

なぜ多変量解析が必要なのか。使わないと、なぜ研究論文になりにくいのか。
多くの実証的研究の目的は、変数間の関連をみて、ものごとの因果関係を明らかにすることにある
そのため、原因となる変数と結果となる変数を測定する尺度を用意し、その関連を明らかにする必要がある
このとき、2つの理由で変数を多く測定したい理由がある

1 測定する尺度の信頼性と妥当性を高めるために多くの変数を用いる

幸福感を測定するには

1) あなたは幸せですか。1はい 2いいえ

という単純なものから

2) 生きていてよかったと思いますか、楽しいと思うことがありますか、などと幸福感を
あらわすようないろいろな表現でたくさんの質問をする

という方法まである

2)のようにたくさん質問したほうが、測定の回数が多く、内容も網羅できて信頼性と妥当性が高い

信頼性 = 誤差が少ない、いつ測っても同じ測定結果になる

・いつどんなときでも幸せと答えるか、・体脂肪計は測るたびに数値が変わらないか

妥当性 = ほんとうにそれが目的としているものを測定しているか、

・幸福感は様々な感じかたがあり内容に幅がある ・体脂肪計はほんとに脂肪を測ってる？電気抵抗でしょ？

多くの変数を似ているもので分類したりまとめたりして得点化して尺度(ものさし、スケール)を作る

2 予測の精度を上げるため目的変数(従属変数)を多くの説明変数(独立変数)で予測する

目的変数を予測するのにより多くの変数で予測したほうが正確になりやすい

1) 年齢 = a × 顔のしわ + 定数 (= 1? 歳: 老化によりしわがで始める年齢)

これよりは

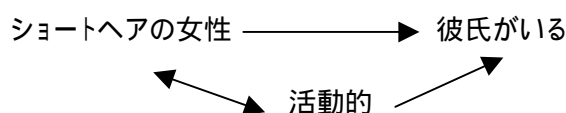
2) 年齢 = a × 顔のしわ + b × 肌の張り + c × 声のしゃがれ具合 + d × 肉のたるみ + … + 定数

のほうが予測しやすい。a、b、c、dは目的変数に対するそれぞれの説明変数の影響力の大きさをあらわす

また、同時に疑似相関(実は直接の関連が無いのに計算すると関連があるように見える)を見抜く、

交絡因子(ある変数とある変数の関連に影響を及ぼしている変数)の存在はないのか？

単相関(相関係数)など1対1の変数間の関連を並べていると、誤りが生じる可能性がある



ショートヘアの女性のほうが彼氏のいる割合が高い(かつて某看護学校で調査したときの結果、ほんと?)

彼氏の有無 = a × 髪の毛の長さ + b × 活動的 + c

実はショートヘアの女性のほうが活動的(勉強もアルバイトも何でも活発)であった。

ショートヘアであることは、活動的であるという影響力を取り除いてもなお、彼氏の有無への影響力を持つか？

3 多変量解析の種類(目的と、変数が量的か質的かで道が分かれる)

1) 多くの変数を似ているものどうして分類したりまとめたりする(信頼性の高い目的変数、説明変数を作成)

量、量、量、量、量… 因子分析(主成分分析)

因子分析の目的=多くの変数がいくつかのことからできているか？分類できるのでは？

その特殊な1つ 主成分分析の目的=多くの変数がある一つのことから測っているか？

2) 予測をして因果関係を考える(目的変数を説明変数から予測させる)

量 = 量、量、量… 重回帰分析

量 = (量)、質、質、(量)… 多元配置分散分析(共分散分析)

質 = 量、質、質… ロジスティック回帰分析

} 一般線形モデル (GLM) へ

3) 1と2を同時に、しかも複雑な因果関係まで…共分散構造分析(または構造方程式モデル)

因子(潜在的な変数)どうしの因果関係を検討、因子間の関連が3角形になっていても相互の影響を予測可能