランスがとれているかが重要であるという見解が一般的になっている。