

# プロフェッショナル統計分析ワークショップ

## 応用編1: DID、PSM、IV


# Q & A

このコースを受講いただいた皆様からの質問の一部とインストラクターからの回答を紹介致します。受講後はお気軽にご質問ください。

(インストラクター: 佐々木亮)



頁	質問	ご回答・佐々木のコメント
4	「インパクト評価では、どのように counterfactual を作り出すかで、評価デザインや統計学的手法が分類されるが、…」→counterfactual について、「仮想現実（介入がなかったとしたらどうなるか）」というご説明でしたが、もう少し具体的に教えてください。単純に介入の有無だけではなく、何を説明変数とするのか、ということも関連するのでしょうか。	Counterfactual（仮想現実）は、単純に、介入を受けなかった場合はどういう状態になっていたかということですね。介入 X という一つの X の有り無しのみが違いということを想定します。実際にはそういう現実が存在しないので、Control Group を特定してそのグループの状況を仮想現実とみなします。RCT ならほぼ完全な仮想現実を作り出せたと言える一歩、そこから離れていくと精度が落ちて行って最終的に事前-事後比較になります。
10	DID の説明で、「単純な事後の 2 群の t 検定よりも、より正確に検定したことになる」とのご説明でしたが、DID をした方がベターなことが分かっているにもかかわらず、実際には出来ない（あ	実際の論文では、単純な事後の 2 群の t 検定の結果を書いて、そのあとに念のため DID という事前-事後の差の 2 群の t 検定を実施しても結果は同じだった、だからより強固に差は有意だと言える、という書き方をしますね。ただし、DID には哲学的な問

	<p>るいは敢えてやらない) というケースは多々あるのでしょうか。(PSM,IVにも関わるのですが、「どこまで正確にやるべきか」という塩梅について伺いたいです。)</p>	<p>題があります。それは、(1) 事前の段階の差が統計学的に有意でないなら「差はない」ということだから、事後の差から引かなくてもいいはずだ、(2) 一方、事前の段階の差が統計学的に有意ならそもそも2群は違う群であり「近似の群を設定するという前提条件を満たすことに失敗している」ということになります。ということで、現在、DIDがもてはやされていることには懸念を感じています。DIDがマッチング(近似群を作る)の失敗を完全に帳消しにするわけではないということです。</p>
10	<p>(上記②に関連して) 2群のt検定と単回帰分析は同じ結果が得られるにも関わらず、(DIDで)後者を使わないことには、理由はありますか。(そもそも前者と後者との大きな違いは、検定の結果、回帰式を得られるかどうか、ということでしょうか。)</p>	<p>DIDの原理を単回帰分析に使わないのかというのは新鮮な指摘ですね。今まで受けたことがございませんでした。おそらく、DIDの原理を使うなら、それぞれのIDごとに事前-事後の引き算をして得た値をYに使うということになると思います。書いていて思いましたが、DIDの原理を単回帰分析にも使えますね。より正確な結果が得られるとも言えるでしょう。なお、2群のt検定では得られずに、単回帰分析すると得られるのがR<sup>2</sup>です。説明力がどれくらい分かるので重要な情報と思います。</p>
10	<p>Excelでの図の作り方について、お薦めの手順があれば具体的に教えてください。</p>	<p>これは特に私独自のテクニックはなくて、世の中の標準的なやり方しか持ち合わせておりません。ただし、分かりやすさの観点からデータ名・データ数値・(可能ならさらに%)をまとめて髭で引き出して表すことをお勧め致します。以下が例となります。</p> 
11	<p>表中の「Simple DID」のSimpleとは何を意味するのでしょうか。日本語訳も教えてください。</p>	<p>Simpleとは特につけなくてもいいですね。そのあとのPSMやIVに比べて単純だという意味だと思います。DIDだけで結構です。</p>
13	<p>表中の「平均処置効果(ATT)」とは「エンドライン時の両群の差」-「ベースライン時の両群の差」のことでしょうか。ATTの正式名称も教えてください。</p>	<p>ATT: Average Treatment effect on the Treatedですね。ATE: Average Treatment Effectとも言うので理解しております。「介入群の効果の平均値」のことですね。なお、経済学などでよく使われますが、評価学や社会学では敢えてこの言葉を使わないのではと思います。</p>
17	<p>正規分布の式について。f(z)=0.4 * 2.72の「f」とは何でしょうか。</p>	<p>fはformulaの略ですね。ただしf(x)とかf(z)と書くとなんかすごいことを書いているような印象を与えるので、次回のテキストからyと書くことにします。</p>

37	「マッチングしても効果はあった」とのご説明でしたが、その効果は下がっても (***)→* )、 「マッチングをした方が良かった (より正確に検定ができたため?) 」と言えるのでしょうか。	マッチングしてから2群のt検定を行うと効果が上がるも下がることもあるでしょう。しかし、いち読者としては、マッチングして (つまり属性を揃えた上で) 効果が下がったなら、その結果は信用できますね。そして、単純な場合とマッチングした場合、どちらでも有意な効果が確認されたということなら二重に信頼できると思います。
41	残差 (error) について、「borrow だけでは income を説明できない複数の要素が残る」とのご説明でしたが、具体的にはどのような要素が考えられるのでしょうか。	それは、内生的なものと外生的なものがあります (Indogenous と Exogenous)。「内生的な要因」は、介入適用を知らされてがぜん「やる気」になる場合 (ネズミなら起きないが人間は知覚できるのでこれが発生する) があります。「外生的な要因」は、分析者が気づいていない第三要因 (季節の変化や経済の変化) の影響ですね。外生的な要因の影響は、RCT なら2群で自動的に同一になっていますが、回帰分析だと分析者が気づかずに取りこぼして影響を受けている場合があります。

ご質問を寄せていただいた皆様、たいへんありがとうございました。この場を借りて御礼申し上げます。今後も随時追加していきたいと存じます。

#### (インストラクター(佐々木)の IV の追加説明)

IV の手続きとして2回回帰分析を行います、その2回目の回帰分析をする時に X の「予測値」を用いましたね。これを聞いて、経済学以外の社会科学分野の研究者は「え！？ X の実測値があるのだからその実測値を使えばいいではないか。‘予測値’はその名のとおり予測に過ぎないのだから不正確でしょう？」と違和感を感じますね。すると経済学者は「いや、予測値の方が正確なんだ」と言うわけです。それを聞いて経済学以外の社会科学分野の研究者は「何言ってんだ、この人たち！！全く理解できない・・・」と思うわけです。するとさらに経済学者が言うわけです。「いいですか、実測値とは世の中のバイアスにまみれにまみれまくった値なのです。そのバイアスを綺麗にそぎ落とした値が予測値なのです。だから予測値こそ正確な値なのです。」と。つまり人間は合理的な存在であり、合理的な判断をすれば予測値を選択するはずだということです。この発想は、他の分野の人たちには理解できませんね。人間世界はそんなに合理的にできていないと・・・(いま特段の合理的理由はないながらもビールを飲みながらこれを読んでいるあなたもきっとそう感じるでしょう)。

つまり、経済学の人たちは「人間は合理的な存在だ」という前提があるのですが、その他の分野の人たちは「人間は合理的な判断だけで動いているのではない」という前提に立っています。評価学も明らかに後者です。経済学者は前者の前提に立って IV を生み出してきましたが、それは他の分野の研究者にとっては思いもつかなかったアイデアだったというわけです。これが、3つの手法の最後に出てくる IV が分かりにくい理由だと理解しております。

ここまでの話を私が一人芝居でもっとビビットにできればよかったのですが、ビデオは私の演技力不足でした。稽古に精進して次の収録でよりいい演技を目指したいと思っております。以上です。

このワークショップの全体構成は以下のようになっています。

- 『プロフェッショナル統計分析ワークショップ』(基礎コース) <https://www.idcj.jp/seminar/statistical-analysis-workshop.html>  
講師：佐々木亮 (ウェスタンミシガン大学評価学博士)、高木桂一 (スタンフォード大学社会学博士)
- 『プロフェッショナル統計分析ワークショップ』(応用コース) <https://www.idcj.jp/pickup/grow/statistical-analysis-workshop-advanced.html>  
インストラクター：佐々木亮 (ウェスタンミシガン大学評価学博士)  
応用コース1：インパクト評価の最新テクニック (DID, PSM, IV)  
応用コース2：構造方程式モデリング (SEM : Structural Equation Modeling)  
応用コース3：インパクト評価のためのサンプルサイズの計算  
応用コース4：メタ分析 (システマティック・レビュー) の計算