



本日の内容

1. イントロダクション
2. 総論 ~ 統計学と研究デザイン
 - ・統計学
 - ・研究デザイン
 } 「誤差」で考える
3. 各論 ~ 2つのトピック
 - ・尺度の利用
 - ・研究デザインの整理

1. イントロダクション

- ・自己紹介
 - ・生物統計学, 疫学とは
- ・看護研究における量的研究
 - ・主流は質的研究?
 - ・量的研究の特徴
- ・本日の内容
 - ・以降の前提

1. イントロダクション

自己紹介

- ・佐伯圭一郎(さいきけいいちろう) 保健学博士

出身: 東京大学医学部保健学科疫学教室(現在の健康総合学科疫学・生物統計学教室)

主たる研究テーマ: 教育・臨床の現場におけるICT活用, 臨床看護研究推進のための支援, 臨地実習前共用試験CBT, 生活習慣と健康

出発点: 「栄養疫学のフィールドワーク」, 「趣味のパソコン」

現職: 大分県立看護科学大学 看護学部看護学科 人間科学講座 健康情報科学研究室 教授

↓

メンバー	教授(佐伯)	~ 疫学・生物統計学
	准教授	~ 医療情報学
	助教	~ (数理)統計学

1. イントロダクション

生物統計学 Biostatistics

- ・単なる「統計学の医学領域への応用」ではない
- ・扱うデータの特性を理解した上で
- ・データの集め方や管理の方法まで
- ・研究計画(データ収集の計画)の段階から
 - ・統計専門家への相談は研究計画段階で
 - ・データを持って行けば何とかしてくれる, 訳ではない
- ・「臨床試験」の分野に専門家が集中

1. イントロダクション

疫学 Epidemiology

- ・集団を対象に
 - ・疾病の分布を知り
 - ・疾病の分布に影響する要因を探る
- ・そのための
 - ・調査計画
 - ・フィールドワーク
 - ・データ解析
- ・生物統計学と重なる部分が多いが
 - ・疫学 → 観察研究から(感染症の疫学)
 - ・生物統計学 → 介入研究から

1. イントロダクション
大分県立看護科学大学

おことわり

- 「論文作成の際のキーポイント」よりも
- 「研究計画のキーポイント」として、より有効
- まったく数式を出さない方が難易度が(双方にとって)アップするので、少し数式が
- データや研究、論文の実例をそのまま出すと支障があるので「雰囲気を残した架空例」を
- その他
 - 滑舌が悪いので、しばしば水分補給をします
 - 脱線しやすい傾向にあるので、視線等で警告を発して下さい

(10)

1. イントロダクション
大分県立看護科学大学

主流は質的研究？

- 看護研究論文の手法を調べると
掲載論文(原著・研究報告等)に占める質的研究
 - A学会誌(2010~2013) 57%(65/115)
 - B学会誌(2010~2013) 59%(88/151)
 - 日本がん看護学会誌は？
(原著・研究報告,2009~2014)
56%(61/109)

以上の数等は、佐伯調べ

(11)

1. イントロダクション
大分県立看護科学大学

混合研究法

- 量的研究
- 質的研究
 - どちらか一方だけではカバーできない領域に
- 関心のある方は
 - <http://jans.umin.ac.jp/news/141215.html>
 - <http://www.jsmmr.org/>

(12)

1. イントロダクション
大分県立看護科学大学

「問い」に対する大きな研究の流れ

- 問い(Clinical Question)に対して
 - 文献研究
 - 独立した研究論文にならなくとも
 - 質的研究
 - 「尺度」の開発
 - 量的研究
 - 関連の分析
 - 介入とその評価
- 順に、または、「問い」の細目毎に

(13)

1. イントロダクション
大分県立看護科学大学

この分野の量的研究の特徴

- 難易度の高い統計手法を利用する研究は少ない
- 「尺度」を用いた研究は多い
 - 量的研究の半分近く(全体の2割弱, 日本がん看護学会誌)
 - 全体の2割近くで「クロンバックのα信頼性係数」
 - 〃 15%ほどで「因子分析」 (A・B学会誌)
- リスクや予後に関する研究は少ない
 - オッズ比やロジスティック回帰, 生存時間分析などはごく稀

(14)

1. イントロダクション
大分県立看護科学大学

以降の前提

- この学会の研究フィールドを考えると
 - 標本数はあまり大きくない
 - 測定する変数は、精神・心理的特性が多い
- この2つが量的研究を考える際の特徴
 - 看護研究の多くの領域に共通
 - 生物統計学のカバーが薄い分野
- 複雑な統計手法は、あまり扱わない

(15)

2. 総論 ～統計学と研究デザイン

- 量的研究のチェックポイント
- 統計学の観点から
 - 統計手法の使い方
 - ミスの例
 - 統計手法の前提と限界
 - 関連と因果
- 誤差 Error
 - 誤差とは
 - 偶然誤差の性質と対処
 - バイアスの性質と対処

2. 総論 ～統計学と研究デザイン

量的研究のチェックポイント

- 細かい点を列挙すると際限がありませんが
 - 執筆要項レベル
 - APAガイドライン(参考図書を参照)など
 - 分析手法
 - JJCOの統計解析結果のレポートに関するガイドラインなど
http://www.oxfordjournals.org/our_journals/jjco/for_authors/jap_guideline.pdf
 - 倫理安全
 - 看護協会, 各学会, 大学で指針が整備, チェックも厳しい
 - 疫学研究に関する倫理指針(文部科学省・厚生労働省)
- 大局的には「誤差の観点で論文の質を」チェック

量的研究における統計解析関連のコメント

- 実際の査読コメントの例
 - 結果等にある「 $p < .05$ 」等の表記は小数点前にも0をつけること
 - 「Kaiser-Meyer-Olkinの標本妥当性の測度」は文献をつけることが望ましい
 - 全対象者数を示すときはnではなくNを用いること
 - 分析方法の最後に使用された統計パッケージを表記してください
 - 一般的に相関係数0.33を中程度の相関と評価するでしょうか
- 比較的軽微なコメントですが
 - N, nや0の表記は掲載済みの論文でもゆらぎが
 - “あたりまえの統計手法”や結果の読み取り(例えば相関の強さ)も, 研究領域で様々

- すこし重いコメントの例として
 - 本研究で用いた研究デザインでは, ■■■とそれに影響すると思われる要因の因果関係を特定することはできない
 - 多重ロジスティック回帰分析は, 研究者が仮定した変数間の関係を予測するものであって因果関係を保証するものではない
- 査読, それ以前に執筆(もっと前に研究計画)段階で
 - 倫理安全については
 - 倫理(安全)の指針が存在
 - 執筆要項(要領, 要綱?)は
 - 基本的には, 文書としてのフォーマットを中心に規程
 - 研究論文の質を担保するためには
 - 研究デザインや統計解析, およびその表記について, 査読者と研究者にある程度同じ視点が

2. 総論 ～統計学と研究デザイン

量的研究の主要チェックポイント

- 研究の意義, 目的の妥当性
- 倫理性
- 対象の設定
 - 目的に合致した設定か
 - 偏りは
 - 対象数は十分か
 - 対照群が必要か
- 測定
 - 適切な測定項目, 測定用具を用いているか
- 統計解析
 - 適切な手法を選んでいるか
 - 解釈を誤っていないか

2. 総論 ～統計学と研究デザイン

チェックを行うには

- 研究の「読み方」のテキストを参考に
 - 例えば
 - 「エビデンスのための看護研究の読み方・進め方」高木廣文, 林邦彦, 中山書店, 2006
 - 「よくわかる看護研究論文のクリティーク: 研究手法別のチェックシートで学ぶ」山川みやえ, 牧本清子編著, 日本看護協会出版会, 2014
- 自分の論文(or研究計画)をチェック
 - 「からく」チェックしすぎないように

2. 総論 ~統計学と研究デザイン

統計学の観点から

- 統計手法を使う理由は
 - データを分かりやすく整理する
 - 記述統計, 作図・作表
 - データに隠されたルールや事実を見つける
 - 比較, 関連, 予測, 分類...
 - 観察された結果の確からしさを判断する
 - 推定, 検定
 - データの蓄積からエビデンスへ!

(19)

2. 総論 ~統計学と研究デザイン

統計手法については

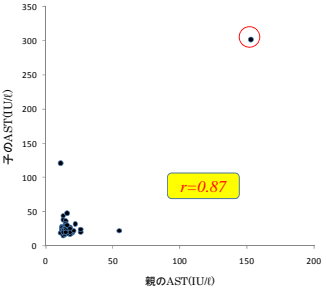
- 手法の使い分け
 - 昨年の研修会(中山和弘先生)
 - 参考Webサイト(青木先生)などを参考に
- 統計解析上の誤り
 - 手法を適切に選択していない
 - 基本的なデータ吟味が不十分
 - 不要な検定の実施
- 解釈の誤り
 - 「有意差がない」は「等しい」ではない
 - 後付けの解釈

(19)

2. 総論 ~統計学と研究デザイン

ミスの例(1)

- 基本的なデータ吟味をきちんと行っていない
 - 相関係数は0.87
 - 1例を除くと -0.14
 - 除かなくても, ケンドールの順位相関係数なら
 - 0.07
 - ピアソンの相関係数の前提は「直線的な関連」
 - 両方の変数が「正規分布からずれている」



(20)

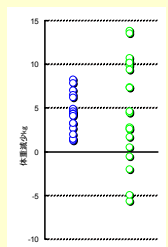
2. 総論 ~統計学と研究デザイン

ミスの例(2)

- 平均値の差の検定で「有意差無し」なら「2群に差が無い」?
 - 検出力が十分である必要はあります
- それ以外に
 - 平均値以外の「分布の特性値」にも注意が必要 → データの吟味を2通りのダイエット法で各20人が減量しました。体重減少(Kg)を2群で比較すると...

	n	平均	S.E.	p
ダイエットA	20	4.32	0.46	0.67
ダイエットB	20	4.92	1.29	

※対応のない2群のt検定による



(21)

仮説検定における2種類の誤り

		真の結果	
		帰無仮説	対立仮説
判断	帰無仮説を受容	○	第二種の誤り
	対立仮説を採択	第一種の誤り	○

- 第一種の誤り: あわてものの誤り
- 第二種の誤り: ぼんやりものの誤り

22

第一種の誤り

- αエラー, 第一種の過誤, とも
- 本当は帰無仮説が正しいのに, “まれな結果だと判断”
 - α=0.05なら, 帰無仮説が正しい場合でも5%で起きる
 - 確率的判断の宿命
 - αを小さく(例えば0.01へ)する
 - どこかで起きている可能性, を忘れない

23

検出力(power)

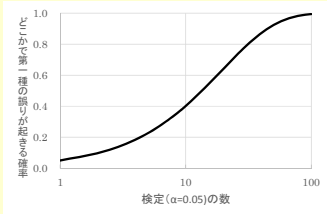
- 対立仮説が正しいときに、正しく対立仮説を採択する(帰無仮説を棄却できる)確率
 - (第二種の誤りを犯さない確率)
- 真の差や関連の大きさ, 標本数つまり標準誤差の大きさに検出力が決まる
- 検出力は0.8程度が望ましい(が・・・)
 - 主に, 標本数を増やすことで検出力を確保する

24

2. 総論 ~統計学と研究デザイン

ミス例(3)

- 研究全体の有意水準の問題
 - 各検定を $\alpha=5\%$ で実施しても
 - 全部で10回検定すれば, $1-(1-0.05^{10})=0.40$



- 数十も検定を行ってれば, どこかに
- 「あわてものの誤り」は存在
- 行き当たりばったりで検定し, 有意なものを無理矢理解釈することは, ミスの元

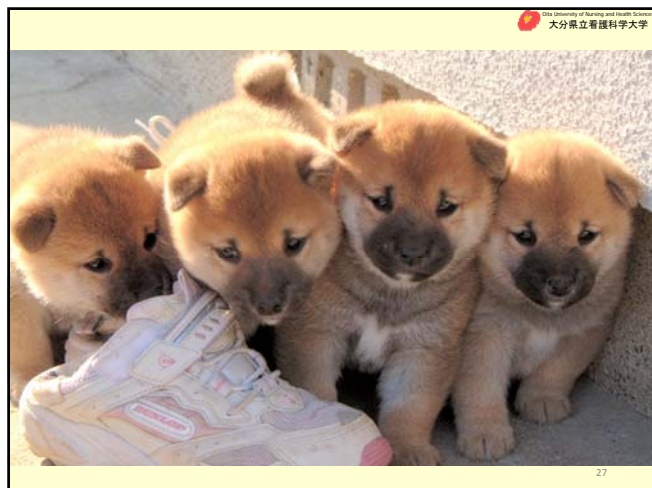
25

2. 総論 ~統計学と研究デザイン

統計手法の前提と限界

- Garbage in, garbage out
- 手法の前提
 - 無作為抽出標本
 - 正規性, etc.
- 限界
 - 平均や相関は, 全体としての傾向
 - すべての個人にあてはまるとは限らない
 - どこまで行っても「確率的」
 - 確率的に, 判断を誤ることもある
 - 非確率的な誤差には(原則的に)対処できない
 - 事実(関連や差異)の確認 \neq 因果を証明

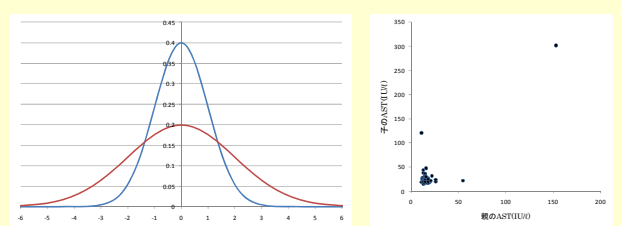
26



2. 総論 ~統計学と研究デザイン

“平均”, “全体の傾向”だけで良いのか

- 看護の対象は「平均的な」事例とは限らない
- 「例外」にも, 十分な対処が必要
- 統計手法が無力なのではなく, 使い手のセンス



28

2. 総論 ~統計学と研究デザイン

関連と因果関係

- 因果関係推定のための視点
 - “Smoking and Health”(1964), 米国公衆衛生総監諮問委員会による5条件
 - 関連の一致性
 - 関連の強固性
 - 関連の特異性
 - 関連の時間性
 - 関連の整合性

疫学のテキスト等を参照して下さい

29

2. 総論 ~統計学と研究デザイン

大分県立看護科学大学

関連の時間性

- 因果関係を論じる際の大前提
- 原因は結果より(作用に必要な時間)前に存在する
 - 当たり前に見えるが、「血圧が高いから減塩している」
- 横断研究の限界
 - パス解析に対する批判
 - 「原因」が時間依存で変化しないものならOK
- 「原因」を過去に求める or 「結果」を未来に

(30)

関連と関係

- 疫学的な関連は、実際には様々な(下の例以外にも)関係により現れる。

a)見かけの関連

b)どちらも共通の原因による結果

c)疾病への中間段階

d)疾病が曝露の原因

大分県立看護科学大学

研究デザインの基本

- 疫学の視点から
 - 観察 vs 実験
 - 横断 vs 縦断
 - 個人単位 vs 集団

(32)

大分県立看護科学大学

例えば

- 新薬の臨床試験は
 - 実験, 縦断, 個人単位
- 患者と非患者の過去の食生活を比較
 - 観察, 縦断, 個人単位
- 都道府県の肺がん死亡率と喫煙率の関連は
 - 観察, 横断, 集団単位
- このような大きな枠組みに
 - 測定の方法や分析手法を加えて「デザイン」に

(33)

2. 総論 ~統計学と研究デザイン

大分県立看護科学大学

誤差 Error

- 誤差 = 真値 - 観測された値
 - 研究結果による判断のレベルから
 - 一つ一つの測定値レベルまで
- 誤差を小さくすること
 - = 研究の質を高めること
 - ≡ 研究の効率を上げること

(34)

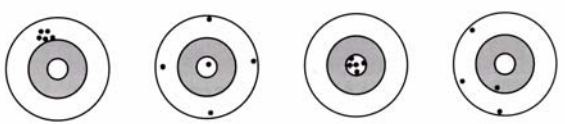
誤差(エラー)への対処

- プログラムのエラー
 - 発見と対処 (デバッグ debug)
- 通信(信号処理)のエラー
 - エラーチェックと訂正のアルゴリズム
- ヒューマン・エラー
- 誤差
 - 検出可能?
 - 予防可能?
 - 補正(訂正)可能?

2. 総論 ~統計学と研究デザイン

誤差の大分類

- ランダムな誤差
 - 偶然誤差
 - 確率誤差
 ≡ 精度, 信頼性の問題
統計学の守備範囲
- ランダムでない誤差
 - 系統誤差
 - 非確率誤差
 - バイアス
 ≡ 正確度, 妥当性の問題
普通の統計学では
「無いこと」が前提



(36)

2. 総論 ~統計学と研究デザイン

偶然誤差への対処

- データの測定誤差を小さくする
 - 信頼性の高い測定道具
 - 精密な測定機器, 良い「尺度」
 - 測定の方法
 - 反復測定, 測定マニュアルとトレーニング
- 統計量の誤差を小さくする
 - 標本数を増やす
 - 適切な統計手法を選択する
- さらに
 - メタアナリシス
 - 事例の追加, 分析のシステム

(37)

2. 総論 ~統計学と研究デザイン

統計的表記

- 標本中の1サンプルの測定値を分解して表現
個人の測定値 = 集団の平均 + 個人の効果 + 誤差

$$x_i = \mu + v_i + e_i$$

- 「個人の効果」は, 集団中の個人の「位置」
- ここでは「誤差」は, 偶然誤差のみを考える

(38)

2. 総論 ~統計学と研究デザイン

数学的性質: 偶然誤差のみを考慮

$$E(x) = \mu + E(v) + E(e)$$

データの平均は, 真の平均に一致するはず
(個人の効果, 誤差の期待値は0)

$$V(x) = 0 + V(v) + V(e)$$

データの分散は, 真の分散と誤差分散の和
・ 誤差の独立性という重大な仮定が必要

意味すること
=「偶然誤差がどれだけ含まれていても, 標本平均の期待値は真の平均」
&「偶然誤差が大きくなるほど, 標本分散は真の分散より大きくなる」

誤差の独立性

- 偶然誤差として単純(?)に扱えるには
 - 誤差と真の値に関連がない
 - バイアスではない, ということ
 - ばらつきの大きさが, 真値と関連しない
 - これは実は難しい条件かも
- 「真の値」は分からないのに?
 - 偶然誤差だけなら「真の値」は推定可能
 - 尺度構成の信頼性などは, 真の値に対する評価

2. 総論 ~統計学と研究デザイン

数学的性質: 偶然誤差のみを考慮

- 観測値の分散は, 真の分散と誤差の分散に分解可能

$$V(x) = 0 + V(v) + V(e)$$

- 意味のある情報と雑音の比
SN比 $\frac{V(v)}{V(e)}$
- 全体の分散に占める真の分散の割合
信頼性係数 $\frac{V(v)}{V(x)}$

2. 総論 ~統計学と研究デザイン

偶然誤差が大きくなる影響

- 偶然誤差なら平均には影響ない
- 標準誤差が大きくなり、推測統計に影響する
 - 信頼区間が広がる
 - 有意になりづらくなる
- 相関係数などの関連の指標が弱められる
 - 相関の希薄化
 - 見かけの相関は真の相関係数にxとyの信頼性係数の幾何平均を乗じたものになる

$$r_{xy} = r \cdot \sqrt{\frac{s_{vx}^2}{s_{vx}^2 + s_{ex}^2} \cdot \frac{s_{vy}^2}{s_{vy}^2 + s_{ey}^2}}$$

xの信頼性係数

2. 総論 ~統計学と研究デザイン

偶然誤差: 反復計測の効果

反復計測(同一対象をn回計測して平均して測定値とする)の効果
もちろんE(e)=0

$$E\left(\frac{1}{n} \sum x\right) = \mu + E(v) + E(e)$$

$$V\left(\frac{1}{n} \sum x\right) = 0 + V(v) + \frac{1}{n}V(e)$$

2. 総論 ~統計学と研究デザイン

系統誤差(バイアス)の整理

・ 例示は、代表的なパターン

- 情報バイアス
 - 出版バイアス (研究結果として報告される情報の偏り)
 - 応答バイアス (回答の偏り)
- 選択バイアス
 - 自己選択(self-selection)バイアス (対象者の偏り)
 - Healthy worker effect
- 交絡 Confounding

・ 偶然誤差 ↑ 結果が“ぼんやり”する
・ 系統誤差 ↑ 逆転する場合や誇張される場合も

2. 総論 ~統計学と研究デザイン

情報バイアスへの対処

- 出版バイアス
 - 文献研究の際に重要
 - 対処は難しい
- 応答バイアス
 - 客観的な測定項目の導入
 - 影響の評価
 - 単純記述なら → 結果の偏り
 - 2群比較なら
 - 2群とも同様の偏りなら → 影響は小さい
 - 2群で偏りかたに差があれば → 交絡の可能性

2. 総論 ~統計学と研究デザイン

選択バイアスへの対処

- 対象の偏り
 - 結果の一般化ができる範囲を明確に
 - 偏りが、注目する結果に影響を与えるか吟味
- 未回収、無回答等の影響
 - 「アンケートに記入もれがあるケースは除外」
 - 分析者自身が、偏りを作り出している可能性
 - 未回収や無回答が、重要な変数と関連する場合
 - 交絡の可能性

→情報バイアスや選択バイアスによる交絡の可能性

偏りの影響と補正の可能性

- 性別が賛否に関連している場合
 - 標本内での性比で、「全体の結果」は変化する
- 男女ごとに結果を見る ⇒ 層別化
- 男女が半々、と想定した結果を算出 ⇒ 標準化

交絡 (confounding)

- 交絡因子
 - 疾病発生に影響する
 - 曝露と関連する
 - 曝露の原因・疾病と曝露の間・疾病の結果 ではない
- 曝露と疾病の関連を歪ませる

48

交絡の例

- 薬Aの方が効果が高い?

		治療成績			
		改善	不変・悪化	計	改善割合
薬	A	70	30	100	0.70
	B	60	40	100	0.60

- 患者の重症度で分けると

※軽症者		治療成績			
		改善	不変・悪化	計	改善割合
薬	A	60	10	70	0.86
	B	30	0	30	1.00

※重症者		治療成績			
		改善	不変・悪化	計	改善割合
薬	A	10	20	30	0.33
	B	30	40	70	0.43

- 重症度が交絡因子として作用していることがわかる

49

層別化

- 交絡に解析時に対応する手法
 - 基本は研究計画でバイアスに対処すること!
- 層別化により、曝露と交絡因子の間の関連を断つ
 - 層毎の効果の指標の要約
 - 重み付け平均、Mantel-Haenszelの方法
 - 層毎の効果に一致した傾向があれば!
- 当然、未知の交絡因子には対応不能

50

交絡への対処

- 介入研究なら
 - 無作為化が最高の予防法
- 観察研究なら
 - 機知の交絡因子
 - 測定項目に加えて解析で対処
 - 層別化、傾向スコア、多重ロジスティックなど
 - 未知の交絡因子
 - 基本的には対処不能
 - マッチングである程度予防できる可能性

(51)

誤差の影響 全体の整理

- 偶然誤差
 - “有意な結果”が得られにくい
 - 必要標本数の増加
 - 標準誤差, 信頼区間が拡大
 - 相関, 関連の希薄化
- 系統誤差
 - 影響は“どう表れるか”決まっていない
 - 群間比較の場合
 - バイアス自体に群間の相違がなければ, 影響は弱い
 - 群間で相違があれば(交絡であれば), 影響は多大
 - 交絡が確認できれば, ある程度対処可能
 - 存在の可能性を想像できても, 確認できなければ
 - 研究計画で対処することのみ可能

(52)

研究結果には誤差が

- 「残っている」と思って, 考察を進めよ
 - 意味の無い「有意差」を強引に解釈しない
 - 第一種の誤り
 - 本当は重要な結果を見つけられなかったかもしれない
 - 第二種の誤り(検出力)
- 交絡の可能性も心に留めて
- ただし, 過剰に心配することはない
 - 研究は間違ってもある
 - 「ウソ」は悪いが, 偶然の誤りは分かってもらえる

(53)

3. 各論 ～(1)尺度の利用

- 尺度とは
 - 尺度の例と基本構造
 - 尺度を利用する
- 尺度の信頼性と妥当性
 - 信頼性の検討
 - クロンバックの α
 - 妥当性の検討
- 尺度を利用した研究において

(54)

3. 各論 ～(1)尺度の利用

尺度とは

- 例えば、心理状態やQOLなどをはかるツール
 - POMS (Profile of Mood States)
 - SF-36 (MOS 36-Item Short-Form Health Survey)
 - STAI (State-Trait Anxiety Inventory)
 - GHQ30 (The General Health Questionnaire) などなど、多数
- 尺度の基本構造
 - 1尺度に複数の項目(質問)
 - 各項目の選択肢には数段階で得点が
 - 全項目の得点の合計が、尺度得点
 - 複数の尺度(下位尺度)と、まとめた尺度(上位尺度)
 - 尺度得点は、量的変数として処理可能

(55)

3. 各論 ～(1)尺度の利用

研究で尺度を利用する

- 他の方法で測定可能か
- 既存の尺度を探す
 - 先行研究
 - 書籍(たとえば心理測定尺度集 I～VI, サイエンス社)
 - 許諾は必要か
- 自分で作成する
 - 手順
 - 信頼性と妥当性

(56)

3. 各論 ～(1)尺度の利用

尺度を作る手順

- 測定する特性を明確にする
- 項目群の候補を作成
- プレテスト実施
 - 回答の偏り(天井効果・床効果等)チェック
 - 信頼性の検討
 - (可能なら)妥当性の検討
- 項目の修正, 調整
- 実用化

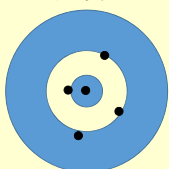
(57)

3. 各論 ～(1)尺度の利用

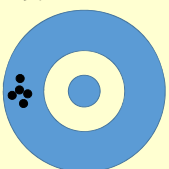
尺度の信頼性と妥当性

- 同じ対象を繰り返し測定した場合のばらつき

尺度A



尺度B



(58)

3. 各論 ～(1)尺度の利用

「尺度」だけでなく

- 私のかかった事例として
 - 生活習慣をはかる尺度




(59)

3. 各論 ～(1)尺度の利用

例 「自己効力感」尺度

- LPC式生活習慣調査(全23尺度)中の1尺度
 - 状況に応じて、適切な判断をすることができる
 - ものごとを計画どおりに、実行することができる
 - まわりの人と協力してやっていくことができる
 - 将来の生活計画を明確にたてることができる
 - いつも目標に向かって努力している
 - 自分の能力(適性)を十分に発揮している
- 各(はい=2, どちらともいえない=1, いいえ=0) 計0～12点
- ちなみに
 - α=0.77
 - 60歳以上女性で平均7.7
 - 「新老人の会」会員集団で9.1という結果

(60)

3. 各論 ～(1)尺度の利用

信頼性の検討

- 内的-一貫性
 - 尺度を構成する項目が、同じ特性を測定しているのか
 - 信頼性係数
 - クロンバックのα(α信頼性係数)が代表的
 - 折半法
- 安定性
 - 同じ対象を繰り返し測定した結果が一致するか
 - 再テスト法(時間をずらして反復測定)
 - inter-observer error(観察者による評価の場合)
 - 相関係数で

※この2つの観点は別もの

(61)

3. 各論 ～(1)尺度の利用

α信頼性係数

- クロンバックのα(Cronbach's alpha)
 - 項目間に相関が高いと「和の分散」が「分散の和」より大きくなることから
- 信頼性をαのみで主張することも多い

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

統計パッケージの出力例
(エクセルでも簡単に計算できます)

信頼性統計量			
	Cronbachの アルファ	項目の数	
	.793	6	

項目合計統計量				
	項目が削除された場合の尺度の平均値	項目が削除された場合の尺度の分散	修正済み項目合計相関	項目が削除された場合のCronbachのアルファ
適切判断	9.57	3.101	.516	.773
計画実行	8.69	2.638	.582	.752
協力実行	8.51	3.357	.399	.794
生活設計	8.83	2.368	.647	.735
目標努力	8.68	2.889	.501	.747
能力発揮	8.96	2.344	.608	.750

尺度がk個の項目から構成されるとき、各項目の分散の総和と尺度得点の分散から計算

(62)

3. 各論 ～(1)尺度の利用

妥当性の検討

- 的(まと)が見えていないと評価困難
- 内容妥当性
 - みんな考えて
- 構成概念妥当性
 - 因子分析の結果に納得
 - 確かに“使える”結果になった
- 基準関連妥当性
 - 真の値、類似の特性、関連する特性、などと相関

(63)

3. 各論 ～(2)研究デザインの整理

- 再度、デザインとは
- 観察研究のデザイン
- 介入研究のデザイン
- デザインを決めるには

(64)

3. 各論 ～(2)研究デザインの整理

再度、デザインとは

- 研究計画における、「方法」の具体的内容
 - 対象の設定
 - 条件
 - 比較群の設定
 - 数
 - 注目する因子
 - 測定手段
 - 介入の有無
 - その他
 - 研究の“時間”
 - 分析の方法 など

(65)

3. 各論 ～(2)研究デザインの整理

例えば、代表例として

- 目標を達成できないことが(ほぼ)確実なデザイン
 - 対象数の不足
 - 検出力が低い
 - 必要な対象数を計画時に
 - 事後に検出力を算出
 - 観察期間の不足
 - 介入の効果や予後を観察する期間の長さ
 - エンドポイントの変更、観察項目の変更
 - 比較群の不存在
 - 「効果」は比較によって確認できる
 - 研究の外部に比較できるデータは

(66)

3. 各論 ～(2)研究デザインの整理

デザインの効用

- マッチング、前後比較デザイン
 - ペアとなるデータの間で、個人の効果が相殺される
 - サンプルサイズを押しえられる
 - ペアの間で、バイアスの程度が同じなら
 - 比較結果にバイアスは少なくなる
- クロスオーバーデザイン
 - 有効だが、適用できる状況が制限される
- 無作為化比較試験(RCT)
 - 理論上、交絡を防げる
 - 介入研究の理想のデザイン

(67)

3. 各論 ～(2)研究デザインの整理

1対1マッチングを行えば

- 対応あるサンプルで「個人の効果」が同じなら

個人の測定値 = 集団の平均 + 個人の効果 + 誤差

$$x_{0i} = \mu_0 + v_i + e_{0i}$$

$$x_{1i} = \mu_1 + v_i + e_{1i}$$
- ペア間の差は、集団間の差(+僅かな誤差)
- 分析手法は、やや複雑になるが
- サンプル数は押さえることができる

(68)

3. 各論 ～(2)研究デザインの整理

どうしても「少数例」

- 研究テーマが重要なものなら、きちんと論文に
 - 原著は難しいが、報告として
- もちろん「誤差を押さえる」努力も
- 冗長な考察や複雑な解析を極力省き
- 生データを丁寧に記述
 - 自分でデータを追加して、原著に
 - **メタアナリシスの素材に**

(69)

3. 各論 ～(2)研究デザインの整理

おわりに ～ 統計家に事後相談するときは


- 統計家は
 - 扱っているテーマに関する専門知識は持たない
 - 統計学の枠組みで質問
- 以下の事項を明快に整理してから相談を
 1. 測定に用いた尺度(量、順序、名義)は
 2. どのデータを同時に組み合わせるか
 3. 注目する変数以外に、影響を与える変数は
 4. 分析によって示したいのは
 - データの分布を示したい
 - 2変数の関連を示したい
 - 複数の変数の、ある変数への影響の大きさを比べる
 - 予測、判別の方法を導きたい
 - 複数の要因間の関連を知りたい

(70)

3. 各論 ～(2)研究デザインの整理

お疲れ様でした

- 説明不足の点や分かりづらい点など、メールでの質問を受け付けます
 - いつでも、何でもすぐに反応、とは行きませんが
- 統計学など、よく分からないことは恥ずかし事ではありません
 - スルーしたり、投げ出すことは良くないですが
- 統計や研究法を専門にしている方に、遠慮無く質問、相談して下さい
 - 相談者が熱心で、意義ある研究なら、ちゃんと対応してくれるはず



(71)