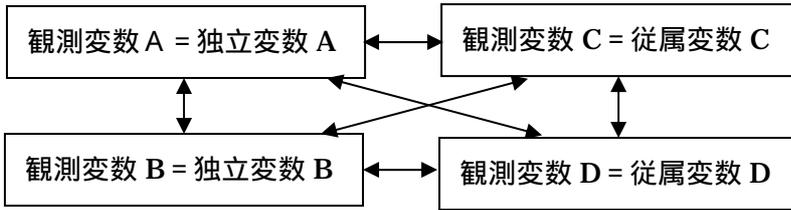
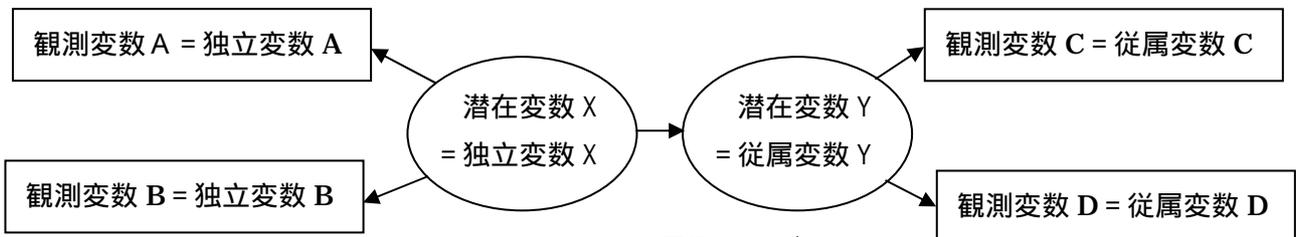


そもそも何をやろうとしている・・・



飽和 (Saturated) モデル

飽和モデル = 全観測変数の関連のしかたをすべての共分散（相関係数）で説明する 当たり前
 それよりも、因果の方向を考えたり、より少ない変数（潜在変数 = 直接観測できない背景にあるもの）間の関連で説明したりしてみる。これが推定（予測）モデルで、



推定（予測）モデル

こんな関連図でどれだけうまく関連を説明できるか、一番うまく説明できるように矢印（影響力）の大きさ（=パス係数）を計算してみるのが目的。推定モデルでの（観測変数 A と潜在変数 X 間のパス）×（潜在変数 X と潜在変数 Y 間のパス）×（観測変数 C と潜在変数 Y 間のパス）が観測変数 A と観測変数 C の相関係数とどの程度一致しているか（差を χ^2 でみる）。推定モデルでは、観測変数と観測変数の間でつながっている道筋の全パスを掛け合わせるとその間の相関をあらわす = パストレーシングルール。

このとき、うまく説明できている程度を見る方法（=適合度）として、飽和モデルにどれだけ近い説明力を持つか検討する方法（飽和モデルとの差を χ^2 でみる）と、その飽和モデルとの離れ具合を、独立モデルと比較して検討する方法（NFI、TLI、CFI など）がある。

独立モデルとは、観測変数間の相関が 0 というまったく観測変数間の関連を説明しようとしなないモデル。



独立 (Independence) モデル

推定モデルは、説明力において飽和モデルより小さく、独立モデルより大きい

おもな適合度の指標

サンプル数、自由度に影響を受ける古典的適合度

2

飽和モデルと推定モデルのずれ具合(分散共分散行列におけるずれの程度の合計)、0なら完全に適合。帰無仮説はこれ。Pが.05より小さく有意だと、帰無仮説を棄却(=だからよくない)。200サンプル以下ほどならOKとも言われる。これを越えると χ^2 が大きくなってすぐ有意だから適さないと。

GFI、AGFI

GFI: 決定係数(飽和モデルでの全分散が推定モデルでの分散でどれだけ説明できたか)、AGFI: 自由度調整済み決定係数、に該当。 .9以上。 .85以上とも。

自由パラメータの数が多(自由度が小さい)と大きくなる = パスを引けば大きくなる
変数が多いと誤差間などを引かないと高くない。使わないコンセンサスはあるという専門家も

自由度 = 飽和モデルのパラメータ数 - 自由パラメータの数

飽和モデルのパラメータ数 = 観測変数の数 × (観測変数の数 + 1) / 2

自由パラメータの数 = パラメータのうち固定しないものの数(誤差 + パス + 共分散 + 分散)

ちなみに 飽和モデルの自由度 = 0 独立モデルのパラメータ数 = 観測変数の数

独立モデルの自由度 = 飽和モデルのパラメータ数 - 独立モデルのパラメータ数)

サンプル数、自由度によらない(それを考慮 = ペナルティを与えるという)適合度

χ^2 / df

自由度の影響を考慮。値が0に近いほどよい。2、3、6未満という話もあるが、明確な基準はない。

RMSEA

$\sqrt{(\chi^2 / df - 1) / (N - 1)}$ $\sqrt{\quad}$ は平方根 square root、Nはサンプル数

自由度も、サンプル数も考慮。 .05よりも小さいとよい。 .1を越えると不適。

信頼区間も算出されるが、小さいほうは0に近いほうがよい

AIC

$2 + k(k - 1) - 2df$ kは観測変数の数 $k(k - 1) - 2df$ は自由パラメータの数の2倍と一致
小さいほど適合度がよい。基準値はなく、値に絶対的な意味はない。同じデータでモデルを比較するのに利用。自由度あるいは自由パラメータを考慮してある。

CAIC

$2 + (1 + \log n)[k(k - 1) - 2df] / 2$ nはサンプル数で、AICにさらにサンプル数の影響を考慮

NFI

.9 ~ .95でOK、.95を越えるとよい。欠点はパスを引けば引くほど高くなるだけだということ。独立モデル = すべての変数が独立(相関が0)というモデル。それに比べてどのくらいよいか。

Null model (独立モデル)の χ^2 - 提案したモデルの χ^2

Null model (独立モデル)の χ^2

TLI または NNFI

NFIの χ^2 の部分を χ^2 / df にし(自由度の影響を考慮) 分母は上の式から1ひいてあるもの

1に近いほどよいが、1以上になる場合もあり、その場合、1と出力される

CFI

NFI の χ^2 の部分を $(\chi^2 - df)$ にしたもの (自由度の影響を考慮) でそれ以外は上と同じ。
やはり .95 以上とか。1 より小さいときは常に TLI より大きくなる

自分が主張したいモデルをほかのモデルとならべて、より多くの適合度の指標、とくに上記後半のものを比較して示す。やはり最適モデルでは、よい適合度は必要だが・・・特に CFI、RMSEA など

参考文献は、豊田秀樹氏、狩野裕氏の関連著書や関連サイトすべて
その他、<http://users.rcn.com/dakenny/causalm.htm> など