

保健統計学  
第12・13回 尺度による測定と因子分析  
教科書 pp. 239 – 285

米倉佑貴(yyonekura@slcn.ac.jp)

聖路加国際大学大学院

看護学研究科看護情報学分野

# 今回の内容

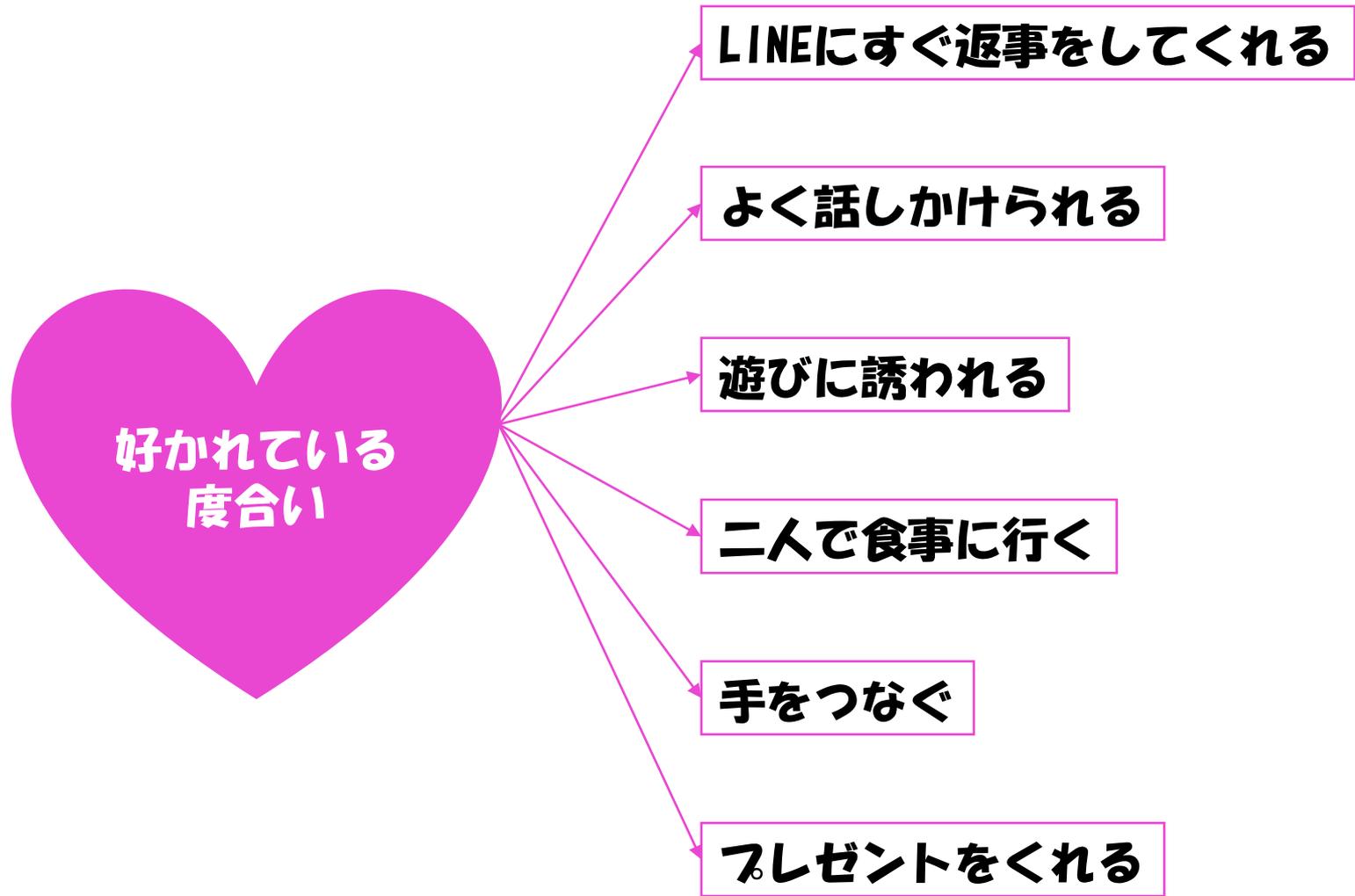
- 尺度による潜在変数の測定
  - 観測変数, 潜在変数とは
  - 古典的テスト理論
  - 尺度の信頼性, 妥当性
- 因子分析
  - 因子分析の概要
  - 探索的因子分析
    - 因子数の決定
    - 因子抽出の方法
    - 因子の回転
  - 確認的因子分析

# 尺度開発と因子分析

# 直接観測できないものを測定する方法

- 測定してデータ化できないものは統計解析で扱えない
  - 目に見えないもの, 触れないものなどはデータ化するのに工夫が必要
    - 心, 態度, 気持ち, QOL, 学力, 健康, etc...
- 測定・観測できる性質・反応から推定する試み
  - 生理学的指標
    - ストレスとコルチゾール, fMRIによる脳機能の測定
  - 質問紙法
    - 抑うつ尺度(CES-D)
  - 投影法
    - ロールシャッハテスト
  - 作業法
- 観測変数と潜在変数
  - 観測変数：直接測定できる変数
  - 潜在変数：直接測定できないが、観測変数によって推定することができる(と考える)変数
- 尺度(スケール)は潜在変数を測定するための道具

# 尺度による測定のイメージ



- 観測できる行動・特性から観測できない概念・特性の程度を推定する

# 尺度による測定の考え方（古典的テスト理論）

## • 古典的テスト理論の考え方

- 質問項目への回答はその項目が測定しようとする概念・特性の真の値と誤差に分けられると考える
- 誤差は平均0, 標準偏差 $\sigma$ の正規分布に独立に従うと考える
- 誤差は平均すると0だから, 同じ特性を反映する**多くの項目を集めて合計すれば真値に近い値を求められる**

## • 古典的テスト理論の測定モデル

$$R_{i1} = f_i + u_{i1}$$

$$R_{i2} = f_i + u_{i2}$$

$$R_{i3} = f_i + u_{i3}$$

⋮

$$R_{ij} = f_i + u_{ij}$$

すべて足し合わせて $f$ について解くと

$$f_i = \frac{\sum R_{ij} + \sum u_{ij}}{j}$$

$R$ : 質問への回答,  $i$ :対象者番号,  $j$ :項目番号,  $f$ : 真の値,  $u$ : 誤差

誤差の平均(期待値)は0なので $\sum u_{ij}$  は項目が多くなれば0に近づく

# 多項目尺度の例(1)

## • K6

- 国民生活基礎調査で用いられる精神健康度の尺度
- うつ・不安障害のスクリーニングに用いられる
  - 5点以上: 何らかのうつ・不安の問題がある可能性
  - 10点以上: 国民生活基礎調査でのうつ・不安障害の基準
  - 13点以上: 重度のうつ・不安の問題がある可能性

過去 30 日の間にどれくらいの頻度で次のことがありましたか

		0点	1点	2点	3点	4点
1	神経過敏に感じましたか	全くない	少しだけ	ときどき	たいてい	いつも
2	絶望的だと感じましたか	全くない	少しだけ	ときどき	たいてい	いつも
3	そわそわ、落ち着かなく感じましたか	全くない	少しだけ	ときどき	たいてい	いつも
4	気分が沈みこんで、何が起これども気が晴れないように感じましたか	全くない	少しだけ	ときどき	たいてい	いつも
5	何をするのも骨折りだと感じましたか	全くない	少しだけ	ときどき	たいてい	いつも
6	自分は価値のない人間だと感じましたか	全くない	少しだけ	ときどき	たいてい	いつも

引用：大野裕他「一般人口中の精神疾患の簡便なスクリーニングに関する研究」  
(平成 14 年度厚生労働科学特別研究事業)

# 多項目尺度の例(2)

## • 日本語版統御感尺度

### – 統御感(sense of mastery)

- 人々が現在自分たちの生活に影響を与えている重要な状況をコントロールすることができるという信念

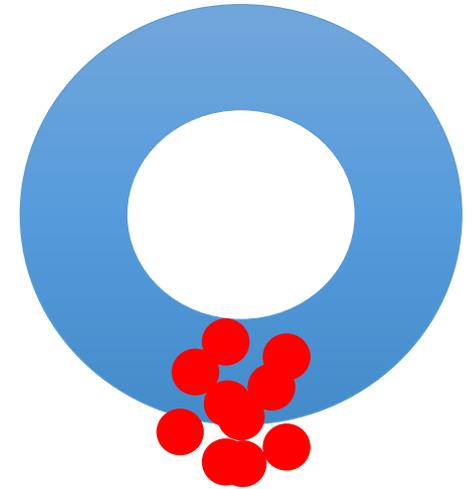
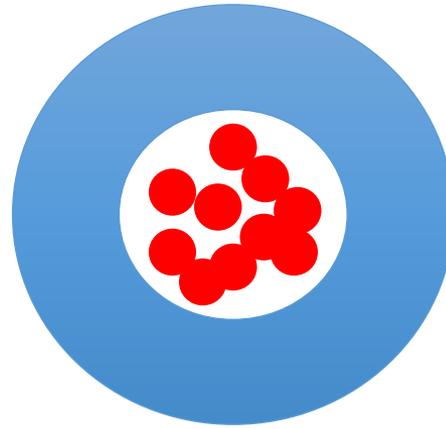
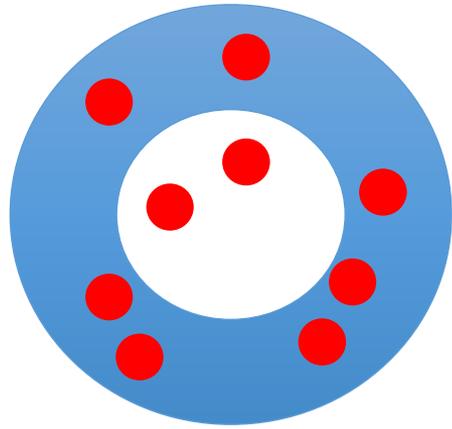
問4 あなたの生活・人生への感じ方についてお聞きます。以下の(A)～(G)の項目について、それぞれどの程度あてはまるかを答えてください。(○はそれぞれ1つずつ)

	とてもあてはまる	ややあてはまる	ややあてはまらない	全くあてはまらない
(A)自分の身に起こることを、コントロールすることができない	1	2	3	4
(B)自分が抱えている問題のいくつかをどうしても解決できない	1	2	3	4
(C)自分の生活や人生の中で大事なことの多くを変えるために、私ができることはほとんどない	1	2	3	4
(D)生活や人生上の問題を解決しようとするとき、よく自分が頼りなく感じる	1	2	3	4
(E)ときどき、生活や人生の中で、周りの人や状況に従わせられているように感じる	1	2	3	4
(F)将来私の身に何が起こるのかは、たいていは、自分次第で決まる	1	2	3	4
(G)自分でやると決めたことは、ほとんどどんなことでもできる	1	2	3	4

# 尺度開発のステップ

1. 測定したい概念を明確にする
2. アイテムプールを作る
3. 回答形式を決める
4. アイテムプールのレビュー
5. 試作版の実施
6. 項目分析, 妥当性, 信頼性の検討
7. (必要があれば)尺度の改訂

# 尺度の性能：信頼性と妥当性



信頼性が低く、妥当性も低い

信頼性が高く妥当性も高い

信頼性は高いが  
妥当性が低い

- 信頼性は測定の再現性を表す  
→ 投げた玉が的の一点に集中して当たる程度
  - 同じ所にあたる（ブレがない）（内的一貫性）
  - 何度やっても同じ結果になる（再現性，再テスト信頼性）
  - 誰が評価しても同じ結果になる（評価者間信頼性）
- 妥当性は尺度によって測りたいものが測れているかを表す  
→ 的の中心に近いところに当たる程度
  - 中心に当たるためには中心に集中してあたらなければいけないので，**信頼性が高くなければ妥当性は高くない**
  - 構成概念妥当性

# 信頼性の指標 (1)

- 内的一貫性 (internal consistency)
  - 尺度を構成する項目が同じものを測定している程度
  - クロンバックのアルファ係数(Cronbach's alpha)が内的一貫性の指標としてよく用いられる
  - 0から1の間を取り, 1に近いほど内的一貫性が高い
    - 0.70~0.95であれば良好であるとされている(Terwee et al., 2007)
  - 計算方法1:  
分散を使う方法

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum \sigma_i^2}{\sigma_y^2} \right)$$

k:項目数,  $\sigma_i^2$ :i番目の項目の分散,  $\sigma_y^2$ :尺度得点の分散

- 計算方法2:  
各項目間相関係数の平均値を用いて計算する方法

$$\alpha = \frac{k\bar{r}}{1 + \bar{r}(k-1)}$$

k:項目数,  $\bar{r}$ :相関係数の平均値

# 信頼性の指標 (2)

- 再テスト信頼性(test-retest reliability)
  - 同じ内容の尺度を同じ人に2回回答してもらい得点の相関をみる
  - 時間的に安定していると考えられるものを測定する際に確認する
  - 級内相関係数が再テスト信頼性の指標として用いられる
    - いろいろな計算方法があるが、二変量混合効果モデルで絶対一致のものを使うのがよいとされている(Koo TK & Li MY, 2006)
    - 級内相関係数の大きさは0.7以上が目安

- 評価者間信頼性(inter-rater reliability)

- 同じ対象を複数人に評価してもらい、その一致度を計算  
→κ係数など

$$\kappa = \frac{\text{観察された一致率} - \text{偶然一致する確率}}{1 - \text{偶然一致する確率}}$$

- κ係数の目安  
0～0.2 : わずか(slight), 0.2～0.4 : まずまず(fair), 0.4～0.6 : 中等度(moderate), 0.6～0.8 : よい(substantial), 0.8～1 : ほぼ完璧(almost perfect)

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics .1977;33:159-174.

# 妥当性

- 妥当性の分類
  - 内容妥当性(content validity), 表面妥当性(face validity)
  - 基準関連妥当性(criterion-related validity)
    - 収束妥当性(convergent validity)、併存妥当性(concurrent validity)、弁別妥当性(discriminant validity)、予測的妥当性(predictive validity)
  - 因子妥当性(factorial validity)
  - 既知集団妥当性(known groups validity)
  - 交差妥当性(cross validity)
- すべてまとめて「構成概念妥当性」と呼ぶこともある
  - 構成概念(construct)
    - 「その存在をとりあえず仮定することによって複雑に込み入った現象を比較的単純に理解することを目的に構成した概念」(豊田秀樹「因子分析入門」)
    - QOL, 健康, 態度, 性格, 学力
- いろいろな種類の妥当性を検討・確認できるほどよい  
→証拠が多いほどよいという考え方。  
妥当性がある, ないという二分法ではない

# 妥当性の分類と確認の方法(1)

- 内容(的)妥当性(content validity)
  - 尺度を構成する項目の内容が尺度で測定しようとしている概念の意味内容と合致している度合い
  - 確認方法
    - 複数の専門家による項目内容の確認  
→統計解析は行わない
    - 多数の専門家に協力を得て、各項目を評価してもらい、評価が一定の水準になるまで修正を繰り返す方法もある(デルファイ法)  
→評価の度数分布や平均などの統計量をもとに評価する
- 表面妥当性(face validity)
  - 測定しようとしている概念を測定できるように見えるかどうか
  - 内容妥当性に似ているが、より簡易・表面的な検討  
→内容妥当性を検討する方がよい
  - 確認方法
    - 専門家等による項目内容の確認

# 妥当性の分類と確認の方法(2)

- 基準関連妥当性(criterion-related validity)
  - 外的基準との関連が理論から予想されるものと合致している度合い
  - 以下のように細分化される
    - 収束(的)妥当性(convergent validity)
      - 同じ概念を測定する既存の尺度と同様に概念を測定できる度合い
      - 既存の尺度の項目を減らした短縮版を作る際などに確認する
    - 併存(的)妥当性(concurrent validity)
      - 尺度により測定しようとしている概念に類似した概念と関連があること
    - 弁別妥当性(discriminant validity)
      - 理論的に関連しないと考えられるものと関連しないこと
    - 予測的妥当性(predictive validity)
      - 尺度で測定する概念が、理論的に予測しうる他の概念を予測できる度合い

# 妥当性の分類と確認の方法(3)

- 基準関連妥当性の確認方法

- 収束妥当性

- **同じ**概念を測定する従来の尺度と関連があることを確認する

- 併存妥当性

- **類似した**概念を測定する従来の尺度と関連があることを確認する

- 弁別妥当性

- 理論的に関係のない概念・変数と関連がないことを確認する

- 予測妥当性

- 理論的に予測しうる概念・変数との関連性を確認する

→相関分析, (重)回帰分析, ロジスティック回帰分析など

# 妥当性の分類と検討の方法(4)

- 既知集団妥当性(known groups validity)
  - 集団間での特性の比較結果が理論的に予想されるものと一致する度合い
  - 確認方法
    - 集団間で尺度で測定した特性値を比較する  
→平均の比較(t検定, 一元配置分散分析, 重回帰分析)
- 交差妥当性(cross validity)
  - 異なる特性を持った集団においても測定したい概念を測定できる度合い
  - 確認方法
    - 異なる集団を対象にしても妥当性の検討が成り立つことを検証する

# 妥当性の分類と検討の方法(5)

- 因子妥当性(factorial validity)
  - 尺度で実際に測定したデータが、測定しようとしている概念の理論的な因子構造と合致している度合い
  - 因子構造
    - 因子: (複数の)質問への回答や観測される行動などに共通して影響を与えるもので、直接観測できないもの
    - 因子構造は因子の数や特定の因子から影響を受ける質問への回答や行動等の対応関係のこと
  - 確認方法
    - 因子構造が理論から導かれる仮説と一致するかを確認する
      - 尺度の因子構造の確認、仮説モデルの構築
        - 仮説モデルに対して**確認的因子分析(検証的因子分析)**を実施(confirmatory factor analysis)  
→こちらが第一選択
        - データからのアプローチ・・・**探索的因子分析**の結果の確認(explanatory factor analysis)  
→検証的因子分析でうまく行かなかった場合、仮説がない場合は探索的因子分析を行う