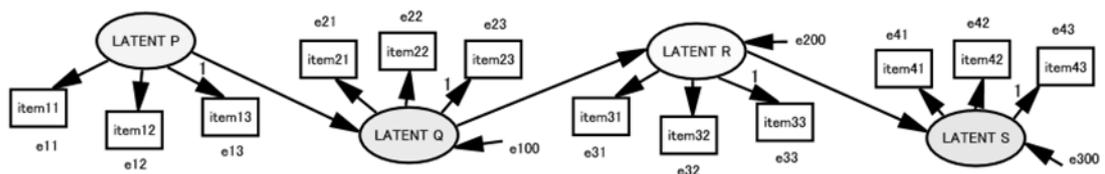


臨床で働きながら研究をしよう

統計の裏技と SPSS の使い方

北村 俊則

北村メンタルヘルス研究所



はじめに

このマニュアルは北村メンタルヘルス研究所が主催する「臨床で働きながら研究をしよう：統計の裏技と SPSS の使い方」という研修会のためのハンドアウトをもとに編集したものです。この研修会は、医療・福祉・教育現場で働いている医師・看護師・助産師・保健師・ケースワーカー・心理士・保育士の皆さんが、御自分の疑問について研究デザインを組んで、データを収集し、パソコンに入力してから、統計法も使って解析し、結果を報告できるようになるためのものです。研究デザイン・使用する尺度・統計法・レポートの書き方は表裏一体です。従来の教科書や参考書は、統計学の専門書、統計ソフトウェアの使用解説マニュアル、研究論文の検索や読み方のガイドブック、論文の書き方や投稿の指導書など、いくつかに分かれていました。臨床に従事しながら研究をして、それを論文にしたり、学会で発表することを考えている人にとって 1 冊ですべての道筋が分かる書籍はなかなかありませんでした。本書の第一の目的はそこにあります。

本書は統計学の専門書ではありません。複雑な数式の叙述はありません。出てくる数式は憶える必要はありません。最初に架空のデータセット（事例は産後うつ病）を準備し、これをもとに研究と統計の基礎を学びます。サンプルデータは北村メンタルヘルス研究所のホームページから無料でダウンロードできます。

また、数学の専門家はだれも教えてくれなかった「裏技」を盛り込んでいます。統計の手法、SPSS の使い方、原稿の執筆、投稿と査読への反応など、様々な場面で「先輩」から伝授される秘伝をお伝えします。さらに各回のテーマに関する先行研究の講義もオマケとしてついています。

- * 自分の臨床データをまとめて学会で発表してみたい
- * 病院・クリニックのサービス向上に役立つ資料を集めてレポートを書いてみたい
- * 日常のデータ収集をもっと科学的にしてみたい
- * 最新研究論文を読んで、その統計結果を読めるようになりたい

こうした希望をお持ちの臨床家が、ご自分の臨床現場の疑問に答える調査ができるようになるのが、本書の目的です。

2013 年 3 月

北村メンタルヘルス研究所

北村俊則

HP: <http://www.institute-of-mental-health.jp/>

本書での使用するソフトウェア

- SPSS 19.0
- AMOS 19.0

北村メンタルヘルス研究所のホームページ (<http://www.institute-of-mental-health.jp>) に本書で用いたサンプルデータが掲載されています。自由にダウンロードしてください。

目次

1	臨床における科学研究と統計 【概論】	1
	1 科学と研究	1
	2 研究計画書	10
	3. バイアス	19
2	変数（変量）と統計法【代表値と散布度】	21
	1 変量と尺度	21
	2 SPSS の構成とデータ入力	27
	3 代表値と散布度	42
	4 変数の変換	63
	5 欠損値の扱いかた	64
3	2人のアセスメントを一致させよう — DSM の大うつ病エピソード診断を中心に 【臨床評価の信頼性】	67
	1 測定の信頼性と妥当性	67
	2 臨床評価の信頼性	70
4	うつ病を見つけよう — EPDS を中心に 【スクリーニング法の妥当性】	89
	1 メンタルヘルス領域におけるスクリーニング	89
	2 スクリーニング検査の妥当性指標	115
	3 自己記入式調査票の検討課題	133
5	グループの特徴を見る 【カイニ乗検定】	139
	1 カイ 2 乗検定	139
	2 フィッシャーの直接確率検定	173
	3 マクニマー検定	180
	4 相対危険度とオッズ比	185

6	危険群の判定 【t test, one-way ANOVA】	197
	1 2グループの平均値の比較：独立したサンプルの t 検定	197
	2 単一グループで2変数の平均値の比較：	217
	3 変数の平均値を比較したい：t 検定の道筋と注意点	223
	4 平均値の差を分散から検定する：2群の一元配置分散分析	231
	5 3つ以上のグループの平均値を比較したい	238
	6 平均値の差の検定のまとめと検討	243
7	重症度の評価 【Pearson 相関係数, 偏相関係数, 内的整合性】	247
	1 順位変数と順位変数の関連を見たい	247
	2 交絡と偏相関係数	263
	3 内的整合性	273
8	産後の抑うつ状態の予測 【重回帰分析と媒介変数】	285
	1 相関から回帰へ	285
	2 平均への回帰	294
	3 回帰分析から重回帰分析へ	298
	4 SPSS による重回帰分析	308
	5 重回帰分析の応用	327
	6 3つの変数の関係：再考	327
	7 moderator	341
	8 mediator	344
9	見える現象から見えない概念へ（1）— PBI の下位尺度を作ろう 【探索的因子分析】	347
	1 見える現象から見えない概念へ	347
	2 探索的因子分析の基礎	354
	3 探索的因子分析の手順	357
	4 SPSS の実際	370
	5 どのような時に探索的因子分析をするのか	380
10	見える現象から見えない概念へ（2）— PBI 下位尺度の頑強性を知ろう 【確認的因子分析】	381
	1 探索的因子分析から確認的因子分析へ	381
	2 AMOS の基礎	385

3 共分散構造分析の実行	402
4 EPDS の因子構造を見てみよう	415
11 うつ病の原因を人生史から探る（1） — 影響の連鎖を知ろう	
【パス解析と構造回帰モデル】	
	423
1 パス解析の基礎	423
2 パス解析の実際	438
3 モデルの改良と比較	445
4 様々なパスモデルと同値モデル	448
5 パス解析から構造回帰分析へ	451
6 さまざまなモデル構築	457
12 統計の結果から発表まで 【臨床への還元】	461
1 多項目評価法の作りかた	461
2 原稿の準備と引用スタイル	471
2 参考書	479
付録	483
索引	499

本書でよく出てくる略号表とギリシャ文字一覧表

略号	正式名	意味
df	degree of freedom	自由度
m	mean	平均
n	number (of participants)	標本数
p		有意水準
s	standard deviation	標準偏差
S^2	variance/mean square	分散 (平均平方)
ss	sum of squares	変動 (平方和)

大文字	小文字	読み方	アルファベット表記
A	α	アルファ	alpha
B	β	ベータ	beta
Γ	γ	ガンマ	gamma
Δ	δ	デルタ	delta
E	ε	イプシロン, エプシロン	epsilon
Z	ζ	ゼータ	zeta, dzeta
H	η	イータ, エータ	eta
Θ	θ	シータ	theta
I	ι	イオタ, アイオタ	iota
K	κ	カッパ	kappa
Λ	λ	ラムダ	lambda
M	μ	ミュー	mu, my
N	ν	ニュー	nu, ny
Ξ	ξ	グザイ, クシー	xi
O	o	オミクロン	omicron
Π	π	パイ	pi
P	ρ	ロー	rho
Σ	σ	シグマ	sigma
T	τ	タウ	tau
Υ	υ	ウプシロン, ユプシロン	upsilon, ypsilon
Φ	ϕ	ファイ, フィー	phi
X	χ	カイ, キー	chi, khi
Ψ	ψ	プサイ, サイ, プシー	psi
Ω	ω	オメガ	omega

1

臨床における科学研究と統計

概論

1 科学と研究

1. 科学を定義してみよう

科学とはなんだろう。科学の対象は自然であったり、社会現象であったりする。そうした対象のうち、たいていはあるひとつの領域に科学は注目する。そして、特定領域に見られるさまざまな事象に何らかの法則があるのだと考えるのが科学の手法である。さらに、何らかの法則が超自然の力（例えば神の摂理）によって決められていると考えるのではなく、合理的な法則が存在すると考えるのである。こうした法則的認識を目指す合理的知識の体系が科学である。そして、合理的知識を積み重ねることで法則を組み立てようとする作業も科学と呼ぶ。

表 1-1. 科学の分類

科学の分類	例
自然科学 natural sciences	物理学, 化学, 生物学, 天文学, 工学
社会科学 social sciences	法学, 政治学, 経済学, 社会学
人文科学 art and humanities	哲学, 美学

では科学にはどんな種類があるのだろうか。次の表に挙げるように、我々が「科学」という用語から連想する学問領域は多数ある（表 1-1）。通常、これらは自然科学、社会科学、人文科学の3つに分けられることが一般的である。この分類は何を基礎にしているのだろうか。たとえば物理学や生物学は、自然現象を対象とした学問領域である。哲学や美学は、外界に存在している自然現象を学問の対象とするのではなく、ヒトの心の中の事柄（「何が正しいことか」「何が美しいことか」など）を学問の対象としている。中間に存在する社会科学に分類される法学、政治学、経済学などは、個々のヒトを学問対象とするのではなく、有機体としてのヒトの集合を学問対象

としている。したがって、自然科学、社会科学、人文科学という3分類は、学問の対象による分類であるといえる。

ところで、これらさまざまな学問領域をそれぞれの学問で使われる研究手法で見てみよう（図 1-1）。「顕微鏡で調べる」や「試験管で実験する」といった手法は、物理学、化学、生物学の研究手法としてすぐ思いつくものである。一方、「美しい作品を鑑賞する」「何が正しいか考える」「歴史的意味を考える」などは、人文科学の手法である。両者を区別している特徴のひとつは、対象を数値で表すことであろう。つまり、物理学、化学、生物学などの研究では、物の重さや大きさ、変化の速さなど、すべて数字で表す。一方、「何が正しいか」「何が美しいか」という手法では数字を使うことはない。自然科学の基礎は対象を数値で表すことであるともいえる。



図 1-1. 科学の対象と研究手法

例えば音楽学は美学の一領域であり、通常は人文科学に分類される。しかし、音楽学はいつも数値を扱わないものでもない。たとえば三和音（例：ドミソの3つの音）がヒトの耳に快に聞こえるか不快に聞こえるかを、3つの音のヘルツを測定して、快に聞こえる和音の法則を求めようとすれば、それは明らかに数値によって法則を求める学問といえる。歴史学を見てみよう。古い建物の建設時期を求めるのに、建築に使われた木材の放射性炭素（¹⁴C）の量を測ることがある。炭素の同位体のうち、時間とともに壊変し別の元素になる炭素 14 の半減期と、本来の炭素 14 の量を推定することから、かなり正確に木造建築物の建築年代を推定できるのである。こうした手法で法隆寺の建築年代を推定する手法も、明らかに数値によって法則を求める学問である。文学を見てみよう。ある一人の作家の作品の内容を、数値で表して、作風の特徴を研究することも行われる。こうしてみると、学問分野の分類は、研究対象というより、その研究手法の特徴で行えるといえよう。三和音のヘルツの組み合わせから快・不快を予測する、木材の放射性炭素から建物の建造年を推定するなど自然科学の手法なのである。自然科学の特徴は、対象が何であっても（太陽、惑星、動物、細胞、音楽、歴史……）、それを数値で表記することである。

さて再度、科学の定義に戻ろう。科学は合理的手法で知識を体系化する営みである。ここで合理的とはどのようなことを指すのであろうか。まずは、論理的整合性が求められる。AはBである。BはCである。したがって

AはCであるといった整合性が科学には求められる。これは哲学の基礎であり、当然に自然科学の論証に際しても求められる。これから読者諸賢が行おうとしている臨床研究は最終段階で報告書や論文として公表されるものである。この際に、論旨の整合性は強く求められる。

科学に求められる合理性を担保するもうひとつの重要な概念は弁証法的思考であろう。ソクラテスの問答法・対話法に始まる哲学的思考である。反対の立場にある2人（あるいは2グループ）が、質問と討論を繰り返す。ヒトが提示するある考えには当然、矛盾が含まれている。この矛盾を一方の者が指摘し、指摘された当事者は考えに修正を施す。理性的思考を刺激し、アイデアを生み出して行くという手法である。1個人のなかでも弁証法的思考は取れる。当初の考えの矛盾点を自ら指摘し、それに自分自身が答えて行くことが科学者に求められる資質である。

こうした思考方法は、臨床研究をスタートする前に、是非とも身につけるべきものである。後で述べるように、研究計画書を書き、実際に集めたデータを解析し、報告書を作成するときに必要な手法である。

2. 基礎研究と臨床研究

医療・看護・福祉領域に働く人々が研究を行う場合、その内容はさまざまである。ここでは便宜的に「基礎研究」と「臨床研究」にわけて考えてみたい。前者は研究者個人あるいは人類全体の「知的好奇心」を満足させることを目的にしている。患者や人類への直接的な恩恵はないが（遠い将来、恩恵を導くかもしれないが）、知の集積に役立つ研究である。一方、後者は近い将来あるいは中程度の将来の患者あるいは人類全体の健康に寄与することを目的として行われる。

臨床家にとっては、行っている研究の成果がどのように臨床に反映されるかを意識することが、研究のポテンシャルを維持する。多忙な臨床業務の合間に研究のための時間を割くのであるから、よほどの動機づけがなければ長時間、興味を持つことは困難であろう。「この研究は誰の役に立つのか？」という疑問を、研究を開始する前に自身に問うてみるのが勧められる。つまり、自身の研究が臨床研究であることを意識することが大切である。もちろん、臨床家が基礎研究をすることに意味がないというものではない。臨床家であるがゆえに行える基礎研究も多い。しかし、本書は、研究に興味を持った臨床家を読者と設定している。

そこで次に、臨床研究の順序について考えてみよう。日々の臨床場面で遭遇した疑問が臨床研究のスタートラインである。ここでAクリニックの助産師グループに登場してもらおう。本書を通じて、研究を行って行く人々である。この助産師たちが、「産後うつ病の患者の中で、自殺企図歴のある者は生活史を聴取すると子ども時代に虐待を受けたと報告する率が高いのではないかと感じた」と感じた疑問が研究の作業仮説となる。

「うつ病患者中で、自殺企図歴のあるものは自殺企図歴のない者に比較して、15歳以前に両親のいずれから心理的・身体的・性的虐待あるいはネグレクトを受けた率が高い」というのが、この仮説になる。そこでAクリニックの助産師たちは「うつ病」「自殺」「児童虐待」などをキーワードとしてネット上で先行研究をブラウズ（検索）するであろう。先行研究を概観しても、自分の仮説を証明あるいは否定する文献に遭遇しなかったとしよう。あるいはいくつかの文献が見つかったも、一定の結論が出ていなかったとしよう。さらに、先行研究はあるものの、すべて欧米の研究であり、日本からの報告は皆無だったとしよう。また、多くの先行研究はあるものの、新たな疑問（例えば、うつ病患者の年齢に対する考察がどの論文でも希薄だった）が発生したとしよう。こうした状況は、Aクリニックの助産師たちの疑問が臨床研究として調査を行う価値があることを示している。

作業仮説を明確にしたAクリニックの助産師グループは、次に研究のデザインについて考える。今回の例でいえば、彼女たちは自分の研究の対象を今後、うつ病の診断がついた患者とした。そして、「15歳以前に両親のいずれから心理的・身体的・性的虐待あるいはネグレクト」の評価方法を明確にし、データを集め、集まった結果を統計解析し、論文や報告書として公表することとなる。仮説を支持する結果を得たとすると、それは例えば「う

つ病患者については虐待歴を含めて子ども時代の生活史を慎重に聴取する」という臨床へのアドバイスとして還元できる（図 1-2）。

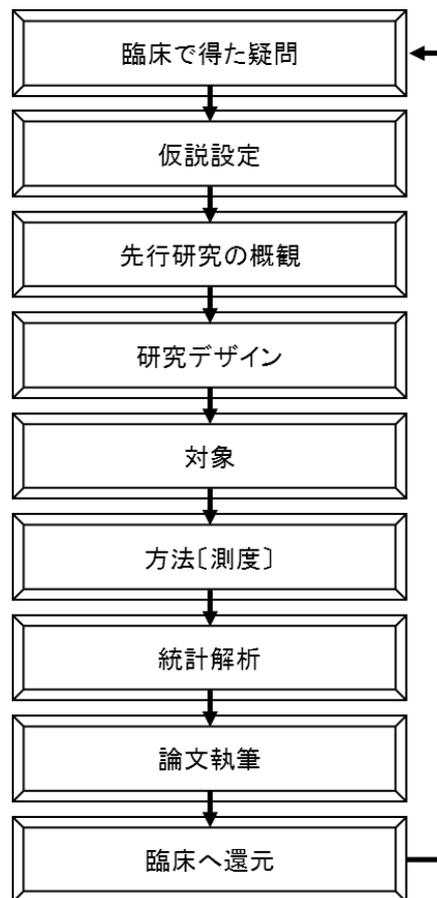


図 1-2. 研究手続きの連鎖

ところで、こうした結果を得たスタッフには新しい疑問が発生する。Aクリニックの助産師グループには、「子ども時代に虐待を受けたほうがうつ病になった時の自殺企図に結びつくのはなぜだろう」という新たな疑問が湧いてきた。ある日、自分が担当することになったうつ病患者の生活史を聞くと、頻回の自殺企図歴があり、加えて小学校時代に父親から受けた心理的虐待が認められた。さらに、対人関係の特徴から境界性パーソナリティが推定できた。そこから、その助産師は「うつ病の際の自殺企図と子ども時代の虐待を受けた体験の関係を結びつけているのは境界性パーソナリティではないか」という疑問を持つようになった。これが新しい疑問に対する仮説の芽となるのである。常に自身の臨床体験に学問的興味の日を向けることで、研究のテーマはいくつも出てくるのである。

では、こうした研究マインドを持った研究者になるには、日々どのような努力が必要であろうか。Aクリニックの助産師たちのように、臨床から出た疑問の解決を文献に求めることが何よりも大切である。親切な先輩が自身の経験から教えてくれることも多い。しかし、個人的経験はバイアスがかかったものであることは疫学が教えるところである¹。先輩・指導者の言葉も、常に科学的批判の精神をもって聞き、決して盲信してはいけない。

もちろん、平素から新しい文献に目を通し、自分の専門領域の標準的教科書は通読し、基本的知識を身につけよう。その内容を臨床に「翻訳」する作業をすることも同様に重要である。論文や教科書を読んだら、それが自分の明日の臨床にどのように活用すればよいか考えよう。

¹ 例えば Cohen, P., & Cohen, J. (1984). The clinician's illusion. *Archives of General Psychiatry*, 41, 1178-1182.

3. 臨床研究の基本は測定である

自然科学は対象を測定し、数値化して法則を求める学問である。臨床研究も自然科学の手法をとる限り、測定と数値化がその基礎となる。おそらく臨床家が現場で取り扱う事象のほとんどは測定可能であろう。メンタルヘルス領域について見ても、精神・心理症状、行動・兆候、パーソナリティ、社会適応、対人関係など、多くの事象が測定の対象となっている。

数値化というと、1, 2, 3・・・といった正の整数や、1.54, 3.76 といった小数を含む数値が連想される。しかし、「あり」「なし」という評価も数値化できる。たとえば性別を考えてみよう。研究対象がヒトであれば男性あるいは女性という性差が記述される。こうした定性的測定も数値としてあらわされる。たとえば男性であれば“1”を与え、女性であれば“2”を与えるといった手法はもっとも通常の方法である。男性を“0”（ゼロ）、女性を“1”としてもよいし、男性を“1000”、女性を“2000”にしてもよい。いずれの方法でも数値をもって性差という対象者の特徴を数字で表しているのである²。

評価をする、つまり対象のある特徴を数値で表し、それを科学研究の出発点とする場合に、その評価の確度が高いことが、研究結果や導き出される結論の質の高低を決定している。本書の第3章と第4章はこのことを詳しくみて行く。

評価・数値化の確度は、信頼性(reliability)と妥当性という2つの概念で決められている。信頼性とは、同一の評価を繰り返して行っても（試験・再試験信頼性 test-retest reliability）、また別の評価者が行っても（評定者間信頼性 interrater reliability）、だいたい一致した結果（評点）が出ることを指す³。一方、妥当性とは本当に評価したいことをその評価法が測っているかを指している。

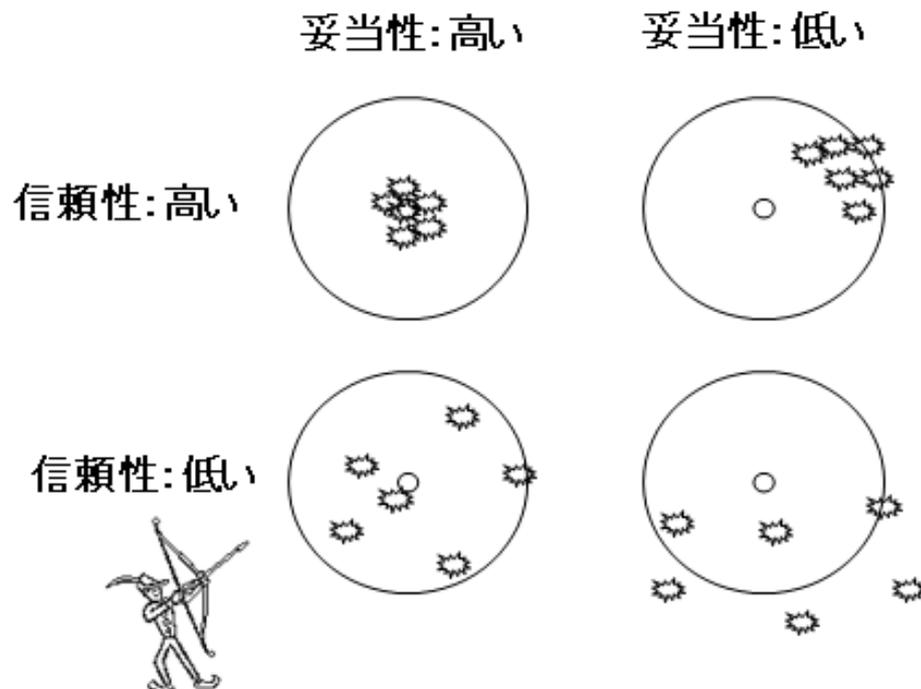


図 1-3. 評価の信頼性と妥当性

² こうした値の割り振りは線形変換 (linear transformation) をしたことであり、たとえば他の変数との相関係数を求めるような場合に同一の結果を得られる。線形変換については第8章参照。

³ 評定者間信頼性と試験・再試験信頼性については第3章でくわしく解説する。

このことを例えば、弓（アーチェリー）で的を射る作業を比喻として考えてみよう（図 1-3）。弓に矢を添えて狙いに射る行為が科学的に評価する行為であると想定しよう。アーチェリーの選手が科学者であると見立ててみよう。的の中心が、まさに研究者が評価しようとしている現象そのものである。選手（評価者あるいは評価法）にはさまざまな者がいる。だいたいの中心に矢が当たるが（妥当性が高い）、射るたびに上下左右にブレたりブレなかったりする（信頼性が低い）者もいる。一方、ブレることはなく、だいたい同一点に矢が当たるが（信頼性が高い）それは的からは外れた場所である（妥当性が低い）者もいる。射るたびに上下左右にブレたりブレなかったりして（信頼性が低い）、なおそれは的からは外れた場所である（妥当性が低い）者もいる。選手（評価法）としてもっとも望ましい（確度が高い）のは、当然、だいたいの中心に矢が当たり（妥当性が高い）、かつ射る時々に上下左右にブレない（信頼性が高い）者である。矢の当たる位置が常に一定であることが信頼性であり、矢が的に当たるのが妥当性である。

評価法の信頼性や妥当性を検討するに当たって、さまざまな評価法をいくつかに分類しておこう（表 1-2）。まず、対象被験者のある特徴（例えば抑うつ状態）がどれほど重症かを評価すること考えてみよう。評価する主体は誰であろうか。もちろん、臨床スタッフや研究スタッフが患者や一般住民を前にして、抑うつ状態の重症度を（客観的に）測ることが行われる。この場合、評価の主体は外部者である。多くはその領域の専門家である。一方、メンタルヘルス領域の評価対象の多くは患者や一般住民の心の中で起きていることがらである。抑うつ気分などがそのよい例である。そこで被験者自らが自身の状態を（主観的に）評価することも行われる。

また、評価法は、その手順がかなり細かく体系だっている場合と、評価するものの自由に任せる場合がある。前者を構造化された評価、後者を非構造化された評価と称する。外部者による構造化された評価が構造化面接（structured interview）である。本人による構造化された評価が自己記入式調査票（self-rating scale, questionnaire）である（構造化された評価法と構造化されていない評価法の間のもので「半構造化」と表現することもある）。本人が非構造化されたフォーマットで記載する自由記載も評価法として活用されることも少なくない。

表 1-2. 評価方法の分類

	測定主体	
	外部者	本人
構造化 structured	構造化面接 structured interview	自己記入式調査票 self-rating scale (questionnaire)
非構造化 unstructured	評価尺度 rating scale	自由記載

信頼性と妥当性の両方が高いことが科学研究の道具として求められる。しかし、ある評価法（例えば、統合失調症で見られる各種の陰性症状）がこれに該当しないこともよくみられることである。もし妥当性は高いが信頼性が低いのであれば、たとえば評価方法を構造化し、その信頼性を向上させることが求められる。たとえば統合失調症で見られる各種の陰性症状の重症度評価を臨床家の直感に任せていれば、その信頼度は低いであろう。そうした場合は、重症度評価の手順を文章化して、だれでも同様な結果に至るように工夫すればよい。信頼性があっても妥当性が低い場合は、概念の見直しをしたり、操作的評価法の見直しが必要になる。もし陰性症状の重症度評価尺度を用いても妥当性に欠けるのであれば、そもそも陰性症状の重症度といった 1 次元の概念規定をすることの可否を検討するべきであろう。

もちろん、信頼性も低く妥当性も低いのであれば、評価手段としてこれ以上、使うわけにはいかない。ただし、科学者として注意すべきは、科学としての価値を見いだせると感じる場合は、信頼性と妥当性の双方を向上させ

る工夫を行うべきである。ケネディ大統領とジョンソン大統領を通じて国防長官であった Robert McNamara が「努力の必要があるのは重要な事柄を測定できるようにすることであり、測定できる事柄を重要にするのではない The challenge is to make the important measurable, not the measurable important」と述べたといわれている⁴。「確度が低いから科学の対象から除外する」のではなく、「確度は低い科学として重要であるから、確度の高い評価法を開発する」ことが、我々のとるべき道であろう。

⁴ Slade, M., & Pribe, S. (2001). Are randomized controlled trials the only gold that glitters? *British Journal of Psychiatry*, 179, 286-287.

2 研究計画書

1. 論文の構成

研究の成果は論文として公表される。公表されないのであれば調査研究は意味をなさない。これから行おうとしている研究は成果の公表をもってゴールとする。したがって、研究をスタートする際は、その終点の形を考えておこう。

通常、論文の構成は以下のようになっている（図 1-4）。

表紙	標題・著者名・短縮した標題・単語数・表の数・図の数
はじめに	先行研究のまとめ 残された問題と今回の研究の目指すもの 仮説設定
方法	対象 調査方法 統計手法
結果	
考察	今回のデータの検討 臨床への意味 研究の限界
謝辞・利益相反など	
文献	
表	
図の説明	
図	
雑誌のホームページに掲載する資料	

図 1-4. 論文の構成

まだひとつもデータが集まっていない状態で論文を書けるわけもない。しかし、表紙、はじめに、方法の3つの部分はデータがなくても書くことはできる。ここを書いたものが研究計画書 (research protocol) である。研究を始める前には研究計画書を必ず書く。研究計画書のないまま研究を開始すれば、かなり高い確率で途中で困難に遭遇し、あるいは頓挫することになる。また、ヒトを対象とする研究であれば倫理委員会 (ethical committee, institutional review board) の承認が必須であるが、研究計画書は倫理委員会に提出するものでもある。研究計画書の骨格は (a) 先行研究のまとめ (b) 仮説設定 (c) 対象 (d) 方法 (e) 統計手法 で構成されている。この4点を次に概観しておこう。

1) 先行研究のまとめ

「うつ病患者の中で、自殺企図歴のある者は生活史を聴取すると子ども時代に虐待を受けたと報告する率が高いのではないかと考えたAクリニックの助産師たちは、自身の疑問に答えられるかもしれないデータを探して、ネット検索を熱心に行った。この時点で重要な先行研究を網羅しておくことは、データ収集が終わり、解析に入る際にたいへん役立つ。

メンタルヘルス領域ではアブストラクト誌として Evidence-based Mental Health が BMJ publishing から刊行されている。大変有用な情報が詰まった雑誌であり、毎号の通読は価値がある。

2) 仮説設定

ここで研究者は自分の研究デザインにおける基準変数 (criterion variable) と説明変数 (explanatory variable) を明確にすべきである。基準変数は従属変数、目的変数、外的基準と言われる。説明変数は独立変数とも言われる (表 1-3)。研究の「理論的背景」を考えてこれを決めて行くのが通常の方法である。

表 1-3. 基準変数と説明変数

基準変数 criterion variable	説明変数 explanatory variable
従属変数 dependent variable/bound variable	独立変数 independent variable/free variable
目的変数	
外的基準	

例えば、1000 人規模の地域住民疫学調査を行ったとしよう。うつ病、パニック障害、アルコール症などの生涯有病率を調査した。さらに年齢と性別も確認した。ここから仮説を設定してみよう。リサーチ・クエスチョンは「うつ病の生涯有病率に性差はあるか?」とした。この場合、作業仮説 (operational hypothesis) は「女性におけるうつ病の生涯有病率は男性におけるうつ病の生涯有病率より高い」となる。そして、基準変数は「うつ病の生涯有病率」であり、説明変数は「性差 (男性・女性)」である。つまり、「XXを説明したい」という表現のXXにあたるものが基準変数 (従属変数・目的変数・外的基準) であり、「△△で説明したい」という表現の△△にあたるものが説明変数 (独立変数) である。従って、リサーチ・クエスチョンと作業仮説は内容的には同一であり、その記述方法が違うのである。「はじめに」の最終パラグラフはリサーチ・クエスチョンあるいは作業仮説のいずれかを明記する。

そこで、周産期メンタルヘルス領域の例をいくつか挙げながら基準変数と説明変数の関係を見てみよう。

まず、仮説「産後 5 日目の抑うつ状態 (マターニティ・ブルー) が高いと産後 1 カ月の抑うつ状態が重症になる」を考えてみよう (図 1-6)。産後 5 日目の抑うつ状態から産後 1 カ月目の抑うつ状態が (ある程度だがしかし統計的に有意に) 予測できるというのが狙いの仮説である。もし、これが確認できれば、産後 1 月目の産後うつ病に対する事前の予防的介入を、産後 5 日目に抑うつ状態 (マターニティ・ブルー) が高い母親に絞り込むことができる。臨床的意味を持つ研究テーマであると言えよう。この研究では、「XXを説明したい」という表現のXXにあたるもの (基準変数) が産後 1 カ月の抑うつ状態であり、「△△で説明したい」という表現の△△にあたるもの (説明変数) が産後 5 日目の抑うつ状態である。このモデルでは 1 つの基準変数を 1 つの説明変数で予測する形をとっている。

基準変数と説明変数

産後5日目の抑うつ感情(マターニティ・ブルーズ)が高いと産後1ヶ月の抑うつ状態が重症になる

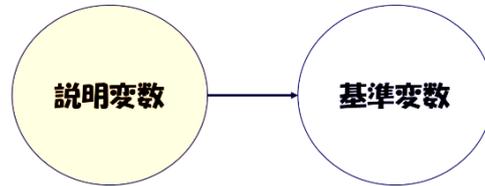


図 1-6. 基準変数と説明変数 (1)

次に、「児童期に受けた父からの養育と母からの養育が愛情の薄いものであるほど産後1ヶ月の抑うつ状態が重症になる」という仮説を考えてみよう(図1-7)。このモデルでは変数が3つ存在する。すなわち、産後1ヶ月の抑うつ状態、児童期の父からの養育、児童期の母からの養育である。そして、「XXを説明したい」という表現のXXにあたるもの(基準変数)が産後1ヶ月の抑うつ状態であり、「△△で説明したい」という表現の△△にあたるもの(説明変数)が、児童期の父からの養育および児童期の母からの養育である。つまり、1つの基準変数を2つの(複数の)説明変数で予測するというのがこのモデルである。このように、1モデル中に説明変数は2つ以上あってもよい。

基準変数と説明変数

児童期に受けた父からの養育と母からの養育が愛情の薄いものであるほど産後1ヶ月の抑うつ状態が重症になる

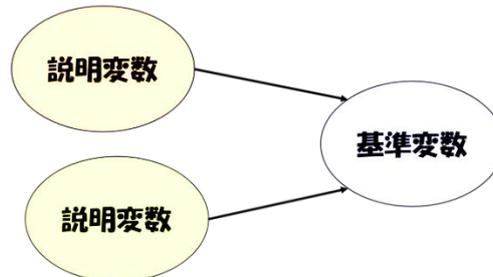


図 1-7. 基準変数と説明変数 (2)

さらに、「妊娠後期の抑うつが強いほど同じ時点の不安が強い。妊娠後期の不安が強いほど同じ時点の抑うつが強い」という仮説を見てみよう(図1-7)。上記2つのモデル違う点はどこだろうか。最初の2つのモデルでは、説明変数が時間的に先行して発生しており(あるいはそう考えるのが受け入れられる)、基準変数が後から発生している。したがって、説明変数が基準変数を「予測する」という表現を用いたのである。ところが、第3のモデルでは、抑うつと不安は同時期(妊娠後期)に測定している。抑うつと不安の間には関係があることが仮説として考えられた。つまり、「抑うつが重症な妊婦は不安も重症だ」「不安が重症な妊婦は抑うつも重症だ」という2つの可能性が同時に存在する。しかし、どちらがどちらを予測しているという因果関係は推定できない。一方が基準変数であり、もう一方が説明変数であり得るが、その逆もあり得るのである。

基準変数と説明変数

妊娠後期の抑うつが強いほど同じ時点の不安が強い。妊娠後期の不安が強いほど同じ時点の抑うつが強い。

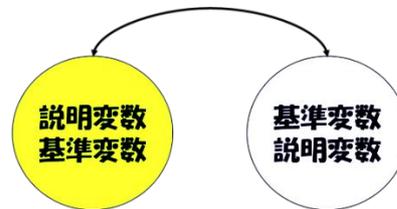


図 1-7. 基準変数と説明変数 (3)

最後のモデルを見てみよう。ここでの仮説は「妊娠初期の不安が強いほど産後 5 日目の抑うつ感情（マターニティ・ブルーズ）が高く、産後 5 日目の抑うつ感情が高いほど産後 1 ヶ月の抑うつ状態が重症になる」である（図 1-8）。継時的に評価しているから、測定時期が早い項目が測定時期が後の方の項目の程度を予測していると推定できる。まず、「XXを説明したい」という表現のXXにあたるもの（基準変数）が妊娠初期の不安であり、「△△で説明したい」という表現の△△にあたるもの（説明変数）が、産後 5 日目の抑うつ感情（マターニティ・ブルーズ）である。さらに加えて、次に産後 5 日目の抑うつ感情（マターニティ・ブルーズ）が今度は説明変数になり、産後 1 ヶ月の抑うつ状態を説明するのである。A ⇒ B ⇒ C という流れ図におけるBを介在変数 (mediator) と呼ぶ。

基準変数と説明変数

妊娠初期の不安が強いほど産後5日目の抑うつ感情(マターニティ・ブルーズ)が高く、産後5日目の抑うつ感情が高いほど産後1ヶ月の抑うつ状態が重症になる

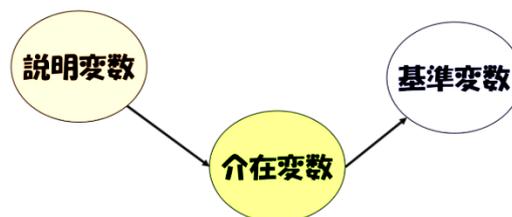


図 1-8. 基準変数と説明変数 (4)

研究計画書の仮説設定は研究者の自由な発想で決められる部分である。この際に、基準変数と説明変数を明記することは以降の作業を容易にする。データが集積したのちに行う統計解析は、こうしたモデルの適否を見る手段である。もし可能なら、上記のようなモデル図（パス図ともいう）を書いて、研究計画書に添付することも賢明な方法である。

ここまで議論すれば、基準変数と説明変数の存在しない研究はありえないことは明らかであろう。実態を調査したという内容であっても、調査者が明記しない（あるいは気づいていない）仮説が存在するのである。

3) 対象

臨床例であれ非臨床例であれ、対象選択の基準と手法、予定人数などを研究計画書に明示する。対象選択方法は結果のバイアス (bias) の大きな源になる。バイアスについては後述。

4) 方法

作業仮説と対象が決まれば、仮説検証できるような測定方法を選ばなければならない。メンタルヘルス領域の各種の評価手技については他書を参照⁵。

5) 統計手法

研究計画書ではすでに述べた研究仮説とそのモデルを検証（証明するか否定するか）するためにどのような統計解析を用いるかを明記する。もし、適切な統計解析法が思いつかないのであれば、研究はそこでいったん留め、再度、全体を見直すべきである。可能性としてあるのは、(a) そもそも統計解析が当てはまらない仮説・モデルであるか (b) 統計解析は可能だが非常に高度な統計を必要とするかであろう。前者であればこれ以上、研究を進めることは無駄である。そもその研究仮説に戻り再検討するべきである。後者であれば、非常に高度な統計を使いこなせる専門家を探し、この段階で共同研究者として参加してもらうべきである。データが出揃ってから統計の専門家に相談するのでは遅いのである。

また、平素から統計学を学び、統計学の教科書・マニュアルにも精通しておこう。もちろん、統計学の講座にも出よう。そして先行研究を読む際にそこで使われている統計手法を精読し、分からなければそこでさらに勉強することで力が付く。

2. 原稿の投稿から出版まで

研究成果を原稿にしたら専門雑誌に投稿する。当然のことであるが、その場合の注意点について事前に知っておこう (図 1-9)。

まず投稿先を選定する。当然のことであるが、結構、難しい。あなたが投稿しようとしている原稿を受け入れてくれる領域はどこかを考えてみよう。たとえば、Aクリニックの助産師のように「うつ病における自殺企図と子ども時代に受けた児童虐待の体験は境界性パーソナリティ傾向が介在する」という内容の原稿であれば、うつ病領域、気分障害領域、児童虐待領域、トラウマ領域、パーソナリティ領域など、複数の領域が考えられる。まず、いずれかの領域を決める。

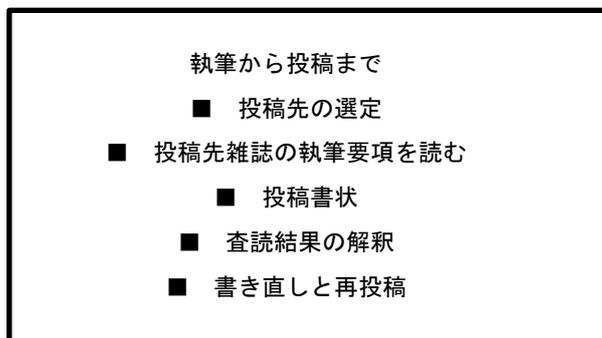


図 1-9. 執筆から投稿まで

⁵ 北村俊則 (1995). 精神症状測定理論と実際—評価尺度, 質問票, 面接基準の方法論的考察—第2版. 海鳴社, 東京.

Aクリニックのグループは気分障害領域で「攻めよう」と決めた。次に彼女たちは、気分障害領域にある雑誌にどのようなものがあるか一覧表示をした（図書館のネットサービスを利用した）。多数の雑誌があるし、見たことのないものもある。困った彼女たちは次に各雑誌のインパクトファクター（impact factor: IF）を調べた。インパクトファクターは自然科学や社会科学領域で、その雑誌の影響度を測る指標で、雑誌の水準を示す一つの目安とされている。その雑誌が過去 2 年間で掲載した論文数で、そこで掲載されたすべての論文がその年に他誌で何回引用されたかを割ることで求められる。つまり評判になった論文を多く掲載する雑誌ほどインパクトファクターが高くなる。インパクトファクターが高い雑誌に掲載されればより多くの読者の目に留まる。しかし、ハードルが高い分、査読審査も厳密になり、投稿しても却下される率が高くなる。却下されれば次の投稿誌を選ばなければいけない。その分、時間もかかってしまう。

投稿誌を決めるにはインパクトファクター以外にも考慮する必要がある。雑誌には編集委員会（editorial board）と編集委員長（Editor-in-Chief）がおり、通常、雑誌の表紙に氏名が記載されている。ここに名前が挙げられている研究者は研究の方向性が似ているものが多い。いわば一学派の研究者で編集委員会が構成されている可能性があるのだ。この学派にとって受け入れ難い論を展開している原稿が受理される可能性は低い。一方、この学派にとって容易に受け入れることができる内容の原稿であれば受理されやすいであろう。Aクリニックのグループは、その学派の代表的研究論文を見ることで、Aクリニックグループのテーマである「うつ病における自殺企図と子ども時代に受けた児童虐待の体験は境界性パーソナリティ傾向が介在する」という話題に興味を示してくれるかを判断してみた。

さらに、雑誌を編集する立場から言えば、たとえ興味深いテーマであったとしても、類似の内容の論文を毎月、掲載するのは読者の興味に合致しないことから、避けたい気持ちになる。すでに発表された論文よりさらに興味深い結果が出ているのであれば別だが、類似内容の論文が、たとえば 4 月号ではロシアからの報告、5 月号ではドイツからの報告、6 月号では韓国からの報告・・・と続けば、さらに日本からの報告を受理したくはなくなる。投稿しようと計画している雑誌の最近の掲載誌の傾向をつぶさに確認することは大切である。

Aクリニックグループは、自身にとって最初の投稿原稿だったので、中程度のインパクトファクターのB誌を選び投稿することにした。次に彼女たちががすることは、これも当然であるが、B誌の執筆要領を読む作業である。執筆要領で注意すべき点はいくつかある。まず、そこで使われている英文が、British English か American English であるかを区別する。パンクチュエーション（英文での文字間の開け方、コンマなどの打ち方、大文字小文字の使い方など）も異なってくる。また、メンタルヘルス系の雑誌のフォーマットとして多く使われているのは American Psychological Association (APA) のマニュアルである。APA format について通曉する必要がある。また医学系では Vancouver style も多く用いられている。特に、文献の引用スタイルが雑誌によって異なるので、ここには十分な注意が必要である。さらに見出しの書き方、図表の掲載の仕方にも注意しよう。そこまで細かいことを気にしなければいけないのかと思う読者もおられるであろう。しかし、一流誌の査読者はその領域の一流の研究者であることを思い返してもらいたい。一流の研究者は強迫傾向の人が多い。強迫傾向の人が気になる些細な点にも配慮することが、投稿者には必要である（書式については第 14 章参照）。

いよいよB誌に投稿する段階までこぎつけたAクリニックグループは、投稿書状を書くことになる。最近ネット投稿のスタイルをとる雑誌が増えている。彼女たちの投稿した原稿は、通常であれば 3 人ほどの査読者が審査する。これまでは編集委員長が編集委員会の内外から査読者を選び、査読を依頼してきた。しかし最近、査読者を選定する新しい方法をとる雑誌がふえてきた。それは、投稿著者に査読してもらいたい研究者を 3 ないし 5 名を挙げさせ、それを参考に編集委員長が査読依頼を出すというものである。投稿著者の指定した研究者に必ず査読が回るというものではない。しかし、それらの研究者に査読が行く可能性は高い。また、査読者候補を挙げないと、ネット投稿の次頁に進まないようになっている雑誌もある。そこで、そうした雑誌に投稿する際は自分の研究についてコメントできる査読者を選んでおかなければならない。原稿で引用した、その領域の第一人者のアドレスはその人の書いた論文の連絡先に書いてあることが多い。そこでこれらの査読可能な研究者の氏名とメールアドレスを、自分の原稿を投稿する際にネット上の指示に従って入力すればよい。

投稿した原稿が無修正で受理されることはほとんどない（私の経験で 1 回だけあった）。ほとんどの場合、査読者がさまざまな意見を述べてくる。まず、査読結果の解釈をしなければならない。査読はほとんど匿名で行われる。しかし、査読者がだれか推察してみよう。多くの場合、その領域の第一人者が査読に当たっているため、おおよその見当がつく。あの査読者ならこういうだろうという推測は立つ。この判断は査読に対する返答を書く上で重要になる。もちろん、査読意見がいつも正論というわけではない。しかし投稿論文の査読は民主主義の通じる場所ではない。投稿者は査読者に歯向かうことはできない。査読者は常に神様なのである。もし、反論があれば考察（Discussion）に記載する。さらに困るのは、複数の査読者で意見が異なることもある。「それでは困る！」という意見を書けるだけの力量が必要になる。

書き直しと再投稿について注意点を述べておこう。投稿している A クリニックグループは人生初めての経験であり、一生の思い出になる作業をしている。しかし、その原稿を査読している査読者はおそらく年間 10 本ほどの投稿原稿を読み、意見を書いている。査読を依頼されることは研究者としての名誉であり、世界中どこでも無償の作業になっている。いくら名誉だといっても、無給労働をさせられ、投稿原稿の結構な数のものは「箸も棒にも引っかけられない」内容であり、ウンザリという気分にもなっている。修正原稿が再投稿されても、その原稿を第 1 頁から最後までまた読み通したいとはあまり思わない。投稿者はまずこのことを頭に入れて、「査読者フレンドリー」な返書を書くように努めよう。

第 1 に、査読意見の項目ごとに修正を行うことを心がけよう。そのためには査読意見を冷静に読むことが大切である。査読意見で批判されると、せっかく書いた原稿を非難されたことから失望、怒り、疑惑などさまざまな感情が出てくる。場合によっては数日おいてから読み直すこともよい。そして、査読意見の項目ごとに分けて加筆修正を行う。しかし、無理な修正要求には応えなくてよい。査読者が単純に誤解していることもある。見解の相違としか言いようのないこともある。この場合は、査読意見を無視するのではなく、「……という意見もある。これは先行研究や今回の所見から否定できる」と考察（Discussion）に述べるようにする。

加筆修正が済んだら、編集委員長あてに再投稿書状を書く。この書面では、査読者の意見を箇条書きに挙げた上で、ひとつずつ訂正部分を掲載し、その説明をつける。また、ここで使う文章は丁寧なものを心がける。英語の公的で丁寧な書法を勉強しておくことも必要である。

3. 研究のタイプ

ところで戦略から臨床研究を見ると、表 1-4 のように 2 つに分かれる。自分がどちらの研究をするのかも自覚しておこう。タイプ A の研究は、まず期待した結果が出るのが分かる、だれでも思いつくような研究で、類似の先行研究も多く、使用できる尺度もすでにそろっており、従って研究にかかる期間も短い。修士号や博士号のように決められた時間内に原稿を仕上げたい場合に向いている low risk low return の研究である。この逆がタイプ B の研究である。

ところが修士課程や博士課程に入った院生がまず考える研究課題は往々にして気宇壮大、その結果は長く歴史に残るであろうテーマを選ぶ傾向がある。意欲満々で入学したのだから当然である。研究したいと思う人は、たいていは壮大な研究計画をもっているものである。それを短時間で完成させたいと思う、あるいは完成できると信じていることが多い。しかし、壮大な研究計画は後日の楽しみに取っておこう。なぜなら、そのまま研究を開始すれば時間内に論文を提出することは不可能になるからである。そこで、気宇壮大な研究課題はそのまま温存し、ただしそれは一生の研究課題とし、この数年でまず何をするかを考え、テーマを設定することが望まれる。本書の読者はこれから論文を書こうと思っている方が多いと推測する。初めての論文を書く場合は、小振りの研究からスタートすることにしよう。上記のタイプ A の研究課題である。そして、リサーチ・クエスチョン（研究仮説）は 1 つか 2 つに絞る。多数のリサーチ・クエスチョンを一つの研究で挙げると收拾がつかなくなる危険性がある。

3 バイアス

1. バイアスの定義

多くは症例対照研究 (case control study) における (第 5 章参照) 推論のいずれかの過程において真実から異なる結論を導き出すような体系的な誤りをバイアスという。以下の例で見ると、研究者が多くは気づかずに研究した結果、出てきた所見の解釈を誤らせる可能性のある事柄である。バイアスには多くの種類があるが、ここではいくつかのものについて解説するにとどめる⁶。

2. ナイマン・バイアス (有病率・罹患率バイアス)

例えば心筋梗塞の既往のある者とその既往のない者を対象して、危険因子として特定の現象 (例: 喫煙, 肥満, 脂質異常症 (高脂血症), 特定のパーソナリティなど) をもって比較したとしよう。「心筋梗塞の既往」はその生涯有病率を聞いている。過去に心筋梗塞がありそのため (おそらく重症であったため) 死亡した者は今回の調査からは抜けてしまう。また生存していたとしても、発症から調査時点までの間に危険因子とした特定の現象 (例: 喫煙, 肥満, 脂質異常症 (高脂血症), 特定のパーソナリティなど) が変化することもある。その結果、本来パーソナリティ傾向 (タイプ A パーソナリティ) が心筋梗塞の危険因子であっても、その関係が消失したり、むしろ反対の傾向が出てくることもある。これをナイマン・バイアス (Neyman bias) あるいは有病率・罹患率バイアス (prevalence-incidence bias) という。

産後うつ病を経験した女性を出産はしたがうつ病を経験しなかった女性に対してあるパーソナリティ尺度を施行して産後うつ病の危険因子を見出そうとする研究において、このバイアスが発生する可能性はある。まず、産後うつ病の女性は調査に消極的になるかもしれない。また、産後うつ病の結果としてパーソナリティ傾向が変化するかもしれない。

3. バークソン・バイアス (入院率バイアス)

入院患者を対象とした研究で、疾患 X と疾患 Y の合併の可能性を評価したとしよう。疾患 X を有する患者はそうでない患者に比べ疾患 Y の発病率が明らかに高かった。しかし、同一の調査を一般人口で行うとほとんど差がなかった。これは、入院するには単に疾患 X を有していることが理由でなく、他に多くの決定要因があり、その一つが疾患 Y の存在であるからである。このバイアスをバークソン・バイアス (Berkson bias) あるいは入院率バイアス (admission bias) という。

ある精神科クリニックで産後うつ病で来院した女性に 15 歳以前に両親のいずれかから死別あるいは離別したかを聞き、対照群として産後うつ病を経験しなかった産婦を産科クリニックで募ったとしよう。15 歳以前の両親との離死別体験と産後うつ病に有意の関連が出て、これはバークソン・バイアス (入院率バイアス) である可能性がある。例えば、産後うつ病の女性の中で 15 歳以前の両親との離死別体験を有する者ほど受診行動を取りやすいとするとこうしたバイアスが発生し得るのである。当然、バークソン・バイアスのため、本来みられるべき関係が消えてしまうことも考えられる。

⁶ Sackett, D. (1979). Bias in analytic research. *Journal of Chronic Disease*, 32, 51-63.

4. 発見信号・バイアス(仮面取りバイアス)

例えば産後うつ病と児童虐待の関係を見る研究で、訪問保健師がうつ病の自己記入式調査票（エジンバラ産後うつ病評価尺度）を実施し、その陽性者に虐待的育児が多いと報告したとしよう。ところが、自己記入式調査票は被験者全員に実施したものの、虐待的育児の有無はそれを疑った時に保健師が詳細に問診をしていた。つまり、自己記入式調査票が陽性の母親について、保健師は念のため育児態度について詳細に聞き出していた。一方、自己記入式調査票が陰性の母親についてはあまり細かく育児の様子を聞いていた。うつ病だけでなく育児についても自己記入式調査票で均等に聞いた場合、両者には何の関連も見出さなかった。これを発見信号・バイアス（detection signal bias）あるいは仮面取りバイアス（unmasking bias）という。

受診者のなかで境界性パーソナリティの者には過去の被虐待の有無を聞き、その他のパーソナリティの者については本人からの自発的発言があった時のみ被虐待経験を聞き出した結果、児童期の被虐待経験と境界性パーソナリティの間に有意の関連を見出したような研究も発見信号・バイアスの結果である危険性をはらんでいる。

2

変数（変量）と統計法 代表値と散布度

1 変数（変量）と統計法

1. 変数（変量）と尺度

1) 変数（変量）

自然科学研究の基礎は現象を観察し、それを数値で表記することである。臨床家が研究を行う場合、観察した事象を数値で表す。たとえば、循環器病の専門家が毎日診察する患者の血圧と脈拍を測定し、それを後日、まとめて解析しようと考えたとすると、血圧（xxx mm 水銀柱）の値や心拍数（毎分 xx 回）が「数値で表された事象」ということになる（表 2-1）。数値で表される個別の事象には「収縮期血圧」「拡張期血圧」「心拍数」などの名称が与えられる。これを変数（variable）という。「収縮期血圧」などの名称を変数名という。変数は変量ともいわれる。

表 2-1. 変数・被験者・標本

患者番号	収縮期血圧	拡張期血圧	脈拍
1	175	88	88
2	166	102	79
3	87	45	92
4	24	84	59
5	180	91	68

被検者

変数

変数

変数

標本の
大きさ

複数の変数がそれぞれ何人かの被験者（ケース）についての測定結果を伴って記された場合、それを標本（sample）という。ある日の内科外来患者 120 名の血圧と心拍数のデータがそろっていれば、それが標本である。被験者の数が 120 名いるので、これを標本の大きさ（sample size）という。被験者数は通常 N（number の省略形）で表示するので、「N の大きさ」などともいうことが多い。

なお、研究全体での標本の大きさを $N = 120$ と大文字の N で記載し、そのうち降圧剤を服用しているものが 110 名いればそれを $n = 110$ と小文字の n で表記することも多い⁷。

2) 尺度

変数が示しうる数値にはいくつか特徴がある。たとえば、血圧は、134, 159, 204, …… などと様々な値を示すが、癌の組織検査の結果は 5 段階しか示さない。患者の性別も変数である。しかし、ここには「男性」と「女性」しかない。こうした特徴で変数を分けることがある。変数の特徴を尺度（scale）という⁸。尺度の特徴で、その変数について使用する統計解析手法が決まってしまう。そのため、ある変数がどのような尺度であるかは、統計解析のもっとも重要なスタートラインになる。

そこでまずは、尺度について解説を行う。尺度には (a) 名義尺度（nominal scale）(b) 順序尺度（interval scale）(c) 間隔尺度（interval scale）(d) 比尺度（ratio scale）がある。

(a) 名義尺度

例えば、青年期の男女の自己効力感を日本、アメリカ、中国、イギリスの 4 か国で共同研究したとしよう。自己効力感尺度の日本語版、英語版、中国語版を用いて、各国で 1000 人以上のデータを集めた。このデータ解析を行うに当たって、被験者の青年がどの国の若者かは重要なポイントになる。そこで、国籍を観察し（アンケートで確認し）、それを数値にしなければならない。しかし、日本、アメリカ、中国、イギリスという国名は数字ではない。EXCEL に入力しようとした研究者は、便宜的に

日本 = 1
アメリカ = 2
中国 = 3
イギリス = 4

と振って、データ入力した。しかし、日本が 1 で、中国が 3 である必然性は何もない。たとえ、

イギリス = 1
中国 = 2
アメリカ = 3
日本 = 4

と入力しても何ら不都合はない。どちらが大きいとか、どちらが先でどちらが後だとかといった分け方ではないからである。国名のアルファベット順に並べても、

China = 1
Japan = 2
United Kingdom = 3

⁷ APA format 参照

⁸ 評価手段のうちで非構造化された観察者による評価法や自己記入式調査票に「〇〇尺度」という名称を与えることが多い。そのため評価測定手段を「尺度」と呼ぶことも多い。

単に並べる便宜であるので統計上の意味をなすものではない。こうした尺度を名義尺度 (nominal scale) という。

名義尺度の含まれる変数には、性別 (男・女)、手術後 5 年目の生死 (5 年目の時点で生きているか死んでいるか)、結婚歴 (独身・既婚・離婚)、支持政党などがある。

(b) 順序 (順位) 尺度

ある大学で 3 年生の英語の試験成績を「優・良・可・不可」の 4 段階で採点入力していた。これを調査解析の対象にしようと考えた。この場合、値の大小、後先 (あとさき) が決まっている。しかし、優と良の間の距離が良と可の間の距離と同一である保証はない。こうした尺度を順序 (順位) 尺度 (ordinal scale/rank scale) という。

優は 1 位、良は 2 位、可は 3 位といった「位」をつけられる特徴を有するのが順序 (順位) 尺度である。そこで EXCEL にデータ入力をする場合、

優 = 1
良 = 2
可 = 3
不可 = 4

と振ることは合理的であり、また

不可 = 1
可 = 2
良 = 3
優 = 4

と振るのも良い。しかし、

優 = 1
可 = 2
良 = 3
不可 = 4

という配点は許されない。なぜなら 4 者には順位があるからである。

順序 (順位) 尺度の含まれる変数には、スポーツ大会の着順、アンケートの 3 段階項目 (「そうだ」「どちらともいえない」「ちがう」) などがある。

(c) 間隔尺度

ある研究者は 3 月の平均気温と桜の開花日の早い遅いとの関係を調査しようと考えた。平均気温は、値の大小があり、さらに 22 度と 25 度の間隔は 27 度と 30 度な間隔と同じである。従って、順序 (順位) 尺度とは異なる。こうした尺度を間隔尺度 (interval scale) という。

(d) 比尺度

ある研究者は治療中のうつ病の患者の体重を測定して、それを研究の対象とした。体重も値の大小があり、2つの値の間隔も比較できるものである。上記の気温と類似している。違う点を挙げれば、体重には絶対的なゼロ点がありうる。体重がマイナスになることはない。一方、気温は氷点下のマイナス温度がある。

体重では比率を計算できる。48 kg の体重は 24 kg の体重の 2 倍である。つまり割り算が可能なのである。しかし、48°Cの気温が 24°Cの気温の 2 倍とはいえない。気温にはスタートラインとなる最低点がないからである。体重のような、加減乗除ができる尺度を比尺度 (ratio scale) という。

(e) 尺度の比較と意味

再度、名義尺度、順序尺度、間隔尺度、比尺度を比べてみよう (表 2-2)。「順位」「等間隔性」「加減」「乗除」の4つの基準で、これら尺度は分類することができる。

表 2-2. 尺度の分類と特徴

意味	名義尺度	順序 (順位) 尺度	間隔尺度	比尺度
大小・上下などの順列	なし	あり	あり	あり
値と値の間の等間隔性	なし	なし	あり	あり
値の連続性	なし	なし	あり	あり
加減	意味ない	意味ない	意味あり	意味あり
乗除	意味ない	意味ない	意味ない	意味あり

ここで注意するのは「血圧は比尺度だ」といった決めつけをすることである。「温度は間隔尺度である」と決めてしまうと、確かに気温は間隔尺度であるが体温は (マイナスはあり得ないので) 比尺度となる。被験者の出身地の都道府県は名義尺度であろう。しかし、季節性感情障害の研究者が被験者の出身都道府県を、その日照時間でランクをつけて評価しようとするのであれば、順序 (順位) 尺度と考えた方が妥当であろう。

【練習問題】 以下の変数は何尺度か述べなさい。

- 身長
- 眼圧
- 血圧
- 肥満度 body mass index (BMI)
- 血中のコレステロール濃度
- 髪の毛の色 (黒, 白, 褐色)
- 血中の白血球数
- 性別 (女性, 男性)
- 生涯飲酒量
- 血液型 (A, B, AB, O)
- 染色体数
- 結婚歴 (未婚, 既婚, 離婚, 死別)

3) 尺度から統計手法へ

ではなぜ、統計学では変数がどの尺度であるかが重要になるのか。統計手法は様々なものがある。どの手法を使うかを定める手順は大きくわけて 2 つある。第 1 に、その統計が何を確認しようとするのかである。つまり、統計の目的である。2 つの事象に相関や関連があることを言いたいのか。ある現象が複数のグループで異なることを言いたいのか。多数の変数の裏に潜む概念を確認したいのか。これが統計の目的であり、目的ごとに使用する統計手法が準備されている。

いまひとつは、扱う変数が正規分布を取るか否かで使用する統計手法が異なる。正規分布については本書の後に述べる。正規分布は、変数の値の分布が左右均等の釣鐘（つりがね）状になっている状態をいう。正規分布を取れるのは、間隔尺度と比尺度である。さらに、測定段階数の多い順序（順位）尺度も正規分布を示しうると見做すこともよく行われる⁹。しかし、名義尺度や測定段階数の少ない（例えば 3 段階以下）順序（順位）尺度には正規分布を見做すことはできない。

このふたつの基準（解析の目的と変数の正規性）の組み合わせで、用いる統計手法がほぼ決まってしまう。正規分布を示す変数に用いる様々な統計手法をパラメトリック・テスト（parametric tests）、正規分布を示さない変数に用いる様々な統計手法をノン・パラメトリック・テスト（non-parametric tests）と呼ぶ（表 2-3）。つまりおびただしい数ある統計解析手法はパラメトリック・テストとノン・パラメトリック・テストに分けられている。

現実には、ノン・パラメトリック・テストは数に限りがあり、（後で勉強するように）使い勝手が悪い。一方、パラメトリック・テストは手法数も多く、使用法も便利である。

表 2-3. 統計法の分類

	パラメトリック	ノン・パラメトリック
相関	積率相関係数 product-moment coefficient	Spearman rank correlation
関連	---	カイ 2 乗検定 chi-squared test
対応のない 2 群の比較	T 検定 t test	Mann-Whitney U test
対応のある 2 群の比較	T 検定 t test	Sign test Wilcoxon matched pairs signed-rank test
3 群以上の比較	一元配置分散分析 one-way analysis of variance (ANOVA)	Kruskal-Wallis test
因子の構造	因子分析 factor analysis	潜在構造分析 latent structure analysis
複数変数による予測	重回帰分析 regression analysis	logistic regression analysis



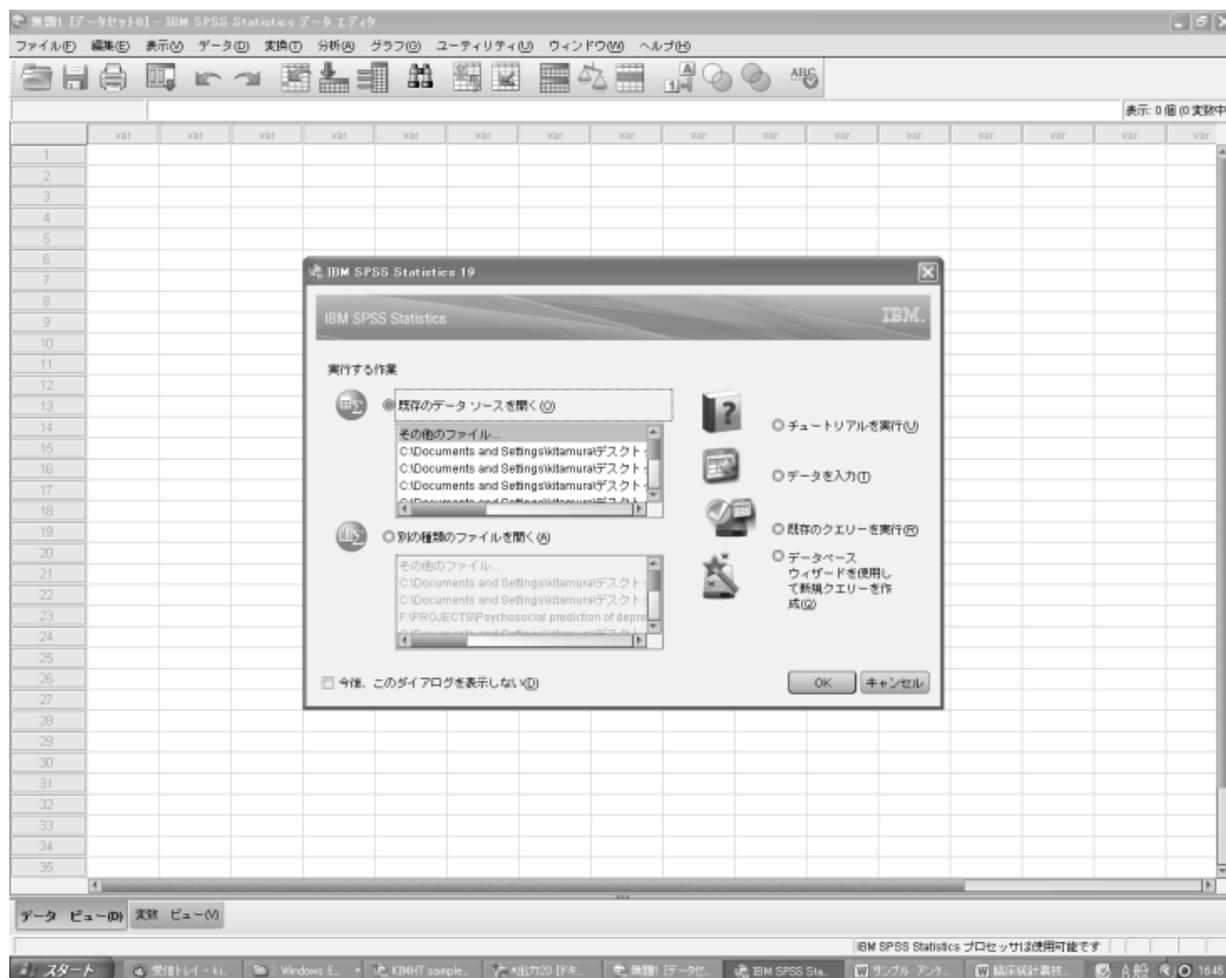
⁹ どの程度の順序（順位）尺度を正規分布を示し得るものと見做すかについてのニュアンスは、利用する統計解析法によっても異なり、さらに専門領域によっても異なる。読者は自分の専門領域の雑誌を読んで「業界の常識」を身につけていただきたい。さらに「業界の常識は、時代によっても大きく変わる。5 年前の常識は現在の非常識であることも多い。

2 SPSS の構成とデータ入力

ここで Statistical Packages for Social Sciences (SPSS) の使い方に話を移そう。かつて、統計計算は手計算の時代があった。やがて電卓を使って計算が行われるようになり、1970 年代ごろから大型電子計算機が大学などで比較的容易に使われるようになってきてから、統計計算専用のソフトウェアが開発・販売されるようになった。そのひとつが SPSS であった。やがて個人が所有できるパーソナル・コンピュータの普及に伴い、SPSS も個人仕様で用いられるようになってきた。最近では Statistical Packages for Social Sciences という表現はされず、SPSS が商品名となっている。

1. SPSS の立ち上げ

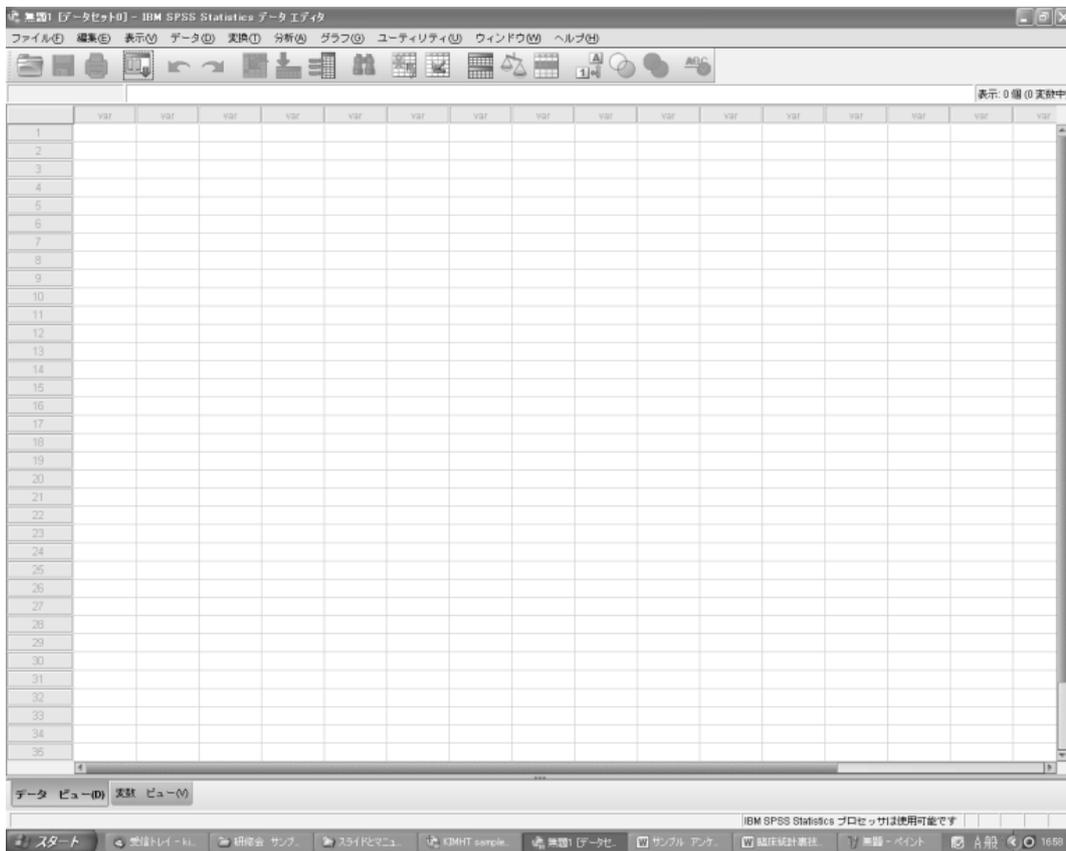
SPSS のプログラムファイルをクリックすると次のような画面がでてくる。



中央の  が 【キャンセル】 ボタンをクリックして以下のダイアログ・ボックスを消す。



以下の画面が出てくる。こうなったらデータを入力できる。



画面左下に、【データ・ビュー (D)】と【変数 ビュー (V)】という指示がある。SPSS には、データ・ビュー画面と変数・ビューという2つの画面がある。画面下のその部分をクリックすることで画面が変更できる。



2. 変数の設定

【幅】 はデータ・ビュー画面でのその変数に当てられた幅を示す。いったん入力したデータであれば、幅の設定が狭くて、データ・ビュー画面で一部が欠けても、データとしては欠けてはいない。

【少数桁数】 はデータ・ビューで表示する際の小数点以下の桁数を示す。「少数桁数」は「2」がデフォルト¹⁰である。この設定にしておくと、例えば「AGE 年齢」の変数に「57」と入力してもデータ・ビューには「57.00」と小数点以下 2 桁まで表示される。これが煩雑であれば「少数桁数」を「0」に設定すれば「57」と表示される。

【ラベル】 は変数の説明である。たとえば HADdep という変数に、HOSPITAL ANXIETY AND DEPRESSION SCALE DEPRESSION SUBSCALE という入力をラベルに入れておくと、シンタックスでは HADdep という入力で済むが、解析結果の出力では HOSPITAL ANXIETY AND DEPRESSION SCALE DEPRESSION SUBSCALE と表示される。ラベルに日本語表記を用いてもその後のシンタックス文入力等で煩雑になることはない。日本語の論文を書く場合はラベルを日本語表記にしておくことも良い。レベルでは変数の内容を詳しく記載することが後の操作を容易にする。

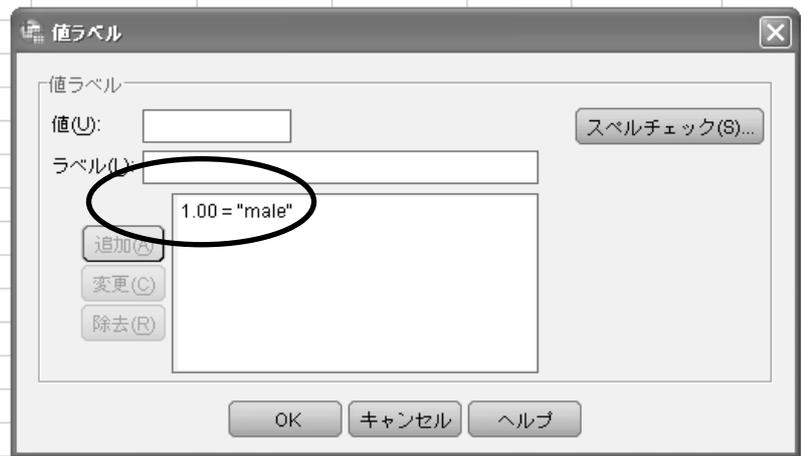
【値】 は、名義尺度等で数値に置き換えて入力した場合に、どの数値が何を表すかを SPSS に知らせるものである。例えば性別は名義尺度である。仮に「男性を 1」「女性を 2」と入力したとしよう。後日、1 と 2 のどちらが男性であったかわずれることもある。

そこで、値のセルにカーソルを移動しクリックする。すると下のような「値ラベル」ダイアログ・ボックスが現れる。値 (U) にまず 1 と入力、次にラベル (L) に male を入力し、追加 (A) をクリックする。



すると 1.00 = “male” という表示が四角い枠内に現れる。同様に、値 (U) に 2 と入力し、ラベル (L) に female と入力し、追加 (A) をクリックすると、2.00 = female という表示が、1.00 = “male”の次の行に加えられる。ラベルは日本語でも可であるので、活用しよう。

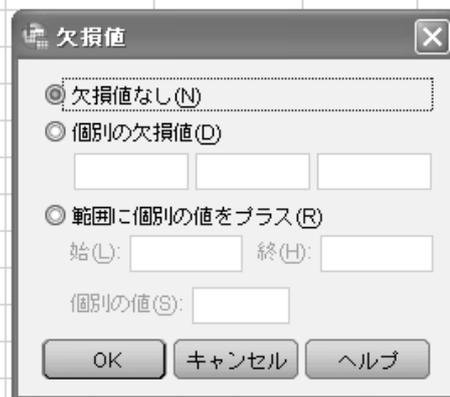
¹⁰省略時の値のこと。プログラムに与えるべき値をユーザーが設定しなかった場合に採られる既定値



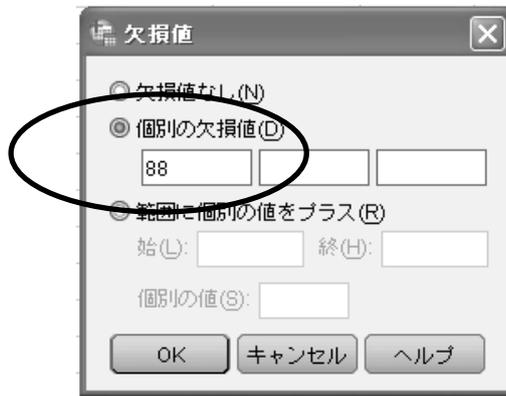
【欠損値】 は特定の値を欠損値 (missing value) と特定する際に用いる。被験者についてアンケートデータが集まったがある項目については被験者が記載をしていなかった場合がこれに該当する。また、面接調査で本来、聞き取るべきだった項目を聞き忘れた場合もこれになる。さらに、「既婚者には結婚期間を年数で問う」という項目について、未婚者からはそもそも回答が得られるものではない。これも欠損値となる。

SPSS では欠損値は「システム欠損値」と「ユーザー指定の欠損値」がある。数値の入力がない場合はシステム欠損値として認識される。一方、たとえば「既婚者には結婚期間を年数で問う」という面接項目につて、未婚者については 88 と入力して区別したとしよう。この場合、このデータセットに限って、88 を「ユーザー指定の欠損値」として SPSS に記憶させなければならない。

欠損値のカラムにカーソルを合わせクリックすることで、「欠損値」というダイアログ・ボックスを出す。システム欠損値のみの場合、「欠損値なし」の前の◎をクリックすればよい。



「ユーザー指定の欠損値」を指定するには、「個別の欠損値 (D)」をクリックし、ここに 88 と入力し、OK ボタンをクリックする。



【列】 はデータ・ビュー画面でのその変数の列幅を規定する。データ・ビューの表示にのみ影響するものである，入力されたデータに影響はない。

【配置】 はデータ・ビューのセルの中で値が左寄せ，右寄せ，中央の3つのうちいずれかを選ぶ。

【尺度】 は，その変数が「スケール (S)」（間隔尺度あるいは比尺度）「順序 (O)」（順序尺度）「名義 (N)」（名義尺度）のいずれであるかを指示するものである。尺度の部分のセルにカーソルを移動しクリックすると，上のようなプルダウン・メニューが出る。そこで「スケール (S)」「順序 (O)」「名義 (N)」のいずれかを選択し，クリックすることで決める。【型】で数値と決めた場合は，尺度をスケールにすることが推奨される。例えば「男性を 1 女性を 2」といった 2 値の変数であっても数値であるので，【型】ではスケールにしておこう（後の統計解析作業が容易になる）。【型】で「文字型」を指定した場合，【尺度】は「名義」になる。



こうして入力されたデータのサンプルを見てみよう。北村メンタルヘルス研究所のホームページ (<http://www.institute-of-mental-health.jp>) から「sample.sav」というデータセットを開けてみよう。

KIMHT sample data 06.sav [データセット1] - IBM SPSS Statistics データエディタ												
ファイル(F) 編集(E) 表示(V) データ(D) 変換(T) 分析(A) グラフ(G) ユーティリティ(U) ウィンドウ(W) ヘルプ(H)												
	名前	型	幅	小数桁数	ラベル	値	欠損値	列	配置	尺度	役割	
1	ID	数値	8	0	PATIENT ID	なし	なし	10	右	スケール(S)	入力	
2	W1Q1	数値	1	0	OWN IMPRESSION OF THE CURRENT PRG...	なし	なし	8	右	スケール(S)	入力	
3	W1Q2	数値	1	0	HUSBAND IMPRESSION	なし	なし	8	右	スケール(S)	入力	
4	W1Q3	数値	1	0	PREGNANCY DESIRED	なし	なし	8	右	スケール(S)	入力	
5	W1Q4	数値	1	0	DRINK	なし	なし	8	右	スケール(S)	入力	
6	W1Q4X1	数値	1	0	DRINK FREQUENCIES	なし	なし	8	右	スケール(S)	入力	
7	W1Q5	数値	1	0	DRINK DURING PREGNANCY	なし	なし	8	右	スケール(S)	入力	
8	W1Q6	数値	1	0	SMOKING	なし	なし	8	右	スケール(S)	入力	
9	W1Q6X1	数値	2	0	SMOKING FREQUENCIES	なし	なし	8	右	スケール(S)	入力	
10	W1Q7	数値	1	0	SMOKING DURING PREGNANCY	なし	なし	8	右	スケール(S)	入力	
11	W1Q8	数値	1	0	MARITAL STATUS	なし	なし	8	右	スケール(S)	入力	
12	W1Q9	数値	2	0	HUSBAND AGE	なし	なし	8	右	スケール(S)	入力	
13	W1Q10	数値	8	2	OWN AGE	なし	なし	8	右	スケール(S)	入力	
14	W1Q11	数値	8	2	OCCUPATION	なし	なし	8	右	スケール(S)	入力	

3. データ入力と保存

変数ビュー画面で使用する変数の仕様を決めたら、画面左下の データ・ビュー (D) をクリックして、データ・ビュー画面に移動する。



そうすると次のような画面が現れる。変数ビューで設定した変数名が第 1 行に横並びに表示されている。もっとも左の列には、上から下に向けて 1 から順に番号が自動的に振られている。これが行番号である。その上には様々なアイコンが並んでいる。内容と使用方法は他書を参照されたい。その上にプルダウン・メニューが並んでいる。

ここからデータの入力に入る。



1: ID	ID	W1Q1	W1Q2	W1Q3	W1Q4	W1Q4X1	W1Q5	W1Q6	W1Q6X1
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									

今回は、某産科病院で妊娠後期、産後 1 カ月目、産後 3 カ月目の 3 時点でアンケート及び助産師による構造化面接を行ったデータ（架空）を用いている。

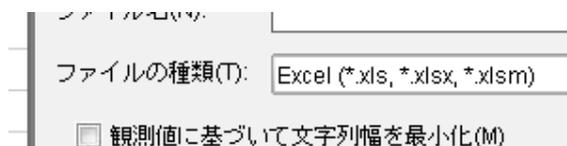
まず、妊娠後期アンケートの質問項目を見よう（図 2-1）（これもホームページに掲載してある）。今回、第 1 波調査（妊娠後期）を“W1”，第 2 波調査（産後 1 カ月目）を“W2”，第 3 波調査（産後 3 カ月目）を“W3”と略記することとする。そこで、第 1 波調査（妊娠後期）の質問項目（Q）の第 1 番に W1Q1 という変数名を設定した。同様に、次の質問項目には W1Q2 という変数名を設定した。各項目には複数の回答選択肢が準備されている。「いずれかひとつを選ぶ」という強制選択方式になっている。妊婦が選択肢のどれかに ○ をつけている。この番号を SPSS データ・ビューの W1Q1 という列の該当するケースのところに入力して行く。



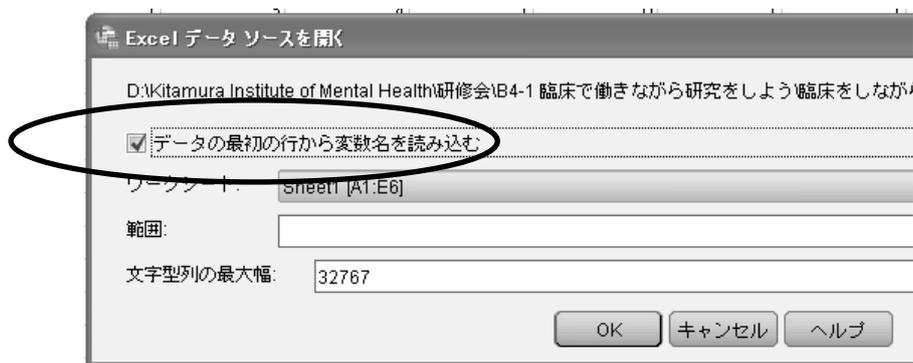
すると「データを開く」のダイアログ・ボックスが現れる。ここで「ファイルの種類 (T)」の右側のボタンをクリックする。



ここで Excel を選択する。



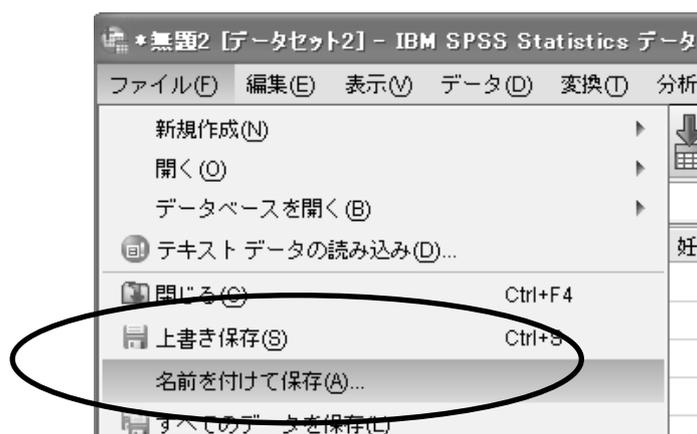
枠内に EXCEL ファイルのファイル名が出てくるので、希望するファイル名を選び、「開く」をクリックする。すると、「Excel データソースを開く」のダイアログ・ボックスが現れる。自動的に「データの最初の行から変数名を読み込む」に がついている。このままにしておく。また選んだ EXCEL ファイルのワークシートについても自動的に第 1 ワークシートが選ばれている。希望するデータがファイル中の第 1 ワークシートにあるのであれば、ここもこのままにしておく。もし、他のワークシートのデータを SPSS に移行したいなら、ここで希望のワークシートを選択して、指示する。



決まったら「OK」をクリックする。すると SPSS のデータ・ビューが現れる。

	患者番号	年齢	妊娠後期EPDS	産後5日目EPSS	産後1ヶ月目EPDS
1	1	22	2	2	8
2	2	31	3	1	3
3	3	28	7	5	15
4	4	18	15	10	7
5	5	26	22	15	5
6					

このデータファイルにはまだファイル名が設定されていないので、左上には「無題」と現れる。忘れないうちに、「ファイル (F)」のプルダウン・メニューから「名前を付けて保存 (A)」を選んで、「名前を付けてデータを保存」のダイアログ・ボックスからファイル名を決めて保存手続きをとる。



さらに変数ビューに移動し、「名前」から「尺度」までが希望通りになっているか確認する。特に、「型」と「尺度」には注意する。

ところで、上記の作業では、EXCEL の 1 行目に書いてある情報が各列の変数名であると SPSS が読み込んでいます。つまり、EXCEL の 2 行目からは実際のデータであると判断している。従って次のような EXCEL 入力を行うと、C 列は SPSS にとっては変数名がなく、また D 列の変数名が「EPDS」となり、その第 1 ケースの

値が「産後 5 日目」という文字情報になってしまう。SPSS に移行した場合、ある列のどこかのマスに文字情報があれば、自動的にその変数の「型」は「文字」になり、以降の「2」や「1」も文字として認識される。

	A	B	C	D	E
1	患者番号	年齢	EPDS		
2			妊娠後期	産後5日目	産後1ヶ月目
3	1	22	2	2	8
4	2	31	3	1	3
5	3	28	7	5	15
6	4	18	15	10	7
7	5	26	22	15	5
8					

この EXCEL ファイルを先ほどと同様に SPSS で読み込むと、そのデータ・ビューは次のようになる。

	患者番号	年齢	EPDS	V4	V5
1	.	.	妊娠後期	産後5日目	産後1ヶ月目
2	1	22	2	2	8
3	2	31	3	1	3
4	3	28	7	5	15
5	4	18	15	10	7
6	5	26	22	15	5
7					

また、すでに述べたように、SPSS の変数名は英文字と数字の組み合わせが使いやすい。EXCEL の変数名を日本語で入れておき、そのデータを SPSS に移行し、その上で SPSS 変数名を和文から英文に変えても良いが、手間がかかりすぎる。特に変数名が多いデータではそうである。従って、当初から EXCEL の変数名を以下のような英文字で表記することが勧められる。

	A	B	C	D	E
1	ID	AGE	W1 EPDS	W2 EPDS	W3 EPDS
2	1	22	2	2	8
3	2	31	3	1	3
4	3	28	7	5	15
5	4	18	15	10	7
6	5	26	22	15	5
7					

EXCEL ファイルで変数名を設定せず、1 行目からケースのデータを入れる場合もある。この場合、「Excel データソースを開く」のダイアログ・ボックスで「データの最初の行から変数名を読み込む」の を外して作業をする。

	A	B	C	D	E
1	1	22	2	2	8
2	2	31	3	1	3
3	3	28	7	5	15
4	4	18	15	10	7
5	5	26	22	15	5
6					

SPSS 上は次のようになる。変数名が少数の場合は良いが、多数になると変数名の設定の誤りが増えるので、避けるべき手法であろう。

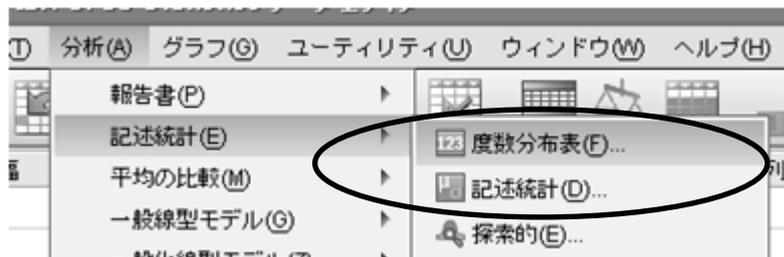
	V1	V2	V3	V4	V5	var
1	1	22	2	2	8	
2	2	31	3	1	3	
3	3	28	7	5	15	
4	4	18	15	10	7	
5	5	26	22	15	5	
6						
7						

3 代表値と散布度

膨大な数のデータをそのまま提示しても、見る者にはその意義は伝わらない。そこでその標本のある変数のデータとしての特徴を、少ない指標で表そうとしよう。これを記述統計量 (descriptive statistic) という。ある変数について実際のデータが標本のなかでどのように分布しているかを探る作業になる。サンプル・データ (sample.sav) から、参加女性の年齢の分布 (distribution) を見てみよう。

1. ヒストグラム

「分析 (A)」をクリックしプルダウン。メニューを開き、ここから「記述統計 (E)」にカーソルを合わせる。さらにそこから現れたプルダウン・メニューから「度数分布 (F)」を選びクリックする。



次のようなダイアログ・ボックスが現れるので、OWN AGE [W1Q10] を選び、ここをクリックすると色が反転。そこで ⇒ をクリックすると、選んだ変数が「変数 (V)」欄に移動する。



ここで「貼り付け (P)」をクリックすると、シンタックス・エディタに以下の文章が出現する。

```
5  
6 FREQUENCIES VARIABLES=W1Q10  
7 /ORDER=ANALYSIS.
```

この 2 行をカーソルを合わせて色を反転させる。

```
9 FREQUENCIES VARIABLES=W1Q10  
10 /ORDER=ANALYSIS.  
11 ▶
```

この状態で、次のアイコンをクリックする。



すると、このシンタックス文の命令が実行される。結果はまず「度数」「有効 = 214」「欠損値 = 6」と出てくる。これは 250 人分のデータで OWN AGE に入力があったのが 241 件、入力がなかったのが 6 件であることを意味している。

統計量

OWN AGE

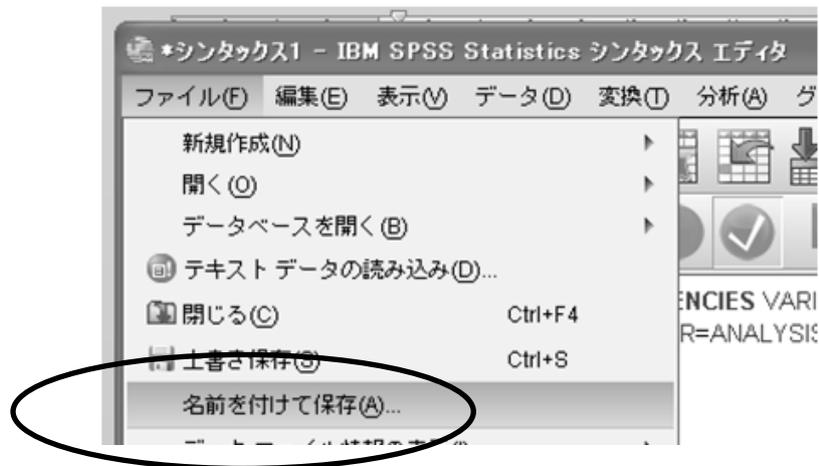
度数	有効	241
	欠損値	6

さらに以下のような各値に何事例があったのかを数値で表す結果が出てくる。

OWN AGE

		度数	パーセント	有効パーセント	累積パーセント
有効	18.00	2	.8	.8	.8
	20.00	2	.8	.8	1.7
	21.00	4	1.6	1.7	3.3
	22.00	3	1.2	1.2	4.6
	23.00	10	4.0	4.1	8.7
	24.00	9	3.6	3.7	12.4
	25.00	13	5.3	5.4	17.8
	26.00	18	7.3	7.5	25.3
	27.00	21	8.5	8.7	34.0
	28.00	17	6.9	7.1	41.1
	29.00	24	9.7	10.0	51.0
	30.00	18	7.3	7.5	58.5
	31.00	22	8.9	9.1	67.6
	32.00	12	4.9	5.0	72.6
	33.00	13	5.3	5.4	78.0
	34.00	11	4.5	4.6	82.6
	35.00	14	5.7	5.8	88.4
	36.00	8	3.2	3.3	91.7
	37.00	5	2.0	2.1	93.8
	38.00	8	3.2	3.3	97.1
	39.00	3	1.2	1.2	98.3
	40.00	1	.4	.4	98.8
	41.00	1	.4	.4	99.2
	43.00	1	.4	.4	99.6
	44.00	1	.4	.4	100.0
	合計	241	97.6	100.0	
欠損値	システム欠損値	6	2.4		
合計		247	100.0		

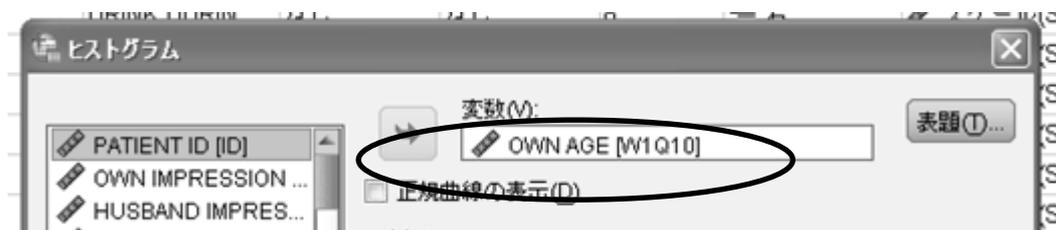
ところで上記の計算をしたシンタックス文は、「ファイル (F)」⇒「名前を付けて保存 (A)」⇒「ファイル名 (N)」⇒「保存」で、保存しておこう。



上記の度数の表は正確であるが見にくい。そこで次に、



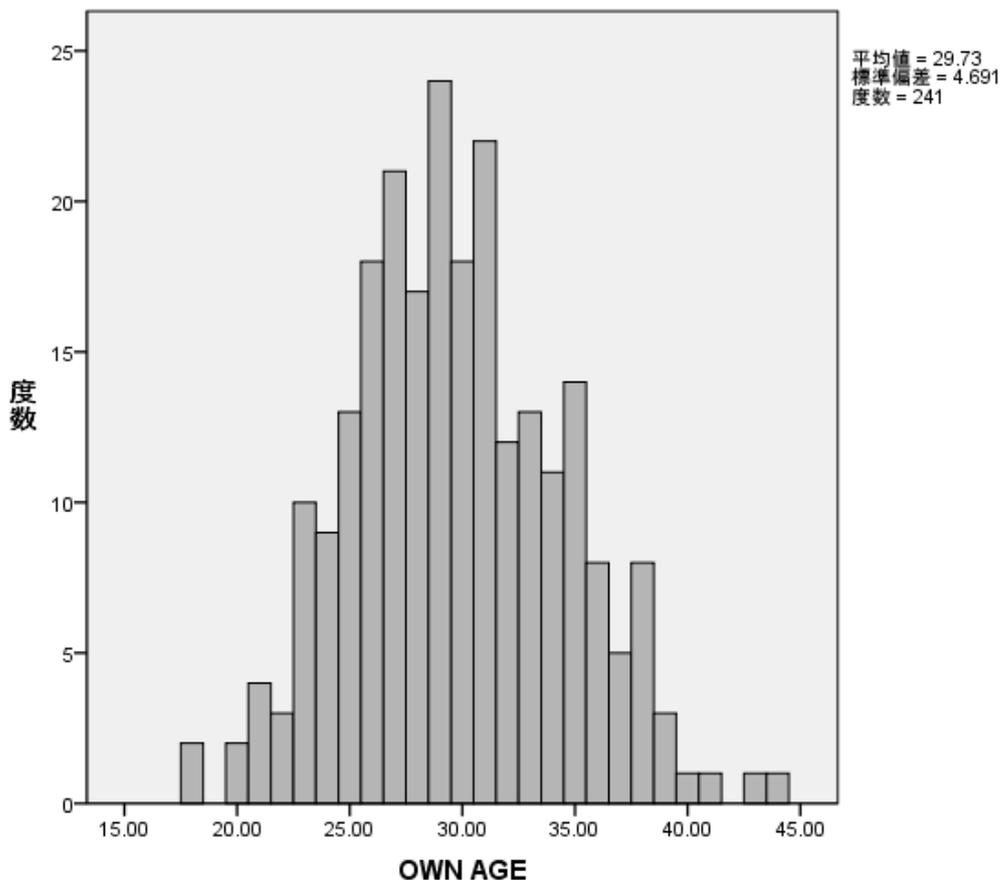
「グラフ (G)」⇒「レガシーダイアログ (L)」⇒「ヒストグラム (I)」を選び、ヒストグラムのダイアログ・ボックスのなかで、左にある変数の一覧表から OWN AGE という変数にカーソルを合わせてクリックすると色が反転するので、さらに⇒をクリックすると、「変数 (V)」のマスの中に、選択した変数 OWN AGE [W1Q10] が移動する。



これを先ほど同様、「貼り付け」を行い、さらにそのシンタックス文

```
2 GRAPH  
3 /HISTOGRAM=W1Q10.  
4
```

を反転、実行する。その結果は以下のように出力される。こうした図をヒストグラム (histogramme) という。



様々な変数のヒストグラムを見るとさまざまな分布を示すことが分かる。

2. 代表値

こうしたバラツキを呈するデータセットを少数の数値で代表させるのが代表値 (average) の考えかたである。代表値として、平均値 (mean), 中央値 (median), 四分位数 (quartile), 百分位 (percentile), 最頻値 (mode) がある。

【平均】 (mean) は、大きさ n のデータセット $x_1, x_2, x_3, x_4, \dots, x_n$ があつた時、

$$m_x = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

で求められる。これは

$$\left(\sum_{i=1}^n x_n\right)/n = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

という式でも表現される。

ここで加算記号である \sum (シグマ) の説明をしておこう。仮に 5 つの数値があり、その合算 (和) を求める計算は、

$$4 + 5 + 47 + 21 + 8 = 85$$

となる。各値を x と表すと、式は

$$x_1 + x_2 + x_3 + x_4 + x_5$$

と書ける。そこで 5 つの変数を一つの一般的表記として x_i で表す。そして、 x_i の i が 1 から 5 までの合算を、

$$\sum_{i=1}^5 x_i$$

と表記する。さらに、値の数の上限を一般化するのであればそれを n とする。上記の式は、

$$\sum_{i=1}^n x_i$$

となる。当然、

$$i < n$$

である。

【中央値】 は、順序 (順位) 尺度、間隔尺度、比尺度で、データをその値の大きさに並べたとき、ちょうど全体の中央に位置するケースの示す値を指す。ケースの数 (標本の多さ) が偶数であれば、 $n/2$ 番目の測定値と $(n/2)+1$ 番目の測定値の示す値の平均をもって中央値とする (例: 100 人分のデータであれば大きい順に 1 ~ 100 まで番号を振り、50 番目の値と 51 番目の値の平均を中央値とする)。中央値より上の値を示すケースと中央値より下の値を示すケースはその数が同等である。標本を、上下、大小、軽重で 2 群する際に役立つことが多い。先ほどのヒストグラムで述べれば、中央値の左右のヒストグラムの面積は等しくなるともいえる。

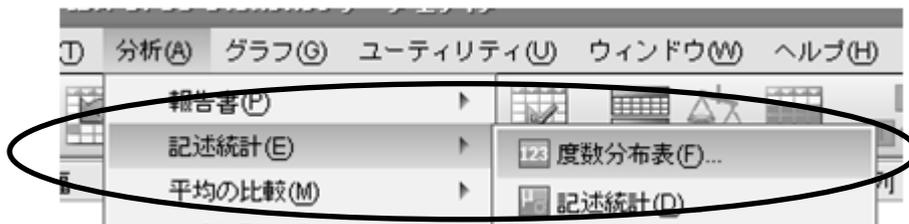
【四分位】 は、順序 (順位) 尺度、間隔尺度、比尺度で、データをその値の大きさに並べ、1/4 番目のデータの値を第 1 四分位、3/4 番目のデータの値を第 3 四分位という。中央値は第 2 四分位である。

【百分位】 は、順序 (順位) 尺度、間隔尺度、比尺度で、データをその値の大きさに並べ、1/100 ずつ分けたときの値をいう。

【最頻値】 は、名義尺度、順序 (順位) 尺度、間隔尺度、比尺度が最も多い測定値ないしカテゴリーをいう。

SPSS で平均値、中央値、四分位数、百分位、最頻値を求めてみよう。

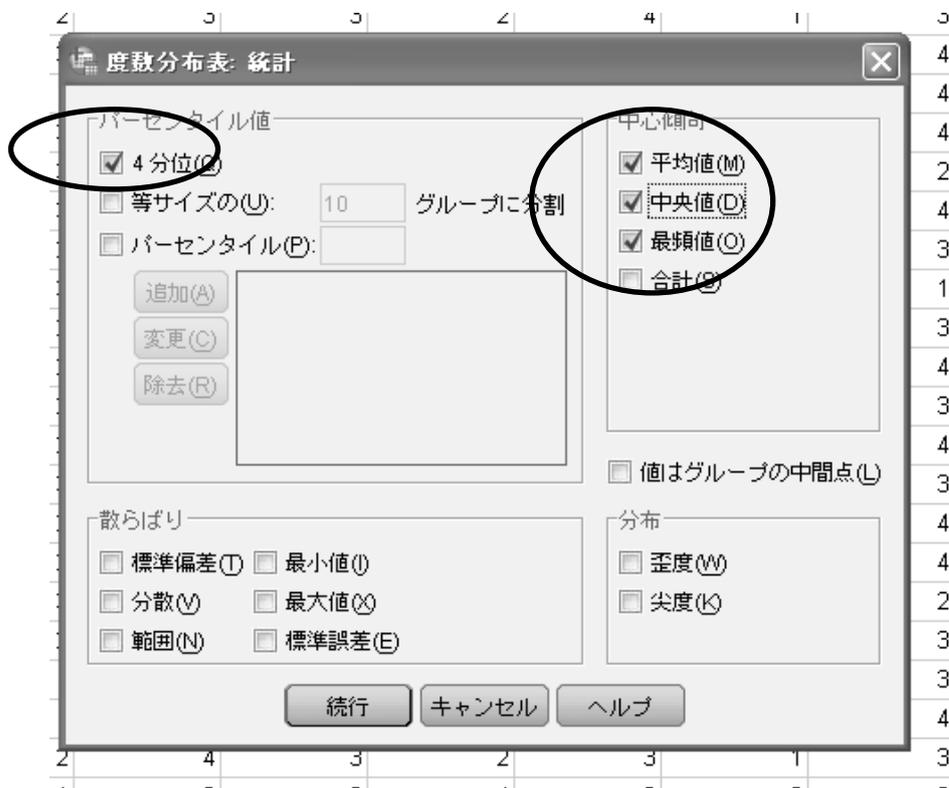
「分析 (A)」をクリックしプルダウン。メニューを開き、ここから「記述統計 (E)」にカーソルを合わせる。さらにそこから現れたプルダウン・メニューから「度数分布 (F)」を選びクリックする。



下のようなダイアログ・ボックスが現れるので、OWN AGE [W1Q10] を選び、ここをクリックすると色が反転。そこで⇒をクリックすると、選んだ変数が「変数 (V)」欄に移動する。



ここで「四分位 (Q)」「平均値 (M)」「中央値 (D)」「最頻値 (O)」のマスをクリックして ✓ を付ける。そして「続行」をクリックし、元のダイアログ・ボックスに戻り、OKをクリックする。



その結果の出力は、次のようになる。

統計量

OWN AGE

度数	有効	241
	欠損値	6
平均値		29.73
中央値		29.00
最頻値		29
パーセンタイル	25	26.00
	50	29.00
	75	33.00

3. 散布度

同一の平均値あるいは中央値を示していても2つの標本のデータのヒストグラムが大きく異なることは稀でない。被験者のデータは通常はバラツキがある。このバラツキの度合いを示す指標が散布度 (measures of dispersion) である。

散布度には分散 (variance) と標準偏差 (standard deviation) がある。パラメトリック・テストに共通して用いられる基本的概念である。

先ほどの某産科病院受診の女性サンプルに戻ってみよう。年齢の平均値 (m) は 29.7 歳であった。そこで各ケースの年齢が平均よりどれほどバラツキがあるか計算してみよう。このバラツキを【偏差】 (deviation) という。 n 人のケースで i 番目のケースの年齢を x_i とすると、その差は

$$d_i = x_i - m$$

である。これを各ケースに当てはめてみよう。

	W1Q10
1	30
2	23
3	31
4	33
5	26
6	29
7	28
8	33
9	30
10	33
11	23

Case 1	$d = 30 - 29.7 = 0.3$
Case 2	$d = 23 - 29.7 = -6.3$
Case 3	$d = 31 - 29.7 = 1.3$
Case 4	$d = 33 - 29.7 = 3.3$
Case 5	$d = 26 - 29.7 = -3.3$
Case 6	$d = 29 - 29.7 = -0.7$
Case 7	$d = 28 - 29.7 = -1.7$
Case 8	$d = 33 - 29.7 = 3.3$
.....	

こうしたバラツキの大きさを数値化したい。しかし、各ケースの偏差を単純に足し合わせるとゼロになる。平均より大きい値の年齢では偏差が正 (プラス) の値を示し、平均より小さい値の年齢では偏差が負 (マイナス) の値を示す。そこで、両者を足し合わせれば打ち消しあってしまうのだ。

そこで、偏差を二乗する。負の値でも二乗すれば正の値になる。各ケースの偏差を二乗し合算（二乗和）した値を【変動（平方和）】（sum of squares）という。変動はこの標本の特定変数（年齢）の値のバラツキの程度を示す指標であるといえる。変動の値が大きければその変数のバラツキは著しく、変動の値が小さければバラツキは僅少である。

$$SS = \sum_{i=1}^n (x_i - m)^2$$

しかし、各ケースの平均との差（偏差）の二乗和の絶対値は、標本の大きさ（ n ）が大きければ必然的に大きくなる。従って異なる n の複数標本間の比較には適さない。そこで、1 ケースごとの変動（平方和）を求める。これが【分散（平均平方）】（variance/mean square）である。

$$s^2 = \frac{\sum_{i=1}^n (x_i - m)^2}{n}$$

この値はじつは誤差が生じやすいことが知られていて、上記の式の n を $(n-1)$ と置き換えて補正することが多い。これを【不偏分散】（unbiased variance）という。SPSS で不偏分散を採用している。

そもそも各ケースの値のバラツキの程度を表現しようとして始まったことである。偏差の概念に戻るべく、分散の平方根を求め、これを【標準偏差】（standard deviation）という。

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n - 1}}$$

平方根の概念を用いると

$$y = x^2$$

であれば

$$x = \sqrt{y}$$

の関係がある。そこで

$$\text{分散} = \text{標準偏差}^2$$

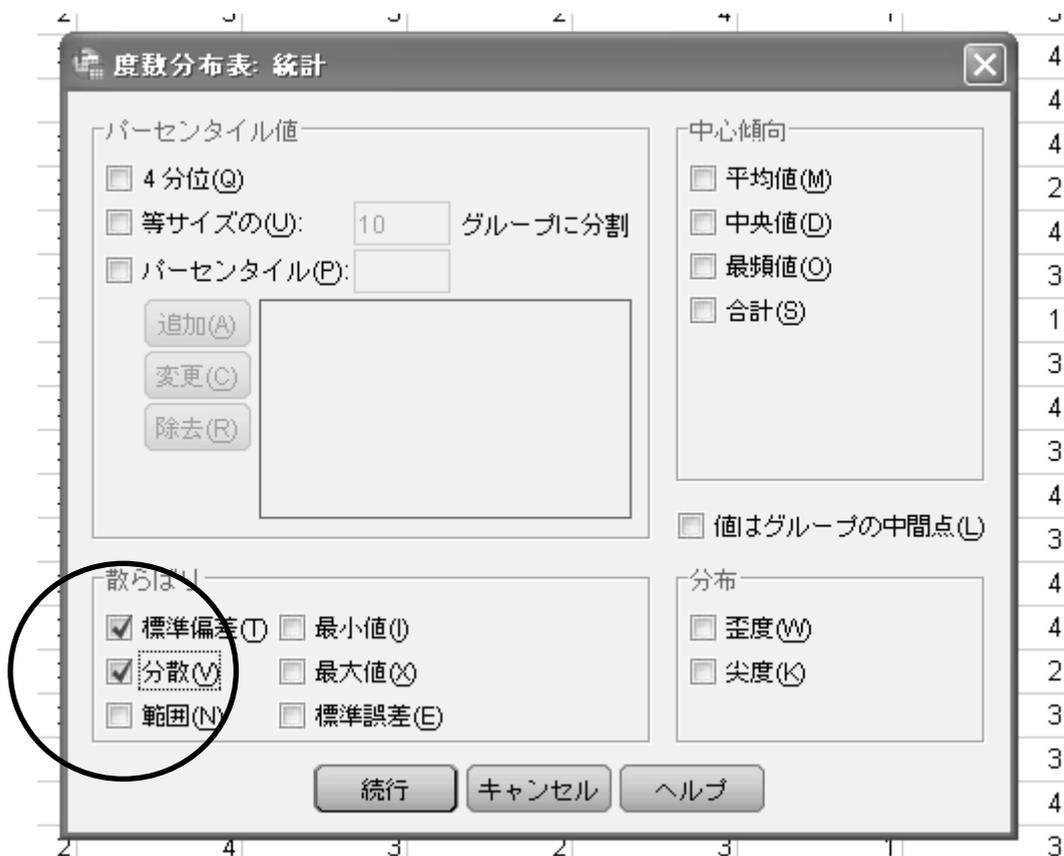
あるは

$$\text{標準偏差} = \sqrt{\text{分散}}$$

となる。

研究発表を行う際は、用いた尺度の得点の平均値と標準偏差を示す。この両者を示すことで、その変数のおよその大きさとバラツキの程度を読者に（そしてその前に査読者に）示すことが出来るからである。

SPSS での散布度を求めるのも、「分析」⇒「記述統計」⇒「度数分布」⇒「統計量」と進み、「度数分布表：統計」のダイアログ・ボックスで、標準偏差、分散をクリックして を入れる。「続行」をクリックし、もとの画面に戻り、OKをクリックする。



次のような統計結果が出力される。

統計量

OWN AGE

度数	有効	241
	欠損値	6
標準偏差		4.691
分散		22.006

4. 正規分布

1) 正規分布とは

我々が扱う多くの変数の部分は、先ほどの年齢分布のような特徴を示す。分布はほぼ左右対称、単峰性、釣鐘型である。これを正規分布 (normal distribution) ガウス分布 (Gaussian distribution) という¹²。身長、体重、体温、知能指数など多くの生命現象の分布が正規分布を示す。そして我々が日常的に使用する多くの統計解析手法は、変数が正規分布を示すことを前提としている。こうした解析法をパラメトリック・テストと呼ぶことはすでに述べた。

正規分布は

¹² ドイツの数学者ガウス (Johann Carl Friedrich Gauss 1777-1855) に由来する

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

という式で表される。ここで、 $\exp(x) = e^x$ 、 π は円周率、 σ^2 は母分散。しかし、この式を覚える必要はない。

正規分布を示すことがわかれば、その変数の分布は平均値と標準偏差を示すことを推定できる。第 1 に、正規分布を示すデータでは、平均値、中央値、最頻値の 3 者が同じ値になる。

平均値 = 中央値 = 最頻値

第 2 に、平均値から 1 標準偏差分以内に、全ケースのうち 68.26 % (約 68 %) が含まれ、さらに平均値から 2 標準偏差分以内に、全ケースのうち 95.44 % (約 95 %) が含まれる (図 2-2)。

【正規分布における平均値と標準偏差の関係】

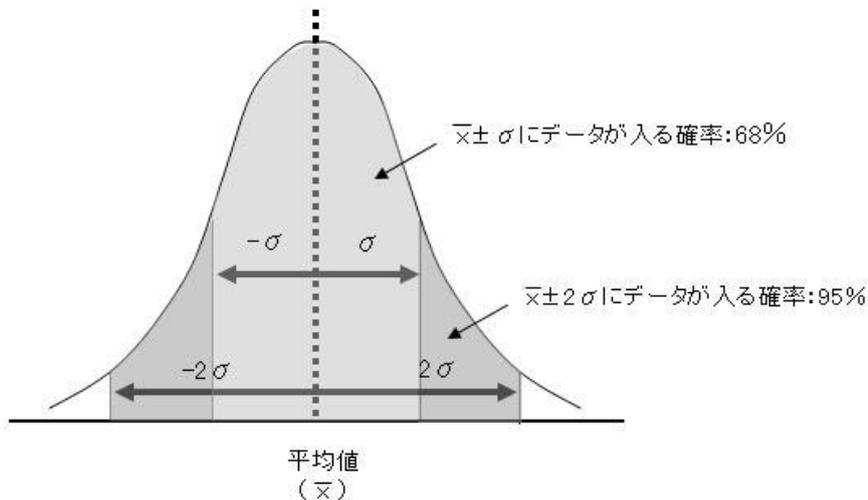


図 2-2. 正規分布

こうした特徴から変数が正規分布を示すことが明らかになれば、平均値と標準偏差の値のみでその変数のデータ分布を (実際にデータにおける各ケースの値やあるいはそのヒストグラムを見なくとも) 推定できる。そのため統計学では大変重要な概念となっている。

臨床研究者が変数を扱う際に重要になるのは、収集したデータが本当に正規分布を示しているかである。もし、その変数とその標本中で正規分布を示すと想定できるのであれば、統計技法としてパラメトリック・テストを用いることができる。しかし、その変数とその標本中で正規分布を示すと言えないのであれば、使える統計技法はノン・パラメトリック・テストとなり、技法の選択肢が狭まり、また使用の便宜性も落ちる。

2) 正規性の検定

では、目の前にある変数のデータが正規分布を示すことを示すことを確認するにはどうすればよいのか。

まずは、【ヒストグラム】を描いてみよう。正規分布から外れている分布はすぐにわかる。

次に【歪度】(skewness) を見てみる。歪度は、データが平均の回りに対称に分布していない度合いを示す尺度である。非対称度ともいう。 x_1, x_2, \dots, x_n というデータで、その平均が m であり、標準偏差が s であれば、

$$\text{skewness} = \frac{\sum_{i=1}^n (x_i - m)^3}{ns^3}$$

から計算される。これも覚える必要はない。

歪度がゼロ (0) であれば、左右対象の分布である。歪度が正の値であれば、分布は右に長く伸びている (図 2-3)。多くのケースは低値に集まっている。こうした分布は比較的多くみられる。

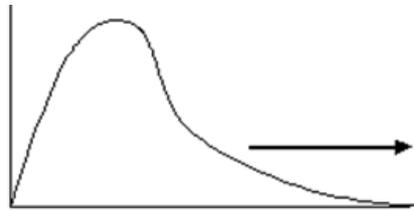


図 2-3. 正の歪度

歪度が負の値であれば、分布は左に長く伸びている (図 2-4)。多くのケースは高値に集まっている。

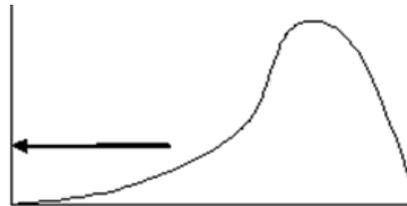


図 2-4. 負の歪度

歪度がゼロに近ければ正規性を考えても問題はない。-0.3 ~ +0.3 の間であれば十分である。

-1.0 以下あるいは +1.0 以上であれば問題が出てくるし、-2.0 以下あるいは +2.0 以上であればパラメトリック・テストを直ちに使うわけには行かない¹³。この場合、何らかの対応を検討しなければならない。その手法については後で述べる。

【尖度】(kurtosis) は、分布や頻度分布のとんがり具合を表す値である。尖度が大きければ鋭いピークと長く太い尾を持った分布であり、尖度が小さければより丸みがかったピークと短く細い尾を持った分布であるという事が判断できる (図 2-5)。

¹³ West, S. G., & Finch, J. F. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues and applications (pp. 56-75). Newbury Park CA: Sage.

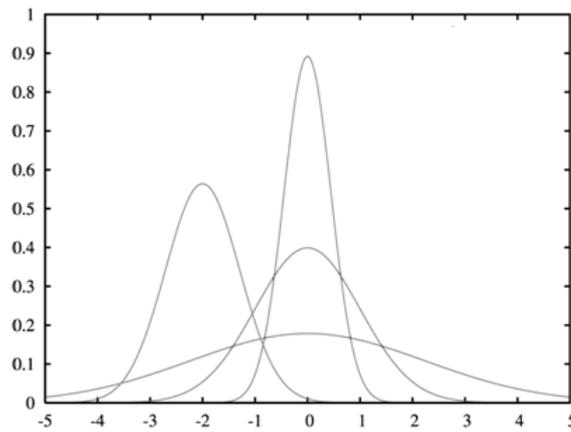


図 2-5. 尖度

図 2-5 には 4 つの変数の分布が示されている。そのうち 3 つの変数はその平均値が同一であるが、そのとんがり具合が大きく異なる。 x_1, x_2, \dots, x_n というデータで、その平均が m であり、標準偏差が s であれば、

$$\text{kurtosis} = \frac{\sum_{i=1}^n (x_i - m)^4}{ns^4} - 3$$

から計算される¹⁴。これも覚える必要はない。

正規分布では尖度がゼロ (0) になる。尖度が正の値であれば、中央の尖りがきつく、左右の裾が狭い分布を表す。尖度が負の値であれば、中央の尖りが緩やかで、左右に裾が広い分布を示す。尖度の絶対値が 4 を超えると正規性を想定するわけにはいなくなる¹⁵。

さらに判断が困る場合は【正規性の検定】を行う。ここには Shapiro-Wilk 検定 (Shapiro-Wilk test) と Kolmogorov-Smirnov 検定がある。Shapiro-Wilk 検定の結果は p 値で表される。 p 値が 0.5 未満である時、「その変数の分布は有意に正規分布に従わない」と判断する。 p 値が 0.5 以上の場合、「その変数の分布は有意に正規分布に従わないとはいえない」と判断する。

SPSS で正規分布の確認をしてみよう。先ほどと同じ被験者の年齢 (AGE) を扱う。まずヒストグラムを描いてみる。先ほどと同様に、「グラフ (G)」⇒「レガシー ダイアログ (L)」⇒「ヒストグラム (I)」と進み、ヒストグラムのダイアログ・ボックスが出たら、変数の一覧表から OWN AGE [W1Q10] をクリックして色を反転させ、それから⇒のアイコンをクリックし、「変数 (V)」に OWN AGE [W1Q10] を移動する。ここで、変数の直下にある「正規性の表示 (D)」の口をクリックして ✓ を入れる。「OK」をクリックする前に「貼り付け」をクリックして、シンタックス文

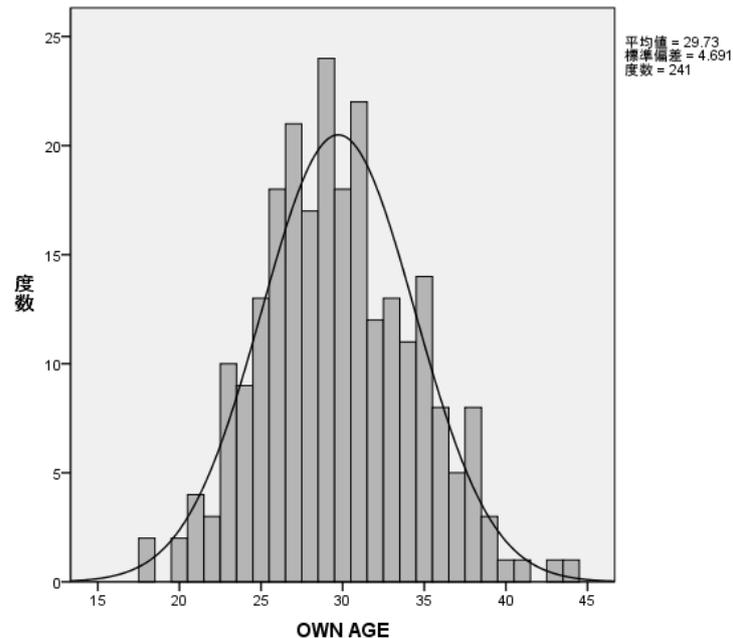
```
5
6 GRAPH
7 /HISTOGRAM(NORMAL)=W1Q10.
8
```

¹⁴ -3 を付けない計算式を挙げている教科書もある。

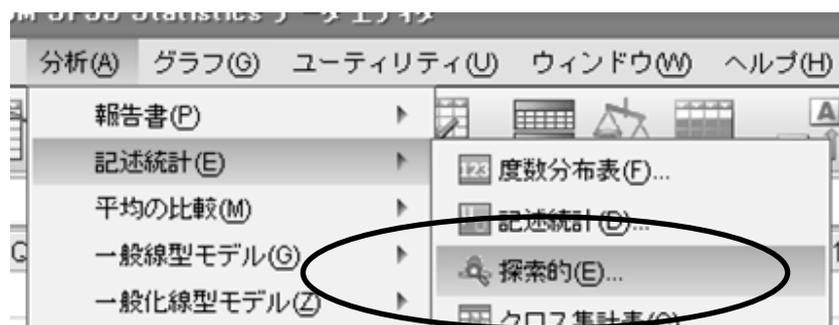
¹⁵ West, S. G., & Finch, J. F. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues and applications (pp. 56-75). Newbury Park CA: Sage.



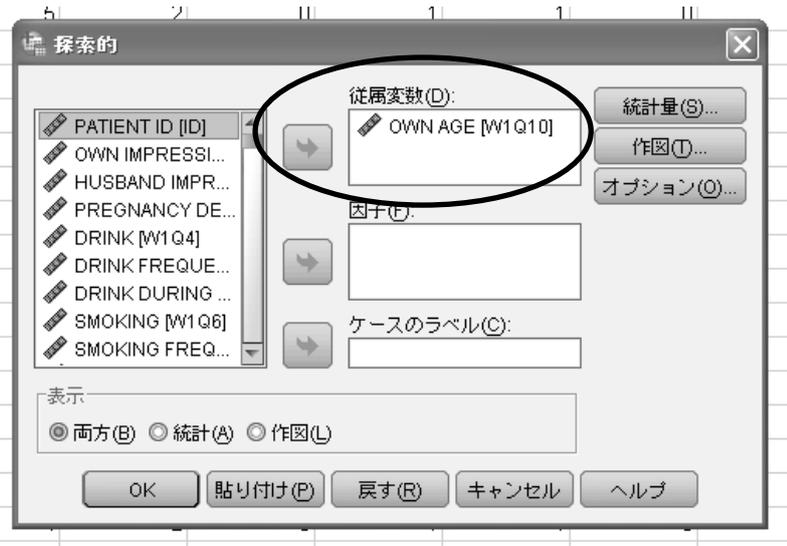
を保存しよう。このシンタックス文をカーソルで選び、すべて色を反転させてからアイコンをクリックする。すると次の結果が現れる。ここには、「正規性の表示 (D)」を指示したおかげで、正規分布の推定値が表記されている。



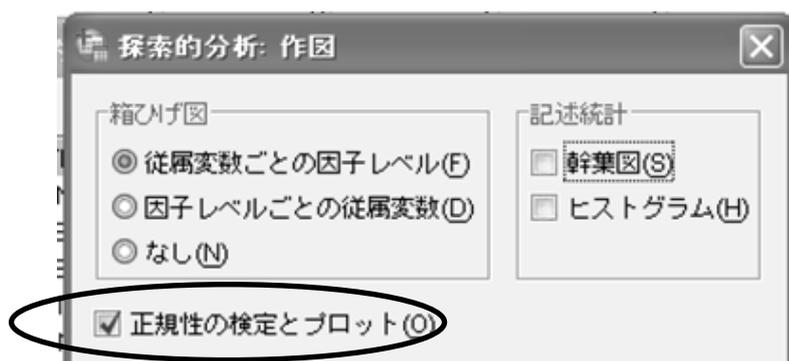
次に、「分析 (A)」をクリックしてプルダウン・メニューを出し、ここで「記述統計 (E)」にカーソルを合わせると自動的にいくつかの下位メニューが現れる。ここから「探索的 (E)」を選んでクリックする。



すると「探索的」のダイアログ・ボックスが現れる。変数の一覧表から OWN AGE [W1Q10] をクリックして色を反転させ、それから \Rightarrow のアイコンをクリックし、「変数 (V)」に OWN AGE [W1Q10] を移動する。



そこで、「作図 (T)」をクリックし「探索的分析：作図」のダイアログ・ボックスを開く。ここでは「正規性の検定とプロット (O)」の口をクリックして を付ける。そして「続行」をクリックし、もとの画面に戻り、「OK」をクリックする。



結果の出力を見てみよう。Shapiro-Wilk の有意確率が 0.145 となっている。ここから「年齢の分布は正規分布に従わないとはいえない」と結論する。なお、有意水準が 0.145 でなく、.145 と記載してあるのは、そもそも有意水準は 1.0 を超えることがなく、まず絶対に 1.0 未満の値であるからゼロ (0) を省略している。これは American Psychological Association の publication manual のフォーマットでもある。

正規性の検定

	Kolmogorov-Smirnov の正規性の検定 (探索的) ^a			Shapiro-Wilk		
	統計量	自由度	有意確率.	統計量	自由度	有意確率.
OWN AGE	.072	241	.004	.991	241	.145

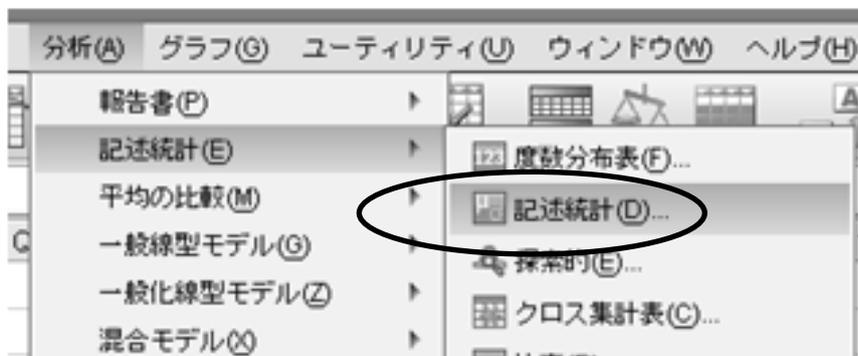
a. Lilliefors 有意確率の修正

さらにこの出力の前には歪度と尖度も示されている。歪度は 0.22, 尖度は -0.098 であり、これも変数分布の正規性を示唆するものである。

記述統計

		統計量	標準誤差
OWN AGE	平均値	29.73	.302
	平均値の 95% 信頼区間		
	下限	29.14	
	上限	30.33	
	5%トリム平均	29.68	
	中央値	29.00	
	分散	22.006	
	標準偏差	4.691	
	最小値	18	
	最大値	44	
	範囲	26	
	4 分位範囲	7	
	歪度	.220	.157
	尖度	-.098	.312

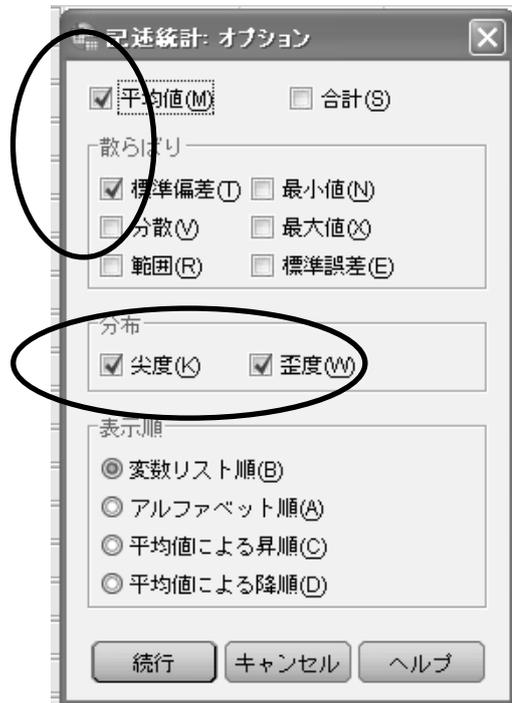
尖度と歪度の求め方は別の経路がある。「分析 (A)」をクリックしてプルダウン・メニューを出し、ここで「記述統計 (E)」にカーソルを合わせると、自動的にいくつかのメニューが現れるので、「記述統計 (D)」を選びクリックして、「記述統計量」のダイアログ・ボックスを開ける。



変数の一覧表から OWN AGE [W1Q10] をクリックして色を反転させ、それから ⇒ のアイコンをクリックし、「変数 (V)」に OWN AGE [W1Q10] を移動する。



次いで「オプション (O)」をクリックして、「記述統計：オプション」のダイアログ・ボックスを開ける。ここでは、尖度、歪度に加え、平均値と標準偏差の口をクリックして ✓ を付ける。それから「続行」をクリックしてもとの画面に戻り、ここで「OK」をクリックして、結果を表示させる。



以下の結果が出力される。

記述統計量

	度数	平均値	標準偏差	歪度		尖度	
	統計量	統計量	統計量	統計量	標準誤差	統計量	標準誤差
OWN AGE	241	29.73	4.691	.220	.157	-.098	.312
有効なケースの数 (リストごと)	241						

5. 代表値と散布度の論文での表示

代表値と散布度について検討を加えた。実際に論文を書くときにはどこまで記載すればよいのだろうか。

論文の結果パートでまず書くべきことは、当該研究で用いた すべての変数 の、本来取り得る範囲、平均値、標準偏差である。さらに正規性が問われそうな場合は、skewness を一覧表示しよう。正規性の検定を要求されることは稀である¹⁶。いくつかの項目を足し合わせるなどして算出する変数がある。たとえば、エジンバラ産後うつ病尺度は 10 項目の得点の合算である。こうした変数を合成変数 (composite variable) という。合成変数を研究

¹⁶ もちろん、査読者による話ではある。

で使用することは多い。この場合、論文の結果パートの最初の一覧表示では合成変数の内的整合性の値（後述）も同時に挙げるのが必須である¹⁷。

平均値はあまり極端に偏っていないことを確認する。天井効果（ceiling effects）や床効果（floor effects）の可能性を排除するためである。あまりに低値あるいは高値に偏っていれば歪度も大きいであろう。

標準偏差があまりに小さい場合、そもそもその値が標本を横断して一定に近いわけであるから、連続量として扱うことについて疑問が発生してしまう。

6. 正規分布に従わない場合の対応

おそらく多くの論文執筆に際して最重要課題のひとつは、変数の分布（名義尺度を除いて）が正規分布を取らない場合の対応であろう。

順序（順位）尺度もその評価段階数が 4 あるいは 5 以上であれば、間隔尺度あるいは比尺度、すなわち連続量を示す変数と同様と見做して、パラメトリック・テストを用いることが出来る。この場合も、正規性の検討は必要である。

以下のような質問項目を考えてみよう（図 2-6）。設問に対する回答選択肢は 4 つ準備されている（「はい、たいていそうだった」から「いいえ、全くそうではなかった」）。しかし、今後、パラメトリック・テスト（例えば因子分析）を使用するつもりであれば（そしてその方が良いに決まっている！）、4 段階尺度でなく、回答選択肢数を 4 から 5 にすることを考えるべきであろう。査読の批判を避けることに加えて、実際に結果の出方が望ましいのである。たとえば、2 件法尺度として作られたテストの回答選択肢を 4 件法にするだけで、合成変数の内的整合性が改善する¹⁸。

- 9 不幸せなので、泣けてきた。
- 1) はい、たいていそうだった
 - 2) はい、かなりしばしばそうだった
 - 3) ほんの時々あった
 - 4) いいえ、全くそうではなかった

図 2-6. 4 件法尺度の例

このことは、調査が終了してからでは遅く、むしろ、調査開始前に調査票を選ぶ基準として回答選択指数（「何件法か？」）を十分考慮する必要がある。自身で調査項目を作成する際も同様の注意が必要である。

調査が終了し、連続量を示す変数の分布をヒストグラムで見ると二峰性だった場合は、何らかの区分点を設定して、あたかも名義尺度であるかのように扱う。二峰性分布を示すデータの区分点の設定の常道はない。研究者のセンスで決めればよい。たとえば、中央値で上下二分するのも一法であろう。

もうひとつの方法は、値が極力正規分布に似た分布を取るよう値の変更を加えることである。たとえば正の方向に skew したデータはその対数を取ってみよう。変数 x が正方向に skew していたなら

$$\text{new}x = \log(x)$$

¹⁷ この事の重要性は共分散構造分析で再度取り上げる。

¹⁸ Kijima, N., Tanaka, E., Suzuki, N., Higuchi, H. & Kitamura, T. (2000). Reliability and validity of the Japanese version of the Temperament and Character Inventory. *Psychological Reports*, 86, 1050-1058.

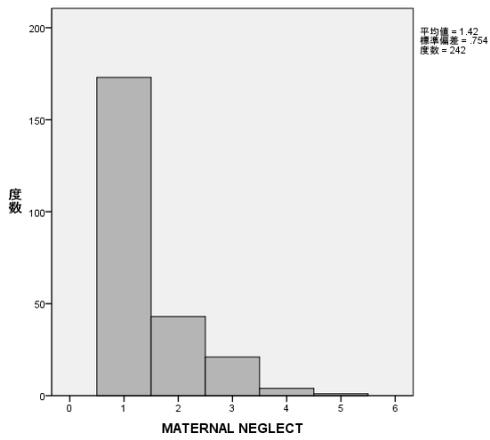
と対数変換をしてみよう。

また、負の方向に skew したデータは、その平方根を求めてみよう。

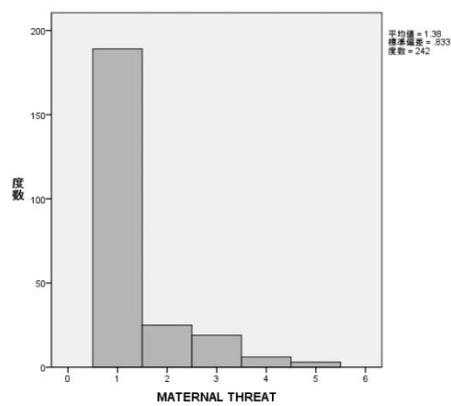
$$\text{newx} = \sqrt{x}$$

実例で見てみよう。サンプル・データでは、15 歳以前に両親から受けた虐待について聞いている。ここでは母から受けた虐待についてその分布を見てみよう（図 2-5）。いずれも、その分布は左側に寄っていて、強く正の方向に skew していた。

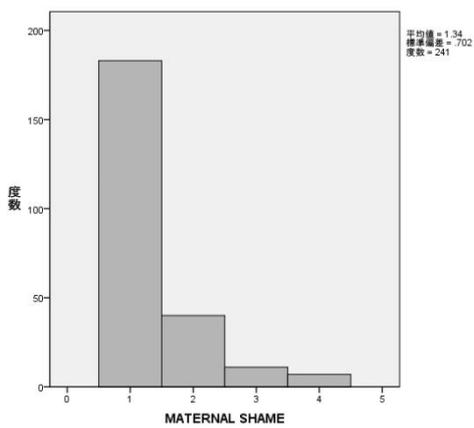
無視



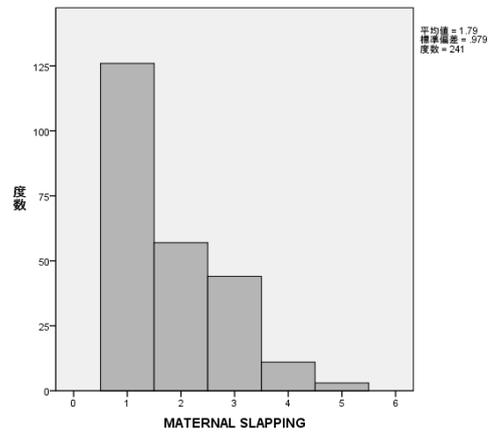
脅迫



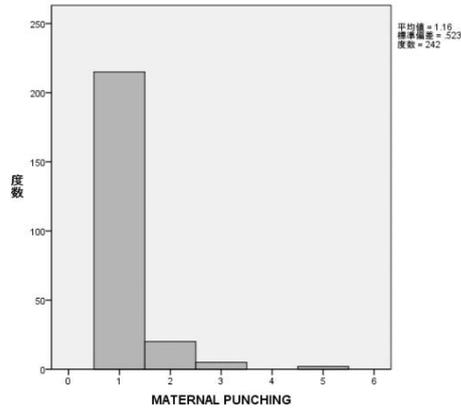
恥じ



平手打ち



こぶし打ち



物でたたく

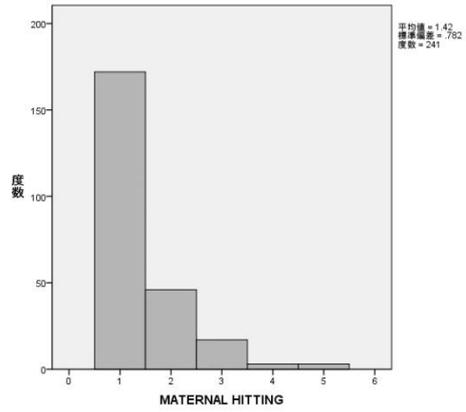
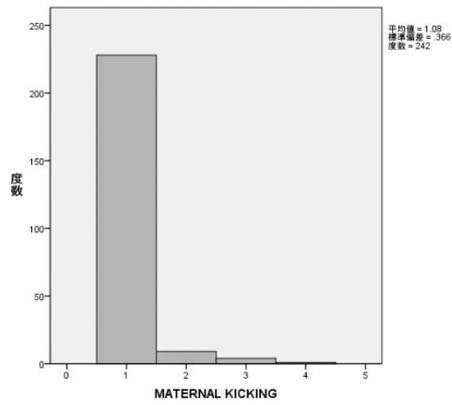


図 2-6. 被虐待体験の頻度分布

蹴る



火傷

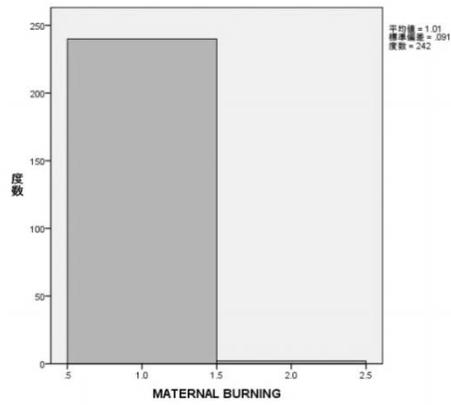


図 2-6. 被虐待体験の頻度分布 (続)

そこで各変数を対数変換させた。そのためのシンタックス文は次のような compute コマンドである。

```

1 COMPUTE NEWW1Q31X1=LN(W1Q31X1).
2 EXECUTE.
3 COMPUTE NEWW1Q31X2=LN(W1Q31X2).
4 EXECUTE.
5 COMPUTE NEWW1Q31X3=LN(W1Q31X3).
6 EXECUTE.
7 COMPUTE NEWW1Q31X4=LN(W1Q31X4).
8 EXECUTE.
9 COMPUTE NEWW1Q31X5=LN(W1Q31X5).
10 EXECUTE.
11 COMPUTE NEWW1Q31X6=LN(W1Q31X6).
12 EXECUTE.
13 COMPUTE NEWW1Q31X7=LN(W1Q31X7).
14 EXECUTE.
15 COMPUTE NEWW1Q31X8=LN(W1Q31X8).
16 EXECUTE.
17

```

その上で、原法と対数変換後の新変数で skewness, kurtosis, Shapiro-Wilk を比較してみると、以下のようになった（表 2-4）。

表 2-4. 被虐待体験の頻度分布と対数変換

	SKEWNESS		KURTOSIS		Shapiro-Wilk	
	原法	対数変換	原法	対数変換	原法	対数変換
邪魔者扱いされたり、「置いていっちゃ うよ」「お前のような子はもういら ない」「うちの子ではない」などと言わ れた	1.91	1.35	3.39	0.41	0.60	0.62
「ごはん抜きよ」「捨てるよ」「部屋に 行きなさい」と脅されたり、あなたが 大切にしていたペットや人に危害を加 えるとか、おもちゃを壊すとか捨てる と言って脅された	2.33	1.74	5.07	1.67	0.53	0.56
侮辱されたり、他の人がいる前で強く 叱られたり馬鹿にされて 恥ずかしい 思いをさせられた	2.25	1.65	4.72	1.47	0.55	0.57
平手で叩かれた	1.08	0.53	0.42	-1.21	0.77	0.77
こぶしで叩かれた	4.60	3.18	25.90	10.21	0.33	0.37
物で叩かれた	2.21	1.39	5.39	0.76	0.60	0.63
蹴られた	5.06	4.30	27.94	18.24	0.24	0.25
煙草などで火傷をおわされた	15.46	15.46	239.00	239.0	0.04	0.04

Shapiro-Wilk では改善はなかったが、skewness と kurtosis では改善を見ている。「火傷」項目はほとんどが経験していないので、正規性の分布に変更する余地が全くなかったことも分かった。おそらく、「火傷」項目は名義変数として扱うことの方が適切なデータなのであろう。

最後に、正規分布に戻すことは断念して、ノン・パラメトリック・テストを用いる方法が残されている。

5 欠損値の扱いかた

1. 欠損値とは

調査を行えばいくつかの項目で回答がないことがある。これを欠損値 (missing value) という。たとえば、サンプル・データ (sample.sav) の Hospital Anxiety and Depression Scale (HADS) の項目の得点を見てみよう。ID 85 のケースの HADS 第3項目から第7項目までが数値の入力がない。また ID 90 の HADS 第5項目にも数値の入力がない。これが欠損値である。

ID	HAD2	HAD3	HAD4	HAD5	HAD6	HAD7	HAD8
83	2	4	1	4	4	1	3
84	1	3	1	3	3	1	3
85	1	4
86	2	4	1	3	4	1	4
87	2	3	1	3	4	2	2
88	2	4	1	3	3	1	3
89	1	4	1	4	4	1	4
90	2	4	1	.	2	1	4
91	3	3	1	4	4	2	3

2. 欠損値があった場合の対応

単一項目の場合、欠損値はそのものとして扱うしかない。たとえば「うつ病の有無」という変数に欠損値があるケースが5つあったとしよう。平均年齢が「うつ病」のあり群となし群で差があるかを検討したいと考えた。しかし、計算をする場合、この5ケースは「存在しない」と見なさいと計算はできない。

ところが複数項目で合成変数を作成する場合を考慮してみよう。たとえば HADS は全部で14項目の設問からできていて、抑うつ7項目、不安7項目の各合算点を抑うつ・不安の総合得点としている。14項目すべてに無回答だった場合は、「そのケースは存在しない」と見做すしかない。しかし、わずか2項目のみ欠損値で、残りの12項目は見たところ正直に回答しているように見えたとしよう。このケースを「存在しない」とするのは「もったいない!」と感じるであろう。特に、比較的多数項目から作成する合成変数の場合は、こうした事態が起こりやすい。そこでの対応には (1) リストワイズ除去法 (2) 項目平均値代入法 (3) ケース平均値代入法 (4) 回帰代替法がある。

3. SPSS での対処

ここでは項目平均値代入法のシンタックス文の例を以下に挙げる。

```
RECODE DEP1,DEP2,DEP4 (1=0)(2=1)(3=2)(4=3)(5=4)
      INTO D1,D2,D4.
RECODE DEP3,DEP5 (5=0)(4=1)(3=2)(2=3)(1=4)
      INTO D3,D5.
```

```

RECODE ANX2,ANX3,ANX5 (1=0)(2=1)(3=2)(4=3)(5=4)
        INTO A2,A3,A5.
RECODE ANX1,ANX2 (5=0)(4=1)(3=2)(2=3)(1=4)
        INTO A1,A2.
COUNT      QSCALE=DEP1,DEP2,DEP3,DEP4,DEP5,
            ANX1,ANX2,ANX3,ANX4,ANX5(0,1,2,3,4).

RMV
/DEP1_1=SMEAN(DEP1)/DEP2_1=SMEAN(DEP2)/DEP3_1=SMEAN(DEP3)
/DEP4_1=SMEAN(DEP4)/DEP5_1=SMEAN(DEP5)/
/ANX1_1=SMEAN(ANX1)/ANX2_1=SMEAN(ANX2)/ANX3_1=SMEAN(ANX3)
/ANX4_1=SMEAN(ANX4)/ANX5_1=SMEAN(ANX5).
COMPUTE      DEPX=DEP1_1+DEP2_1+DEP3_1+DEP4_1+DEP5_1.
COMPUTE      ANXX=ANX1_1+ANX2_1+ANX3_1+ANX4_1+ANX5_1.
IF           (QSCALE EQ LT 8)DEPX=99.
IF           (QSCALE EQ LT 8)ANXX=99.
MISSING VALUES  DEPX,ANXX(99).
VAR LABELS    DEPX 'DEPRESSION MEANSUB'/
              ANXX 'ANXIETY MEANSUB'
              QSCALE 'NUMBER OF SCALE ITEMS ANSWERED'.

```