

# 「0 と 1 のあいだ」いっしょに学ぶ統計学的解析の基礎 <sup>馬場塾オンライン特別企画</sup>

2025年7月1日19:00~開講

目次

	ータから「意味」を引き出す体験的学びへ セッション 1 の概要:オッズとオッズ比	1 1 1 1
R	<b>と RStudio の活用</b> <ol> <li>R のダウンロードとインストール</li> <li>RStudio のダウンロードとインストール</li> <li>インストール後の確認</li> <li>インストール後の確認</li> <li>Studio の基本的な使いかた</li> </ol>	2 2 2 3 3
セ	<ul> <li><b>ッション 1:オッズ比</b></li> <li>1) 確率とオッズの違い</li> <li>2) 相対リスクとオッズ比</li> <li>相対リスク</li> <li>オッズ比</li> <li>3) RStudio を活用した例題</li> <li>i) パッケージの導入と準備</li> <li>ii) データ読み込み</li> <li>iii) データの前処理</li> <li>iv) 分割表の作成</li> <li>iii) ニーレー</li> </ul>	$\begin{array}{c} 4 \\ 4 \\ 5 \\ 5 \\ 6 \\ 7 \\ 7 \\ 8 \\ 9 \\ 10 \end{array}$
セ	ッション2:フィッシャーの正確確率検定       1         1)独立性の検定       1         2) x2 検定における期待値       1         3) フィッシャーの正確確率検定のしくみ       1         i) 分割表のセル値の全パターンを列挙する       1         ii) 赤玉・白玉の抽選モデル       1         iii) 親測値とそれ以上に極端なパターンの確率を合計する (p 値の計算)       1         ii) パッケージの導入と準備       1         ii) データ読み込み       1         ii) データの前処理       1         v) Fisher の正確確率検定       1	11 11 12 13 14 16 17 17 17 18 19 20

## データから「意味」を引き出す体験的学びへ

このたびのハンズオンセミナーでは,統計学の基本概念を実際のデータ操作を通じて 体得することを目指します.対象は,胃X線検査の撮影や読影,診断などの実務に携わ りながら,「統計の読み書き」に苦手意識を感じている初学者の方々です.本セミナー では,Excel形式のデータをR言語で読み込み,分割表を作成し,オッズ比や信頼区間, Fisherの正確確率検定といった基本的な統計指標を実際に計算する一連のプロセスを, 講師とともにステップバイステップで体験していただきます.

本セミナーの目標は,統計学を「知識」としてではなく,「**手を動かし,目で確かめ**, 意味を読み取る」実践的な力として身につけることです.Rを用いたデータ操作に不安 がある方も,事前に必要なパッケージの準備から始め,講師が丁寧にナビゲートします のでご安心ください.医療現場や研究の場で「統計が語るストーリー」を理解し、活用 する第一歩を、ぜひ本セミナーで踏み出していただければ幸いです.

### セッション1の概要:オッズとオッズ比

セッション1では二重造影像の面積率を中央値で「大きい/小さい」に分け、その大小 とある手技(両肩当て左向き)との関係を分割表で可視化し、オッズ比の計算を通じて 差異を読み解きます.

#### セッション2の概要:Fisherの正確確率検定

セッション2では、スケッチ前後での画像の「造影効果」や「描出範囲」の変化をテーマに、フィッシャーの正確確率検定を用いて「見ための違い」が統計的に有意かどうかを評価します.これにより、実験前後の変化を信頼をもって解釈する力を養います.

#### セッション 3 の概要: 95% 信頼区間

セッション3では、2×2表から算出したオッズ比に対して、信頼区間の計算や対数オッズ比の標準化を行います。信頼区間の背後にあるt分布および正規分布の考え方や、推定値の不確実性をどのように評価・表現するかについても解説します。

## R とRStudio の活用

### 1) R のダウンロードとインストール

統計解析を行うためのソフトとして, R と RStudio を使います. R は解析エンジンで, RStudio はその操作を便利にするための「作業画面」です. 両方をインストールして使います.

- 1. 以下の URL を開きます: https://cran.r-project.org
- 2. 上部のメニューから「Download R for (Windows / macOS / Linux)」をクリックします.
- Windows  $\neg \# \rightarrow \lceil \text{Download R for Windows} \rfloor$
- macOS  $\neg \# \rightarrow \neg$  Download R for macOS
- 3. 次に表示されるページで、インストーラー (.exe や.pkg ファイル) をダウンロードします.
- ダウンロードしたファイルをダブルクリックし、表示される指示にしたがってインストー ルを進めます。
- ・基本的に「次へ」や「同意する」をクリックしていけば OK です.

### 2) RStudio のダウンロードとインストール

- 1. 以下の URL を開きます: https://posit.co/download/rstudio-desktop/
- 2.「RStudio Desktop」の下にある「Download」ボタンをクリックします.
- 3. OS (Windows、macOS など) に応じたインストーラーをダウンロードします.
- 4. ダウンロードしたインストーラーを実行し、画面の指示にしたがってインストールします.

### 3) インストール後の確認

- R と RStudio のインストールが終わったら、まず RStudio を開いてみましょう.
  - 1. 画面が 4 つの領域に分かれたウィンドウが表示されれば成功です.
  - 2. 画面下の「Console (コンソール)」に1+1 と入力して Enter を押してみましょう.
  - 3. 正しく動作していれば, [1]の下に [1] 2 と表示されます.

### 4) よくある質問

Q.R と RStudio はどちらを先にインストールすればよいですか?

A. R を先に, RStudio を後にインストールしてください.

Q. RStudio だけでは使えないのですか?

A. RStudio は R が動作するための環境なので, R 本体が必要です.

#### 5) RStudio の基本的な使いかた



#### 図 1: RStudio の画面

RStudio は 4 つの画面 (ペイン) に分かれていて, 左上の「スクリプトペイン」にコードを 書きます. コードを実行するときは, 実行したい行を選択し, 上の「Run」ボタンをクリックす るか、Ctrl + Enter (Mac では Cmd + Enter)を押すと, 左下の「コンソールペイン」に実行結 果が表示されます. 右上には変数の一覧, 右下にはグラフやファイル一覧などが表示されます.

## セッション 1:オッズ比

オッズ比 (odds ratio, OR) とは、ある要因 (たとえば喫煙) の有無によって、特定の出来事 (たとえば病気) の起こりやすさに差があるかを比較するための指標です.2つのグループで「起 こる」「起こらない」の数からオッズとオッズ (起こった数 ÷ 起こらなかった数) を求め、それ らのオッズの比を計算することで、出来事の起こりやすさの違いを評価します.オッズ比が1で あれば、要因の有無による差はないことを意味します.1より大きければ、その要因がある方が 出来事が起こりやすく、1より小さければ起こりにくいことを示します.オッズ比は症例対照研 究でよく使われ、要因と結果の関連の強さを簡潔に表現する統計指標です.

#### 1) 確率とオッズの違い

私たちは日常的に「確率○○%」という表現に慣れています.たとえば、「ある病気にかかる確 率が 20%だった」というとき、これは 100 人のうち 20 人がその病気にかかったことを意味し ます.一方で、オッズ (odds) は、ある事象が起こった数と起こらなかった数の比(割合の比) として定義されます.つまり、

同じことを確率を使って表すと、次のようになります.

オッズ = 
$$\frac{b35 + b}{b35} = \frac{b}{1-p}$$

たとえば:

- 100人のうち 20人が病気(発症)した場合
   → 起こった数 = 20,起こらなかった数 = 80
   → オッズ = 20 ÷ 80 = 0.25
- ・同じ情報を確率で表すと

   → 起こる確率 = 0.2,起こらない確率 = 0.8
   → オッズ = 0.2 ÷ 0.8 = 0.25

このように、オッズは「数」でも「確率」でも同じ値になります. なぜなら、確率とは「全体に対する割合」であり、全体の人数は分子と分母に共通して打ち消されるからです.

### 2) 相対リスクとオッズ比

相対リスクとオッズ比は、どちらも「ある特徴(例:喫煙)があるかどうか」と「ある事象 (例:肺がん)が起きるかどうか」の関連の強さを表す指標ですが、この2つには意味や使い方 に違いがあります.



図 2: 相対リスクとオッズ比

#### 相対リスク

相対リスク(relative risk, RR)は、ある出来事が起こる確率(リスク)を2つのグループで 直接比べたものです.つまり、相対リスクは、ある特徴をもつ人たち(例:喫煙者)と、特徴を もたない人たち(例:非喫煙者)を観察し、それぞれのグループで実際に事象がどのくらい起こ ったか(発生率)を比べることで求めます.これは「直接的に」確率を比べる方法です.

図 2 の例では,ある特徴をもつ人の中で事象が起きたのが 3 人,起きなかったのが 17 人で, 発生率は 3 / (3 + 17) = 3/20.特徴をもたない人の発生率は 1 / (1 + 19) = 1/20. これらの比, すなわち 3/20 ÷ 1/20 = 3 が相対リスクです.「3 倍発生しやすい」というように,直感的に解 釈しやすいのが特徴です.

表 1: オッズ比と相対リスクの比較

指標	意味	解釈しやすさ	注意点
相対リスク	確率の比	直感的で分かりやすい	症例対照研究では使えない
オッズ比	オッズの比	やや分かりにくい	確率が高いときは相対リスクと大きくずれる

#### オッズ比

オッズ比は、すでに事象が起きた人と起きていない人に分けて、その人たちがどれくらい特徴 をもっていたかを調べます.これは「間接的に」確率を推定する方法で、主に過去を振り返る研 究(症例対照研究)で用いられます.図2の例では、「ある特徴有り」のうち「ある事象発生」 が3人、「ある事象未発生」が17人で、オッズは3/17.「ある特徴無し」では、「ある事象発生」 が1人、「ある事象未発生」が19人で、オッズは1/19. オッズ比は、3/17÷1/19=3×19/ 1×17 ≒ 3.35 となります.

事象の発生がまれな場合(発症率が低いとき)は、相対リスクとオッズ比は似た値になります が、事象が多くなると両者はかけ離れた値になることがあります。そのため、研究のデザインや 事象の頻度に応じて、どちらを使うべきかを考えることが重要です。

> オッズ比 = ある特徴を持つグループに,ある事象が起こるオッズ ある特徴を持たないグループに,ある事象が起こるオッズ

### 3) RStudio を活用した例題

ここでは、頭低位で撮影する腹臥位二重造影正面位像において椎骨の左側に現れる二重造影領 域の面積と撮影手技に関する研究例を取り上げます。「両肩当て左向き手技」が「左肩当て右向 き手技」と比べ、二重造影領域の面積がどのくらい広く現れるかを表現する指標として、オッズ 比を算出してみます。生データ(data.xlsx)は以下の URL からダウンロードしてください。

https://ixsen.com/wp-content/uploads/2025/07/data.xlsx

```
i) パッケージの導入と準備
```

```
パッケージのインストール:
```

```
install.packages("readxl", repos = "https://cloud.r-project.org")
install.packages("dplyr", repos = "https://cloud.r-project.org")
```

- 1. install.packages() とは、R に追加機能をインストールするコマンドです.
- 2. "readxl" は、Excel ファイル (.xls, .xlsx) を読み込むためのパッケージです.
- 3. 引数 repos で, パッケージをダウンロードする CRAN のミラーサイト (リポジトリ) を 指定しています.
- 4. "dplyr"は、データ操作(整形・抽出・集計など)を行うためのパッケージです.

#### 利用準備:

library(readxl)
library(dplyr)

- 1. library() は、パッケージを R に読み込んで追加機能を使うためのコマンドです.
- library()を実行する前に、一度だけ install.packages() でインストールしておく必要があります. R を再起動した場合や新しいセッションでは、再び library() を実行しないと使えません.

ii) データ読み込み

Excel ファイルの読み込み:

df1 <- read\_excel("data.xlsx")</pre>

- 1. 「<-」は読み込んだデータを df1 という名前のオブジェクト (データフレーム) に代入する ための代入演算子です. 以後, df1 を使ってデータの集計・分析・加工ができます.
- 2. read\_excel() は readxl パッケージに含まれている関数であり、ファイルの中身(表形式の データ)を R に読み込んでデータフレームに変換します.

 R で Excel ファイルや CSV ファイルなどのデータを読み込む場合,データファイルを R スクリプトと同じフォルダ (ディレクトリ) に保存しておくと便利です.同じフォルダに 置いておくことで,ファイル名だけで読み込むことができファイルの場所を探す必要がな くなります.

データの表示:

View(df1)

- 1. View() は, R においてデータフレームを表形式 (スプレッドシートのような形式) で開く 関数です.
- 2. view (小文字) と書くとエラーになります.

iii) データの前処理

二重造影像の面積率 (S 値) を中央値で二分する:

```
df2 <- df1 %>%
mutate(面積率 = if_else(S >= median(S), " 大きい", " 小さい"))
View(df2)
```

- 1.「%>%」はdplyr パッケージの追加機能であり、「そして」「次に」のような意味を持ち、コードを読みやすくつなげるために使います. つまり、dfl というデータフレームを、次の処理に渡す、という意味です.
- 2. mutate() は新しい列を作る関数で、今回は「面積率」という名前の列を df1 に追加しています.
- 3. if\_else(条件, "大きい", "小さい") は, 条件式が TRUE なら "大きい"、FALSE なら "小さい" を返します.
- 4. S >= median(S) は, S 列の値が, S の中央値以上なら TRUE、そうでなければ FALSE を 返します.
- 5. mutate() と if\_else() は, dplyr パッケージに含まれているため、事前に library(dplyr) が 必要です.

iv) 分割表の作成

両肩当て左向き手技とS値の大小:

```
cross1 <- addmargins(table(df2$両肩当て左向き,df2$面積率))
rownames(cross1)[rownames(cross1) == "Sum"] <- " 合計"
colnames(cross1)[colnames(cross1) == "Sum"] <- " 合計"
cross1 <- cross1[c(" 実施", " 未実施", " 合計"), c(" 大きい", " 小さい", " 合計")]
cross1
```

##

##		大きい	小さい	合計
##	実施	107	43	150
##	未実施	43	107	150
##	合計	150	150	300

- 1. table() は, 2 つの変数でクロス集計(分割表)を作ります. ここでは,「実施」「未実施」× 「大きい」「小さい」の表ができます.
- 2. addmargins() は, table() の結果に合計を追加し, 分割表の行・列に「Sum」の行・列を追加します.
- rownames()で、3つの行名("実施"・"未実施"・"Sum")を取り出し、各行名が "Sum" か どうかをチェックする条件式:rownames(cross1) == "Sum"で "Sum(英語)"という行名 だけ "合計(日本語)"に書き換える処理をしています。"Sum"のままで良ければ、省略し ても構いません。
- colnames() で、3 つの列名 ("大きい"・"小さい"・"Sum") を取り出し、各列名が "Sum" かどうかをチェックする条件式: colnames(cross1) == "Sum" で "Sum(英語)" という列 名だけ "合計 (日本語)" に書き換える処理をしています。"Sum"のままで良ければ、省略 しても構いません。
- 5. 「==(ダブルイクオール)」は, 条件式の左辺と右辺が等しいかどうかを比較し, 結果は TRUE か FALSE で返されます.「=(イクオール)」は「<-」と同じ代入演算子であり,「==」とは意 味が異なります.
- 6. table() や addmargins() の結果は、行や列の順序が自動で並んでしまうため、のちの計算 処理や報告書用に並べ替えたいことがあります. そこで、cross1[c("実施"、"未実施"、"合 計")、c("大きい"、"小さい"、"合計")] によって、順序を明示的に指定し、行と列の順番を再 構成しています.

v) オッズ比

オッズ比:

$$OR = \frac{a/b}{c/d} = \frac{a \cdot d}{b \cdot c}$$

値の代入とオッズ比の算出:

a <- cross1[" 実施", " 大きい"] b <- cross1[" 実施", " 小さい"] c <- cross1[" 未実施", " 大きい"] d <- cross1[" 未実施", " 小さい"] OR <- (a \* d) / (b \* c) OR

## [1] 6.191996

- 1. オブジェクト [文字列 A, 文字列 B] の […] の中は, 行と列を指定するためのインデックス です. ここでは, 分割表 cross1 の 1 行 1 列目 ("実施"かつ"大きい")を a に代入, 1 行 2 列目を b に代入, 2 行 1 列目を c に代入, 2 行 2 列目を d に代入しています.
- 2. OR <- (a \* d) / (b \* c) コマンドでオッズ比を計算し, OR コマンドでオッズ比を表示します.

### セッション 2:フィッシャーの正確確率検定

フィッシャーの正確確率検定は、2つのカテゴリ変数(たとえば「薬を飲んだ/飲まない」と 「症状が改善した/しなかった」)の関係を調べるための検定方法です。特に、データの件数が少 ないときに有効です。通常の $\chi^2$ (カイ二乗)検定では、大きなサンプルサイズを前提としてい ますが、フィッシャー検定はすべての可能な結果の組み合わせを計算し、「今のデータが偶然に 出る確率(p値)」を正確に求めます。これにより、「2つの変数に関連があるかどうか」を判断 できます。R では fisher.test() 関数で簡単に実行できます。

#### 1) 独立性の検定

フィッシャーの正確確率検定と x<sup>2</sup>(カイ二乗)検定は、いずれも二つの変数のあいだに関連 があるかどうかを調べるための方法です。特に、2×2の分割表で使われることが多く、どちらも 「独立性の検定」として分類されますが、その考え方と適用条件には重要な違いがあります。

x<sup>2</sup>検定は、観測されたデータと、もし二つの変数が独立であると仮定したときに期待される データ(期待値)との差をもとに、その差が偶然とは思えないほど大きいかどうかを評価する方 法です.「もし独立ならこうなるはず」という期待値とのズレを調べる検定、と言い換えること もできます. 一方、フィッシャーの正確確率検定は、その名のとおり、期待値を用いずに、分 割表で観測された値と同じ行・列の合計をもつすべての 2×2 表を列挙して、観測された表と同 じかそれ以上に極端な結果が出る確率を正確に計算する方法です.このとき、確率は超幾何分布 にもとづいて求められます.サンプル数が少ない場合でも信頼できる結果が得られるため、小標 本のときや、ある特定の効果があるかどうかを片側検定で調べたいときに有効です.

つまり、 x<sup>2</sup> 検定は大まかな目安を知りたいときに便利で、サンプルサイズが十分に大きいと きに適しています.一方、フィッシャーの正確確率検定は、小さいデータでも正確な結果を得た いときに使われます.分析の目的やデータの規模に応じて、これらを適切に使い分けることが大 切です.

#### 2) x2 検定における期待値

たとえば肺癌と喫煙の関係を調べたとき、表2のような結果になったとします.このとき、喫 煙と肺癌が無関係(独立)なのかを  $\chi$  2検定で検討するには、観測された分割表の値(観測値: 表 2)と、喫煙と肺癌が独立だと仮定したときに期待される値(期待値:表 3)とを比較します. 期待値は「行の合計 × 列の合計 ÷ 全体の合計」で計算します.「喫煙者かつ肺癌」の期待値は、 喫煙者の合計と肺癌の合計を掛けて全体で割ったものです.観測値と期待値のあいだに大きなず れがあれば、喫煙と肺癌のあいだに独立でない関係がある可能性が示唆されます.

ただし, χ<sup>2</sup> 検定は,ある程度十分なデータがあることを前提にした近似的な検定方法です. 特に,期待値が5より小さいセルがあると,近似がうまく働かず,検定の結果が不正確になるお それがあります.このような場合には,フィッシャーの正確確率検定がよいとされています.

表 2: 観測値

	肺	癌	
喫煙	あり	なし	行合計
あり	3	17	20
なし	1	19	20
列合計	4	36	40

表 3: 期待値(独立の仮定のもと)

喫煙	あり	なし	行合計
あり	2	18	20
なし 列合計	$2 \\ 4$	18 36	20 40

### 3) フィッシャーの正確確率検定のしくみ

フィッシャーの正確確率検定は、2×2 分割表において「行と列の合計(マージン)が固定され た条件下で、観測されたデータが偶然に生じたものといえるかどうか」を正確な確率に基づいて 評価する方法です.ここでは、胃憩室と胃底腺ポリープの有無について行った単純な調査例(図 3)を用い、検定のしくみを順を追って説明します.

		J	リスト	· (5	ĭΞ·	—)						2×2 5	〉割表					
	氏	名	年齢	性	穹窿部 限局性 胃丸	後壁の 突出像 夏 <b>室</b>	<sup>穹窿部</sup> の小さ <b>胃底腹</b>	and/or 体部 な輪状影 <b>えポリープ</b>										
No1	ОШ	O男	59	男	あ	り	な	L			<b>.</b> _ <b>.</b> = <b>.</b> _ <b>.</b> _	胃底腺术	ピリープ					
No2	ОШ	0子	34	女	な	し	あ	り	_					<b>肖</b> 想至		あり	なし	
No3	〇島	O美	64	女	あ	b	あ	り		あ	り	1(33.3%)	2 (66.7%)	3				
No4	O⊞	〇郎	48	男	な	L	あ	り		な	し	$2^{(66.7\%)}$	1 (33.3%)	3				
No5	O尾	0子	52	女	あ	b	な	し					_					
No6	O谷	〇司	53	男	な	し	な	L		合	計	3	3	6				

図 3: 結果を集計し分割表を作成する

i) 分割表のセル値の全パターンを列挙する

最初に行うのは、行と列の合計(マージン)を固定したまま、分割表の中に入りうるセル値の すべてのパターンを洗い出すことです。今回の例では、「胃底腺ポリープの有無」と「胃憩室の 有無」の2つのカテゴリからなる2×2分割表(図4左)を用います。いずれのカテゴリも、各群 3名ずつ、合計6名のデータがあり、表の各行・列の合計はすでに決まっています。

このとき重要なことは、一つのセル(たとえば左上の a)の値が決まれば、他の 3 つのセルの 値は行と列の合計に従って自動的に決まる、という点です。つまり、a がとりうる値は 0~3 の 4 通りで、それぞれに対応するパターン(A~D)が決まります (図 4 右).

- ・パターン A: a=0
- ・パターン B: a = 1
- ・パターン C: a=2
- ・パターン D: a = 3

このように、固定された行と列の合計のもとで考えられるすべての分割表パターンを明示的に 列挙することが、フィッシャー検定のしくみを理解するための最初のステップです.ここで、実際に観測されたデータはパターン B となります.



図 4: 分割表の行列の合計を満たすようなセル (区画) の値の全パターンを洗い出す

ii) 赤玉・白玉の抽選モデル

つぎに行うのは,パターン A~D のそれぞれがどれほどの確率で生じるかを計算することで すが,その前に,赤玉・白玉の入った箱から無作為に玉を取り出すという「抽選」(図 5)を想定 すると良いでしょう.玉の選ばれかたの確率は,もともと箱に入っている玉の色と数によって決 まるからです.

観測値であるパターン B について見てみましょう.これは、胃憩室がある 3 名と胃憩室がない 3 名のうち、胃底腺ポリープがあった例 3 名中 1 名に胃憩室があったパターンです.この状況は、赤玉 3 個と白玉 3 個の計 6 個の玉が箱に入っており、その中から 3 個の玉を取り出したとき、1 個が赤玉である場合の確率を考えることと同じです.取り出した玉の 1 個が赤玉のとき必ず白玉は 2 個になりますし、箱のなかには赤玉が 2 個、白玉が 1 個、必ず残ります.この抽選モデルでも、ひとつのセルの値が決まれば他の 3 つのセルの値は自動的に決まる、ことを実感できます.

## パターンBの分割表が得られる確率

		箱から取り出した玉	箱のなかに残った玉	もともとの箱のなかの玉
赤	玉	<b>4</b> •	2 • •	3 • • •
白	玉	$2$ $\bigcirc$	1 〇	$3 \circ \circ \circ$
		取りと	出した後	取り出す前

### は、赤玉・白玉の取り出し試行確率の求めかたと同じ

図 5: 赤玉・白玉の抽選モデル

確率とは、ある事象が起こる「可能性の大きさ」を表す数値で、通常 0 以上 1 以下の値で表 されます、確率の定義は、起こりうるすべての場合の数(分母)のうち、その事象が起こる場合 の数(分子)の割合です。このときの分母は 6 個の玉の中から 3 個を取り出す場合の組合せの 数です。一方、分子は、赤玉 3 個のなかから 1 個を選ぶ組合せの数と白玉 3 個のなかから 2 個 を選ぶ組合せの数の積(かけ算)になります (図 6).

ここで, 分子がなぜ「赤玉3個の中から1個を選ぶ組合せの数」と「白玉3個の中から2個 を選ぶ組合せの数」の積になるのかを考えてみましょう。この確率は、「赤玉が1個、白玉が2 個」という特定の組み合わせが選ばれる場合の数を数えています.つまり、「赤1・白2」という 構成が何通りの選び方で実現されるかを数える必要があります。

赤玉が3個,白玉も3個あるとき,それぞれの玉には識別できるラベル(「赤 a」「赤 b」「赤 c」「白 d」「白 e」「白 f」)がついていると考えます.これは,「同じ赤玉」でも区別できる個体と して扱うという前提があるからです。このとき、「赤玉を1個、白玉を2個」選ぶ方法は、

- 赤玉3個から1個を選ぶ組合せ(3通り)と、
- 白玉3個から2個を選ぶ組合せ(3通り)

のすべての組み合わせの対が考えられます。つまり、

 $3 \times 3 = 9$ 

通りある, ということになります. 以上のことから, 分母は 20, 分子は 9 となり, パターン B の観測値が得られる確率は 0.45 と算出されます。



#### -ンBの分割表が得られる確率 パタ-

図 6: 確率の求めかた

各パターンの確率は次のように求めます. 分子は赤玉から a 個, 白玉から 3 - a 個を選ぶ組合せの数で, 分母は 6 個から 3 個を選ぶ全体の組合せ数です:

- $\mathscr{N}\mathscr{P} \mathscr{V} \wedge (a = 0)$ :  $\frac{\binom{3}{3} \times \binom{3}{0}}{\binom{6}{2}} = \frac{1 \times 1}{20} = 0.05$
- ・ パターン B (a = 1):

$$\frac{\binom{3}{2} \times \binom{3}{1}}{\binom{6}{3}} = \frac{3 \times 3}{20} = 0.45$$

・ パターン C (a = 2):

$$\frac{\binom{3}{1} \times \binom{3}{2}}{\binom{6}{3}} = \frac{3 \times 3}{20} = 0.45$$

• パターン D (a = 3):

$$\frac{\binom{3}{0} \times \binom{3}{3}}{\binom{6}{3}} = \frac{1 \times 1}{20} = 0.05$$

このようにして,全体で起こりうるすべての分割表に対し,その出現確率を組合せの理論に基づいて求めることができました。すべてのパターンに対応する確率を計算し,それらを足し合わせると1になります.これは,全ての可能な組合せを網羅していることを意味しています.

#### iii) 観測値とそれ以上に極端なパターンの確率を合計する (p 値の計算)

最後に行うのは,観測された分割表が,どれほど「珍しい」かを表す p 値の計算です. ここで の p 値とは,「観測されたパターンと同程度,またはそれ以上に極端なパターンが起こる確率の 合計」です. 今回の観測値はパターン B (a = 2) で,その確率は 0.45 でした. この値と同じ かそれより小さい確率を持つパターンをすべて合計します. パターン B が得られる確率ではな いことに注意してください.

- パターンA (0.05):Bより小さい
- ・ パターン B (0.45): 観測値
- パターン C (0.45):同じ確率
- パターン D (0.05): B より小さい

したがって, p 値は

$$0.05 + 0.45 + 0.45 + 0.05 = 1.00$$

これは, 観測されたような分割表がまったく珍しくないことを意味しており, このデータを通 常の有意水準(たとえば 0.05)と比較すれば,「胃憩室と胃底腺ポリープに関連がある」とはと てもいえないという結論になります.

#### 4) RStudio を活用した例題

ここでは、咽頭癌の X 線写真のスケッチ経験の前後で、咽頭部の「描出範囲」と「造影効果」に 変化があったかどうか、すなわちスケッチ学習という介入と写真の画質との関係を分析するため に、フィッシャーの正確確率検定を用いて p 値を算出してみます. 生データ(X\_ray\_4040.xlsx) は以下の URL からダウンロードしてください.

```
https://ixsen.com/wp-content/uploads/2025/07/X_ray_4040.xlsx
```

```
i) パッケージの導入と準備
```

パッケージのインストール:

```
install.packages("readxl", repos = "https://cloud.r-project.org")
install.packages("dplyr", repos = "https://cloud.r-project.org")
```

- 1. install.packages() とは、R に追加機能をインストールするコマンドです。
- 2. "readxl"は、Excelファイル(.xls, .xlsx)を読み込むためのパッケージです.
- 3. 引数 repos で, パッケージをダウンロードする CRAN のミラーサイト (リポジトリ) を 指定しています.
- 4. "dplyr"は、データ操作(整形・抽出・集計など)を行うためのパッケージです.

#### 利用準備:

library(readxl)
library(dplyr)

- 1. library() は、パッケージを R に読み込んで追加機能を使うためのコマンドです.
- library()を実行する前に、一度だけ install.packages() でインストールしておく必要があります. R を再起動した場合や新しいセッションでは、再び library() を実行しないと使えません.

ii) データ読み込み

Excel ファイルの読み込み:

df <- read\_excel("X\_ray\_4040.xlsx")</pre>

- 1.「<-」は読み込んだデータを df1 という名前のオブジェクト (データフレーム) に代入する ための代入演算子です.以後, df1 を使ってデータの集計・分析・加工ができます.
- 2. read\_excel() は readxl パッケージに含まれている関数であり、ファイルの中身(表形式の データ)を R に読み込んでデータフレームに変換します.

 R で Excel ファイルや CSV ファイルなどのデータを読み込む場合,データファイルを R スクリプトと同じフォルダ (ディレクトリ) に保存しておくと便利です.同じフォルダに 置いておくことで,ファイル名だけで読み込むことができファイルの場所を探す必要がな くなります.

```
iii) データの前処理
```

```
日本語化と順序化:
df_jp <- df %>%
  mutate(
    <sup>*</sup>介入の段階<sup>*</sup> = factor(
     recode(phase, "before" = " スケッチ前", "after" = " スケッチ後"),
     levels = c(" スケッチ後", " スケッチ前")
    ),
    <sup>*</sup>造影効果<sup>*</sup> = factor(
     recode(as.character(contrast), "1" = " 良", "0" = " 不良"),
     levels = c(" 良", " 不良")
    ),
    `描出範囲` = factor(
     recode(as.character(depiction), "1" = " 良", "0" = " 不良"),
     levels = c(" 良", " 不良")
    )
 )
```

- 1.「%>%」はdplyr パッケージの追加機能であり、「そして」「次に」のような意味を持ち、コードを読みやすくつなげるために使います.つまり、df というデータフレームを、次の処理に渡す、という意味です.
- 2. mutate() は新しい列を作る関数で、ここでは「介入の段階」・「造影効果」・「描出範囲」という3つの名前の列を一度に df に追加しています.
- 3. mutate() は, dplyr パッケージに含まれているため、事前に library(dplyr) が必要です.
- 4. recode() は, 値の置き換えを行う関数で, 読み込んだ Excel データの"before" を "スケッ チ前" に、"after" を "スケッチ後" に変換しています.
- 5. recode(as.character(…)) は、元の変数 (contrast, depiction) が数値 (0/1) なので、これ を文字列に変換したあと、"0"を"不良"、"1"を"良"に変換するためのコードです.
- 6. factor(…, levels = c("スケッチ後", "スケッチ前"))は、順序を明示することで、グラフや表の並び順や検定を行う際の基準を定めるためのコードです。ここでは、"スケッチ後"→ "スケッチ前"の順番で並べるように指定しています。

iv) 分割表の作成

スケッチ前後の造影効果の変化:

```
cross1 <- addmargins(table(df_jp$介入の段階, df_jp$造影効果))
rownames(cross1)[rownames(cross1) == "Sum"] <- " 合計"
colnames(cross1)[colnames(cross1) == "Sum"] <- " 合計"
cross1
```

##

##		良	不良	合計
##	スケッチ後	36	4	40
##	スケッチ前	19	21	40
##	合計	55	25	80

- 1. table() は、2 つの変数で分割表を作ります. ここでは、「スケッチ後」「スケッチ前」× 「良」「不良」の表ができます。
- 2. addmargins() は, table() の結果に合計を追加し, 分割表の行・列に「Sum」の行・列を追 加します.
- 3. rownames() で, 3 つの行名 ("スケッチ後"・"スケッチ前"・"Sum") を取り出し, 各行名 が "Sum" かどうかをチェックする条件式: rownames(cross1) == "Sum" で "Sum(英語)" という行名だけ "合計 (日本語)" に書き換える処理をしています. "Sum" のままで良けれ ば、省略しても構いません.
- 4. colnames() で、3 つの列名 ("良"・"不良"・"Sum") を取り出し、各列名が "Sum" かどう かをチェックする条件式: colnames(cross1) == "Sum" で "Sum(英語)" という列名だけ "合計(日本語)"に書き換える処理をしています。

スケッチ前後の描出範囲の変化:

```
cross2 <- addmargins(table(df_jp$介入の段階, df_jp$描出範囲))
rownames(cross2)[rownames(cross2) == "Sum"] <- " 合計"
colnames(cross2)[colnames(cross2) == "Sum"] <- " 合計"
cross2
```

40

40

80

## 良 不良 合計 ## スケッチ後 35 5 ## スケッチ前 22 ## 18 合計 57 ## 23

1. スケッチ前後の造影効果の変化の分割表を cross1 に格納したので, 描出範囲の変化の分

割表は cross2 に格納しています.

2. cross1 や cross2 のような変数名は、自分がわかりやすい名前であれば任意のもので構い ません.

v) Fisher の正確確率検定

スケッチ前後の造影効果の変化:

```
fisher.test(cross1)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: cross1
## p-value = 0.001542
## alternative hypothesis: two.sided
```

- 1. fisher.test() は、特にデータ数が少ない場合や、セルの期待度数が小さい場合に適した検定 方法であり、分割表に基づいて2つのカテゴリ変数に関連性(独立性)があるかどうかを 検定します.
- 2. データ数が多く分割表の期待度数が5以上であれば、同じ独立性検定であるカイ二乗検定 も利用可能です.
- 3. R でカイ二乗検定を行う場合は、chisq.test(cross1) と入力し実行してください.

スケッチ前後の描出範囲の変化:

fisher.test(cross2)

```
##
## Fisher's Exact Test for Count Data
##
## data: cross2
## p-value = 0.03286
## alternative hypothesis: two.sided
```

1. R でカイ二乗検定を行う場合は、chisq.test(cross2) と入力し実行してください.

2. 造影効果と描出範囲の評価が互いに影響し合っている場合 (交絡)には、介入の段階と造影効果あるいは介入の段階と描出範囲の単変量解析であるフィッシャーの正確確率検定では、真の関係を評価できません。そこで、目的変数をスケッチの段階 (前/後)、説明変数を造影効果と描出範囲とした多変量解析 (ロジスティック回帰モデル)を行えば、両者の関係を同時に評価できることになります。ただしこのモデルは「造影効果や描出範囲の情報から、ス

ケッチ前か後かを予測する」という意味ですから、一般的な解釈とは逆になります.した がって、ロジスティック回帰で得られた結果は、変数間の統計的な関連の強さの指標であ ることを理解しておく必要があります.興味があれば、以下のコードを実行してください.

ロジスティック回帰分析のコード:

```
df_jp <- df_jp %>%
  mutate(造影効果 = relevel(造影効果, ref = " 不良"),
        描出範囲 = relevel(描出範囲, ref = " 不良"))
model <- glm(`介入の段階` ~ 造影効果 + 描出範囲,
            data = df_jp, family = binomial)
result <- summary(model)$coefficients[-1, ]</pre>
or <- exp(coef(model))[-1]</pre>
ci <- exp(confint.default(model))[-1, ]</pre>
result table <- data.frame(</pre>
  項目 = rownames(result),
 P 值 = signif(result[, "Pr(>|z|)"], 3),
  オッズ比 = round(or, 3),
  `95%CI 下限` = round(ci[, 1], 3),
  `95%CI 上限` = round(ci[, 2], 3)
)
rownames(result_table) <- NULL</pre>
print(result_table)
##
          項目
                    P値 オッズ比 X95.CI下限 X95.CI上限
## 1 造影効果良 0.000963 0.123
                                     0.035
                                                0.427
## 2 描出範囲良 0.019000 0.233
                                     0.069
                                                0.787
```

## セッション 3:95% 信頼区間

95%信頼区間とは,統計学的な推定において「本当の値がこの範囲にあるだろう」と言える区間のことです. 信頼区間は,平均だけでは伝えきれない「どれくらい確かか」を示す重要な指標であり,研究結果の信頼性を表現したり判断したりするためにも欠かせません.

### 1) 点推定と区間推定

統計学における推定には、大きく分けて「点推定」と「区間推定」という2つの方法がありま す. これらはいずれも、調査や実験から得られたデータをもとに、母集団の性質を推しはかるた めの方法 (図 7) であり、X 線所見から肉眼所見や組織所見を推定することとよく似ています.



母平均	母分散	標本平均	不偏分散
母相関係数	母回帰係数	標本相関係数	標本回帰係数
母比率	母オッズ比	標本比率	標本オッズ比
母中央値	母相対リスク	標本中央値	標本相対リスク

図 7: 部分から全体を推定する

点推定とは、母集団のある特性(たとえば平均や割合)を、ひとつの値で推定する方法です. たとえば、ある撮影法によって検査した 100人に着目し、その効果を示す数値データを取得し たとしましょう.その効果の平均が「50」であれば、「母集団における効果の真の平均も 50 で あろう」と考えるのが点推定です.この方法はシンプルでわかりやすく、調査結果を要約するの に便利です.しかし、実際のデータにはばらつきがあるため、ひとつの値だけで「本当の値」を 言い切るのは難しく,どのくらい正確なのかがわかりにくいという弱点があります. そこで登 場するのが,区間推定です.これは「本当の値(母数)はこの範囲にあるだろう」と,ある幅を もった区間で推定する方法です.さきほどの例で言えば,「効果の平均は 50 だが,その 95% 信 頼区間は 45 から 55 である」といったように,数値の範囲を示します.区間推定では,推定値 の「信頼性」をより正確に伝えることができますから,医療や社会調査など,結果の信頼性が重 視される分野では,点推定だけでなく信頼区間を併せて報告することが一般的です.



図 8: 信頼区間の意味・定義 -BMI と走行距離の調査をふまえて-

### 2) 信頼区間の意味・定義

区間推定における「信頼水準」は、95% や 99% といった形で示されます.これは、「同じ方法 で何度も調査を行えば、95% の割合で真の値を含む区間が得られる」という意味です.注意す べき点は、特定の信頼区間が「真の値を含む確率が 95% である」という意味ではなく、「繰り 返したときの頻度」として解釈するのが正しい理解です.また、信頼区間の幅は、データのばら つきやサンプルサイズに影響されます.標本の数が少なかったり、ばらつきが大きかったりする と、信頼区間は広くなり、推定の不確かさが大きいことを意味します.逆に標本数が多く、ばら つきが小さければ、信頼区間は狭くなり、より正確な推定ができていると判断できます.

このように、点推定は「中心的な値」を、区間推定は「その値の信頼性」を示すものであり、両 者を併せて理解することが大切です。図8には次節で取り上げる「BMIと走行距離の調査」を ふまえて、信頼区間の意味・定義を示します。統計学的推定では、平均や割合といった数値を示 すだけでなく、その背景にある不確かさを信頼区間として表すことで、誠実で信頼性の高い情報 提供が可能になります。

### 3) 信頼区間の算出のしくみ

#### Wald 法 (ウォルド法):

標本平均などの点推定値は、母集団の中から一部を抽出した結果ですから、調査のたびに少し ずつ異なる値になります.このばらつきを見積もるために、「標準誤差 (Standard Error)」とい う指標が使われます.これは、推定値がどれくらいブレるかを数値化したもので、標本の標準偏 差をサンプルサイズの平方根で割ることで求められます.



図 9: 中心極限定理

多くの場合,点推定値は「正規分布」に従うと考えられています.とくにサンプルサイズが 大きければ,中心極限定理(図9)により,たとえもとのデータが非正規分布でも,平均や割合な どの推定値は正規分布に近づく性質があります.この性質を利用して,信頼区間は次のように計 算されます.

#### 推定值 ± 1.96 × 標準誤差

この式は,標準正規分布において約 95% の確率でデータが ±1.96 の範囲に収まるという事実 に基づいています (図 10). この方法は Wald 法(ウォルド法)と呼ばれ,信頼区間の計算に広 く用いられています.



図 10: 標準正規分布における 95% 信頼区間

### オッズ比の信頼区間と Woolf 法:

オッズ比(odds ratio)の場合も、信頼区間を算出することで、推定の信頼性を評価することができます.オッズ比は以下の式で求められます:

$$OR = \frac{a/b}{c/d} = \frac{a \cdot d}{b \cdot c}$$

ただし, a, b, c, d は 2×2 の分割表におけるセルの度数です. オッズ比は 0 以上の正の数値 をとり,分布が非対称(右に長い)になるため,そのままでは正規分布とみなせません. そこで, オッズ比の対数(log)をとることで,その分布を正規分布に近づける工夫がされます. これが対 数オッズ比(log odds ratio)です:

$$\log(\mathrm{OR}) = \log\left(\frac{a \cdot d}{b \cdot c}\right)$$

この対数オッズ比に対する標準誤差は、以下の式で近似されます:

$$SE\left(\log(\mathrm{OR})\right) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

この標準誤差と正規分布をもとに信頼区間を計算する方法が,先ほどの Wald 法の応用であ り,オッズ比の推定にも適用されます.この計算法には特別な名前があり,Woolf 法(ウルフ 法)と呼ばれています.Woolf 法は,オッズ比の対数変換とその標準誤差に基づいて信頼区間を 算出する方法で,とくに症例数が中程度以上ある場合に有効です.

Woolf 法に基づく 95% 信頼区間は以下の手順で求めます.

1. オッズ比 OR を計算

- 2. log(OR) を求める
- 3. 標準誤差 SE を求める
- 4. 信頼区間 (log スケール) を計算

 $\log(OR) \pm 1.96 \times SE$ 

5. 指数変換 exp() を使ってもとのスケールに戻す

 $\exp(\log(OR) \pm 1.96 \times SE)$ 

この方法によって、オッズ比の 95% 信頼区間 [OR lower, OR upper]が得られます。実務で はこの Woolf 法がもっとも一般的であり、医学・疫学・社会調査など多くの分野で用いられて います。

### 限界と注意点:

Woolf 法にも注意すべき点があります. とくに, a, b, c, d のいずれかが小さい(とくに5 未満)場合,この近似は不安定になり,信頼区間が適切でなくなる可能性があります. このよう な場合には、フィッシャーの正確確率検定など、より厳密な方法が使われることもあります. こ のように,信頼区間の算出は推定値の分布や性質を理解することが不可欠で,とくにオッズ比の ような非対称な統計量では,対数変換と Woolf 法が重要な役割を果たします. 次節では,実際に R と RStudio を使い,この手順にしたがってオッズ比と 95% 信頼区間を計算してみましょう.

### 4) RStudio を活用した例題

ここでは A 製薬東京営業所員の体格と月別走行距離のデータ (図 11) を用いて,低 BMI 群の 走行距離が短距離群であるオッズを基準とし高 BMI 群の走行距離が短距離群であるオッズの倍 率,つまりオッズ比とオッズ比の 95% 信頼区間を算出してみます.



図 11: 営業所員の体格と月別走行距離の関係を示した散布図 (●:高 BMI, ○:低 BMI)

### i) データ入力と分割表の作成

データ入力:

- a <- 17
- b <- 6
- c <- 11
- d <- 20
  - 1. 集計後の値(度数)が判明している場合には、それぞれの値をもとに 2×2 の分割表をつく ることができます.
  - 2. a~d は以下のような意味です.
  - ・a:高BMIかつ短距離走行の人数
  - ・ b:高 BMI かつ長距離走行の人数
  - ・ c:低 BMI かつ短距離走行の人数
  - ・ d: 低 BMI かつ長距離走行の人数

- 3. オッズ比を算出する際は、度数の配置に重要な意味があります. ここでのオッズ比は、「低 BMI 群」に比べて「高 BMI 群」は、どれくらい短距離走行である傾向が強いか?を表す 指標になります.
- OR = 1: BMI 群と走行距離には関係がない
- OR > 1:高 BMI 群の方が短距離走行である傾向が強い
- OR < 1:高 BMI 群の方が短距離走行である傾向が弱い

BMI と走行距離の分割表:

cross

##		短距離群	長距離群
##	高BMI群	17	6
##	低BMI群	11	20

- matrix() は、c(a,b,c,d) で与えられたベクトルから行列(マトリクス)を作成する関数です. 引数nrow = 2 で 2 行の行列を作る指定をし、byrow = TRUE で行単位でデータを配置す るよう指定しています.
- dimnames = list(…) で行名と列名を付けます. c("高 BMI 群", "低 BMI 群") で行 (BMI 群) に名前を付け, c("短距離群", "長距離群") で列 (走行距離群) に名前を付けるよう指定しています.

ii) オッズ比と対数オッズ比の標準誤差

オッズ比:

(再掲) OR = 
$$\frac{a/b}{c/d} = \frac{a \cdot d}{b \cdot c}$$

OR <- (a \* d) / (b \* c) OR

## [1] 5.151515

対数オッズ比:

(再掲) 
$$\log(OR) = \log\left(\frac{a \cdot d}{b \cdot c}\right)$$

log\_OR <- log(OR)
log\_OR</pre>

## [1] 1.639291

- 1. OR <- (a \* d) / (b \* c) コマンドでオッズ比を計算し, OR コマンドでオッズ比を表示します.
- log() は、自然対数(底が e の対数)をとる関数です.オッズ比そのものは 0~∞ の値であ り分布が歪んでいますが、その対数の範囲は -∞~∞ になり正規分布に近づくので推定や 検定に使いやすくなります.

対数オッズ比の標準誤差:

(再揭) 
$$SE(\log(OR)) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

SE\_log\_OR <- sqrt(1/a + 1/b + 1/c + 1/d) SE\_log\_OR

## [1] 0.6053092

- 1. 対数オッズ比の分散は、1/a + 1/b + 1/c + 1/d で近似することができます.
- 2. 標準誤差とは, 推定統計量(平均・割合・オッズ比など)のブレ具合のことで, 分散の平方 根で求めることができます. R ではsqrt()を利用します.

iii) 95% 信頼区間

95% 信頼区間の算出:

(再掲) 
$$\log(OR) \pm 1.96 \times \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$
  

$$\Rightarrow OR_{lower}, OR_{upper} = \exp\left(\log(OR) \pm 1.96 \times \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}\right)$$

CI\_lower <- exp(log\_OR - 1.96 \* SE\_log\_OR) CI\_upper <- exp(log\_OR + 1.96 \* SE\_log\_OR)

- 1. 標準正規分布では約 95% の値が平均 ±1.96 の範囲に収まるので, 推定値の周囲 ±1.96× 標準誤差の範囲を信頼区間として採用します.
- 2. 対数オッズ比をexp()\_ 指数関数で、もとのオッズ比に書き戻します.

オッズ比と 95% 信頼区間の表示:

cat(" オッズ比と 95% 信頼区間 (外れ値を除く 54 例) =",
 round(OR, 3), "[", round(CI\_lower, 3), ",", round(CI\_upper, 3), "]\n")

## オッズ比と95%信頼区間(外れ値を除く54例) = 5.152 [ 1.573 , 16.873 ]

# 索引

addmargins, 9, 19 as.character, 18 byrow, 28 chisq.test, 20 colnames, 9, 19 dimnames, 28 dplyr, 7, 8, 17, 18 Excel, 7, 17 exp, 30 factor, 18 FALSE, 8 fisher.test, 11, 20 Fisher の正確確率検定, 1, 11, 20 if\_else, 8 install.packages, 7, 17 levels, 18 library, 7, 17 list, 28 log, 29 matrix, 28 median, 8 mutate, 8, 18 nrow, 28 p值,11 R, 2 read\_excel, 7, 17 readxl, 7, 17 recode, 18

rownames, 9, 19 RStudio, 2 sqrt, 29 table, 9, 19 TRUE, 8 View, 8 Wald 法, 24 Woolf 法, 26 イクオール,9 オッズ,1,4 オッズ比, 1, 4, 27, 28 カイ二乗検定,11,20 カテゴリ変数,11 クロス集計,9 スプレッドシート,8 ダブルイクオール,9 データフレーム, 7, 8, 17 パッケージ,7 フィッシャーの正確確率検定,11 マトリクス,28 ロジスティック回帰,21 中心極限定理,24 交絡,20 介入,17 代入演算子, 7, 17 信頼区間, 1, 22 信頼性,22 信頼水準,23 分割表,9

分散,29

割合,22 区間推定,22 単変量解析,20 基準,18 多変量解析,20 対数オッズ比,28,29 対数変換,26 平均,22 度数,27 引数,28 指数変換,26 推定統計量,29 有意水準,16 期待值,11 期待度数,20 標準正規分布,24 標準誤差, 24, 28, 29 正規分布,24 母数,23 母集団,22 点推定,22 独立性の検定,11 症例対照研究,6 目的変数,20 相対リスク,5 確率,4 観測値,11 説明変数,20 超幾何分布,11

# いっしょに学ぶ統計学的解析の基礎 「0 と 1 のあいだ」

\_\_\_\_\_

2025 年	7月	1	日	初	版	発	行	
2025 年	7月	14	日	第二	版	発	行	
著者	書音	田諸	灾	柗□	山純也	J	數納優利	之
企画協力	J 富	樫聖	子	佐藤	泰清二		安藤健-	
運営協力	」藤	原	猛	丸山	山福乡	Ē		
発行所	馬場	塾オ	ンラ	っくこ	/			
URL	http	s:/	/ixs	en.o	com/			

