

医学論文のための

研究 デザインと

統計 解析

査読者が
教える

●著.....森本 剛



中山書店

● 目 次 ●

はじめに——本書の目的	iii
本書の利用法	v

1 章 臨床研究と統計解析

1.1 臨床研究と統計解析	1
1.2 統計解析ソフトウェアについて	4

2 章 これだけは覚えておきたい統計基礎

2.1 変数とは	7
(1) 連続変数	9
(2) 順序変数	13
(3) 名義変数	13
(4) 2 元変数	14
2.2 標準偏差と標準誤差	16
2.3 統計解析の原則	20
2.4 ダミーコード化	23

3 章 データ収集とデータクリーニング

3.1 データセットの作成	30
3.2 欠損（欠測）値の扱い方	38
3.3 データクリーニング	40
3.4 入力規則	42

4 章 臨床研究で使われる統計基礎

4.0 推定とは	53
4.1 検定とは	54
4.2 検定の根本的な考え方	55
4.3 t 検定	59
4.4 χ^2 検定	64

4.5	もう一度ここで推定	68
4.6	検定・推定結果を臨床的に評価する	72
4.7	交絡因子	74
4.8	多変量モデル（多変量解析）	78
5 章	臨床研究のデザインのしかた	84
5.1	研究デザインの概念	85
(1)	横断研究	86
(2)	コホート研究	88
(3)	ランダム化比較対照試験（RCT）	91
(4)	症例対照研究	96
5.2	研究デザインの選択	98
5.3	判定・評価の重要性	101
5.4	判定・評価の信頼性・妥当性	102
5.5	評価のタイミング	104
5.6	研究デザインのまとめ	108
6 章	研究計画書（プロトコル）作成	111
6.1	なぜ研究計画書（プロトコル）か？	112
6.2	研究計画書（プロトコル）に必要な要素	115
(1)	概要	118
(2)	目的・背景	120
(3)	研究デザイン	120
(4)	適格基準	121
(5)	説明と同意	122
(6)	採用手続きと割り付け	125
(7)	アウトカムの評価	126
(8)	有害事象の評価・報告	127
(9)	観察項目とスケジュール	128
(10)	研究費・COI	129
(11)	研究組織	131
(12)	各種委員会	132
6.3	研究計画書（プロトコル）作成の要点	132

7章 論文化に向けて 135

7.1 標本患者集団を明確に 135

7.2 図表の選び方 139

Topic マーカー研究と診断特性 144

1. マーカー研究 144

2. 診断特性 146

3. ROC 曲線 150

4. 選別点（カットオフポイント）の設定 152

5. マーカー研究のデザイン 154

おわりに 159

索引 161

Column

統計解析ソフトウェアの操作による変数の扱い方 27

個人データの扱いにはくれぐれも注意を 31

データ入力時は半角英数でセッティングを 36

欠損（欠測）値の取り扱い 39

普段の診療の質を上げましょう！ 51

両側検定と片側検定の違い 61

オッズ比とハザード比 72

多重比較 80

傾向スコア（Propensity score）を用いた解析 83

生存解析 90

サンプルサイズ計算 93

非劣性試験 95

研究期間は3年くらいを目安に 108

ケースシリーズ・症例報告 110

臨床研究をきちんと行うと測定装置のエラーもわかる？ 134

Table 1 139

サブグループ解析 143

2

これだけは覚えておきたい 統計基礎

私が勝手に、医学統計の世界で3大宗教と呼んでいるものがある。初心者が「とにかくやってしまうベスト3」である。

連続変数教：とにかく何でも連続変数として処理する

平均値±SE教：とにかく何でも平均値とSE（標準誤差）で示す（基礎実験を経験されていた方に多い）

多変量解析教：とにかく多変量解析を行った結果が正しいと考える

これら3つの手法を駆使して、データに何らかの関係性があることを示すのはたやすい。しかし、それを臨床にフィードバックするときにきちんと意味づけられる人は少ない。

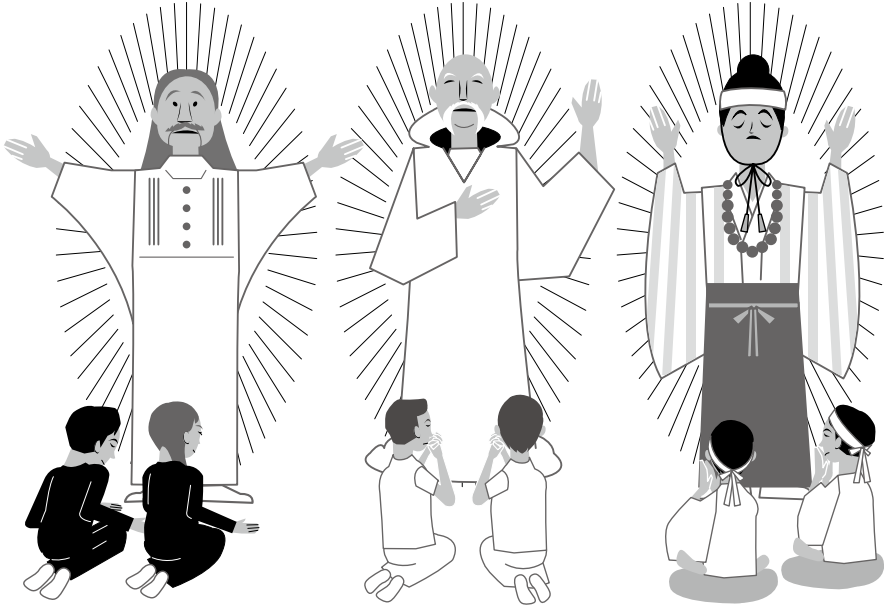
本章では、本書で統計を理解するための必要最小限の知識として、変数の種類とその表現方法について解説する。

3大宗教

連続変数 教

平均値±SE 教

多変量解析 教



初心者がとにかくやってしまうベスト3

2.1 変数とは

図 2.1 はある論文から持って来た表 (Table) である。見ての通り、すべて数字で表されている。患者によって変わる数字、すなわち変数である。臨床研究では基本的には数値になったデータ、つまり変数を扱う。したがって解析のときだけでなくデザインの段階から、変数を常に意識しておくことが大切である。

変数には、大きく分けて、(1) 連続変数、(2) 順序変数、(3) 名義変数、(4) 2 元変数の 4 種類がある (表 2.1)。本節ではこの 4 つそれぞれの特徴について解説したい。

Table 1. Patient, Lesion, and Procedural Characteristics

	Everolimus-Eluting Stent N=1597	Sirolimus-Eluting Stent N=1600	P
Patient characteristics			
Age, y	68.9±9.7	69.3±9.6	0.33
Age ≥75 y, n (%)	494 (31)	509 (32)	0.59
Male sex, n (%)	1238 (78)	1217 (76)	0.33
Body mass index	24.2±3.6 (1580)	24.3±3.5 (1586)	0.5
Coexisting condition, n (%)			
Hypertension	1269 (79)	1290 (81)	0.41
Diabetes mellitus	726 (45)	713 (45)	0.61
Insulin-treated diabetes	175 (11)	163 (10)	0.48
Treated with oral medication only	353 (22)	381 (24)	0.25
Treated with diet therapy only	198 (12)	169 (11)	0.1
Dyslipidemia	1189 (74)	1204 (75)	0.6
ESRD (eGFR<30 mL/min per 1.73 m ²) not on hemodialysis	33 (2.1)	39 (2.4)	0.48
Hemodialysis	93 (5.8)	80 (5.0)	0.3
Atrial fibrillation	110 (6.9)	123 (7.7)	0.38
Anemia (hemoglobin <11.0 g/dL)	198 (12)	206 (13)	0.69
Chronic obstructive pulmonary disease	44 (2.8)	30 (1.9)	0.1
Malignancy	100 (6.3)	112 (7.0)	0.4
Cardiac risk factor, n (%)			
Current smoker	332 (21)	326 (20)	0.77
Family history of coronary artery disease	253/1375 (18)	262/1365 (19)	0.59
Previous myocardial infarction	471 (29)	496 (31)	0.35
Previous stroke	180 (11)	162 (10)	0.29
Heart failure	209 (13)	207 (13)	0.9
Peripheral vascular disease	143 (9.0)	137 (8.6)	0.7
Previous percutaneous coronary intervention	757 (47)	811 (51)	0.06
Previous coronary artery bypass grafting	6 (3.9)	99 (6.2)	0.003

(Kimura T et al. *Circulation*. 2012;126:1225-1236)

図 2.1 ある論文の Table

表 2.1 変数の種類

種類	長 所	短 所
連続変数	統計学的な検出力は最大 事象を鋭敏に検出できる	欠損値が発生しやすい
順序変数	事象を比較的鋭敏に検出できる	//
名義変数	臨床的にわかりやすい 数字では表現できない事象が表現できる	そのままでは統計処理ができない
2 元変数	欠損値を少なくできる 解析結果の解釈が容易である	鋭敏さに欠ける

(1) 連続変数

臨床現場における多くのデータは連続変数である。身長や体重、BMI、血圧から始まり、血液検査の項目もすべて連続変数である。また「時間」も連続変数であるがゆえに、「何でも連続変数として扱ってしまう」初心者（連続変数教）が多いのうなずける。

統計学的な検出力は、事象を鋭敏に表現する連続変数がいちばん大きい。一方で連続変数を分割したり、群分けしたりして、情報量を減らせば減らすほど検出力が小さくなる。ただし、「関連性を見る」という点において検出力が大きく、より低い p 値を出すことができるが（「 p 値」に関しては 4.1 節「検定とは」〈p.54〉参照）、逆に臨床的な意味づけをすることが難しいことも多い。例えば、年齢 1 歳あたりのリスク比が 1.002 で有意だとしても（「有意」に関しても 4.1 節「検定とは」〈p.54〉参照）、それをどのように診療に結びつけるか、という解釈は困難である。実際に連続変数を使って出した結果については、理解のしやすさやメッセージの伝わりやすさを踏まえる必要がある。

連続変数は欠損値が発生しやすいことが欠点である。また連続変数を扱うときは、常に分布を確認することが大切であり（図 2.2）、分布の形はデータによって様々で、分布によって解析方法も異なってくる。図 2.2 ① のように真ん中を中心にほぼ左右対称のお椀をひっくり返した形（正規分布）になっているものもあれば、図 2.2 ② のように右側に長く尾を引く分布になっている場合もある。入院日数などは、5 日目くらいにピークがあって、そこから右へなだらかに長く下りて行くカーブを描く。ちなみに左側に長く尾を引くような分布は、臨床研究ではあまり見掛けない。それから夫婦岩のように大きな山からちょっと離れたところにピョンと山ができているグラフもある（図 2.2 ③）。教育データなどで見かける。

図 2.2 ① の分布は平均値（mean）と中央値（median）がほぼ一緒だと判断できる（図 2.3 ①）。中央値とは全データを大きい順（小さい順でも構わない）に並べ、ちょうど真ん中（50% の位置）にある値を指す。標本数を n として、標本数が奇数ならちょうど真ん中 $\left(\frac{n+1}{2}\right)$ となる値をとればよく、偶数ならば真ん中の 2 つ $\left(\frac{n}{2} \text{ と } \frac{n}{2} + 1\right)$ の平均値となる。平均値と中央値が同じならば、分布を代表する値（代表値）は平均値であるとして問題ない。

一方で図 2.3 ② のグラフは平均値と中央値が離れているように見える。左右対

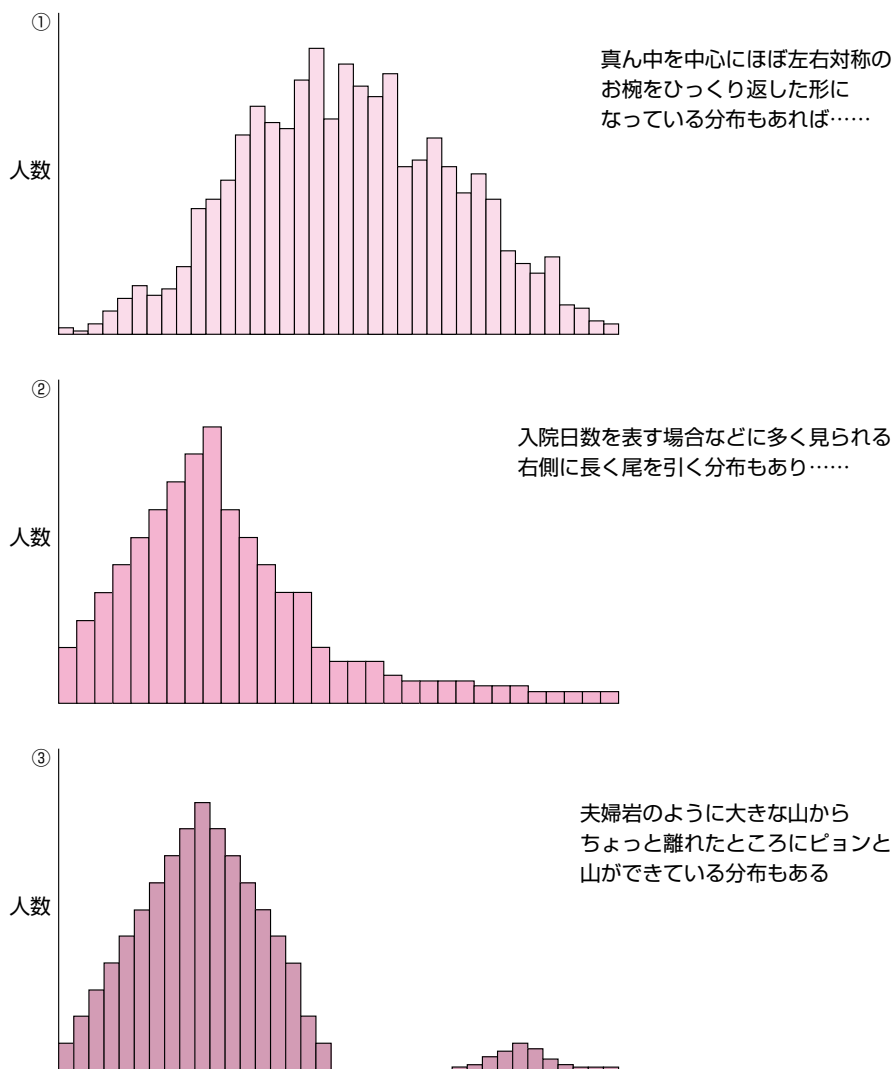


図 2.2 連続変数の分布あれこれ

称ではない偏った分布で、平均値が片方（この場合は右）に引っ張られ、分布の中で中心的な値から外れている。このような分布を代表する値（代表値）には、中央値のほうがふさわしい。なお最頻値（mode、分布の山の最も高いところの値）はほとんどの臨床研究では使われない。集計をすることが目的とはならないからと考えられる。

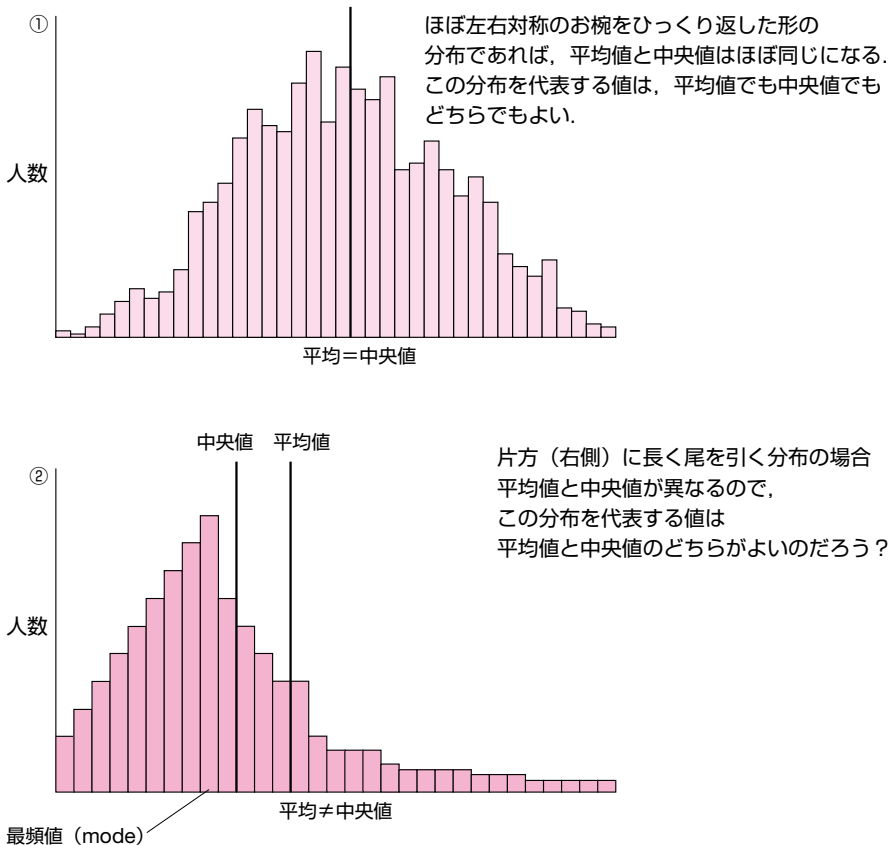


図 2.3 連続変数の分布による代表値の違い

図 2.3 ② のような右に長い分布は、年収の分布も有名である。横軸が年収で縦軸がその人数としてヒストグラムを描くと、少人数だけれども、大多数の人から桁外れに高収入の人が右に長く存在するため、その人たちに引っ張られて平均値も山の高いところから右へずれる。したがって年収の全体の平均値を提示されても、大多数の人がピンとこないのは当然で、多くの人（山の高いところ）は平均値よりもずっと左に位置しているからである。

平均値 $\pm$ SD（標準偏差，2.2 節「標準偏差と標準誤差」〈p.16〉参照）で表せるのは、平均値と中央値がほぼ一致し、お椀を逆さにしたような形の分布（正規分布）をとるデータが原則である。図 2.3 ② のような「平均値 $\neq$ 中央値」のデータ