

# 観察研究における感度分析の勧め 実践編

・・・観察研究に基づく意思決定に関わる全ての方へ・・・

日本製薬工業協会  
データサイエンス部会  
2016 年タスクフォース 2

2019 年 4 月

## 初めに

観察研究においてバイアスへの対処は重要です．このため、種々の研究のデザインや解析方法がもちいられています．本書であつかう内容は解析方法にあたりますが、層別解析やモデルをもちいた解析などの、いわゆる「調整解析」がテーマではありません．むしろ、こうした解析では調整が難しい、あるいは不可能なバイアスの影響を評価するために行われる定量的バイアス分析（quantitative bias analysis）がテーマです．こうした用語に馴染みがない読者も少なくないことを考慮して、本書では、こうした方法を感度分析（sensitivity analysis）と呼んでいます．

本書の[入門編](#)（2017 年 10 月）では、とりあえず感度分析を行ってみること、その意義や可能性を大まかに理解することを重視していました．一方、この実践編では、感度分析に対する理解をさらに深めるとともに、実際に行えるようになることを目指しています．とはいえ、こうした手法には、使ってみることで理解が深まり、さらにくわしい内容を学びたくなるといった側面もあります．このため、この実践編でも、読み進める順序によっては「とりあえず行ってみる」こともできるよう、全体を 2 部にわけました．読者の興味や関心におうじて第 1 部から第 2 部へ、また第 2 部から第 1 部へと自在に行き来することで、感度分析に対する理解が深まり、適切な運用につながることを願っています．

また実践編でも、必要な計算を「とりあえず行ってみる」ことができるツールを用意しました．入門編でも紹介した Lash 等の教科書（入門編：REFERENCE 14）でもちいられている Excel シートを参考に、新たに作成した計算ツール ESATJ がそれです．とくに本書の第 2 部で、このツールをもちいて、感度分析の方法や計算例を提示していきます．必要に応じ、本書を公開している日本製薬工業協会の WEB ページから ESATJ を入手してください．

## 本書であつかう範囲

[入門編](#)では、観察研究における感度分析の意味や意義を把握していただくことを重視し、なるべく数式はもちいず、感度分析の大まかな手順と基本的な考え方について紹介しました。

この実践編では、入門編ですでに紹介した方法についても、重複をいとわず、その理論的な背景を紹介しました。例えば、感度分析のモデルを導く過程についてもくわしく解説しています。こうした理論的、もしくは教科書的な内容を重点的に紹介したのが本書の第1部です。内容をよく理解しないまま、こうした方法を運用するには抵抗がある、できれば基礎からしっかり学んでおきたい、という読者には第1部から順に読んでいくことをお勧めします。なお、第1部では、読者のヒントになればと以下を設けました。

**補足：**タイトルが「補足」ではじまるセクションは、先に進むうえで必須の内容ではありません。関心がない読者はとばして読むことも可能です。

**参考：**理解をふかめるうえで、参考になるかもしれない事項を取り上げました。必要に応じて、ご参照ください。

一方、第2部では、入門編ではくわしく紹介できなかった確率的感度分析を第4章で、また感度分析に必要なバイアスの情報の設定方法、感度分析を行う目的を含む準備から結果の報告までを第5章で、そして第6章では2つの感度分析の実施例を紹介しています。入門編の内容はだいたい理解できたので、その先をはやく知りたい、という読者は、まず第2部から読みはじめ、必要に応じて第1部を参照するとよいかもしれません。例えば、確率的感度分析について、仮にくわしい計算内容が把握できなかったとしても、その方法がもつ意義や考え方については理解できるかもしれません。また、計算ツールを適切にもちいることができれば、具体的にどんな結果が得られ、そこからどのような解釈が導かれるのか、その概略を体験することができます。読者の関心が、計算それ自体ではなく、それをもちいるための情報収集や結果の解釈、またそれらにもとづく判断や意志決定にある場合は、まず第2部にざっと目をとおして頂くとよいかもしれません。第2部でしめされている事例や、実践にともなう課題、またそれらに対するアプローチなどを大まかに把握することが、こうした読者にとっては有用だと考えられるからです。

以上の構成を念頭におきながら、本書を存分にご活用ください。

## 対象とする読者 …観察研究に基づく意思決定に関わるすべての方へ…

実践編では、製造販売後調査等も含め、実務として観察研究に従事している方や、安全監視として日常的に研究論文を評価していながら、こうした感度分析にはあまり馴染みがなかった、という読者を対象としています。

例えば、実践編第1部は、こうした分析を実務として行うであろう統計解析担当者も対象としています。一方、実践編第2部では、実際に感度分析を行えるよう、また統計解析担当者の側から計画者に対してより具体的な提案ができるよう、感度分析の具体的な手法にも踏み込んで解説を行い、課題解決の手がかりとなる論文も紹介しています。

診療情報データベースをもちいて薬剤疫学研究を行っておられる方だけでなく、そうした薬剤疫学研究の結果を日常的に評価しておられる方も、是非とも本書をご一読ください。

# 第 1 部

---

## 目次

1	実践的な感度分析に向けて .....	1
1.1	バイアス再考 .....	1
1.1.1	感度分析が対象とする「バイアス」とは? .....	1
1.1.2	観察研究の標的（ターゲット）となる試験 .....	1
1.1.3	補足：いわゆる「選択バイアス」について .....	2
1.2	バイアスの分類と感度分析 .....	4
1.2.1	バイアスの分析と分類 .....	4
1.2.2	バイアスの分類と感度分析の方法 .....	6
1.3	誤差と感度分析 .....	7
1.3.1	系統誤差と偶然誤差 .....	7
1.3.2	系統誤差と感度分析 .....	8
1.3.3	偶然誤差と感度分析 .....	9
1.3.4	感度分析に対する留意点 .....	10
1.4	感度分析を行うおもな方法 .....	10
1.4.1	単純な感度分析とその応用 .....	10
1.4.2	確率的感度分析とその応用 .....	11
2	本書であつかう効果指標と数値例 .....	13
2.1	効果指標 .....	13
2.1.1	単純な分割表と効果指標 .....	13
2.1.2	単純な分割表を対象とすることの意義 .....	14
2.1.3	単純な分割表にもとづく感度分析の応用 .....	14
2.2	数値例 .....	15
3	単純な感度分析とその応用 .....	17
3.1	交絡に対する感度分析 .....	17
3.1.1	感度分析の方針 .....	17
3.1.2	感度分析のためのパラメータと分割表のデータ .....	17
3.1.3	効果指標の修飾が想定される場合 .....	18
3.1.4	効果指標がリスク比やオッズ比の場合 .....	19
3.1.5	数値例 .....	20

3.1.6	交絡に対する感度分析と系統誤差.....	21
3.2	誤分類に対する感度分析.....	22
3.2.1	差別的誤分類と従属的誤分類.....	22
3.2.2	補足：非差別的で独立な誤分類は「性質がよい」のか？.....	23
3.2.3	感度分析のためのパラメータ.....	24
3.3	曝露の誤分類.....	24
3.3.1	感度分析の方針.....	24
3.3.2	感度分析のためのパラメータと分割表のデータ.....	25
3.3.3	計算例.....	27
3.3.4	補足：他のパラメータ設定.....	28
3.3.5	不可能な値.....	28
3.3.6	補足：データベース研究における曝露の誤分類.....	29
3.4	アウトカムの誤分類.....	31
3.4.1	感度分析の方針.....	31
3.4.2	感度分析のためのパラメータと分割表のデータ.....	31
3.4.3	計算例.....	33
3.4.4	稀なアウトカムに対する誤分類の影響.....	34
3.4.5	アウトカムの誤分類によってバイアスが生じない場合.....	35
3.4.6	アウトカムの誤分類をめぐる留意点.....	35
3.5	交絡要因の誤分類.....	36
3.5.1	感度分析の方針.....	36
3.5.2	感度分析のためのパラメータと分割表のデータ.....	37
3.5.3	計算例.....	37
3.5.4	補足：効果指標の修飾が想定される場合.....	38
3.5.5	補足：交絡要因、及び共変量の非差別的誤分類.....	40
3.5.6	補足：さらに複雑な誤分類.....	40
3.6	選択バイアスに対する感度分析.....	42
3.6.1	選択バイアスに対する調整解析.....	42
3.6.2	感度分析の方針.....	42
3.6.3	予備的な計算例.....	43
3.6.4	選択確率による方法.....	44
3.6.5	数値例における選択確率.....	45
3.6.6	選択確率と効果指標の関係.....	45
3.6.7	観察されなかった層の分割表を求める方法.....	46

3.6.8	数値例における観察されなかった層の分割表.....	46
3.6.9	補足：2つの方法の関連と、他のパラメータ設定.....	47
3.7	ケースコントロール研究の場合.....	48
3.7.1	コホート研究とのおもな相違点.....	48
3.7.2	交絡.....	49
3.7.3	曝露の誤分類.....	50
3.7.4	ケースコントロール研究におけるアウトカムの誤分類.....	51
3.7.5	選択前の集団でのアウトカムの誤分類.....	52
3.7.6	選択バイアス.....	53
3.8	人-時間を分母とした発生率を含む効果指標の場合.....	55
3.8.1	発生率の差、及び比.....	55
3.8.2	交絡.....	55
3.8.3	曝露 人-時間 の誤分類.....	56
3.8.4	アウトカムの誤分類.....	57
3.8.5	選択バイアス.....	58
3.9	多次元感度分析.....	60
3.9.1	交絡.....	60
3.9.2	アウトカムの誤分類.....	62
3.9.3	多次元感度分析の有用性と限界.....	63
	第1部のまとめ.....	64
	参 考 [本文中の対応箇所を括弧内に表示].....	65
	参考 1- 1 new user ではなく、prevalent user を試験に含めた場合 [ 本文：1.1.3 ] .....	65
	参考 1- 2 意図した介入からの逸脱、及び介入後の欠測によるバイアス [本文：1.2.2、 3.6.1].....	67
	参考 1- 3 臨床試験等で行われる統計解析における「感度分析」 [ 本文：1.3.3 ]....	68
	参考 2- 1 WTCR に対する研究の概略 [ 本文：2.2、3.1.5 ].....	69
	参考 3- 1 交絡の補正：リスク差（効果指標の修飾がある場合） [ 本文：3.1.3 ].....	70
	参考 3- 2 交絡の補正：リスク比 [ 本文：3.1.4、3.1.5、3.7.2、3.8.2 ].....	71
	参考 3- 3 交絡の補正：オッズ比（コホート研究の場合） [ 本文：3.1.4、3.7.2 ]..	72

参考 3- 4 交絡の補正：回帰モデルによる感度分析、E-value と感度分析 [ 本文：2.1.3、3.1.6 ]	73
参考 3- 5 交絡の補正：別のパラメータをもちいる場合 [ 本文：3.1.6 ]	75
参考 3- 6 差異的で独立な誤分類、非差異的で従属な誤分類 [ 本文：3.2.1 ]	76
参考 3- 7 曝露の誤分類の補正：差異的誤分類 [ 本文：3.3.2、3.3.4、3.7.3 ]	78
参考 3- 8 非差異的で独立な曝露の誤分類 [ 本文：3.2.2、3.3.3、3.3.5 ]	78
参考 3- 9 アウトカムの誤分類の補正：差異的誤分類 [ 本文：3.4.2、3.7.4 ]	80
参考 3- 10 交絡要因の誤分類の補正：差異的誤分類 [ 本文：3.5.2、3.5.3 ]	81
参考 3- 11 交絡要因に対する独立で非差異的誤分類の影響 [ 本文：3.5.5 ]	81
参考 3- 12 選択バイアスの補正：観察されなかった層の分割表を求める方法 [ 本文：3.6.7 ]	82
参考 3- 13 対象者選択によるバイアスが生じない場合（十分条件） [ 本文：3.6.7 ]	83
参考 3- 14 別のパラメータによる選択バイアスの補正 [ 本文：3.6.4、3.6.6、3.6.9 ]	83
参考 3- 15 アウトカムの誤分類の補正：ケースコントロール研究 [ 本文：3.7.5 ]	85
参考 3- 16 アウトカムの誤分類：アウトカム発生までの時間やハザード等の補正 [ 本文：3.8.4 ]	86
第 1 部の参考文献	90

## 1 実践的な感度分析に向けて

この1では、本書であつかう感度分析の範囲や、それと関連した概念について概略をしめす。よりよい実践のためには、もちいる方法の理論的な背景や限界について理解を深めることも重要である。感度分析では、方法の選択から結果の解釈まで、その多くが解析者の手に委ねられているからである。

とくにバイアスという概念や、それにまつわる用語には様々な定義や解説がある。読者によっては、あの教科書にはこう書いてあった、わたしはこう習った、とさまざまな不満や不安を感じるかもしれない。そこで、重要な用語や概念についてあらためて整理し、解説を補足した。生じうる誤解をさける意味でも、まず1に目をとおして頂きたい。

### 1.1 バイアス再考

#### 1.1.1 感度分析が対象とする「バイアス」とは？

入門編でも述べたように、感度分析とは、ある観察研究で生じている可能性があるバイアスの影響を定量的に評価する方法である。入門編では「その研究の限界が、そのままバイアスの原因となりうる」と述べた。

ただし、バイアスと言っても、量的な感度分析が対象とするのは、ある具体的な観察研究の主要な効果指標、つまりリスク差やリスク比、オッズ比等の推定値におけるバイアスである。この研究で、この効果指標を推定したいとき、と対象を絞って検討する。

バイアスが小さい研究は **妥当性 (validity)** が高い研究である、と呼ばれることがある<sup>1</sup>。研究の妥当性などと言うと、さらに漠然としたイメージを抱いてしまうかもしれない。しかし、量的な感度分析では、ある研究である効果指標をバイアスなく推定できるか、という意味での研究の妥当性を問題にする。

#### 1.1.2 観察研究の標的 (ターゲット) となる試験

では、妥当性が高い観察研究とはどのような研究だろう。この点を検討する際、あくまでも架空の想定として「もし理想的なランダム化比較試験 (RCT) を行うことができたとしたら」と考えることがある。あるいは逆に「このデータベースで利用可能なデータをもちいて、どのような RCT なら模倣 (emulate) できそうか」と考えながら観察研究をデザインする。こうした、ある観察研究の標的 (ターゲット) とみなせる理想的な試験のことを **ターゲット試験 (target trial)** と呼ぶことがある<sup>2, 3, 4</sup>。あくまでも科学的な意味での「理想」なので、倫理性や実施可能性は度外視する。

研究の計画段階では、ターゲット試験を念頭におくことで、どのようにデータをもちいれば理想に近づけることができるか、選択・除外基準、比較したい治療方針、その割り付け手順、追跡期間、調べたいアウトカム、解析対象集団、また解析計画といった個々の要素について具体的に検討することができる。また、すでに完了した研究については、実際の観察研究がどれほどターゲット試験と異なっているか、やはり個々の要素について具体的に検討することができる。すると、ターゲット試験からの偏差が、そのままバイアスの原因となりうることがわかる<sup>2</sup>。このように、ターゲット試験を想定することは、より妥当



性の高い観察研究をデザインするうえで有用であるし、また観察研究で生じるバイアスをより具体的に検討するうえでも有用だと考えられている。

こうした考え方をバイアスの評価にもちいる試みの1つとして **Cochrane Collaboration** が提供している **Risk Of Bias In Non-randomised Studies of Interventions**（以下、ROBINS-I）というツールがある<sup>3</sup>。提供されているチェックリストにそって検討していくと、その観察研究で懸念されるバイアスや、そのリスクの程度が評価できる。本書では、このツールを例にとり、その内容を大まかに紹介した（1.2）。

このような観点からは、いわゆる一般化可能性（**generalizability**）の問題、ときに適用可能性（**applicability**）や移送可能性（**transportability**）等とも呼ばれる問題を、バイアスとは区別して考える必要がある<sup>4</sup>。例えば、合併症がない患者に限定して行った研究の結果を、そのまま合併症をもつ患者集団にあてはめる（**apply**）、もしくは移送する（**transport**）ことはできないだろうし、合併症の有無をとわず、両者をあわせた集団全体へと一般化する（**generalize**）ことも難しいかもしれない。しかし、そのせいで、その研究が対象としている集団、すなわち合併症がない患者集団における推定値にバイアスが生じるわけではない。つまり、上述したような意味でのバイアス評価において焦点があたっているのは、あくまでもその研究が対象としている患者集団の内部での研究の妥当性、すなわち内的妥当性（**internal validity**）なのである<sup>3, 4</sup>。

以上のように「ある観察研究の、ある効果指標の推定値におけるバイアス」と対象を絞ったとしても、具体的にどのようなバイアスが問題になるのか、イメージしにくいかもしれない。しかも困ったことに、バイアスという概念やそれにまつわる用語には複数の、ときに矛盾する定義や解説が散見される。入門編と同様に、本書でも、バイアスについて述べる際は因果推論にもとづく構造的な定義（**structural definition**）、すなわち関連する要因の相互関係を因果グラフによって分析し、そのグラフの構造からバイアスを分類する方法になるべく従うようにした<sup>5</sup>（例えば入門編の付録1を参照）。このため、読者によっては、以前どこかで学んだバイアスの定義や分類と異なる、と違和感を覚えるかもしれない。

とくに混乱が生じやすいのが「選択バイアス」かもしれない。例えば、一般化可能性が損なわれることを「選択バイアス」と呼ぶことがある。しかし、すでに述べたように、ある研究の内的妥当性を問題にする限り、これを「バイアス」と呼ぶ必要はない。こうした混乱をさけるため、以下で説明を補足した。

### 1.1.3 補足：いわゆる「選択バイアス」について

どこかで以下のような解説を見たことがあるかもしれない。

例えば、高齢者を対象に健康に関する調査を行うとする。対象者の募集を病院で行うかクリニックで行うかによって、対象者の疾患や重症度が異なるため、おなじ高齢者でも対象集団の特性が異なる。…このように、対象者を決める時点で生じる偏りを選択バイアスという。

この記述を参考に、こんな例を考えてみよう。すなわち、ある病院で募集した対象者と、あるクリニックで募集した対象者について「健康に関する」アウトカムを比較することが、その調査の目的だとしよう。すると確かに、これらの集団間では「対象者の疾患や重症度

が異なる」ことだろう。しかし、仮に対象者全員について疾患や重症度等、アウトカムの発生に影響を及ぼす要因のデータが入手できたとしよう。また、これらの要因は、対象者が病院やクリニック等の医療機関を選ぶ際にも影響を及ぼしていると仮定しよう。この場合、これらの要因は、医療機関の選択（病院かクリニックか）とアウトカムの発生、双方の原因にあたる要因（共通の原因）である。したがって、入門編でも紹介したように、この比較で生じる可能性があるバイアスを交絡と呼ぶことができる。

また、やはり **Cochrane Collaboration** が提供している、ランダム化比較試験のバイアスを評価するツール（**Cochrane Risk of Bias Tool for Randomized Controlled Trials**、略称 **RoB**）の旧版や関連する解説等では「選択バイアス」の定義として以下がしめされていた<sup>6</sup>。

選択バイアスとは、比較したい集団間のベースライン特性の系統的な差をしめす。ランダム化の比類のない強みは、もしそれが成功裡に行われれば、参加者を介入に割り付ける際に生じうる選択バイアスを防ぐことにある。

この記述なども、例えば「比較したい集団間のベースライン特性に系統的な差が生じることを交絡と呼ぶ。ランダム化の比類のない強みは、もしそれが成功裡に行われれば、交絡を防ぐことにある」と書きかえられるだろう。

その後、ランダム化比較試験のバイアス評価ツール（**RoB**）については、観察研究のバイアス評価ツール（**ROBINS-I**）に用語をあわせる形で内容が修正された。その際、ランダム化比較試験でも問題となる「バイアス」を、構造的な定義によって整理する作業が行われている<sup>7</sup>。

いずれにせよ、交絡に相当するバイアスの一部を「選択バイアス」と呼ぶことは現実に、また頻繁に行われている。こうした用語や概念の混乱は、かなり以前から指摘されており「選択バイアスを交絡から区別すること（**distinguishing selection bias from confounding**）」という節を設け、くわしく解説している教科書もあるほどである<sup>1</sup>。

こうした現状を考えると、どんなバイアスをなんと呼ぼうとかまわないではないか、と感じる読者もいるだろう。しかし感度分析では、バイアスが発生するメカニズムを構造的に把握し、そのメカニズムにふさわしい方法を選ぶことも重要である。例えば、実質的に交絡にあたるバイアスを選択バイアスとみなして感度分析を行う、といったことは避けたい。このため本書では、もっぱら実践的な理由から、構造的な定義を採用している。こうした定義にしたがうと、選択バイアスは以下の条件下で生じる<sup>8</sup>。

- ・ その研究で関連性を調べたい2つの変数（多くの場合、曝露や介入といった原因に相当する変数と、アウトカム等の結果に相当する変数）の双方が、対象者の参加や選択と関連している場合。

この条件は、研究のどの段階でバイアスが生じるかに応じ、さらに以下の2つに分けることができる。

- (1) これら2つの変数（原因と結果）の双方が、研究初期の対象者の参加や選択の有無と関連している場合。
- (2) これら2つの変数（原因と結果）の双方が、曝露や介入を開始してからの追跡の有無と関連している場合。

前者は入門編であつかった「ケースコントロール研究における不適切な患者選択」に相当し、後者は「アウトカムの欠測（追跡不能による選択バイアス）」に相当する（例えば、入門編の付録 1 を参照）。またコホート研究について、これら 2 つに相当する事例をごく一般的に述べると、以下ようになる。いずれも、すでに紹介した、バイアス評価ツール ROBINS-I でしめされている内容である。

### (1) 介入前：試験の対象者選択によるバイアス

適格な対象者の一部や、一部の対象者の介入開始直後の追跡期間、もしくはアウトカムの一部を除外することが、介入ともアウトカムとも関連していると、比較したい介入の効果が等しくても、介入とアウトカムの間に関連性が生じる。一例として、ある介入をあらたに開始した対象者（new user）ではなく、すでにそれが開始されている対象者（prevalent user）を試験に含めた場合。

### (2) 介入後：欠測によるバイアス

試験に含まれ、介入を受けた対象者の追跡後の情報が欠測することで生じるバイアス（例えば、予後要因の影響により、介入群間で追跡不能の程度が異なる場合）：介入状況や交絡要因等、他の変数が欠測となった対象者を除外することで生じるバイアス。

以上のうち、(2) については「アウトカムの欠測（追跡不能による選択バイアス）」とほぼ同内容である。一方、(1) については、ここで述べられている「除外」の内容がイメージしにくいかもしれない。参考までに、文中で例示されている、new user ではなく prevalent user を試験に含めた場合について、説明をくわえた（参考 1- 1）。

以上、とくに誤解が生じやすい「選択バイアス」について解説を補足した。以下ではさらに視野をひろげ、どのようなタイプのバイアスに対してどのような感度分析を行うのか、その概略を見ておこう。

## 1.2 バイアスの分類と感度分析

### 1.2.1 バイアスの分析と分類

ある観察研究のある推定値のバイアスを評価する際、具体的にどのようなバイアスが懸念されるのだろうか。まずは質的な意味でのバイアス分析を行う必要がある。一般的には、因果グラフをもちいて関連する要因（変数）の相互関係や、その測定方法について、なるべく具体的に視覚化して検討することが推奨されている<sup>9</sup>。一方、すでに紹介したように、ある手順にそってバイアスのリスクを評価するツールも提案されている。以下では、すでに紹介した ROBINS-I というツールを参照し、観察研究のどのような側面に注目してバイアスを評価するのか、またその結果をどのように感度分析に結びつけるのか、その概略を紹介しよう。

まず ROBINS-I では、バイアスが生じうる時期を介入前、介入時、介入後という 3 つに分けて整理している。また時期ごとにバイアスが発生する構造を考慮したうえで、全部で 7 つのドメインにバイアスを分類している。その概要を以下にしめた（Table 1- 1）。

Table 1- 1 ROBINS-I におけるバイアスのドメイン

ドメイン	解 説
介入前	バイアスのリスク評価が、ランダム化試験の場合と明確に異なる
交絡によるバイアス	1つ、もしくは複数の予後因子（関心があるアウトカムを予測する要因）が、ベースラインで受ける介入をも予測する要因である場合、ベースラインにおける交絡が生じる。 ROBINS-I は、時間依存性交絡にも対処することができる。ただし、時間依存性交絡は、対象者が、比較したい介入の一方から他方に移り、かつベースライン後の予後因子がベースライン後に受けた介入に影響を与える場合に生じる。
試験の対象者選択によるバイアス	適格な対象者の一部や、一部の対象者の介入開始直後の追跡期間、もしくはアウトカムの一部を除外することが、介入ともアウトカムとも関連していると、比較したい介入の効果が等しくても、介入とアウトカムの間に関連性が生じる。一例として、ある介入をあらたに開始した対象者（new user）ではなく、すでにそれが開始されている対象者（prevalent user）を試験に含めた場合。
介入時	バイアスのリスク評価が、ランダム化試験の場合と明確に異なる
介入の分類によるバイアス	介入状況（介入の有無等）を差別的に、あるいは非差別的に誤分類することで生じるバイアス*。 非差別的誤分類は、アウトカムとは関連がない介入状況の誤分類であり、多くの場合、介入効果に関する推定値を、関連性がないという方向に偏らせる。 差別的誤分類は、介入状況の誤分類がアウトカムや、アウトカムのリスク要因と関連している場合に生じ、バイアス発生の引き金となりやすい。
介入後	バイアスのリスク評価が、ランダム化試験の場合とかなり重複する
意図した介入からの逸脱によるバイアス	検討したい介入と対照とする介入の間で、提供されたケアに、意図した介入からの逸脱をしめす、系統的な違いがあるせいで生じるバイアス。 このドメインに対するバイアス評価は、関心がある介入効果に依存する。すなわち割り当てられた介入の効果に関心があるのか、あるいは開始した介入を遵守した場合の効果に関心があるのかに依存する**。
欠測によるバイアス	試験に含まれ、追跡が開始された対象者の、その後の情報が欠測することにより生じるバイアス（例えば、予後要因の影響により、介入群間で追跡不能の程度が異なる場合）。またこうした対象者のうち、その後の介入状況や交絡要因等、他の変数が欠測となった対象者を除外することで生じるバイアス。
アウトカムの測定によるバイアス	アウトカムのデータを測定する際、差別的な、あるいは非差別的な誤差が生じることにより生じるバイアス。こうしたバイアスは、アウトカムの評価者が、介入状況を知っていたり、また介入群間で異なる方法でアウトカムを評価すると生じる。また、測定誤差が、介入状況や効果と関連している場合にも生じる。
報告する結果を選択することによるバイアス	導きたい結論に依存して、メタアナリシス（や他の統合）に含める推定値を除外する、といった形で報告する結果を選ぶこと。

\* 誤分類の差索性・非差索性については 3.2.1 参照。

\*\* ROBINS-I では「割り当てられた介入の効果」を intention-to-treat 効果、また「開始した介入を遵守した場合の効果」を per-protocol (PP) 効果と呼んで区別している。

各ドメインの詳細な内容と、バイアスのリスクを評価するためのチェックリストについては、公開されているガイダンスも参照してほしい<sup>4</sup>。このガイダンスは、すでに終了した研究のバイアスを評価するうえでも有用だが、バイアスが小さい、つまり妥当性が高い観察研究を立案するうえでも参考になるだろう。また、バイアスが発生するメカニズムについて、さらに理解を深めるうえでも有用かもしれない。

### 1.2.2 バイアスの分類と感度分析の方法

ROBINS-I のドメインを概観すると、介入前と介入時のドメインには、ランダム化を行わない観察研究ならではのバイアスが列挙されている。一方、介入後のドメインには、ランダム化比較試験でも生じるバイアスが含まれている。

ある観察研究でこれらのバイアスが懸念されるとき、それを感度分析ではどのようにあつかえばよいのだろうか。入門編では、ごく大まかに、未知・未測定 of 交絡要因、誤分類、選択バイアス、と 3 つのタイプのバイアスを取り上げ、それぞれに対する感度分析の方法を概説した。実のところ、これ以外の方法を追加する必要はない。7 つのドメインに分類されているバイアスに対し、これら 3 種類の方法を以下のように対応させることができる (Table 1-2)。

Table 1-2 バイアスのドメインと感度分析の方法

時期	ドメイン	感度分析の方法
介入前	交絡によるバイアス	未知・未測定 of 交絡要因
	試験の対象者選択によるバイアス	選択バイアス
介入時	介入の分類によるバイアス	介入 (曝露) の誤分類
介入後	意図した介入からの逸脱によるバイアス	本書では対象外
	欠測によるバイアス	選択バイアス 本書では対象外
	アウトカムの測定によるバイアス	アウトカムの誤分類
	報告する結果を選択することによるバイアス	本書では対象外

表にしめたように、7 つのドメインのうち 3 つのドメイン、すなわち「意図した介入からの逸脱によるバイアス」、「欠測によるバイアス」、及び「報告する結果を選択することによるバイアス」を本書では対象外とした。

これらのうち、「意図した介入からの逸脱によるバイアス」と「欠測によるバイアス」はランダム化比較試験でも生じるバイアスであり、ICH E9 (R1) 「臨床試験のための統計的原則補遺 臨床試験における estimand と感度分析 (案)」で述べられている「中間事象 (intercurrent event)」をめぐる問題とも重なる。とは言え、やや高度な内容が含まれるため本書では対象外とし、関心がある読者のために参考記事を設けた (参考 1-2)。

また「報告する結果を選択することによるバイアス」は、複数のアウトカムに対する結果や、複数の解析方法をもちいた結果、また複数の部分集団から得た結果のうち、研究者がある特定の結果だけを選択的に報告することにより生じる。こうしたバイアスを避けるには、まず観察研究の計画書に主要な解析計画を明記することが必要である。さらに、観察研究においても解析計画を事前に、また詳細に文書化することも有益であろう<sup>10</sup>。とはいえ、量的な感度分析において、こうしたバイアスを対象とすることは現時点では一般的でない。

## 1.3 誤差と感度分析

### 1.3.1 系統誤差と偶然誤差

しばしば指摘されるように、量的な感度分析により、系統誤差 (systematic error) によって生じる不確かさを量的に見積もることができる<sup>9</sup>。関連して、以下の点が強調されることが少なくない<sup>16</sup>。

系統誤差は、偶然誤差 (random error) より大きい可能性があり、また現実にもそうであることが多い。その影響を評価しそこなうと悲惨な結果をもたらす可能性がある。この問題は、規模の大きな研究や複数の研究の併合によってさらに悪化する。というのは、こうした研究では、規模の大きさが偶然誤差を小さくするが、その結果、偶然誤差は全体からみれば小さな構成要素に過ぎないからである。こうした研究で「統計的有意性」や信頼区間に注目するのは無意味であり、たんに系統誤差の産物 (artifacts of systematic error) に過ぎないものを、まるで現実の因果関係の反映のようにみなし、それに注目しようと心に決めるようなものである。

ここで言及されている「統計的有意性」や信頼区間は、その研究の主要な効果指標、つまりリスク差やリスク比、オッズ比等に対する統計的仮説検定が有意か、あるいはそれらの信頼区間が 0 や 1 を含んでいるか、といった内容である。こうした主要な解析結果が、たんなる「系統誤差の産物」かもしれないとは、いったいどういうことだろう。

こうした指摘の背景には、効果指標の推定値が以下のような構造をもつという想定がある。

$$(\text{推定値}) = (\text{真の値}) + (\text{系統誤差}) + (\text{偶然誤差})$$

後に見るように、効果指標によっては、こうした加法的な構造ではなく

$$(\text{推定値}) = (\text{真の値}) \times (\text{系統誤差}) \times (\text{偶然誤差})$$

といった乗法的な構造を考える方が自然である。例えば、リスク差については加法的な構造を、またリスク比やオッズ比については乗法的な構造を想定すると考えやすい。いずれにせよ、わたしたちの目にうつる推定値が、実はこうした内部構造をもっているのはいかという想定が、感度分析の背景にはある。

上式の **真の値** という項は、現実の因果関係、つまり真のリスク差や真のリスク比を表している。わたしたちが本当に知りたいのは、この値である。理想的なターゲット試験を行うことができれば、この値を、かなり正確に知ることができるかもしれない。



また、**系統誤差**という項は、これまでバイアスと呼んできた量であり、現実の研究ではこうした誤差が系統的に生じる懸念がある。

**偶然誤差**は、まったくランダムに生じる誤差であり、対象者数を増やすほど小さくなる。例えば、介入や曝露の有無をランダムに定めても、群間で集団の特徴が異なることはある。しかし、ランダム化によって生じる集団間の差や、そのせいで生じる推定値の偏りは、あくまでもランダムな誤差、つまり偶然誤差とみなすことができる。ランダム化する対象者数を増やすことができれば、こうした差や偏りは生じにくくなる。一方、ランダム化を行わないことで生じる差や偏りは、いくら対象者数を増やしても小さくならない。むしろ、こうした差や偏りが、より正確に、小さな偶然誤差とともに観察されるだけである。

こうした観点から、上で引用した文章をあらためて読みなおしてみよう。まず、規模の大きな研究を行ったり、複数の研究を併合することで、対象者数を十分に大きくすることができれば、偶然誤差の影響は小さくなる。すると上でしめした式は、近似的に

$$(\text{推定値}) \cong (\text{真の値}) + (\text{系統誤差}), \text{または } (\text{推定値}) \cong (\text{真の値}) \times (\text{系統誤差})$$

と表すことができる。多くの研究者は、自分たちの研究から得た推定値は、真の値に近いはずだと信じていることだろう。しかし、系統誤差が疑われる状況では、仮に真のリスク差が0、あるいは真のリスク比が1でも、系統誤差の値がそのまま推定値として観察される。その研究から得られた推定値は、まさに「系統誤差の産物」かもしれない。このため、系統誤差を含む推定値を、あたかも真の値のようにみなして様々な判断を行うと「悲惨な結果」をもたらしかねない。

では、感度分析を行うことで、この状況がどう変わるのだろうか。

### 1.3.2 系統誤差と感度分析

入門編でも概観したように、感度分析では、関心がある指標（割合、率、リスク差、リスク比、オッズ比等）のバイアスを補正した値、すなわち補正済み指標を求めるのだった。すなわち、アウトカムの発生状況をしめす指標（割合、率等）を補正し、これらの差や比に相当する指標、すなわち曝露や介入の効果をしめす効果指標（リスク差、リスク比、オッズ比等）の値を補正する計算が、感度分析では重要だった。

こうした計算を行うため、バイアスをもたらす量をパラメータとして定め、その値と効果指標の値を結びつける、数学的なモデルを構成した。以下では、バイアスを補正した効果指標の値を、簡単に補正值と呼ぶことにしよう。後に見るように、例えば交絡の影響を補正したリスク差を求める際、以下の計算を行う。

$$(\text{リスク差の補正值}) = (\text{未補正のリスク差}) - (\text{モデルによって定められる値})$$

この計算は、実質的には以下の計算と同内容である。

$$(\text{真の値}) \cong (\text{推定値}) - (\text{系統誤差})$$

もちろん、系統誤差が生じるメカニズムを表すモデルや、それに代入するパラメータの値が正しいとは限らない。つまり、系統誤差の大きさが未知であることに変わりはない。しかし、仮にパラメータの値がこれこれだとしたら、という想定のもとで系統誤差を除去し、バイアスを補正した効果指標の推定値を求めることができる。

以上のように感度分析では、系統誤差と推定値の関係を、モデルを通じて表現する。では、偶然誤差については、どのようにあつかうのだろうか。

### 1.3.3 偶然誤差と感度分析

通常、行われる統計解析では、ほとんど暗黙のうちに系統誤差がないことを想定している。あるいは、それも偶然誤差のうちに含まれている、などと考える。すなわち

$$\text{通常の統計解析における想定： } (\text{推定値}) = (\text{真の値}) + (\text{偶然誤差})$$

薬剤の開発中に行われる臨床試験では、系統誤差の発生を、試験デザイン上のさまざまな工夫や計画的なモニタリング等を通じ、可能な限り最小化することが求められる<sup>17</sup>。もちろん、この理想がいつも達成されるとは限らない。しかし、それに近い状態が実現することを、暗黙のうちに仮定している。このため、臨床試験等で行われる統計解析では、おなじ「感度分析」という用語で、本書とやや異なる内容を差すことが少なくない。参考のため、臨床試験等における統計解析で問題となるバイアスと、それに対する感度分析について、かんたんに補足した（参考 1- 3）。

さて、観察研究では、あるいはそれに対して行われる感度分析では、すでに述べたように「系統誤差もあるはずだ」と考える。すなわち

$$\text{感度分析における想定： } (\text{推定値}) = (\text{真の値}) + (\text{系統誤差}) + (\text{偶然誤差})$$

これまで述べてきたように、感度分析では、推定値と系統誤差の関係をモデル化することで、実質的に

$$(\text{真の値}) \equiv (\text{推定値}) - (\text{系統誤差})$$

といった計算により補正值を求める。だとしたら、さらに歩を進め

$$(\text{真の値}) = (\text{推定値}) - (\text{系統誤差}) - (\text{偶然誤差})$$

といった形で補正值を求めることもできそうだ。

後で見ると、偶然誤差も考慮した感度分析では、概ね以上のような方法で補正值を求める。つまり、モデルをもちいていったん系統誤差を除去し、さらに偶然誤差にあたる量も考慮して補正值を求める。

系統誤差だけでなく偶然誤差も考慮することで、わたしたちが漠然と真の値に近いのだろうと考えている推定値より、さらに外れた値が得られるかもしれない。しかし、こうした計算を重ねることで、得られた研究結果を過信することなく、いったん「頭を冷やす」ことができる。より確からしい結論を導くには、どんなエビデンスが欠けているのか、具体的な指針を得ることもできるかもしれない。

1.3.1 の冒頭で「量的な感度分析により、系統誤差（systematic error）によって生じる不確かさを量的に見積もることができる」と述べた<sup>9</sup>。この文章には続きがあり、すべて引用すると以下ようになる。



量的な感度分析により、系統誤差（systematic error）によって生じる不確かさを量的に見積もることができ、研究結果に対する過信（overconfidence）に対抗することができ、また将来の研究を導くことができる。

あたりまえのことだが、感度分析の結果を過信することも本末転倒である。感度分析の留意点についても簡単にふれておこう。

### 1.3.4 感度分析に対する留意点

入門編でも述べたように、解析者があるバイアスに対して抱いている想定は、感度分析のために定めるモデルと、それに含まれるパラメータに代入する値によって表現される。このため、パラメータに極端な値を代入すれば、ある意味、どんな補正值でも得ることができる。

パラメータの値やその分布について適切な情報が得られない場合、最終的には解析者の判断や意見、すなわち知識や経験に裏付けられた推測（educated guess）によって、それらを定めるほかない<sup>8, 9, 16</sup>。つまり感度分析では、通常の統計解析以上に、解析者の判断とそれにもとづくインプットが求められる。こうした感度分析の特徴から、以下の点に留意すべきである<sup>16</sup>。

- ・ 感度分析により、因果関係の有無について、通常の統計解析以上のことが示せるわけではない。感度分析の実施により、あると考えられた関連性がないことが示唆されたり、その逆が示唆されるかもしれない。しかし、こうした結果から、直ちに関連性の有無が確立されるわけではない。
- ・ 感度分析の結果をしめす際、どのようなモデルをもちいたか、またパラメータにどのような値（または分布）をもちいたかを明示することが重要である。感度分析を行ったところ、このような結果が得られた、と述べるだけでは不十分である。このモデルにこのような値（または分布）をもちいたら、という条件を常に明確に提示すべきである。

以上の点は、他の解析者が行った感度分析の結果を解釈する際にも留意すべきである。解析者がもちいたパラメータの値や分布を変更すれば、おそらくまた別の結果が得られるだろう。他の解析者が計算を再現できるよう、感度分析の方法をくわしく提示することの重要性が、こうしたことから伺える。

次に、想定される系統誤差や偶然誤差をどのように感度分析に反映させるのか、本書で取り上げる感度分析の全体像を素描しておこう。

## 1.4 感度分析を行うおもな方法

### 1.4.1 単純な感度分析とその応用

入門編でも紹介したように、すべての感度分析の出発点は、ある指標のバイアスの原因として想定される要因や条件をパラメータで表し、それらの値と観察されたデータから補正值を求める数学的なモデルを構成することである。こうしたモデルに含まれる個々のパラ

メータに 1 つの値を代入して補正値を 1 つ求める、というのが入門編で紹介した **単純な感度分析** (Simple Sensitivity Analysis) だった。

この方法を少しだけ拡張したのが、**多次元感度分析** (Multidimensional Analysis) である。計算方法そのものは単純な感度分析とおなじだが、個々のパラメータに複数の値を代入することで、バイアスの影響をより多角的に検討することができる。このパラメータの値をこの範囲で変化させると、補正値がこの範囲で変化する、といった傾向を大まかに把握することができる。

#### 1.4.2 確率的感度分析とその応用

さらに進んだ方法として **確率的感度分析** (Probabilistic Sensitivity Analysis、略称 PSA) があり、本書の重要な主題の 1 つである。この方法では、以下の 2 種類の不確かさを、ともに考慮することができる。

- ・ 系統誤差の方向や大きさに対する不確かさ
- ・ 偶然誤差による、効果指標の推定値がもつ不確かさ

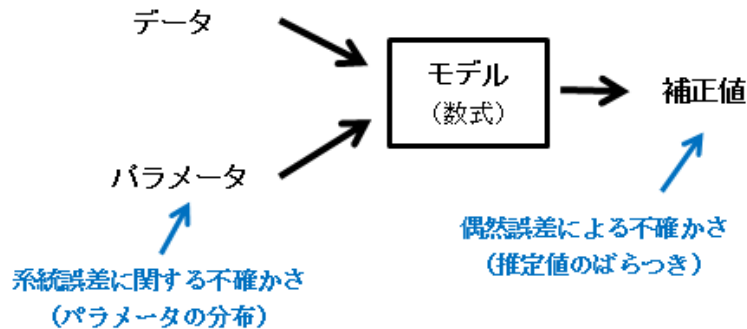
前者は、基本的には、解析者が系統誤差に対して抱いている想定の不確かさである。モデルに含まれるパラメータの値をいくつに定めればよいか、あまり明確な確信が持てない。こうした主観的な意味での不確かさを、確率的感度分析では、個々のパラメータに分布をもたせることで表現する。例えば「このパラメータの値は概ねこの値だと考えられるが、可能性としては、最小から最大まで、これくらいの範囲ではないかと想定される」といった解析者の想定を、ある分布として表現する。そして、この分布からランダムに抽出した値をパラメータに代入し、補正値を求める。これは、すでに述べた  $(\text{真の値}) = (\text{推定値}) - (\text{系統誤差})$  という計算に相当する。

一方、後者は、効果指標のばらつきによって表される、もっと客観的な不確かさである。確率的感度分析では、上述した手順でランダムに抽出した値をパラメータに代入し、いったん補正値を求めたうえで、さらに効果指標のばらつきに相当する量を引く。こうして、系統誤差だけでなく、偶然誤差も考慮した補正値を求める。これは、すでに述べた  $(\text{真の値}) = (\text{推定値}) - (\text{系統誤差}) - (\text{偶然誤差})$  という計算に相当する。

以上の計算をくり返すことで、補正値の分布が得られる。ただし、前者の、系統誤差に対する補正だけを行うと、解析者が系統誤差に対して抱いている不確かさのみを反映した補正値の分布が得られる。一方、前者の補正にくわえ、後者の、偶然誤差に対する補正も行うと、系統誤差と偶然誤差、双方による不確かさを反映した補正値の分布が得られる。当然ながら、前者より後者のほうが、よりばらつきの大きな分布が得られる。

まとめると、確率的感度分析を行うことで、系統誤差と偶然誤差という 2 種類の不確かさを、量的に統合することができる (Figure 1-1)。なお、確率的感度分析を 1 回行くと、その解析のなかでは補正値の計算を複数回行うことになる。しかしこれは、あくまでも 1 つの効果指標に対する、1 つのバイアスを対象とした反復計算であり、その結果として補正値の分布が 1 つ得られる。

Figure 1- 1 補正值に「不確かさ」を反映させる



最後に、複合バイアスモデリング（Multiple Bias Modeling）では、以上で述べた確率的感度分析を複数のバイアスに対して行う．と言っても、特別にあたらしいことを行うのではなく、個々のバイアスに対する確率的感度分析を順次行うことになる．例えば、誤分類について補正し、選択バイアスについて補正し、さらに交絡について補正し、最後に偶然誤差の影響を考慮する、といった手順で確率的感度分析をくりかえす．すなわち、複合バイアスモデリングでは、1つの効果指標に対する、複数のバイアスを対象とした反復計算を行い、その結果として補正值の分布が1つ得られる．

Table 1- 3 感度分析の方法の概略

解析の方法	パラメータ の持ち方	対象とする バイアスの数 (解析ごとに)	得られる結果	偶然誤差 を考慮する か？	計算の 難易
<b>単純な感度分析</b> Simple Sensitivity Analysis	パラメータご とに1つの値	1回に1つ	補正值が1つ	通常、 しない	易
<b>多次元感度分析</b> Multidimensional Analysis	パラメータご とに複数の値	1回に1つ	補正值の範囲	しない	易
<b>確率的感度分析</b> Probabilistic Sensitivity Analysis (PSA)	パラメータご とに1つの 分布	1回に1つ	補正值の度数 分布	できる	難
<b>複合バイアス モデリング</b> Multiple Bias Modeling	複数のパラメ ータの（同 時）分布	1回に複数	補正值の度数 分布	できる	難

まとめると、個々の感度分析にもちいるモデルはおなじでも、パラメータの持ち方（入力する値が1つか複数か、また分布をもちいるか）や、偶然誤差を考慮するかどうか、また1つのバイアスのみを考慮するか複数のバイアスを同時に考慮するか、等々により様々な感度分析が考案されている．こうした方法の違いと共通点を、感度分析の名称ごとに Table 1- 3 にまとめた<sup>9</sup>．それぞれの内容については以降で紹介していくが、いったんモデルを定めると、様々な運用が可能なこと、またいずれの方法をもちいるにせよ、単純な感度分析が基本となることを強調しておきたい．

## 2 本書であつかう効果指標と数値例

### 2.1 効果指標

#### 2.1.1 単純な分割表と効果指標

実践編でも、曝露や介入、アウトカム、交絡要因、いずれも 2 値の場合をおもにあつかう。例えば、コホート研究の場合、以下のような分割表に要約できるデータについて考える (Table 2-1)。

Table 2-1 介入・曝露の有無とアウトカムの有無に関する分割表

アウトカム	介入・曝露	
	あり ( $E = 1$ )	なし ( $E = 0$ )
あり ( $D = 1$ )	$A_1$	$A_0$
なし ( $D = 0$ )	$B_1$	$B_0$
合計	$N_1$	$N_0$

この表にあるように、本書では、介入や曝露を  $E$ 、アウトカムを  $D$  で表す。また、表には書かれていないが交絡要因を  $C$  で表す。これらの変数の値は、ありの場合を 1、なしの場合を 0 とする。さらに、アウトカムありの場合の度数を  $A$ 、なしの場合を  $B$ 、また合計を  $N$  で表す。

また本書では、以下の効果指標の推定値に対する感度分析をおもに取り上げる<sup>18</sup>。

$$\text{リスク差: } \widehat{RD} = \frac{A_1}{N_1} - \frac{A_0}{N_0}, \text{Var}(\widehat{RD}) = \frac{A_1 B_1}{N_1^2 (N_1 - 1)} + \frac{A_0 B_0}{N_0^2 (N_0 - 1)},$$

$$\text{リスク比: } \widehat{RR} = \frac{A_1 / N_1}{A_0 / N_0}, \text{Var}[\ln(\widehat{RR})] = \frac{1}{A_1} - \frac{1}{N_1} + \frac{1}{A_0} - \frac{1}{N_0},$$

$$\text{オッズ比: } \widehat{OR} = \frac{A_1 / B_1}{A_0 / B_0} = \frac{A_1 B_0}{A_0 B_1}, \text{Var}[\ln(\widehat{OR})] = \frac{1}{A_1} + \frac{1}{B_1} + \frac{1}{A_0} + \frac{1}{B_0}.$$

上式でしめしたように、アウトカムに対する介入や曝露の効果を表す際、リスク差については「曝露ありー曝露なし」という向きで差をとり、リスク比とオッズ比については「曝露あり/曝露なし」という向きで比をとる。また以下では、混乱が生じない限り推定値をしめす記号  $\hat{\phantom{x}}$  を省略する。

本書では、これらの表記をもちいて様々なバイアスの補正について述べていく。その際、観察されたデータとパラメータの値から補正值をもとめるモデルを数式で表していくが、それらの大半は、本書でもしばしば引用している以下の教科書でしめされているものと同内容である。

- Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd edn. Philadelphia, PA: Lippincott Williams&Wilkins, 2008.

- Lash TL, Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data. New York: Springer, 2009.

なお、これらの教科書では網羅されていない研究デザインや効果指標も本書ではあつまっている。しかし、そのような場合でも、基本的な考え方や数式を導く手順はいずれも似かよっている。このため、本書でしめした数式や計算方法のすべてに対して、その都度、参照文献をしめすことは避けた。かわりに、おもな考え方や数式を導出する過程を可能なかぎり巻末の「参考」に記載した。感度分析を実践するうえでは、ときに解析者がみずからモデルを定めることも必要である。しかし参考に記載したように、その内容はさほど複雑ではない。理解が不十分なまま感度分析を行うことを避ける意味でも、一度は手を動かし、以上の教科書や本書でしめされている数式の意味や内容を確認してみることを薦めたい。

## 2.1.2 単純な分割表を対象とすることの意義

実際の観察研究において、モデルをもちいた様々な解析が行われている現状を考えると、単純な分割表を対象を絞るのは不自然に思えるかもしれない。しかし、これから述べるように、こうした内容に絞っても感度分析の基本的な考え方を知ることができる。また、感度分析の実践という観点からも、あまり複雑な構造のデータに対して感度分析を行うことは得策ではない。

感度分析では問題を単純化して考えることが少なくない。これには計算を簡略化したいという意図にくわえ、得られた結果の解釈を容易にしたいという意図がある。というのは、こうした感度分析の目的は、より厳密なバイアス補正を行い、より正確な推定値を得ることではないからである。むしろ、バイアスに対するある想定のもとで推定値がどれほど変化するか（あるいは変化しないか）大まかに調べるのがまず重要である。こうした観点から行われる単純化として、以下のようなアプローチ、あるいは考え方がある。

- 複数の交絡要因が疑われたり、対象者選択が多段階に行われたり、また測定誤差が複数の過程で生じた可能性がある場合でも、1つにまとめて考えることがある。すなわち、1つの交絡要因で代表させ、1回で選択が行われたとみなし、また測定全体を通じて1度だけ誤差が生じたとみなすことがある。
- 複数のバイアスが相互に関連している可能性がある場合でも、個々のバイアスを独立なものとして、つまり互いに無関係なものとしてみなすことがある。
- 個々のバイアスについても、適切なら、より単純なモデルを採用し、計算を簡略化することがある。
- 単純な分割表に要約できる（未調整の）解析結果と、回帰モデル等をもちいて調整した結果が大きく異ならない場合、単純な分割表に対して感度分析を行うことがある。

このように問題を単純化していくと、単純な分割表にもとづく感度分析でも、意外に適応範囲がひろいことがわかる。

## 2.1.3 単純な分割表にもとづく感度分析の応用

より複雑な構造をもつデータに対する感度分析も、結局は、単純な分割表に対する感度分析の応用、もしくは拡張とみなすことができる。以下は、その代表的な例である。



- ・ **層別による感度分析**

交絡要因等の値を参照して分割表を層別し、まず層ごとの分割表に対して感度分析を行う。さらに層ごとの結果を併合し、試験全体の結果に対する感度分析を行う。

- ・ **回帰モデルをもちいた感度分析**

この方法は、未知・未測定 of 交絡要因に対する感度分析でもちいられることが多い。回帰モデル等に含まれる要因の影響を考慮した上で、さらに（それらと独立な）未知・未測定 of 交絡要因の影響を補正した推定値を求める。その計算の中身は、モデルに含まれる、介入や曝露の効果を表す係数（パラメータ）の値をまず補正し、補正した係数の値から効果指標の推定値を求める、というものである。係数の補正にもちいられる式は、単純な分割表に対する式と類似している。

- ・ **レコードレベルの感度分析**

対象者の特性（年齢、性別、重症度等）によって誤分類の程度が異なることが想定される場合にもちいられることが多い。この方法は、層別による感度分析をさらに極端にした場合、つまり対象者ごとに層を設け、対象者という層ごとに（レコードレベルで）感度分析をくりかえす方法とみなすこともできる。

いずれの方法も、やや発展的な内容を含むため、本書でくわしく述べることは避けた。ただし、その概略を該当箇所や参考等で簡単に紹介した。

以上で概観したように、単純な分割表にもとづく感度分析の適応範囲は意外にひろい。また、それが基礎となつて、より発展的な方法が築かれている。いきなり複雑な構造のデータに対して感度分析を試みるのではなく、まず単純な分割表に対して、あるいは単純な分割表に要約できるように問題を単純化したうえで、感度分析を行ってみることを薦めたい。

## 2.2 数値例

本書の 3 では、おなじ数値例をもちいる。出典は、2001 年 9 月 1 日以降、ニューヨークのワールド・トレードセンター（WTC）周辺で救助や復旧、清掃等にあたつた人たちの健康情報を収集したレジストリー、World Trade Center Health Registry（WTCHR）をもちいた研究である<sup>19</sup>。この研究で得られた結果のなかから、とくに 9 月 11 日当日の粉じんや破片の雲（dust cloud）への曝露（以下、粉じんへの曝露）と、その後のぜん息の発症（以下、ぜん息）に関するデータを以下にしめした（Table 2-2）。

WTCHR の事例では、効果指標としてオッズ比（曝露あり / なし）をもちいており、事件当日の急激かつ大量な粉じんへの曝露だけでなく、現地での作業日数や防塵マスクの使用等との関連性も示唆されている。粉じんへの曝露とぜん息の関連性については、未調整のオッズ比 1.77（95%信頼区間：1.55, 2.01）にくわえ、性別や年齢等の影響を調整したオッズ比 1.28（95%信頼区間：1.09, 1.50）が報告されている。本書では、このデータを数値例としてもちい、リスク差やリスク比についても取りあつかう。また、この表の信頼区間は 2.1.1 でしめした式により再計算したものである。

Table 2- 2 WTCHR からの数値例：粉じんへの曝露とぜん息の関連性

ぜん息	粉じんへの曝露	
	あり ( $E = 1$ )	なし ( $E = 0$ )
あり ( $D = 1$ )	$A_1 = 476$	$A_0 = 450$
なし ( $D = 0$ )	$B_1 = 9,286$	$B_0 = 15,536$
合計	$N_1 = 9,762$	$N_0 = 15,986$
発生割合	4.9%	2.8%
リスク差	2.06% (95%CI: 1.56%, 2.56%)	
リスク比	1.73 (95%CI: 1.53 , 1.97 )	
オッズ比	1.77 (95%CI: 1.55 , 2.02 )	

この WTCHR の事例に対し、別の研究者が感度分析を行っており、おもにアウトカムの誤分類と選択バイアスの影響を検討している<sup>20</sup>。本書では、あくまでも数値例としてこの事例を取り上げていくが、感度分析の意図や内容を理解するうえで参考になるようなら、こうした先行研究でしめされている見解や考察についても簡単に紹介する。参考までに、以上で参照した 2 つの論文の概略を要約した（参考 2- 1）。

なお、本書の第 2 部では、ケースコントロール研究の数値例も参照している。第 1 部 3 でケースコントロール研究に対する感度分析について述べる際、この事例についてもごく簡単に紹介した（3.7）。

### 3 単純な感度分析とその応用

この3では、交絡、誤分類、選択バイアスの順に、単純な感度分析によって補正值を求める方法を概説する。入門編の内容と重複するが、実践編では、数式の背景にある考え方や、そこから導かれる運用上の留意点についても、さらに踏み込んで述べていく。また、単純な感度分析の応用として、多次元感度分析についてもその概略を紹介する。

#### 3.1 交絡に対する感度分析

##### 3.1.1 感度分析の方針

未知、あるいは未測定 of 交絡要因のせいで、効果指標の推定値にバイアスが生じている懸念がある状況である。もし、こうした要因のデータがあれば、例えば層別を行い、バイアスの影響を調整することができる。入門編でも紹介したように、感度分析では、未知・未測定 of 交絡要因についても「仮にこのような交絡要因があったとしたら」という想定のもとで層別を行う。

WTCHR の事例について考えてみよう。この事例について、とくに未知・未測定 of 交絡要因の関与が疑われているわけではなかった<sup>18, 20</sup>。そこで以下では、あくまでも仮想的な数値例として、そのような要因の影響が懸念される場合の感度分析について考えてみよう。

まず、すでにしめた分割表を、この研究では調べることができなかった未知 of 交絡要因の有無によって層別できたとしよう (Table 3-1)。ただし、交絡要因の有無を各度数 ( $A$  や  $N$ ) 右下の、2 つめの添え字で表している。また、データとしては得られていない、層別後の度数を青字でしめた。

Table 3-1 交絡要因の有無で層別した分割表

アウト カム： ぜん息	合 計 (元のデータ)		未知 of 交絡要因			
			$C = 1$		$C = 0$	
	曝露：粉じん		曝露：粉じん		曝露：粉じん	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	$A_1 = 476$	$A_0 = 450$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
$D = 0$	$B_1 = 9,286$	$B_0 = 15,536$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
合計	$N_1 = 9,762$	$N_0 = 15,986$	$N_{11}$	$N_{01}$	$N_{10}$	$N_{00}$

この表で「合計 (元のデータ)」とある部分からリスク差やリスク比を求めると、交絡を無視した解析となり、推定値にバイアスが生じる懸念がある。そこで、右側にしめされている層別した表をもちいてバイアスを補正した値、すなわち補正值を求めたい。これが、未知・未測定 of 交絡要因に対する感度分析の大まかな方針である。

##### 3.1.2 感度分析のためのパラメータと分割表のデータ

層別した分割表を得るには、交絡要因に関する量的な想定をパラメータの値として与える必要がある。効果指標としてリスク差を考え、以下のようにパラメータを設定してみよう。



- ・  $RD_{DC}$ ：未知の交絡要因（ $C$ ）とぜん息（ $D$ ）の関連性をしめすリスク差（交絡要因あり－交絡要因なし）。ただし、このリスク差が、曝露の有無によらず等しいことを仮定する。
- ・  $P_{C1}$ ：曝露あり（ $E = 1$ ）の集団で未知の交絡要因あり（ $C = 1$ ）の対象者が占める割合。
- ・  $P_{C0}$ ：曝露なし（ $E = 0$ ）の集団で未知の交絡要因あり（ $C = 1$ ）の対象者が占める割合。

これらのパラメータを Table 3-1 の度数と結びつけてみよう。まず、最初のパラメータ、すなわち交絡要因の有無によるリスク差については、曝露の有無別に

$$RD_{DC} = \frac{A_{11}}{N_{11}} - \frac{A_{10}}{N_{10}} = \frac{A_{11}}{N_{11}} - \frac{A_1 - A_{11}}{N_1 - N_{11}}, \text{ 及び } RD_{DC} = \frac{A_{01}}{N_{01}} - \frac{A_{00}}{N_{00}} = \frac{A_{01}}{N_{01}} - \frac{A_0 - A_{01}}{N_0 - N_{01}}$$

という関係が得られる。また、残り 2 つのパラメータ、すなわち交絡要因ありの対象者の割合については、曝露の有無別にそれぞれ

$$P_{C1} = \frac{N_{11}}{N_1}, \text{ 及び } P_{C0} = \frac{N_{01}}{N_0}$$

となる。以上から、層別後の度数を、層別前の度数（元のデータ）とパラメータをもちいて表すことができる。例えば、交絡要因ありの層については

$$A_{11} = P_{C1}\{A_1 + (1 - P_{C1})RD_{DC}N_1\}, A_{01} = P_{C0}\{A_0 + (1 - P_{C0})RD_{DC}N_0\}, \\ N_{11} = P_{C1}N_1, N_{01} = P_{C0}N_0.$$

また、交絡要因なしの層については

$$A_{10} = (1 - P_{C1})(A_1 - P_{C1}RD_{DC}N_1), A_{00} = (1 - P_{C0})(A_0 - P_{C0}RD_{DC}N_0), \\ N_{10} = (1 - P_{C1})N_1, N_{00} = (1 - P_{C0})N_0.$$

こうして、層別後の度数を求めることができたので、交絡要因ありの層のリスク差は

$$\frac{A_{11}}{N_{11}} - \frac{A_{01}}{N_{01}} = \frac{A_1 + (1 - P_{C1})RD_{DC}N_1}{N_1} - \frac{A_0 + (1 - P_{C0})RD_{DC}N_0}{N_0} = \frac{A_1}{N_1} - \frac{A_0}{N_0} - (P_{C1} - P_{C0})RD_{DC}.$$

また、交絡要因なしの層のリスク差は

$$\frac{A_{10}}{N_{10}} - \frac{A_{00}}{N_{00}} = \frac{A_1 - P_{C1}RD_{DC}N_1}{N_1} - \frac{A_0 - P_{C0}RD_{DC}N_0}{N_0} = \frac{A_1}{N_1} - \frac{A_0}{N_0} - (P_{C1} - P_{C0})RD_{DC}$$

となる。以上から、交絡要因の有無によらず、層ごとのリスク差は等しくなり、交絡の影響を補正したリスク差  $RD$  は以下ようになる。

$$RD = \frac{A_{11}}{N_{11}} - \frac{A_{01}}{N_{01}} = \frac{A_{10}}{N_{10}} - \frac{A_{00}}{N_{00}} = \frac{A_1}{N_1} - \frac{A_0}{N_0} - (P_{C1} - P_{C0})RD_{DC} \quad \dots \quad \text{式 3-1}$$

### 3.1.3 効果指標の修飾が想定される場合

以上の計算で、交絡要因の有無によらずリスク差が等しくなったのは、交絡要因とぜん息の関連性をしめすリスク差  $RD_{DC}$  が曝露の有無によらず等しいと仮定したからである。

しかし、交絡要因の有無によってリスク差が異なると考えられる場合もあるかもしれない。ある要因の値により、リスク差やリスク比等の値が異なるとき、その要因により効果指標が修飾を受ける、あるいはその要因による効果指標の修飾（effect measure modification）がある等と言う。また、こうした要因を修飾要因（effect modifier）と呼ぶことがある。

効果指標の修飾は、それ自体ではバイアスではない。むしろ、こうした要因がないか、個々の研究のなかで積極的に探索したい場合もある。ところが、こうした要因をもつ対象者の分布が曝露の有無により異なると（アウトカムに影響を及ぼす要因の分布が比較したい集団間で異なるため）交絡が生じる可能性がある。

こうした場合は、パラメータの数を増やして対応する。例えば、交絡要因とぜん息の関連性をしめすリスク差を

- ・ 曝露あり（ $E = 1$ ）のとき： $RD_{DC1}$
- ・ 曝露なし（ $E = 0$ ）のとき： $RD_{DC0}$

と2つのパラメータをもちいて表す。この場合は、交絡要因の有無別に異なるリスク差が得られる（参考 3- 1）。層別した分割表から、Mantel-Haenszel の方法をもちいて層を併合したリスク差を求めることができる。また、標準集団を適当に定め（例えば、曝露ありの集団、曝露なしの集団、あるいは全体集団）、交絡要因の分布を標準集団の分布にそろえた場合のリスク差、すなわち標準化リスク差を求めることもできる。具体的な計算方法等については 3.5.4 を参考にしてほしい。

### 3.1.4 効果指標がリスク比やオッズ比の場合

リスク比やオッズ比に対して感度分析を行いたい場合は、最初のパラメータ、すなわち交絡要因とぜん息の関連性をしめす指標を変更すると、同様な計算により層別後の度数や層ごとの効果指標を求めることができる。

まずリスク比の場合は、最初のパラメータを以下のように定める。

- ・  $RR_{DC}$ ：未知の交絡要因（ $C$ ）とぜん息（ $D$ ）の関連性をしめすリスク比（交絡要因あり / 交絡要因なし）。ただし、このリスク比は曝露の有無によらず等しいと仮定する。

すると、層別後のリスク比、つまり交絡の影響を補正したリスク比  $RR$  は

$$RR = \frac{A_{11}/A_{01}}{N_{11}/N_{01}} = \frac{A_{10}/A_{00}}{N_{10}/N_{00}} = \frac{A_1/A_0}{N_1/N_0} \cdot \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}P_{C1} + 1 - P_{C1}} \quad \dots \quad \text{式 3-2}$$

となる（参考 3- 2）。

また、オッズ比の場合は、最初のパラメータを以下のように定める。

- ・  $OR_{DC}$ ：未知の交絡要因（ $C$ ）とぜん息（ $D$ ）の関連性をしめすオッズ比（交絡要因あり / 交絡要因なし）。ただしこのオッズ比は、曝露の有無によらず等しいと仮定する。

すると、層別後のオッズ比、つまり交絡の影響を補正したオッズ比  $OR$  は近似的に

$$OR = \frac{A_{11}B_{01}}{B_{11}A_{01}} = \frac{A_{10}B_{00}}{B_{10}A_{00}} = \frac{A_1B_0}{B_1A_0} \cdot \frac{OR_{DC}P_{C0} + 1 - P_{C0}}{OR_{DC}P_{C1} + 1 - P_{C1}} \quad \dots \quad \text{式 3-3}$$

となる（参考 3- 3）。

以上でしめしたのは、いずれも、未知の交絡要因とぜん息の関連性が曝露の有無によらず等しいと想定した場合の関係である。このため、層別後のリスク比やオッズ比が層によらず等しくなる。この想定を変えたい場合は、リスク差の場合と同様にパラメータを増やして対応する（参考 3- 2、及び参考 3- 3）。

### 3.1.5 数値例

以上で求めた関係を、リスク比について WTCHR の数値例にあてはめてみよう。

まず、未知の交絡要因によるぜん息のリスク比を 2 とする（ $RR_{DC} = 2$ ）。さらに、交絡要因ありの対象者の割合が、曝露ありの集団で 90%、なしの集団で 50%とする（ $P_{C1} = 0.9$ 、 $P_{C0} = 0.5$ ）。この想定は、ぜん息に対するリスクが交絡要因ありの対象者で高く、かつそのような対象者が曝露ありの集団に多いせいで、曝露によるぜん息のリスクが過大評価されているのではないか、という内容である。これらのパラメータの値を対応する数式

（3.1.2、及び参考 3- 2）に代入すると、未知の交絡要因の有無別に層別した分割表が得られる（Table 3- 2）。

Table 3- 2 交絡要因の有無別に層別した分割表（ $RR_{DC} = 2$ 、 $P_{C1} = 0.9$ 、 $P_{C0} = 0.5$ ）

アウト カム： ぜん息	合計 (元のデータ)		交絡要因			
			$C = 1$		$C = 0$	
	曝露：粉じん		曝露：粉じん		曝露：粉じん	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	476	450	450.9	300	25.1	150
$D = 0$	9,286	15,536	8,334.9	7,693	951.1	7,843
合計	9,762	15,986	8,785.8	7,993	976.2	7,993
リスク比	1.73		1.37		1.37	

あるいは、こうした分割表を経由することなく、式 3-2 をもちいてリスク比の補正値を求めてもよい。すなわち

$$RR = \frac{A_1}{N_1} \bigg/ \frac{A_0}{N_0} \cdot \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}P_{C1} + 1 - P_{C1}} = 1.732 \cdot \frac{2 \cdot 0.5 + 1 - 0.5}{2 \cdot 0.9 + 1 - 0.9} = 1.37.$$

交絡要因の有無によらず、リスク比は 1.37 であり、交絡のせいでリスク比が関連性を強める方向に偏っていた（1.37 → 1.73）可能性が示唆される結果となっている。

以上は、まったく仮想的な数値例だが、元の研究でしめされていた調整済みオッズ比は 1.28 と、1.37 に近い値であった（2.2、及び参考 2- 1）。また、この研究で行われた部分集団解析の結果によると、作業を開始した日によってぜん息のリスクが 1.5 倍～3 倍ほど異なっていた。とくに、当日から作業に従事した対象者は全体の約 40%を占め、こうした対象者は、曝露ありの集団により多く含まれていたと考えられる。したがって、ぜん息のり

スクを2倍ほど増加させる要因が、曝露の有無により約2倍異なる割合で分布しているという想定は、さほど不自然ではないのかもしれない。

### 3.1.6 交絡に対する感度分析と系統誤差

あらためて、交絡によって生じたバイアスを補正する過程を振り返ってみよう。まず、リスク差の補正にもちいた式3-1には、層別後の、すなわち補正後のリスク差  $RD$  と、層別前の、すなわち未補正のリスク差  $(A_1/N_1 - A_0/N_0)$  がともに含まれていた。未補正のリスク差を  $RD^*$  で表すと、式3-1は以下のように書ける。

$$RD = RD^* - (P_{C1} - P_{C0})RD_{DC}$$

この式は、1.3.2 でしめした

$$(\text{真の値}) \equiv (\text{推定値}) - (\text{系統誤差})$$

という式に対応している。言い換えれば、未補正のリスク差  $RD^*$  は、交絡のせいで、あるいはそれを無視した解析のせいで、真のリスク差  $RD$  から、項  $(P_{C1} - P_{C0})RD_{DC}$  の大きさだけ偏っている。

リスク比やオッズ比についても同様な関係がなりたつ。未補正の推定値に \* を付けて表すと、式3-2、及び式3-3は、以下のように書ける。

$$RR = RR^* / \frac{RR_{DC}P_{C1} + 1 - P_{C1}}{RR_{DC}P_{C0} + 1 - P_{C0}}, \quad OR \equiv OR^* / \frac{OR_{DC}P_{C1} + 1 - P_{C1}}{OR_{DC}P_{C0} + 1 - P_{C0}}.$$

これらは、いずれも

$$(\text{真の値}) \equiv (\text{推定値}) / (\text{系統誤差})$$

という式に対応している。系統誤差の影響が加法的か乗法的かという違いはあるが、よく似た方法で系統誤差を除去した推定値を求めていることがわかる。同様な関係は、回帰モデルによって得られた結果に対する感度分析でももちいられる（参考3-4）。また、同様な関係を不等式で評価することで、もっと一般的な場合に拡張することができる<sup>22</sup>。例えば、アウトカムの発生に何倍の影響をもつ未知の交絡要因が、曝露の有無により何倍異なると分布していると、帰無仮説が正しいのに（例えば  $RR = 1$  なのに）効果指標が観察値と等しくなるか（例えば  $RR^* = 2$  となるか）、という値（E-value<sup>23</sup>）を求め、感度分析に役立てる方法が提案されている（参考3-4）。

交絡については、以上で見たように、未補正の推定値と系統誤差に相当する項を分離することができる。このため、未補正の推定値とパラメータの値のみをもちいて、交絡の影響を補正することが可能である。すなわち、層別した分割表を経由することなく、上の関係式をそのままもちいて感度分析を行うことができる。

さらに、交絡については、もっと別のパラメータをもちいた感度分析も考えられる（参考3-5）。パラメータを変えると、当然ながら補正式も変わる。これらは数学的には同値だとしても、パラメータの値を定めるときの考えやすさ、想定のしやすさ、また計算の手間等の面で多少の違いがある。どのようにパラメータを設定するかは事例ごとに、また解析

者によっても異なるかもしれない。このように、感度分析の方法には自由度があり、それだけに解析者の意図や判断が問われる場面が少なくない。以下でも、複数の方法が考えられる場合は、説明や情報を補足する。

## 3.2 誤分類に対する感度分析

### 3.2.1 差別的誤分類と従属的誤分類

効果指標の推定値を求めるためには、曝露やアウトカム、また交絡要因に関するデータが必要である。これらの値に誤差が含まれていると、それをもちいて計算した推定値にもバイアスが生じる懸念がある。入門編でもふれたように、こうしたデータの誤りは測定誤差の一種であり、曝露やアウトカムの有無といったカテゴリーデータに対する測定誤差を分類誤差 (classification error)、または**誤分類** (miss classification) と呼ぶ。本書では、おもに誤分類と呼ぶことにする。

曝露やアウトカム、また交絡要因のデータに誤差がふくまれていたとしても、それがランダムに生じているならバイアスは生じないのではないかと考えたい。実際、とくに問題があるとされているのは、以下に示した、ランダムでない誤分類、つまり他の要因の影響を受けて系統的に生じる誤分類である。

- **差別的誤分類 (differential misclassification)**  
ある要因の誤分類が、それ以外の要因の**真の値**の影響を受けて生じるとき、その誤分類は差別的 (differential) であると言う。例えば、曝露の誤分類が、アウトカムなしの場合に多く生じる、といった場合である。逆に、ある要因の誤分類が、それ以外の要因の真の値と無関係に (独立に) 生じているとき、その誤分類は非差別的 (non-differential) であると言う。なお、本書では differential を「差別的」と訳したが、ほかにも様々な訳語がもちいられている (参考 3- 6)。
- **従属的誤分類 (dependent misclassification)**  
ある要因の誤分類が、それ以外の要因の**誤分類**の影響を受けて生じるとき、その誤分類は従属 (dependent) であると言う。例えば、曝露とアウトカムの双方に誤分類が生じていて、両者に関連性がある場合である。逆に、こうした関連性がないとき、その誤分類は独立 (independent) であると言う。

差別的誤分類が、ある要因の測定誤差と、それ以外の要因の真の値との関係 (測定誤差 vs 真の値) であるのに対し、従属的誤分類は、ある要因の測定誤差と、それ以外の要因の測定誤差との関係 (測定誤差 vs 測定誤差) である。差別的か非差別的か、また従属か独立かによって、全部で 4 つの場合がある。簡単な事例を参考として紹介した (参考 3- 6)。

以上のように場合をわけると、ある要因に対する非差別的かつ独立な誤分類は、他の要因と切り離して考えることができるので、直観的に考えると「性質がよい」ように思える。実際、非差別的で独立な誤分類にはこの印象を裏付けるような性質がある。しかしながら、とりあえず結論だけ述べると、比較的「性質がよい」とされる非差別的で独立な誤分類についても十分な注意が必要である。というのは、次節で補足するように、ある試験で非差異性、もしくは独立性が厳密に成立している保証はなく、わずかな差異性や従属性が無視



できないバイアスをもたらす懸念があるからである。したがって、ある研究のバイアスの評価を、こうした一般的な性質に頼って済ませる、あるいは済ませたつもりになるのは危険である。個々の研究に即して、それぞれにバイアスの影響を評価することの重要性を、あらためて強調したい。

関心がある読者のために、比較的「性質がよい」とされる非差別的で独立な誤分類について一般的に知られている事項を、以下で大まかに紹介した。

### 3.2.2 補足：非差別的で独立な誤分類は「性質がよい」のか？

もっとも単純な場合として、曝露もアウトカムも 2 値で、その一方が非差別的で独立な誤分類を受けている場合を考えよう。このとき以下がなりたつ（参考 3- 8）。

- ・ 曝露とアウトカム（真の値）に関連性がない場合、非差別的で独立な誤分類を受けている曝露やアウトカム（測定値）をもちいて効果指標を計算すると、平均的には、関連性がないことをしめす推定値が得られる。すなわち、曝露とアウトカムに関連性がないという**帰無仮説**が正しいとき、それを覆す結果は生じにくい。このことを「**帰無仮説のもとでのバイアス（bias under the null）**がない」と言うことがある。
- ・ 曝露とアウトカム（真の値）に関連性がある場合、非差別的で独立な誤分類を受けている曝露やアウトカム（測定値）をもちいて効果指標を計算すると、平均的には、関連性がないことをしめす方向にバイアスした推定値が得られる。すなわち、曝露とアウトカムに関連性があるという**対立仮説**が正しいとき、バイアスは生じるが、その方向性を予測することができる。このことを「**対立仮説のもとでのバイアス（bias under the alternative）**はあるが、それは帰無仮説方向へのバイアス（**bias toward the null**）である」と言うことがある。

以上の性質、すなわち帰無仮説のもとではバイアスがなく、対立仮説のもとではバイアスがあるにせよその方向が予測できることから、非差別的で独立な誤分類は「性質がよい」と言われることが多い。あるいは「独立である」という条件を暗黙のうちに仮定しつつ、非差別的な誤分類は問題が少ない、あるいはこうした誤分類では帰無仮説方向へのバイアスしか生じない、と強調されていたこともあった<sup>8, 16</sup>。

しかし、こうした性質は、例えば曝露の水準が 3 以上の場合には成り立たない<sup>24, 25</sup>。また、非差別的性、もしくは独立性が近似的にしか成り立たない場合、わずかと思える差別的性や従属性により関連性を強める方向にバイアスが生じたり、関連性の方向が逆転することがある<sup>26</sup>。

さらに、以上で述べたことは、いずれも期待値についてなりたつ性質であることに留意したい。つまり、非差別的で独立な誤分類によって平均的には関連性を弱める方向にバイアスが生じるのだとしても、それが個々の研究で実現しているとは限らない。例えば、ある方法で曝露を分類したときの感度が、アウトカムの有無によらず 90%だとする。おなじ方法で、真に曝露ありの対象者をアウトカムの有無ごとに 10 人ずつ分類したとき、正しく曝露ありに分類される対象者が、アウトカムの有無によらずぴったり 9 人になるとは限ら

ない。このように、曝露分類に関する感度が真に非差別的でも、その研究ではアウトカムの有無により異なる感度が実現しているかもしれない。

以上から、比較的「性質がよい」とされる、非差別的で独立な誤分類についても十分な注意が必要である。また、このことから、誤分類の一般的な性質に頼ってバイアスの評価を済ませるのではなく、個々の研究に即した形で誤分類の影響を評価することが重要である。そのための方法を、以下で見たい。

### 3.2.3 感度分析のためのパラメータ

以下では、2 値の曝露、アウトカム、及び交絡要因の誤分類について述べる。その際、誤分類を受けた値（測定値）を、その要因を表す記号にアスタリスクを付けて表す。例えば、曝露、アウトカム、及び交絡要因の真の値  $E$ 、 $D$ 、及び  $C$  に対し、その測定値を  $E^*$ 、 $D^*$ 、及び  $C^*$  で表す。2 値の要因については、真の値  $X$  とその測定値  $X^*$  に対して 4 つの組み合わせが考えられる。入門編でも述べたように、これら 4 つの組み合わせを、診断における偽陽性や偽陰性等と同様な考え方で、以下のように対応させる（Table 3- 3）。

Table 3- 3 真の値と測定値の関係

真の値： $X$	測定値： $X^*$		合計
	1	0	
1	Tp	Fn	Tp+Fn
0	Fp	Tn	Fp+Tn
合計	Tp+Fp	Fn+Tn	

Tp : True positive、Fn : False negative、Fp : False positive、Tn : True negative.

さらに、感度（Se）、特異度（Sp）、陽性的中率（PPV）、陰性的中率（NPV）を、以下のように定義し、感度分析の際にパラメータとしてもちいる。

$$Se = \frac{Tp}{Tp + Fn}, Sp = \frac{Tn}{Fp + Tn}, PPV = \frac{Tp}{Tp + Fp}, NPV = \frac{Tn}{Fn + Tn}$$

また、差別的な誤分類をあつかう場合は下付きの記号をもちいる。例えば曝露分類の感度を、アウトカムありのとき  $Se_A$ 、アウトカムなしのとき  $Se_B$  と表す。またアウトカム分類の感度を、曝露ありのとき  $Se_1$ 、曝露なしのとき  $Se_0$  と表す。

以下では、曝露の誤分類、アウトカムの誤分類、交絡要因の誤分類の順に、単純な感度分析の方法を紹介しよう。

## 3.3 曝露の誤分類

### 3.3.1 感度分析の方針

WTCHR の事例について考えよう。この事例では、曝露の有無を「2001 年 9 月 11 日、戸外で、WTC の崩落によって生じた粉じんや破片の雲のなかにいましたか」という質問によっ

て調べていた。このため、回答者によって「粉じんや破片の雲」に対する解釈が異なるのではないか、という懸念が指摘されている<sup>27</sup>。この研究に対して行われた感度分析でも、曝露の誤分類の可能性を指摘し、検証研究の必要性を示唆しているが、実際に感度分析を行っているわけではない<sup>28</sup>。以下では、あくまでも仮想的な数値例として、曝露の誤分類の影響を検討してみよう。

曝露の分類が誤っていると、研究から得られた分割表のデータも誤っているはずである。すると、この測定値による分割表とは別に、真の値による分割表があるはずである。この、いわば真の分割表を、適切なパラメータを定めることで復元してみよう、というのが曝露の誤分類に対する感度分析の大まかな方針である。WTCHR のデータについても、誤分類がない場合の分割表が復元できたとしよう (Table 3-4)。ただし、データとして得られていない度数を青字でしめした。

Table 3-4 曝露に対する誤分類の分割表

アウトカム： ぜん息	曝露の誤分類			
	なし		あり (元のデータ)	
	曝露 (真の値) : 粉じん		曝露 (測定値) : 粉じん	
	$E = 1$	$E = 0$	$E^* = 1$	$E^* = 0$
$D = 1$	$A_1$	$A_0$	$A_1^* = 476$	$A_0^* = 450$
$D = 0$	$B_1$	$B_0$	$B_1^* = 9,286$	$B_0^* = 15,536$
合計	$N_1$	$N_0$	$N_1^* = 9,762$	$N_0^* = 15,986$

### 3.3.2 感度分析のためのパラメータと分割表のデータ

以下では、アウトカムの有無によらず曝露の誤分類の程度が等しい場合、つまり非差別的誤分類を想定し、この数値例について調べてみよう。なお、曝露の誤分類について検討する際、アウトカムには誤分類がないことを仮定する。このとき、感度分析に必要なパラメータは以下の2つである。

- Se : 曝露分類の感度。
- Sp : 曝露分類の特異度。

いま、アウトカムについては誤分類がないことを想定しているので、例えばアウトカムありの人数の合計は、曝露の誤分類の有無によらず等しいはずである。すなわち、 $A_1 + A_0 = A_1^* + A_0^*$  になりたつ。したがって分割表のデータは、曝露の誤分類により、いわば列方向にのみ移動する。また、感度と特異度の定義から、アウトカムありの人数 ( $A_1$  と  $A_0$ ) については、以下になりたつはずである。

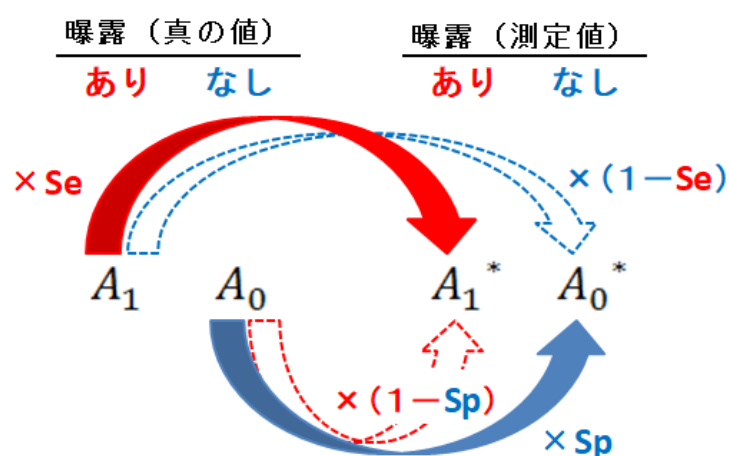
- 真に曝露ありの対象者  $A_1$  人は
  - 確率 Se (感度) で正しく分類され、測定値も曝露ありになる ( $A_1^*$  に含まれる)。
  - 確率  $1 - \text{Se}$  ( $1 - \text{感度}$ ) で誤って分類され、測定値は曝露なしになる ( $A_0^*$  に含まれる)。
- 真に曝露なしの対象者  $A_0$  人は



- 確率  $Sp$  (特異度) で正しく分類され、測定値も曝露なしになる ( $A_0^*$  に含まれる) .
- 確率  $1 - Sp$  ( $1 - \text{特異度}$ ) で誤って分類され、測定値は曝露ありになる ( $A_1^*$  に含まれる) .

参考までに、以上の関係にしたがって分割表のデータ (例えばアウトカムありの場合の度数) が移動する様子を図示した (Figure 3-1) .

Figure 3-1 曝露の誤分類による度数の移動



以上の関係を数式で表すと、以下が得られる.

$$A_1^* = Se A_1 + (1 - Sp) A_0, A_0^* = (1 - Se) A_1 + Sp A_0.$$

アウトカムなしの場合にも同様な関係がなりたつので

$$B_1^* = Se B_1 + (1 - Sp) B_0, B_0^* = (1 - Se) B_1 + Sp B_0.$$

これらを解くことで、誤分類がない分割表のデータが得られる. すなわち

$$A_1 = \frac{Sp A_1^* - (1 - Sp) A_0^*}{Se + Sp - 1}, A_0 = \frac{Se A_0^* - (1 - Se) A_1^*}{Se + Sp - 1},$$

$$B_1 = \frac{Sp B_1^* - (1 - Sp) B_0^*}{Se + Sp - 1}, B_0 = \frac{Se B_0^* - (1 - Se) B_1^*}{Se + Sp - 1}.$$

さらに、合計についても

$$N_1 = \frac{Sp N_1^* - (1 - Sp) N_0^*}{Se + Sp - 1}, N_0 = \frac{Se N_0^* - (1 - Se) N_1^*}{Se + Sp - 1}.$$

こうして、誤分類がない分割表のデータが得られたので、曝露の誤分類の影響を補正した効果指標を求めることができる. 例えば、曝露の誤分類の影響を補正したリスク比  $RR$  は

$$RR = \frac{A_1/N_1}{A_0/N_0} = \frac{Sp A_1^* - (1 - Sp) A_0^*}{Se A_0^* - (1 - Se) A_1^*} \bigg/ \frac{Sp N_1^* - (1 - Sp) N_0^*}{Se N_0^* - (1 - Se) N_1^*} \quad \dots \quad \text{式 3-4}$$

また、リスク差 $RD$ は

$$RD = \frac{A_1}{N_1} - \frac{A_0}{N_0} = \frac{SpA_1^* - (1 - Sp)A_0^*}{SpN_1^* - (1 - Sp)N_0^*} - \frac{SeA_0^* - (1 - Se)A_1^*}{SeN_0^* - (1 - Se)N_1^*} \quad \dots \quad \text{式 3-5}$$

となる。

交絡に対する感度分析と事情が異なるのは、効果指標ごとに異なるパラメータを設定する必要がないことである。一方、式 3-4 や式 3-5 がしめしているように、誤分類の場合は、未補正の効果指標と系統誤差に相当する項を、わかりやすい形に分離することができない。このため、誤分類に対する感度分析では、かならず分割表を経由して補正值を求めることになる。このことは、単純な感度分析を行う際には問題とならない。しかし、確率的感度分析を行う際には、このせいで計算の量や手順が増えることがある。

なお、アウトカムの有無により曝露の誤分類の程度が異なる場合、すなわち差別的誤分類が想定される場合は、パラメータの数こそ増えるものの、ほぼ同様な考え方で誤分類がない分割表のデータや、効果指標の補正值を求めることができる（参考 3- 7）。

### 3.3.3 計算例

以上で求めた関係を、数値例にあてはめてみよう。

粉じんへの曝露がありだった対象者のうち 90%は正確にありと報告していると仮定すると感度は 0.9 である（ $Se = 0.9$ ）。一方、曝露なしだった対象者のうち 20%の対象者は、実際には粉じんを遠方から目視しただけなのに、その内部にいたと錯覚して曝露ありと報告をしていたと仮定し、特異度を 0.8 としてみよう（ $Sp = 0.8$ ）。これらの値を対応する数式（3.3.2）に代入すると、以下の分割表が得られる（Table 3- 5）。

なお、曝露の誤分類の影響を補正したリスク比を求めるだけなら、式 3-4 をもちいて

$$RR = \frac{0.8 \cdot 476 - (1 - 0.8) \cdot 450}{0.9 \cdot 450 - (1 - 0.9) \cdot 476} \bigg/ \frac{0.8 \cdot 9762 - (1 - 0.8) \cdot 15986}{0.9 \cdot 15986 - (1 - 0.9) \cdot 9762} = 2.365 \dots$$

としてもよい。

Table 3- 5 曝露の誤分類を補正した分割表（ $Se = 0.9$ 、 $Sp = 0.8$ ）

アウトカム： ぜん息	曝露の誤分類			
	なし		あり（元のデータ）	
	曝露（真の値）：粉じん $E = 1$	$E = 0$	曝露（測定値）：粉じん $E^* = 1$	$E^* = 0$
$D = 1$	415.4	510.6	476	450
$D = 0$	6,173.7	18,648.3	9,286	15,536
合計	6,589.1	19,158.9	9,762	15,986
リスク比	2.37		1.76	

この例では、曝露の誤分類により、リスク比の値が関連性を弱める方向に偏っていた（2.37 → 1.76）可能性が示唆されている。この数値例では、アウトカムによらず曝露の誤分類の程度が等しいこと、また曝露の誤分類が他の要因とは無関係であることを想定している。すなわち、非差別的で独立な誤分類を想定している。このため、関連性を弱める方向にバイアスが生じているのは自然な結果であると考えられる（3.2.2、及び参考 3- 8）。

### 3.3.4 補足：他のパラメータ設定

感度と特異度ではなく、陽性的中率と陰性的中率をパラメータとして補正値を求めることができる（参考 3- 7）。しかし、曝露の誤分類については、陽性的中率と陰性的中率をもちいた補正が行われることは稀である。というのは、感度と特異度がもつばら分類の正確さ（真の値から見たときの）を表す量であるのに対し、陽性的中率と陰性的中率は、さらに曝露の程度（曝露ありの割合）にも依存しているからである<sup>16</sup>。

曝露の程度は部分集団により、また時期により様々に異なる可能性がある。このため、他の研究から得られた陽性的中率や陰性的中率は、あまり参考にならないかもしれない。仮に内部検証研究を行い、その研究における陽性的中率と陰性的中率を調べることができたとしても、その対象者が本研究の対象者からランダムに選択されている場合を除けば、交絡要因やアウトカムの有無別に調べる必要があるかもしれない。というのは、これらの要因の組み合わせ（交絡要因の有無×アウトカムの有無）ごとに、曝露の程度が異なる可能性があるからである。

### 3.3.5 不可能な値

この数値例では生じなかったが、パラメータ（感度  $Se$  や特異度  $Sp$ ）の値によっては、誤分類の影響を補正した分割表の度数（ $A_1$ 、 $A_0$ 、 $B_1$ 、 $B_0$  等）が負になることがある。例えば、すでにしめたように、真に曝露ありでアウトカムありの対象者数  $A_1$  は

$$A_1 = \frac{SpA_1^* - (1 - Sp)A_0^*}{Se + Sp - 1}$$

と表すことができた（3.3.2）。分母の  $Se + Sp - 1$  については、かなり極端な誤分類が生じない限り負になることはない（参考 3- 8）。一方、分子が正になるのは

$$Sp > \frac{A_0^*}{A_1^* + A_0^*}$$

がなりたつ場合である。このため、特異度  $Sp$  の値が小さすぎると、 $A_1$  の値が負になる可能性がある。同様なことが他の度数についても言えるので、感度と特異度が取りうる値には制限がある。このことは、これから述べるアウトカムの誤分類や交絡要因の誤分類についても同様である。つまり、誤分類に対する感度分析では、パラメータの値を定める際、不可能な値が生じないように注意する必要がある。

### 3.3.6 補足：データベース研究における曝露の誤分類

データベース研究では、データとして利用可能な薬剤処方等の情報をもちいて曝露群を定めるのが一般的である。このとき、関心がある薬剤への曝露について生じうる誤分類として、例えば以下が考えられる<sup>28</sup>。

- ・ 曝露あり（真の値）なのに、誤って曝露なし（測定値）に分類される：  
例えば関心がある薬剤が、データベースではカバーされていない医療施設や薬局等で処方されていると、データ上は曝露なしに分類される。
- ・ 曝露なし（真の値）なのに、誤って曝露あり（測定値）に分類される：  
例えば関心がある薬剤が処方されたものの服薬されなかった場合や、薬剤が間欠的に服薬されているせいで研究者が想定しているほど服薬されていない場合であっても、データ上は曝露ありに分類される。

もっとも単純なケースとして、1つの薬剤 *X* の使用の有無により曝露群を定める場合について考えてみよう。例えば、以下のように2つの群を定めたとする。

- ・ 薬剤 *X* 使用群
- ・ 薬剤 *X* 未使用群

このような単純なケースでも、群や観察開始日にあたる時点（**time zero**、または **index date**）の定義が適切でないと、効果指標の推定値にバイアスが生じる。こうした場合も、その研究のターゲット試験を具体的に想定してみることが有益であり、ひろく一般的にもちいられている定義だから適切だとは限らない<sup>28</sup>。以下では、こうした問題には立ち入らず、群や観察開始日の定義が適切に行われていることを仮定しよう。また、薬剤 *X* がもちいられる可能性がある患者集団が、選択除外基準等により適切に定められていることも仮定しよう。

この単純なケースでは、薬剤 *X* 未使用群には、まったく薬剤を使用しなかった対象者だけでなく、もっと他の薬剤、例えば薬剤 *X* と同様な効果が期待できる薬剤 *Y* を使用した対象者も含まれているかもしれない。しかし、対象者が適切に選択されていれば、すくなくとも誤分類は、薬剤 *X* の使用・未使用という2つの値の間でしか生じない。

ところが、研究の関心が薬剤 *X* の使用と未使用の比較ではなく、薬剤 *X* と *Y* の比較にあるのだとしよう。たんに薬剤が1つ増えただけのようにも思えるが、はたしてそうだろうか。

比較的かんたんなケースとして、両者を併用する可能性がない場合について考えよう。ここで関心があるのが薬剤 *X* と *Y* の比較だとしても、誤分類が発生するメカニズムを考えると、それ以外の、いわば第3のカテゴリーを想定する必要がある。すなわち、誤分類は、以下の3つの曝露状況のあいだで生じる。

- ・ 薬剤 *X* 使用群
- ・ 薬剤 *Y* 使用群
- ・ 薬剤 *X* と *Y*、いずれも未使用群

なお「いずれも未使用群」には、まったく薬剤を使用していない対象者だけでなく、薬剤 *X* や *Y* と同様な患者集団に対してもちいられる、もっと別な薬剤 *Z* を使用した対象者も含

まれているかもしれない。このとき、以上 3 つの曝露状況（真の値）が、正しく、または誤って分類され、以上 3 つのいずれかの値（測定値）が観察される。このため、3 つの真の値が 3 つの測定値のいずれかに分類される確率、すなわち 9 個の分類確率を考えれば、こうした誤分類の発生状況をモデル化することができる。なお、曝露の水準が 3 以上の場合の誤分類については、後に述べる行列補正と呼ばれる方法で一般的にあつかうことができる（3.5.6）。

さて、研究の手順によっては、薬剤  $X$  や  $Y$  の使用がデータ上で確認できた対象者のデータしか収集していないかもしれない。すると「いずれも未使用群」の対象者のデータは利用できない。したがって以上のようなモデルが想定されたとしても、それに即した感度分析を行うことはできないかもしれない。

さらに、薬剤  $X$  と  $Y$  が併用される可能性がある場合、可能性としては以下の 4 つの群が考えられる。

- ・ 薬剤  $X$  と  $Y$  の併用群
- ・ 薬剤  $X$  のみ使用群
- ・ 薬剤  $Y$  のみ使用群
- ・ 薬剤  $X$  と  $Y$ 、いずれも未使用群

この場合、4 つの真の値が 4 つの測定値のいずれかに分類されるので、16 個の分類確率を考えることになる。ただし、この場合も「いずれも未使用群」の対象者のデータが利用できなければ、こうしたモデルをもちいて感度分析を行うことはできない。

なお、以上でしめした薬剤  $X$  と  $Y$  に関する誤分類について、両者の併用がある場合とない場合の分類確率をまとめて例示すると下表のようになる（Table 3-6）。

Table 3-6 薬剤  $X$  と  $Y$  の誤分類に関する分類確率

真の値： ( $X, Y$ )	測定値：( $X^*, Y^*$ )			
	(あり, あり)	(あり, なし)	(なし, あり)	(なし, なし)
(あり, あり)	$p_{11}$	$p_{12}$	$p_{13}$	$p_{14}$
(あり, なし)	$p_{21}$	$p_{22}$	$p_{23}$	$p_{24}$
(なし, あり)	$p_{31}$	$p_{32}$	$p_{33}$	$p_{34}$
(なし, なし)	$p_{41}$	$p_{42}$	$p_{43}$	$p_{44}$

ただし、薬剤  $X$  と  $Y$  の併用群：(あり, あり)、薬剤  $X$  のみ使用群：(あり, なし)、薬剤  $Y$  のみ使用群：(なし, あり)、薬剤  $X$  と  $Y$ 、いずれも未使用群：(なし, なし)、等と対応している。また、真の値がいずれかの測定値に分類されるので、行方向の確率の総和は 1 である。さらに、点線で囲った部分が薬剤  $X$  と  $Y$  の併用がない場合、すなわち (あり, あり) がない場合にあたる。この場合は、点線内部で行方向の総和が 1 となるように分類確率を定める。

ところで、実際の曝露状況はさらに複雑かもしれない。例えば、併用を想定する必要はなく、むしろ 2 剤の投与順序（薬剤  $X \rightarrow$  薬剤  $Y$ 、及び薬剤  $Y \rightarrow$  薬剤  $X$ ）が問題となるケースもある。たんに薬剤  $X$  や  $Y$  をもちいた対象者ではなく、それらの new user に関心があ

る場合は、投与順序も考慮して群を定め、それらのどの群の間で誤分類が生じうるかを検討すべきかもしれない。

さらに、以上ではアウトカムの有無によらない曝露の誤分類（非差別的誤分類）を想定した。しかし、これが疑わしければ、アウトカムの有無別に分類確率を定めることになるため、考慮すべきパラメータ（分類確率）はさらに増えることだろう。

いずれにせよ、現実には生じる曝露の誤分類のメカニズムは想像以上に複雑かもしれない。ただし、実際に感度分析をおこなう際、どこまで精密な、あるいは複雑なモデルをもちいるかは、また別の問題である。感度分析の意図や目的、また利用可能なデータにおうじ、問題を単純化することも実践上は有益である（2.1.2）。

また、以上ではふれなかったが、曝露の誤分類により、曝露の開始時期や曝露期間の情報が不正確になると、結果的に曝露人-時間のデータが不正確になる。すると、人-時間を分母とした発生率を含む効果指標にバイアスが生じる可能性がある（3.8）。

## 3.4 アウトカムの誤分類

### 3.4.1 感度分析の方針

アウトカムの誤分類についても、実際に研究から得られた分割表とは別に、分類の誤りを補正した真の分割表があるはずだ、と考えることができる。WTCHR の事例についても、アウトカムの誤分類がない分割表を復元してみよう（Table 3-7）。ただし、データとして得られていない度数を青字でしめした。なお、アウトカムの誤分類について検討する際、曝露には誤分類がないことを仮定する。

Table 3-7 アウトカムに対する誤分類の分割表

アウトカムの 誤分類	アウトカム： ぜん息	曝露：粉じん	
		$E = 1$	$E = 0$
なし	真の値		
	$D = 1$	$A_1$	$A_0$
	$D = 0$	$B_1$	$B_0$
	合計	$N_1$	$N_0$
あり (元のデータ)	測定値		
	$D^* = 1$	$A_1^* = 476$	$A_0^* = 450$
	$D^* = 0$	$B_1^* = 9,286$	$B_0^* = 15,536$
	合計	$N_1^* = 9,762$	$N_0^* = 15,986$

### 3.4.2 感度分析のためのパラメータと分割表のデータ

WTCHR の事例では、アウトカムであるぜん息の有無を、対象者自身の報告にもとづいて定めていた。しかし、この研究には消防士や警察官、軍関係者等も参加しており、こうした対象者にとって健康状態の悪化は失職につながりかねない。このため、こうした対象者によるアウトカムの報告は正確でない懸念があった。一方、こうした対象

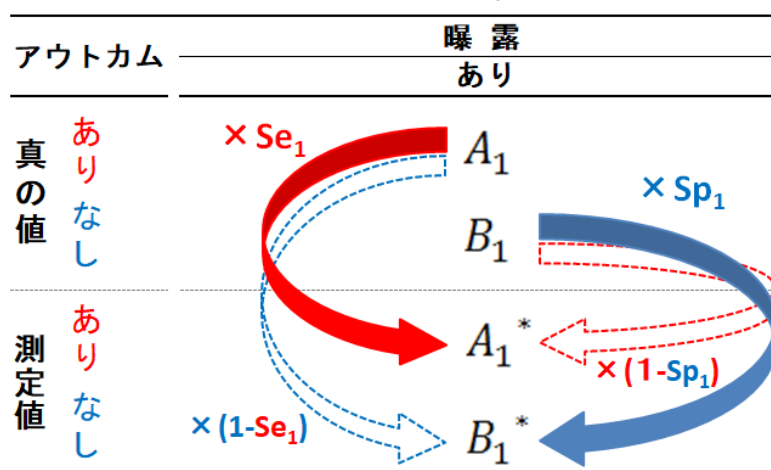
者ほど、事件当日の作業に従事しており、結果的に粉じんへの曝露を受けた可能性が高い<sup>20</sup>。こうして、曝露ありの対象者ほどアウトカムの報告が不正確である可能性、すなわち差別的誤分類が疑われた。

こうした状況を考慮して、アウトカムの誤分類について以下のパラメータを定めよう。

- $Se_1$  : 曝露ありの場合の、アウトカム分類の感度
- $Sp_1$  : 曝露ありの場合の、アウトカム分類の特異度
- $Se_0$  : 曝露なしの場合の、アウトカム分類の感度
- $Sp_0$  : 曝露なしの場合の、アウトカム分類の特異度

いま、曝露には誤分類がないことを仮定しているので、例えば曝露ありの人数の合計は、アウトカムの誤分類の有無によらず等しいはずである。つまり、 $A_1 + B_1 = A_1^* + B_1^* = N_1$ 、及び  $A_0 + B_0 = A_0^* + B_0^* = N_0$  がなりたつ。このため、分割表のデータは、いわば行方向にのみ移動する。感度と特異度の定義から、こうした度数の移動を図示すると以下のようなになる (Figure 3-2)。

Figure 3-2 アウトカムの誤分類による度数の移動  
(曝露ありの場合)



図から、誤分類が生じると（感度や特異度が1未満だと）発生割合がどの方向に偏るか、大まかな傾向が把握できる。

感度  $Se$  が1未満 →  $A$  が減少し、 $B$  が増加する → 発生割合が減少する

特異度  $Sp$  が1未満 →  $B$  が減少し、 $A$  が増加する → 発生割合が増加する

つまり、感度や特異度が低いせいで生じるバイアスは、方向としては逆になる。実際にはこれらが総合されて発生割合にバイアスをもたらすのだが、その際、アウトカムの頻度も影響する。例えばアウトカムが稀な場合、つまりアウトカムありの度数  $A$  と比較して、アウトカムなしの度数  $B$  が圧倒的に大きい場合、特異度の方が大きく影響し、



発生割合を増加させる方向にバイアスが生じる．この点については、数値例とともに後ほど確認しよう．

図でしめした関係を数式で表すと、以下が得られる．

$$A_1^* = Se_1 A_1 + (1 - Sp_1) B_1, B_1^* = (1 - Se_1) A_1 + Sp_1 B_1,$$

$$A_0^* = Se_0 A_0 + (1 - Sp_0) B_0, B_0^* = (1 - Se_0) A_0 + Sp_0 B_0.$$

これらを解くことで、誤分類がない分割表のデータが得られ、さらにアウトカムの誤分類の影響を補正した効果指標を求めることができる（参考 3- 9）．例えば、アウトカムの誤分類の影響を補正したリスク比  $RR$  は

$$RR = \frac{A_1/N_1}{A_0/N_0} = \frac{A_1^* - (1 - Sp_1)N_1^*}{A_0^* - (1 - Sp_0)N_0^*} \bigg/ \frac{(Se_1 + Sp_1 - 1)N_1^*}{(Se_0 + Sp_0 - 1)N_0^*} \quad \dots \quad \text{式 3- 6}$$

また、リスク差  $RD$  は

$$RD = \frac{A_1}{N_1} - \frac{A_0}{N_0} = \frac{A_1^* - (1 - Sp_1)N_1^*}{(Se_1 + Sp_1 - 1)N_1^*} - \frac{A_0^* - (1 - Sp_0)N_0^*}{(Se_0 + Sp_0 - 1)N_0^*} \quad \dots \quad \text{式 3- 7}$$

となる．

### 3.4.3 計算例

以上で求めた関係を、数値例にあてはめてみよう．

すでに述べたように、この事例では、ぜん息になっても正確に報告しない対象者が曝露ありの場合に多く含まれている可能性があった（3.4.2）．逆に、ぜん息がないのに、あえて「あり」と報告する可能性は、曝露ありの対象者ではむしろ低いと考えられる．つまり、アウトカムの分類について、感度は曝露ありの対象者で低く（ $Se_1 < Se_0$ ）、特異度は曝露ありの対象者で高い（ $Sp_1 > Sp_0$ ）ことが想定される．この事例に対して行われた感度分析では、さらに、ニューヨーク市消防局が行った調査（ぜん息に関する消防士本人の報告と医師による診断の比較）も参照している<sup>20</sup>．こうした情報を踏まえて設定されたパラメータの値の組み合わせのなかから、以下を取り上げよう．

- ・  $Se_1 = 0.85$ ：曝露ありの場合の、アウトカム分類の感度
- ・  $Sp_1 = 1.00$ ：曝露ありの場合の、アウトカム分類の特異度
- ・  $Se_0 = 0.90$ ：曝露なしの場合の、アウトカム分類の感度
- ・  $Sp_0 = 0.98$ ：曝露なしの場合の、アウトカム分類の特異度

つまり、感度については曝露ありの方が 5% 低く、また特異度については曝露ありの方が 2% 高い、という想定である．

これらの値を、対応する数式（参考 3- 9）に代入すると、アウトカムの誤分類を補正した分割表が得られる（Table 3- 8）．この例では、アウトカムの差別的誤分類により、リスク比の値が関連性を弱める方向に偏っていた（6.19 → 1.76）可能性が示唆されている．



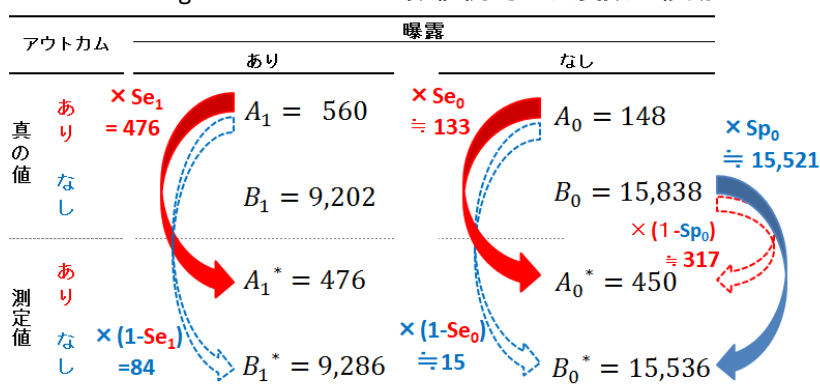
Table 3-8 アウトカムの誤分類を補正した分割表  
( $Se_1 = 0.85$ 、 $Sp_1 = 1.00$ 、 $Se_0 = 0.90$ 、 $Sp_0 = 0.98$ )

誤分類		アウトカム：ぜん息	曝露：粉じん		リスク比
			<i>E</i> = 1	<i>E</i> = 0	
なし	真の値	<i>D</i> = 1	560	148	6.19
		<i>D</i> = 0	9,202	15,838	
		合計	9,762	15,986	
あり (元のデータ)	測定値	<i>D</i> * = 1	476	450	1.76
		<i>D</i> * = 0	9,286	15,536	
		合計	9,762	15,986	

### 3.4.4 稀なアウトカムに対する誤分類の影響

以上で求めた、誤分類の影響を補正した分割表（Table 3-8）が正しいと仮定し、誤分類のせいで分割表のデータが移動する過程をふりかえってみよう。誤分類なしの表から、誤分類ありの表のどこへ何人移動したか、参考までに図示した（Figure 3-3）。

Figure 3-3 WTCHR の数値例での、度数の移動



この数値例では、曝露ありのときの特異度 ( $Sp_1$ ) が 1 なので、曝露ありの対象者については、感度 ( $Se_1 = 0.85$ ) のみの影響でアウトカムありの対象者が 85%に減少している ( $560 \times 0.85 = 476$ )。逆にアウトカムなしの対象者は増加するので、結果的に、曝露ありの群では発生割合が減少する方向に偏りが生じている。

一方、曝露なしの対象者では、むしろアウトカムありの対象者が約 3 倍に増加している ( $148 \rightarrow 450$ )。つまり、曝露なしの群では、感度が低いことで生じるアウトカムありの対象者の減少を、特異度が低いことで生じるアウトカムありの対象者の増加がはるかに上回っている。こうして、曝露なしの群では発生割合が増加する方向に偏りが生じている。

以上から、誤分類のせいで曝露ありの群では発生割合が減少し、逆に曝露なしの群では発生割合が増加したため、リスク比を大きく減少させる偏りが生じたことになる。

このように考えると、補正前後でリスク比が大きく異なるのは、曝露の有無による 5%の感度の違い (0.85 と 0.90) よりも 2%の特異度の違い (1.00 と 0.98) に起因する。特異度の差は 2%とわずかに思えるかもしれない。しかし、この事例ではアウトカムが稀なため、アウトカムなしの対象者が多数を占める。このため感度より特異度のほうが、より大きくバイアスに寄与したものと考えられる。この事例に限らず、関心があるアウトカムの発生が稀な場合は、わずかに思える特異度の差異性が、効果指標に大きな偏りをもたらすことがあるので注意したい。

この事例における感度と特異度の影響については、さらに多次元感度分析により検討する (3.9.2)。

### 3.4.5 アウトカムの誤分類によってバイアスが生じない場合

アウトカムの誤分類については、仮に誤分類があったとしても効果指標の推定値にバイアスが生じない、もしくはその方向が予測できる場合がある。それは、特異度が曝露の有無によらず 1 ( $Sp_1 = Sp_0 = 1$ ) で、かつ感度が非差異的 ( $Se_1 = Se_0 = Se < 1$ ) な場合である。

このとき、リスク比については式 3-6 から

$$RR = \frac{A_1^* - (1 - 1)N_1^*}{A_0^* - (1 - 1)N_0^*} \bigg/ \frac{(Se + 1 - 1)N_1^*}{(Se + 1 - 1)N_0^*} = \frac{A_1^*/N_1^*}{A_0^*/N_0^*} = RR^*$$

となり、誤分類の影響を補正したリスク比  $RR$  と、未補正のリスク比  $RR^*$  が一致する。

また、リスク差については式 3-7 から

$$RD = \frac{A_1^* - (1 - 1)N_1^*}{(Se + 1 - 1)N_1^*} - \frac{A_0^* - (1 - 1)N_0^*}{(Se + 1 - 1)N_0^*} = \frac{1}{Se} \left( \frac{A_1^*}{N_1^*} - \frac{A_0^*}{N_0^*} \right), \text{ よって } RD^* = Se \cdot RD < RD$$

となる。つまりリスク差については、常に関連性を弱める方向にバイアスが生じる。例えば、感度が 0.8 なら、誤分類後のリスク差は、真のリスク差の 80%の大きさになる。

なお、同様な条件下で、オッズ比や発生率比 (incidence rate ratio) についても、関連性を弱める方向にバイアスが生じることが知られている。ただし、アウトカムが稀なら、オッズ比はリスク比と近い値を取るので、バイアスの影響は小さくなるし、アウトカムの誤分類が発生率の分母である人-時間にあたえる影響が小さければ、発生率比についても大きなバイアスは生じない<sup>16</sup>。

### 3.4.6 アウトカムの誤分類をめぐる留意点

以上で概観したように、アウトカムが稀な場合は特異度の影響が大きく、また特異度が 1 なら効果指標にバイアスが生じない場合があるのだった。これらを踏まえると、ある方法でアウトカムの有無を調べ、それが「あり」に分類されたとき、確かに (真に) アウトカムありであること、つまり偽陽性でないことを十分に確認することの重要性が伺える。

例えば、こうした検討が十分に行われ、偽陽性がないこと、すなわち  $FP = 0$  であることが確認できたとしよう。すると特異度が 1 であること、すなわち

$$Sp = \frac{Tn}{Fp + Tn} = \frac{Tn}{0 + Tn} = 1$$

が導かれる (3.2.3) . ただし、既に述べたように、曝露の有無によらず特異度が 1 でも例えばリスク差にはバイアスが生じる懸念がある (3.4.5) . またリスク比についても、特異度が 1 だけでは不十分だった. もう 1 つの条件、すなわち感度が曝露の有無によらず等しいこと、つまり感度が非差別的であることも重要であり、2 つの条件がともに成立しないと、リスク比にもバイアスが生じるのだった (3.4.5) .

2 つめの条件、すなわち感度が非差別的であることは、アウトカムを調べる方法が曝露の有無によらず同様なら、基本的には保障されるように思える. しかし、例えばデータベースから、曝露の有無によらずまったくおなじ条件でアウトカムの有無を判断したとしても、それだけで誤分類の非差別性が保障されるわけではない. 例えば、ある新薬で治療を開始する際、主治医によっては臨床検査の種類や頻度を増やし、より慎重な経過観察を行うかもしれない. こうした、健康を追求する行動 (health seeking behavior) が新薬群でより多く認められると、新薬群でより多くのアウトカムが観察される可能性がある. すると結果的に、アウトカムの誤分類の程度も曝露の有無により異なる可能性がある<sup>28</sup>.

さらに、以上で紹介した性質も、あくまでも期待値に対する議論であることに注意したい. 例えば、リスク比に対して以上で述べたような性質が平均的になりたつのだとしても、個々の研究でそれが実現しているとは限らない. こうした一般的な性質は、よりバイアスの少ない研究を行う上でのよい指針となりうる. しかし、こうした一般的な性質のみに頼って、個々の研究のバイアスを評価するのは早計である. 例えば、アウトカムの分類に偽陽性がないこと、すなわち特異度が 1 であることが保障できれば十分であり、少しばかり感度が低くても妥当性が高い研究ができるはずだ、などと考えるのは短絡的かもしれない.

## 3.5 交絡要因の誤分類

### 3.5.1 感度分析の方針

2 値の交絡要因に誤分類が生じている場合も、曝露やアウトカムの誤分類と基本的な考え方はかわらない. ただし、曝露とアウトカムについては誤分類がないことを想定する.

WTCHR の事例について、計算例として、未知の交絡要因の有無別に層別した分割表をしめした (Table 3- 2) . 以下では、このデータが実際に得られたものと仮定しよう. ただし、度数が整数となるよう四捨五入したデータをもちいる. また、この交絡要因の値は誤分類を受けており、すでにしめた層別解析の結果は誤っている可能性がある、と考えよう.

すると、測定値として得られている、誤分類を受けた交絡要因の値 ( $C^*$  の値) とは別に、誤分類がない正しい値 ( $C$  の値) があるはずである. また、この正しい値をもちいて層別すると、あらたな分割表が得られるはずである (Table 3- 9) . これまでと同様に、この真の分割表を復元してみよう. ただし、データとして得られていない度数を青字でしめた.

Table 3-9 交絡要因に対する誤分類の分割表

誤分類	ぜん息	合 計		交絡要因			
		粉じんへの曝露		$C = 1$		$C = 0$	
		$E = 1$		$E = 1$		$E = 1$	
		$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
なし	$D = 1$	$A_1 = 476$	$A_0 = 450$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
	$D = 0$	$B_1 = 9,286$	$B_0 = 15,536$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
				$C^* = 1$		$C^* = 0$	
				$E = 1$	$E = 0$	$E = 1$	$E = 0$
あり	$D = 1$			$A_{11}^*$ = 451	$A_{01}^*$ = 300	$A_{10}^*$ = 25	$A_{00}^*$ = 150
	$D = 0$			$B_{11}^*$ = 8,335	$B_{01}^*$ = 7,693	$B_{10}^*$ = 951	$B_{00}^*$ = 7,843

### 3.5.2 感度分析のためのパラメータと分割表のデータ

交絡要因の誤分類についても、感度や特異度をパラメータとしてもちいる。ただし以下では、誤分類がアウトカムの有無（真の値）によらず等しい非差別的な誤分類を想定し、以下の2つのパラメータを設定する。

- ・  $Se$  : 交絡要因の分類に関する感度
- ・  $Sp$  : 交絡要因の分類に関する特異度

いま、曝露についてもアウトカムについても誤分類がないと想定しているので、例えば曝露ありで、かつアウトカムありの人数の合計は、交絡要因の誤分類の有無によらず等しいはずである。つまり、 $A_{11}^* + A_{10}^* = A_{11} + A_{10} = A_1$ 、 $A_{01}^* + A_{00}^* = A_{01} + A_{00} = A_0$ 、 $B_{11}^* + B_{10}^* = B_{11} + B_{10} = B_1$ 、 $B_{01}^* + B_{00}^* = B_{01} + B_{00} = B_0$ 等がなりたち、分割表のデータは、行も列もおなじまま、層間でのみ移動する。

こうした度数の移動を、感度と特異度の定義と結びつけることで、誤分類のない交絡要因で層別した分割表が得られる（参考 3- 10）。

### 3.5.3 計算例

こうして得られた関係を、数値例にあてはめてみよう。

この数値例はまったく架空の内容だが、アウトカムの有無によらず感度を  $0.97$  ( $Se_A = Se_B = Se = 0.97$ )、特異度を  $1.0$  ( $Sp_A = Sp_B = Sp = 1$ ) としてみよう。これらの値を対応する数式（参考 3- 10）に代入すると、交絡要因の誤分類を補正した分割表が得られる（Table 3- 10）。

補正後の分割表を見ると、交絡要因ありの場合のリスク比は **1.38**、なしの場合のリスク比は **0.86** と、交絡要因の有無によりリスク比の方向が逆になっている。したがって、この要

因が実際に効果指標を修飾しているのだとしたら、誤分類により、それが見えなくなっていたことになる。

Table 3- 10 交絡要因の誤分類を補正した分割表 (Se = 0.97、Sp = 1.0)

誤分類	アウトカム： ぜん息	交絡要因					
		合 計		C = 1		C = 0	
		曝露：粉じん		曝露：粉じん		曝露：粉じん	
		E = 1	E = 0	E = 1	E = 0	E = 1	E = 0
なし	D = 1	476	450	464.9	309.3	11.1	140.7
	D = 0	9,286	15,536	8592.8	7930.9	693.2	7605.1
	合計	9,762	15,986	9057.7	8240.2	704.3	7745.8
	リスク比	1.73 (MH: 1.33)		1.38		0.86	
あり		C* = 1		C* = 0			
		E = 1	E = 0	E = 1	E = 0	E = 1	E = 0
あり	D = 1	451	300	25	150		
	D = 0	8,335	7,693	951	7,993		
	合計	8,786	7,993	976	7,993		
	リスク比	MH: 1.37		1.37		1.37	

MH: Mantel-Haenszel の方法で層を併合したリスク比

一方、実際には効果指標の修飾はなく、層によらず共通の効果を仮定してよい場合は、補正後のリスク比（層を併合したリスク比）は 1.33 であり、観察されたリスク比 1.37 は、誤分類によりわずかに大きな値に偏っていたことになる。

以下では、この事例と関連して、効果指標の修飾について補足した。

### 3.5.4 補足：効果指標の修飾が想定される場合

未知・未測定の影響要因について効果指標の修飾が想定される場合の感度分析について、すでに 3.1.3 で概略を述べた。その際、層別した分割表から Mantel-Haenszel の方法を用いて層を併合した効果指標を求めることができるし、標準集団を適当に定め交絡要因の分布を標準集団の分布にそろえて効果指標を求めることもできることに言及した。これらの方法は、いずれも交絡要因と曝露の有無ごとに得られる発生割合を、重みをもちいて均す方法である。以下ではおもに標準化について、つまり関心がある集団を標準集団として定め、修飾要因の分布を、あるいは修飾要因でもある交絡要因の分布を、標準集団の分布にそろえて効果指標を求める方法について概観する。

例として、交絡要因の誤分類を補正した分割表（Table 3- 10：上側の表）を用いて、曝露ありの集団、また曝露なしの集団、さらにこれらを合計した全体集団を標準集団とした場合の標準化リスク比を求めてみよう。実際には、これらの集団ごとに交絡要因の分布は異なっている。しかし、この分布を、標準集団として選択したひとつの集団の分布にそろ

えることで、交絡要因の影響をもそろえる（標準化する）ことができる．こうした計算に必要な重みと発生割合をまとめて表に示した（Table 3- 11）．

Table 3- 11 曝露・交絡要因の値別の重みと発生割合

曝 露	交絡 要因	対象者数 (標準化の重み)	交絡要因の分布 (あり・なしの割合)	発生割合 *
あ り	あり	$N_{11}$	$p_1 = \frac{N_{11}}{N_{11} + N_{10}} = 0.928$	$\frac{464.9}{9057.7} = 0.0513$
	なし	$N_{10}$	$1 - p_1 = 0.072$	$\frac{11.1}{704.3} = 0.0157$
な し	あり	$N_{01}$	$p_0 = \frac{N_{01}}{N_{01} + N_{00}} = 0.515$	$\frac{309.3}{8240.2} = 0.0375$
	なし	$N_{00}$	$1 - p_0 = 0.485$	$\frac{140.7}{7745.8} = 0.0182$
合 計	あり	$N_{11} + N_{01}$	$p = \frac{N_{11} + N_{01}}{N_{11} + N_{01} + N_{10} + N_{00}} = 0.672$	$\frac{476}{9862} = 0.0488$
	なし	$N_{10} + N_{00}$	$1 - p = 0.328$	$\frac{450}{15,986} = 0.0282$

\* Table 3- 10（上側の表）の数値をそのままもちいているため、発生割合の分母や分子が小数となっている．

この表から、例えば曝露ありの集団では、交絡要因ありの割合が  $p_1 = 0.928$ 、なしの割合が  $1 - p_1 = 0.072$  である．したがって、曝露ありの集団全体での発生割合  $0.0488$  は、交絡要因の有無別の発生割合（交絡要因ありで  $0.0513$ 、なしで  $0.0157$ ）をこの割合で合成したときに観察される発生割合である、とも解釈できる．すなわち

$$p_1 \cdot 0.0513 + (1 - p_1) \cdot 0.0157 = 0.928 \cdot 0.0513 + 0.072 \cdot 0.0157 = 0.048 \dots$$

一方、曝露なしの集団では、交絡要因の分布は異なるが（割合の比は  $p_0:1 - p_0$ ）、仮に曝露なしの集団でも、交絡要因の分布が曝露ありの集団とおなじ（ $p_1:1 - p_1$ ）だったとすると、曝露なしの集団の発生割合は

$$p_1 \cdot 0.0375 + (1 - p_1) \cdot 0.0182 = 0.928 \cdot 0.0402 + 0.072 \cdot 0.0197 = 0.0361 \dots$$

となる．したがって、曝露ありの集団を標準集団としたときの標準化リスク比は

$$\frac{p_1 \cdot 0.0513 + (1 - p_1) \cdot 0.0157}{p_1 \cdot 0.0402 + (1 - p_1) \cdot 0.0197} = \frac{0.048 \dots}{0.0361 \dots} = 1.349 \dots$$

となる．同様な計算により、曝露なしの集団を標準集団としたときの標準化リスク比

$$\frac{p_0 \cdot 0.0513 + (1 - p_0) \cdot 0.0157}{p_0 \cdot 0.0402 + (1 - p_0) \cdot 0.0197} = 1.210 \dots$$

を求めることができる．さらに、集団全体を標準集団としたときの標準化リスク比も



$$\frac{p \cdot 0.0513 + (1 - p) \cdot 0.0157}{p \cdot 0.0402 + (1 - p) \cdot 0.0197} = 1.271 \dots$$

と求められる。

この数値例では、曝露ありの集団では交絡要因ありが約 90% ( $p_1$ )、また曝露なしの集団では交絡要因ありが約 50% ( $p_0$ ) と、曝露の有無により交絡要因の分布は大きく異なる。しかし、どちらの分布をもちいてもリスク比は概ね 1.2 から 1.3 程度であり、大きく評価が異なることはなさそうである。しかし、こうしたことが常になりたつわけではない。その研究で関心がある集団（曝露ありの集団、なしの集団、全体集団等）に応じて適切に標準集団を定め、対応する感度分析の結果を参照するとよいだろう。

### 3.5.5 補足：交絡要因、及び共変量の非差別的誤分類

3.5 では、データとして得られている交絡要因の分類が誤っている場合について述べてきた。しかし同様なバイアスは、解析にもちいる共変量の分類が誤っている場合にも生じる。

すでに述べたように、曝露の有無によらず修飾要因の分布が等しければ、少なくともこの要因による交絡は生じないはずである (3.1.3)。ところが、この要因に誤分類が生じると、誤分類後の修飾要因の分布が曝露の有無により異なるかもしれない。つまり、真の値がわかれば、この修飾要因による交絡はないはずなのに、誤った値を参照することで見かけ上の交絡が生じる可能性がある。また、この状況で、修飾要因は誤分類なく観察されていて曝露にのみ誤分類が生じていたとしよう。するとこの場合も、曝露の誤分類のせいで、修飾要因の分布が曝露の有無により異なって見えるかもしれない。するとやはり、真の値を参照すれば存在しないはずの交絡が、曝露の誤分類により生じる可能性がある。

こうした現象は、効果指標の修飾について検討したい場合にも悪影響を及ぼす可能性がある。つまり、ある共変量が修飾要因の場合、曝露または修飾要因の非差別的誤分類により、実際には存在しないはずの効果指標の修飾が見かけ上あらわれたり、逆に存在するはずの修飾が見えなくなることがある<sup>30</sup>。前々節で見た数値例 (Table 3-10) でも非差別的誤分類を想定していたが、要因の有無によりリスク比が逆転するという本来あったはずの効果指標の修飾が、誤分類のせいで見えなくなっていた、とも解釈できる結果であった。

なお、2 値の交絡要因に対する独立で非差別的な誤分類についても、バイアスの方向を予測することができる。参考までに、その概略をしめした (参考 3-11)。

### 3.5.6 補足：さらに複雑な誤分類

以上では、差別的な誤分類といっても、ある 1 つの要因の値 (真の値) のみに依存する比較的単純な場合をみつかった。しかし、曝露の誤分類が、アウトカムだけでなく交絡要因の値にも依存する等、もっと複雑な場合も考えられる。また、誤分類が曝露にもアウトカムにも依存する従属な場合や、曝露水準が 3 以上の場合等も考えられる。

こうした複雑な状況に対処するため行列計算に帰着させる方法があり、行列補正 (matrix adjustment) と呼ばれている<sup>16</sup>。これまで参考等でしめしてきた数式も、この方法の特殊な場合とみなせる。



例えば、3.4.2 でしめしたアウトカムの差別的誤分類における、真の度数 ( $A_1$ 、 $B_1$ 、 $A_0$ 、 $B_0$ ) と測定された度数 ( $A_1^*$ 、 $B_1^*$ 、 $A_0^*$ 、 $B_0^*$ ) の関係は、感度や特異度を成分とする行列をもちいて以下のように表すことができる。

$$\begin{pmatrix} A_1^* \\ B_1^* \\ A_0^* \\ B_0^* \end{pmatrix} = \begin{pmatrix} Se_1 & 1 - Sp_1 & 0 & 0 \\ 1 - Se_1 & Sp_1 & 0 & 0 \\ 0 & 0 & Se_0 & 1 - Sp_0 \\ 0 & 0 & 1 - Se_0 & Sp_0 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \\ A_0 \\ B_0 \end{pmatrix}$$

ただし、この行列の成分は、真の値がどのような確率でどの値に分類されるかを表しているので、いずれも 0 から 1 までの値をとる。また、例えば  $A_1$  に含まれる対象者が、 $A_1^*$ 、 $B_1^*$ 、 $A_0^*$ 、または  $B_0^*$  のいずれかに分類される確率の合計は 1 になるので、列ごとの成分の総和は 1 になる。この例でも、例えば 1 列目の総和は  $Se_1 + (1 - Se_1) = 1$  となっている。

また、これはアウトカムの誤分類の例なので、曝露なしの場合の度数 ( $A_0$ 、 $B_0$ ) が曝露ありの度数 ( $A_1^*$ 、 $B_1^*$ ) に分類されることはない。このことが、対応する行列の成分 (1 行 3 列目から 2 行 4 列目までの 4 つの成分) がすべて 0 であることによって表現されている。一方、従属な誤分類を想定する場合は、曝露なしから曝露ありといった誤分類も生じうる。したがって、これらの成分が 0 になるとは限らない。こうして、4 行 4 列の行列により、差別的で従属な誤分類についても一般的にあつかうことができる。

このように、理論的には、複雑な誤分類もあつかえる方法が考案されている。しかし、複雑な誤分類になるほど、具体的にどのようにパラメータの値、言い換えれば上の行列の成分を定めればよいのか、参照できるデータが見あたらないかもしれない。

## 3.6 選択バイアスに対する感度分析

### 3.6.1 選択バイアスに対する調整解析

もし十分なデータが利用可能なら、選択バイアスの影響を調整した解析が可能かもしれない。例えば、コホート研究において対象者選択によるバイアスが懸念される場合、選択されなかった対象者についても一部のデータが利用可能であったとする。それらのデータをもちいて、ある要因が患者選択に及ぼした影響や、その要因と曝露やアウトカムの関係について十分な検討ができたとする。こうした場合なら、選択バイアスの影響を調整した解析が可能かもしれない。

こうした方法は、いずれも選択バイアスの構造的な定義にもとづいている<sup>5</sup>。例えば、関連する要因の相互関係を因果グラフによって分析し、対象者選択と関連する要因を考慮した解析を行う。その際、参考 1- 2 で述べたような方法をもちいることができるかもしれない。あるいは、以降で紹介する感度分析を行うこともできる。実際にこうした方法をもちいて選択バイアスの影響を検討した研究として、デンマーク国家出生コホート<sup>32</sup>に対する事例が参考になるかもしれない。様々な時点で行われた一連の検討を事例として紹介しながら、選択バイアスの構造的な定義や、代表的な解析方法の概略を紹介した論文は、選択バイアスに対する理解を深めるうえでも有用かもしれない<sup>33</sup>。

なお、実際には、以上で述べたようなデータが利用可能なことは稀かもしれない。また、上述したような分析や解析を行うこと自体、現時点では一般的とはいえない。以下でも、こうした検討や解析が難しいので感度分析を行いたい、という状況を想定しよう。

### 3.6.2 感度分析の方針

コホート研究の場合、選択バイアスが生じる状況として 2 つの場合があった (Table 1- 2) 。すなわち、介入前の「試験の対象者選択によるバイアス」と、介入後の「欠測によるバイアス」である。

WTCHR の事例でも、調査に参加しなかった対象者を考慮した感度分析や、欠測を考慮した感度分析が行われている。例えば、WTCHR のデータをもちいた別の研究によれば、救助や復旧に係った労働者やボランティアのうち、雲状の粉じんや破片への曝露があったと回答した対象者は 30,525 人であった<sup>34</sup>。一方、数値例として参照した研究では 25,748 人のデータがもちいられており、欠測を理由に 4,777 人のデータが除外されている。また、別の研究によれば、救助や復旧に係ったのは 91,469 人であったと推定されている<sup>35</sup>。この値をもちいると 60,944 (= 91,469 - 30,525) 人は、WTCHR に参加しなかったことになる。

以下では、調査に参加しなかった対象者について考えよう。まず、WTCHR に参加せず、結果的に研究対象として選択されなかった対象者についても曝露とアウトカムのデータが得られていたと仮定する。すると対象者選択の有無別に、曝露とアウトカムの関連性をしめす分割表が得られるはずである (Table 3- 12) 。

Table 3- 12 対象者選択の有無別の分割表

アウト カム： ぜん息	対象者選択				合計	
	あり (C = 1)		なし (C = 0)			
	曝露：粉じん		曝露：粉じん		曝露：粉じん	
	E = 1	E = 0	E = 1	E = 0	E = 1	E = 0
D = 1	A <sub>11</sub> = 476	A <sub>01</sub> = 450	A <sub>10</sub>	A <sub>00</sub>	A <sub>1</sub>	A <sub>0</sub>
D = 0	B <sub>11</sub> = 9,286	B <sub>01</sub> = 15,536	B <sub>10</sub>	B <sub>00</sub>	B <sub>1</sub>	B <sub>0</sub>
合計	N <sub>11</sub> = 9,762	N <sub>01</sub> = 15,986	N <sub>10</sub>	N <sub>00</sub>	N <sub>1</sub>	N <sub>0</sub>

実際にデータが得られているのは、対象者選択あり (C = 1) の層のデータだけであるが、この表全体を復元することを試みよう。

### 3.6.3 予備的な計算例

この事例に対して行われた感度分析によれば、WTCHR に参加しなかった対象者たちの曝露とアウトカムについて、何をどのように想定すればいいのかわかな情報は得られていない<sup>20</sup>。一方、この研究では、選択バイアスの影響を例示するため、WTCHR に参加しなかった対象者の数をアウトカムと曝露の有無別に以下のように想定している（合計 60,944 人）。

- ・ ぜん息ありの対象者数：曝露ありのとき 472 人、曝露なしのとき 472 人
- ・ ぜん息なしの対象者数：曝露ありのとき 20,000 人、曝露なしのとき 40,000 人

これらの数値をもちいると、参加しなかった対象者をふくむ集団全体での分割表が得られる (Table 3- 13)。

合計欄が、すべての対象者が研究に含まれていたとしたら知り得たであろうデータをしめしている。この数値例では、対象者選択によりリスク比が過小評価されていた可能性が示唆されている (1.92 → 1.73)。

Table 3- 13 選択バイアスを補正した分割表

アウト カム： ぜん息	対象者選択				合計	
	あり (C = 1)		なし (C = 0)			
	曝露：粉じん		曝露：粉じん		曝露：粉じん	
	E = 1	E = 0	E = 1	E = 0	E = 1	E = 0
D = 1	476	450	472	472	948	922
D = 0	9,286	15,536	20,000	40,000	29,286	55,536
合計	9,762	15,986	20,472	40,472	30,234	56,458
リスク比	1.73		1.98		1.92	

この感度分析では、まったく恣意的に対象者数を定めたようにも思え、その背景にどのような想定があるのか把握しにくいかもしれない。以下では 2 つの方法、つまり選択確率による方法と、観察されなかった層の分割表を求める方法を通じ、以上の内容を再検討してみよう。

### 3.6.4 選択確率による方法

すでに述べたように、救助や復旧に係ったのは 91,469 人であったと推定されており、そのうち 30,525 人が WTCHR に参加し、曝露の有無について回答を残していた。いったん欠測例 (4,777 人) を度外視すると、対象者となりえた人々の集団、すなわち対象者選択の源泉となった集団にふくまれる 86,692 (= 91,469 - 4,777) 人のうち、25,748 (= 30,525 - 4,777) 人が対象者として選択されたことになる。つまり、対象者選択の源泉となった集団、すなわち選択前の集団から約 30% の対象者が WTCHR に参加した。

現実とは異なるが、もし曝露やアウトカムの有無によらず、また他の要因も参照することなく、選択前の集団から無作為に 30% の対象者を選択したとする。この場合はランダム・サンプリングとみなせるので、選択バイアスの懸念は小さいだろう。しかし実際は、曝露とアウトカムの有無別に異なる割合で、すなわち偏った対象者選択が行われていた可能性がある。こうした観点から、選択バイアスに対する感度分析を行う際、以下のようにパラメータを設定することができる。

- ・  $S_{A_1}$  : 曝露あり ( $E = 1$ )、アウトカムあり ( $D = 1$ ) の対象者が選択される確率
- ・  $S_{A_0}$  : 曝露なし ( $E = 0$ )、アウトカムあり ( $D = 1$ ) の対象者が選択される確率
- ・  $S_{B_1}$  : 曝露あり ( $E = 1$ )、アウトカムなし ( $D = 0$ ) の対象者が選択される確率
- ・  $S_{B_0}$  : 曝露なし ( $E = 0$ )、アウトカムなし ( $D = 0$ ) の対象者が選択される確率

また、アウトカムなしの対象者に対してパラメータを定めるかわりに、以下のように定めてもよい。

- ・  $S_{A_1}$  : 曝露あり ( $E = 1$ )、アウトカムあり ( $D = 1$ ) の対象者が選択される確率
- ・  $S_{A_0}$  : 曝露なし ( $E = 0$ )、アウトカムあり ( $D = 1$ ) の対象者が選択される確率
- ・  $S_{N_1}$  : 曝露あり ( $E = 1$ ) の対象者が選択される確率
- ・  $S_{N_0}$  : 曝露なし ( $E = 0$ ) の対象者が選択される確率

どちらを用いても本質的な違いはないが (参考 3- 14)、以下では後者の選択確率をもちいることにしよう。すると選択前の集団における分割表のデータ ( $A_1$ 、 $A_0$ 、 $N_1$ 、 $N_0$ ) と、その集団から選択された対象者における分割表のデータ ( $A_{11}$ 、 $A_{01}$ 、 $N_{11}$ 、 $N_{01}$ ) を、以下のように結びつけることができる。

$$A_{11} = A_1 \cdot S_{A_1}, A_{01} = A_0 \cdot S_{A_0}, N_{11} = N_1 \cdot S_{N_1}, N_{01} = N_0 \cdot S_{N_0}.$$

したがって、実際には観察されていない、選択前の集団での分割表のデータを、以下のよう求めることができる。

$$A_1 = \frac{A_{11}}{S_{A_1}}, A_0 = \frac{A_{01}}{S_{A_0}}, N_1 = \frac{N_{11}}{S_{N_1}}, N_0 = \frac{N_{01}}{S_{N_0}}.$$

ただし、選択前の集団の大きさ ( $N = N_1 + N_0$ ) は、この例では 86,692 人と固定されているので、以上の選択確率を取りうる値には制限がある。例えば、以上の関係から

$$S_{N_0} = \frac{N_{01}}{N - N_{11}/S_{N_1}} \quad \dots \quad \text{式 3-8}$$

となる。したがって  $S_{N_1}$  と  $S_{N_0}$  については、どちらか一方を定めればよい。また、選択確率が 0 から 1 までの値となるように定める必要もある。例えば、式 3-8 から、右辺の分母が  $N_{01}$  より大きな値となるように  $S_{N_1}$  を定める必要がある。

### 3.6.5 数値例における選択確率

すでにしめた数値例について、以上で定義した選択確率を求めると、以下のようになる (Table 3-12、Table 3-13)。

- $S_{A_1} = A_{11}/A_1 = 476/948 = 0.502$
- $S_{A_0} = A_{01}/A_0 = 450/922 = 0.488$
- $S_{N_1} = N_{11}/N_1 = 9762/30234 = 0.323$
- $S_{N_0} = N_{01}/N_0 = 15986/56458 = 0.283$

つまり、この数値例では、アウトカムありの対象者の選択確率は約 50%と、集団全体での選択確率より 20%ほど高い。また、アウトカムの有無別に見ると、曝露ありの対象者の選択確率が、曝露なしの対象者より約 2~3%ほど高い。すなわち、アウトカムの有無別に、また曝露の有無別にも、ある一定の傾向をもった対象者選択が行われていたのではないかとこの数値例では想定していることになる。

アウトカムありの対象者が、アウトカムなしの対象者より多く選択される、という状況はケースコントロール研究のようである。しかし、曝露ありの対象者ほど多く選択されているので、これをケースコントロール研究とみなしたとしても、やはり偏った対象者選択が行われていたことになる。

### 3.6.6 選択確率と効果指標の関係

次に、選択確率が効果指標に及ぼす影響を見ておこう。以上でしめた関係から、選択バイアスの影響を補正した元の集団でのリスク比  $RR$  と、この研究で選択された対象者から得られるリスク比  $RR^* = (A_{11}/N_{11})/(A_{01}/N_{01})$  には、以下の関係がある。

$$RR = \frac{A_1/N_1}{A_0/N_0} = \frac{(A_{11}/S_{A_1})/(N_{11}/S_{N_1})}{(A_{01}/S_{A_0})/(N_{01}/S_{N_0})} = RR^* \left( \frac{S_{A_1}/S_{N_1}}{S_{A_0}/S_{N_0}} \right)^{-1} \quad \dots \quad \text{式 3-9}$$

すなわち、この場合も

$$(\text{真の値}) \doteq (\text{推定値})/(\text{系統誤差})$$

という形に、バイアスを含んだ推定値と系統誤差にあたる項を分離することができる。

また、この関係から、例えば以下の条件（十分条件）が満たされれば、リスク比にバイアスが生じないことがわかる。

$$S_{A_1} = S_{N_1} \text{ かつ } S_{A_0} = S_{N_0}$$

この条件は、曝露群ごとに、アウトカムの有無によらず選択確率が等しいこと ( $S_{A_1} = S_{B_1}$  かつ  $S_{A_0} = S_{B_0}$ ) をも意味している (参考 3-14)。例えば、マッチングを行うコホート研究のように、曝露の有無により対象者選択の確率が異なっていたとしても、それぞれ

の群のなかで、アウトカムの有無によらず対象者が選択されていれば、発生割合もリスク比もバイアスなく推定できるはずである。

さらに、式 3-9 から、系統誤差にあたる量  $RRR$  をあらためてパラメータとして定めて感度分析を行うことも考えられる。すなわち

$$RRR = \frac{S_{A_1}/S_{N_1}}{S_{A_0}/S_{N_0}} \text{ とおくと、 } RR = RR^*/RRR.$$

この方法はパラメータが 1 つだけでよいので、より複雑な感度分析を行う際に有用である。

### 3.6.7 観察されなかった層の分割表を求める方法

以上で述べたのは、選択前の集団と、実際に選択された対象者の集団を、選択確率によって結びつける方法だった。以下では、選択されなかった集団に注目し、どのような対象者が研究に含まれなかったのか、参加しなかったのか、と考えていく。こうした想定を量的に表したいとき、いくつかの方法が考えられるが、例えば以下のようにパラメータを定めることができる。ただしこの場合も、選択前の集団の大きさ ( $N$ ) は既知とする。

- $O_{C0}$  : 対象者選択なし ( $C = 0$ ) の集団での曝露オッズ。すなわち、この集団で、曝露ありの対象者が、曝露なしの対象者の何倍多いかという比  $N_{10}/N_{00}$ 。
- $IP_{C0}$  : 対象者選択なし ( $C = 0$ ) の集団での、曝露なしの場合のアウトカムの発生割合  $A_{00}/N_{00}$ 。
- $RR_{DE0}$  : 対象者選択なし ( $C = 0$ ) の集団での、曝露とアウトカムの関連性をしめすリスク比  $(A_{10}/N_{10})/(A_{00}/N_{00})$ 。

このようにパラメータを定めると、対象者選択なしの集団、つまり実際には観察されなかった集団における分割表のデータが得られる。すると、対象者選択ありの集団となしの集団の合計、つまり選択前の集団における分割表も得られるので、それをもちいて選択バイアスの影響を補正した効果指標を求めることができる (参考 3- 12) 。

また、こうした検討により、曝露の有無別の発生割合が対象者選択の有無によらず等しい場合は選択バイアスが生じないことがわかる (参考 3- 13) 。自明のことではあるが、ある要因の有無によらずアウトカムの発生状況が等しければ、その要因で層別を行い、1 つの層の対象者だけを選択してもバイアスは生じないはずである。

### 3.6.8 数値例における観察されなかった層の分割表

すでにしめた数値例 (Table 3- 12、Table 3- 13) に対し、以上で定義したパラメータの値を計算すると以下ようになる。なお、実際に観察された、対象者選択ありの集団におけるパラメータの値を、すぐ右 [括弧内] にしめた。

- $O_{C0} = N_{10}/N_{00} = 20472/40472 = 0.506$  [  $O_{C1} = 0.611$  ]
- $IP_{C0} = A_{00}/N_{00} = 472/40472 = 0.0117$  [  $IP_{C1} = 0.0281$  ]
- $RR_{DE0} = (A_{10}/N_{10})/(A_{00}/N_{00}) = (472/20472)/(472/40472) = 1.98$  [  $RR_{DE1} = 1.73$  ]



この数値例では、研究に参加しなかった、結果的に対象者として選択されなかった集団では、曝露オッズがやや小さく、曝露なしの群における発生割合が低く、またリスク比が高いと想定していたことがわかる。この想定が適切なら、リスク比が高い集団を除外した、もしくは選択しなかったせいで、真の値（選択前の集団でのリスク比 1.92）より低いリスク比が観察されたことになる。

### 3.6.9 補足：2つの方法の関連と、他のパラメータ設定

選択バイアスの影響を補正する2つの方法、すなわち選択確率による方法と、観察されなかった層の分割表を求める方法は、アプローチこそ異なるものの、基本的にはおなじこと、つまり選択前の集団の分割表を復元することを目標としている。

入門編でも述べたように、選択バイアスは、曝露とアウトカムの双方と関連した、両者の「共通の結果」とみなせる要因によって集団を層別する、あるいはその要因によって条件付ける（conditioning）ことで生じる。言い換えれば、交絡要因でない要因による不適切な層別のせいで、かえってバイアスが生じる。この不適切な層別を、いわば解除し、元の分割表にもどす作業がこれまで紹介した2つのアプローチだとも言える。

事例として取り上げたのは、WTCHRに参加しなかった対象者を考慮した感度分析であり、この事例では選択前の集団の大きさを見積もることができた。しかし、状況によっては、選択前の集団の大きさや、そこでのアウトカムや曝露の頻度がまったく不明な場合もある。こうした場合、パラメータの値や、その組み合わせとして想定される内容が多岐にわたり、どのように考えたらよいか苦慮するかもしれない。

コホート研究において、対象者選択によるバイアスが懸念されるとき、以上で紹介した2つの方法のどちらをもちいるかは、状況に応じて解析者が判断することになる。懸念されるバイアスの内容を表現するのに、どの方法が考えやすいか、またパラメータの値を設定しやすいか、といったことが判断材料になるかもしれない。また、2つの方法のいずれについても、さらに別のパラメータ設定が考えられる（参考 3- 14）。さらに、この事例に対して行われた感度分析がそうだったように、観察されなかった層の度数をとりあえず定めてみる、といったアプローチもある<sup>20</sup>。解析者の意図や、選択バイアスに対する想定を考慮して、アプローチしやすい方法を選ぶとよいだろう。

### 3.7 ケースコントロール研究の場合

ケースコントロール研究に対する感度分析としては、教科書や論文等で頻繁に取り上げられている事例がある。それは、ある変圧器組立工場で癌による死亡が多かったことから、そこであつかわれていたおもな化学物質（7種類）への曝露と種々の癌（14種類）による死亡との関連性を検討したケースコントロール研究に由来する<sup>36</sup>。教科書や論文等によく取り上げられているのは、フェノールホルムアルデヒドやポリビニルホルマールを主成分とする化学合成樹脂への曝露と肺癌による死亡の関連性をしめすデータに対する感度分析である<sup>16, 37</sup>。ひとつの事例に対し、単純な感度分析から確率的感度分析まで広範な方法がもちられているので、関心がある読者には、この事例も参考にすることを薦めたい。また、本書の第2部でも、この事例をあつかっている。

#### 3.7.1 コホート研究とのおもな相違点

以下では、対象者選択の源泉となった集団、すなわち選択前の集団から、アウトカムあり（ $D = 1$ ）の対象者は全員ケースとして採用し、アウトカムなし（ $D = 0$ ）の対象者は、その一部をコントロールとして選択するようなケースコントロール研究について考える。

この場合に得られるデータは、すでにコホート研究についてしめた単純な分割表（Table 2-1）と似た内容になる。ただし、リスク比やリスク差の分母となるデータ、つまり選択前の集団の人数（ $N_1$ 、 $N_0$ ）は未知かもしれない。また、アウトカムなしの場合の度数（ $B_1$ 、 $B_0$ ）も、選択前の集団における人数ではなく、そこから選択された一部の対象者、すなわちコントロールの人数を表していることに留意しよう（Table 3-14）。

Table 3-14 介入・曝露の有無とアウトカムの有無に関する分割表：  
ケースコントロール研究の場合

アウトカム	介入・曝露	
	あり（ $E = 1$ ）	なし（ $E = 0$ ）
あり（ $D = 1$ ）	$A_1$	$A_0$
なし（ $D = 0$ ）	$B_1$	$B_0$
選択前の集団	$N_1$	$N_0$

さて、コントロールの選択確率を曝露ありの集団については $S_{N_1}$ （ $= B_1/N_1$ ）、また曝露なしの集団については $S_{N_0}$ （ $= B_0/N_0$ ）と表し、コントロール選択が曝露と無関係に適切に行われていること、すなわち $S_{N_1} = S_{N_0}$ 、あるいは

$$\frac{B_1}{N_1} = \frac{B_0}{N_0}$$

がなりたつことを仮定する。この場合、つまりコントロール選択が適切に行われていれば、仮にリスク比の分母（ $N_1$ 、 $N_0$ ）が未知であったとしても

$$RR = \frac{A_1/A_0}{B_1/B_0} = \frac{A_1/N_1}{A_0/N_0} = \frac{A_1/B_1}{A_0/B_0} = \frac{A_1B_0}{A_0B_1} = OR$$

がなりたち、オッズ比を通じてリスク比をバイアスなく推定することができる。

この関係が示唆しているように、ケースコントロール研究におけるコントロールの人数 ( $B_1$ 、 $B_0$ ) は、コホート研究における分母 ( $N_1$ 、 $N_0$ ) の代りであると考えられる。このため、コホート研究のリスク比についてしめた関係式の多くは

$$N_1 \rightarrow B_1, N_0 \rightarrow B_0$$

という置き換えにより、ほぼそのまま、ケースコントロール研究におけるオッズ比の関係式に書きかえることができる。以下では、このような観点から、ケースコントロール研究における感度分析の方法を概観する。

### 3.7.2 交絡

コホート研究の場合、未知の交絡要因に関する量的な想定をパラメータの値として与え、それらをもちいて交絡要因の有無別に層別した分割表を求めることが感度分析の大まかな方針だった (3.1.1)。

ケースコントロール研究の場合も、求めたい分割表はコホート研究の場合 (Table 3-1) と似た内容になる (Table 3-15)。ただし、アウトカムなしの場合の度数が、コントロール選択によって得られた一部の対象者のデータである点が、コホート研究の場合とは異なっている。

Table 3- 15 交絡要因の有無で層別した分割表：  
ケースコントロール研究の場合

アウトカム	合計 (元のデータ)		未知の交絡要因			
			$C = 1$		$C = 0$	
	曝露		曝露		曝露	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	$A_1$	$A_0$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
$D = 0$	$B_1$	$B_0$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$

ケースコントロール研究の場合、例えば以下のようにパラメータを設定することが考えられる。

- $OR_{DC}$  : 未知の交絡要因 ( $C$ ) とぜん息 ( $D$ ) の関連性をしめすオッズ比 (交絡要因あり / 交絡要因なし)。ただし、このオッズ比は、曝露の有無によらず等しいと仮定する。すなわち

$$OR_{DC} = \frac{A_{11}/B_{11}}{A_{10}/B_{10}} = \frac{A_{01}/B_{01}}{A_{00}/B_{00}}.$$

- $P_{C1}$  : コントロール群 ( $D = 0$ ) において、曝露あり ( $E = 1$ ) の集団で未知の交絡要因あり ( $C = 1$ ) の対象者が占める割合。すなわち  $P_{C1} = B_{11}/B_1$ 。
- $P_{C0}$  : コントロール群 ( $D = 0$ ) において、曝露なし ( $E = 0$ ) の集団で未知の交絡要因あり ( $C = 1$ ) の対象者が占める割合。すなわち  $P_{C0} = B_{01}/B_0$ 。

コホート研究との違いは、交絡要因とアウトカムの関連性をしめす効果指標がリスク差やリスク比からオッズ比になること、また交絡要因の分布を想定する際、集団全体 ( $N_1$ 、 $N_0$ ) に対してではなくコントロール群 ( $B_1$ 、 $B_0$ ) に対して考えることである。

以上のようにパラメータを定めると、例えば交絡要因ありの層の度数は

$$A_{11} = \frac{OR_{DC} P_{C1} A_1}{OR_{DC} P_{C1} + 1 - P_{C1}}, A_{01} = \frac{OR_{DC} P_{C0} A_0}{OR_{DC} P_{C0} + 1 - P_{C0}}, B_{11} = P_{C1} B_1, B_{01} = P_{C0} B_0.$$

したがって、交絡の影響を補正したオッズ比  $OR$  は

$$OR = \frac{A_1 B_0}{A_0 B_1} \cdot \frac{OR_{DC} P_{C0} + 1 - P_{C0}}{OR_{DC} P_{C1} + 1 - P_{C1}}.$$

さらに、元のデータから計算した、未補正のオッズ比  $OR^*$  との関係は

$$OR = OR^* / \frac{OR_{DC} P_{C1} + 1 - P_{C1}}{OR_{DC} P_{C0} + 1 - P_{C0}}.$$

などとなる。これらは、いずれもコホート研究のリスク比に関する関係と同様である（式 3-2、及び参考 3-2）。対応する式を比較すると、いずれも  $N_1 \rightarrow B_1$ 、 $N_0 \rightarrow B_0$  という置き換え、及びこれにともなう置き換え  $RR_{DC} \rightarrow OR_{DC}$  によって得られる関係であることがわかる。

また、これらの関係は、コホート研究におけるオッズ比について、アウトカムが稀な場合に近似的になりたつ関係と一致している（式 3-3、及び参考 3-3）。さらに、コホート研究のオッズ比については、交絡要因ありの対象者が占める割合 ( $P_{C1}$ 、及び  $P_{C0}$ ) の定義を、集団全体 ( $N_1$ 、及び  $N_0$ ) に対する割合から、アウトカムなしの対象者 ( $B_1$ 、及び  $B_0$ ) に対する割合に変更した場合も、おなじ関係が得られる（参考 3-3）。

さらに、交絡要因の影響が曝露の有無により異なることが想定される場合についても、曝露の有無により異なるオッズ比を定めると、層ごとに異なるオッズ比が得られる。それらを Mantel-Haenszel の方法をもちいて併合したり、標準化にもちいることができる点も、コホート研究の場合と同様である。

### 3.7.3 曝露の誤分類

曝露の誤分類については、コホート研究の場合と同様に、例えば以下のパラメータを定めて感度分析を行うことができる。

- $Se_A$  : アウトカムありの場合の、曝露分類の感度
- $Sp_A$  : アウトカムありの場合の、曝露分類の特異度
- $Se_B$  : アウトカムなしの場合の、曝露分類の感度
- $Sp_B$  : アウトカムなしの場合の、曝露分類の特異度

これらのパラメータの値と、誤分類を受けている分割表のデータ ( $A_1^*$ 、 $A_0^*$ 、 $B_1^*$ 、 $B_0^*$ ) から、誤分類がない分割表のデータ ( $A_1$ 、 $A_0$ 、 $B_1$ 、 $B_0$ ) を求めることができる（参考 3-7）。それらをもちいて、曝露の誤分類の影響を補正したオッズ比  $OR$  を求めると

$$OR = \frac{A_1 B_0}{A_0 B_1} = \frac{\{Sp_A A_1^* - (1 - Sp_A) A_0^*\} \{Se_B B_0^* - (1 - Se_B) B_1^*\}}{\{Se_A A_0^* - (1 - Se_A) A_1^*\} \{Sp_B B_1^* - (1 - Sp_B) B_0^*\}}.$$

以上は、アウトカムの有無により曝露の誤分類の程度が異なる、差別的誤分類を想定した補正式である。一方、非差別的誤分類を想定して、 $Se_A = Se_B = Se$ 、また  $Sp_A = Sp_B = Sp$  と置いてみよう。すると、上式は

$$\begin{aligned} OR &= \frac{\{Sp A_1^* - (1 - Sp) A_0^*\} \{Se B_0^* - (1 - Se) B_1^*\}}{\{Se A_0^* - (1 - Se) A_1^*\} \{Sp B_1^* - (1 - Sp) B_0^*\}} \\ &= \frac{Sp A_1^* - (1 - Sp) A_0^*}{Se A_0^* - (1 - Se) A_1^*} \bigg/ \frac{Sp B_1^* - (1 - Sp) B_0^*}{Se B_0^* - (1 - Se) B_1^*} \end{aligned}$$

と変形できるが、この式は、コホート研究におけるリスク比の補正式（式 3-4）と対応している。すなわち、式 3-4 に対し、 $N_1 \rightarrow B_1$ 、 $N_0 \rightarrow B_0$  という置き換えを行うと上式が得られる。ケースコントロール研究における曝露の誤分類についても、コホート研究におけるリスク比の分母を、コントロール群の度数に置き換えて考えればよい。

### 3.7.4 ケースコントロール研究におけるアウトカムの誤分類

アウトカムの誤分類についても、コホート研究の場合と同様に、例えば以下のパラメータを定めて感度分析を行うことができる。

- $Se_1$  : 曝露ありの場合の、アウトカム分類の感度
- $Sp_1$  : 曝露ありの場合の、アウトカム分類の特異度
- $Se_0$  : 曝露なしの場合の、アウトカム分類の感度
- $Sp_0$  : 曝露なしの場合の、アウトカム分類の特異度

ただし、ここで問題になるのが、以上の感度や特異度がどの集団における感度や特異度か、という点である。ケースコントロール研究では、例えばアウトカムありの対象者は全例をケースとして、またアウトカムなしの対象者はその一部をコントロールとして選択する。このため、対象者を選択する前の集団（選択前の集団）と、対象者として選択された集団（選択後の集団）とでは、アウトカムの分布が大きく異なっている。つまり選択後の集団のほうが、アウトカムありの対象者が占める割合はるかに大きい。このため、ケースコントロール研究におけるアウトカムの誤分類については、どちらの集団における感度と特異度を想定するかによって感度分析の方法が異なる。

まず、選択後の集団における感度と特異度をパラメータとして設定する場合は、曝露の誤分類の場合と同様に、コホート研究における関係式を、ほぼそのまま適用することができる。例えば、アウトカムの誤分類の影響を補正したオッズ比  $OR$  は

$$OR = \frac{A_1 B_0}{A_0 B_1} = \frac{\{A_1^* - (1 - Sp_1) N_1^*\} \{B_0^* - (1 - Se_0) N_0^*\}}{\{A_0^* - (1 - Sp_0) N_0^*\} \{B_1^* - (1 - Se_1) N_1^*\}} \quad \dots \quad \text{式 3-10}$$

となるが、これはコホート研究におけるアウトカムの誤分類に対して得られた関係を、そのまま適用することで得られる（参考 3-9）。ただし、選択後の集団においてアウトカムの分類がどれほど正確か、他の研究等を参照しても、参考になる情報が得られないかもしれない。というのは、アウトカムの定義が完全におなじだとしても、またおなじ集団から

コントロールを選択したとしても、後にみるように、研究間でコントロール選択の確率が異なれば、感度と特異度についても異なった値が得られるからである。

ところで、ケースやコントロールを選択する際、なんらかの方法で定義されたアウトカムを参照することが多い。データベースをもちいたケースコントロール研究では、さまざまな条件を組み合わせるアウトカムを定義することが多く、こうした定義は試験ごとに異なるかもしれない。この段階で誤分類が生じると、つまりケースやコントロールの選択にもちいる「アウトカム」の定義が不正確だと、どんな問題が生じるだろう。

例えば、1000 人の集団から、ある定義にもとづいて選択したケース 6 人のうち、4 人は真にケースだったが 2 人はコントロールに分類すべき対象者だったとする。また、コントロール候補とみなされた残り 994 人のなかに、実はケースが 2 人含まれていたとしよう。この候補のなかからコントロールが 20 人選択され、真のケースは含まれていなかった（研究 1）。

一方、別の研究者がおなじ定義をもちいてケースとコントロールを選択したところ、ケースについてはおなじ結果が得られた。一方、コントロールは 40 人選択され、このなかには真のケースが 1 例含まれていた（研究 2）。これら 2 つの研究の結果を下表にまとめた（Table 3- 16）。

Table 3- 16 真のアウトカムと定義上のアウトカムの関係

真の アウトカム	選択前の集団			研究 1 の集団			研究 2 の集団		
	定義上のアウトカム			定義上のアウトカム			定義上のアウトカム		
	あり	なし	合計	あり	なし	合計	あり	なし	合計
あり	4	2	6	4	0	4	4	1	5
なし	2	992	994	2	20	22	2	39	41
合計	6	994	1000	6	20	26	6	40	46

この例では、選択前の集団、研究 1 の集団、研究 2 の集団の順に、感度は 67%、100%、80% となる。つまり、選択後の集団（研究 1 や 2 の集団）での感度は、選択前の集団での感度よりかなり高い。一方、特異度は 99%、91%、95% となり、選択後の集団で低くなっている。この違いは、コントロール選択によって得られるデータ（定義上のアウトカムが「なし」の列のデータ）が研究ごとに異なるせいで生じる。つまり、選択後の集団の感度や特異度は、コントロール選択の程度（確率）により異なった値をとる。

逆に、選択前の集団での感度や特異度を調べた研究があったとしても、その結果をそのまま式 3-10 に代入してはならない。この式では、選択後の集団での感度や特異度をもちいているからである。この点をさらに確認するため、選択前の集団に遡って考えてみよう。

### 3.7.5 選択前の集団でのアウトカムの誤分類

以下では、選択前の集団における感度や特異度をこれまでと同様な記号で表す。また、選択前の集団からケースは全例選択し、コントロールはアウトカムなしの対象者から確率  $S$  でその一部を選択する。つまり、コントロール選択が、曝露の有無によらず等しい確率で行われることを想定する。



このとき、アウトカムの誤分類の影響を補正したオッズ比  $OR$  は

$$OR = \frac{A_1 B_0}{A_0 B_1} = \frac{\{S \cdot Sp_1 A_1^* - (1 - Sp_1) B_1^*\} \{Se_0 B_0^* - S(1 - Se_0) A_0^*\}}{\{S \cdot Sp_0 A_0^* - (1 - Sp_0) B_0^*\} \{Se_1 B_1^* - S(1 - Se_1) A_1^*\}}$$

となる（参考 3- 15）．前節でしめした式 3- 10 との違いは、コントロール選択の確率  $S$  を感度分析のパラメータとして考慮する必要があることである．

また、選択後の集団における感度や特異度が、選択前の集団における感度や特異度とどのような関係にあるのか調べてみよう．選択後の集団における感度と特異度を、あらためて  $Se_c$ 、及び  $Sp_c$  と表すことにする．簡単のため、非差別的誤分類を想定し、選択前の集団における感度と特異度は  $Se$ 、及び  $Sp$  と表すことにする．このとき、両者には以下の関係がある（参考 3- 15）．

$$Se_c = \frac{Se}{Se + S(1 - Se)}、Sp_c = \frac{S \cdot Sp}{1 - Sp + S \cdot Sp}.$$

上式から、コントロール選択の確率  $S$  の値によって、感度と特異度が選択の前後で大きく変化する可能性があることがわかる．例えば、選択前の集団からコントロールを 1%

（ $S = 0.01$ ）選択すると、上式から、選択前の集団での感度（ $Se$ ）が 0.99、0.95、0.90 のとき、選択後の集団での感度（ $Se_c$ ）は約 0.999 となる．選択前の値と比較すると、かなり高い値である．一方、選択前の集団での特異度（ $Sp$ ）が 0.99、0.95、0.90 のとき、選択後の集団での特異度（ $Sp_c$ ）は、それぞれ 0.50、0.16、0.08 となり、著しく低下する．したがって、選択前の集団に対して調べた感度や特異度を、ケースコントロール研究によって得られた選択後の集団での感度や特異度であるかのようにみなして式 3- 10 に代入すると、極めて不正確な補正值が得られる懸念がある．

また、全例を選択する場合（ $S = 1$ ）、感度や特異度が選択の前後で変わらないことも上式から確認できる．

なお、選択前の集団が不明確である等の理由から、コントロール選択の確率  $S$  が不明確な場合は、その値を解析者が何らかの想定を置いて定める必要がある．

### 3.7.6 選択バイアス

コホート研究については、選択バイアスに対する感度分析として、選択確率による方法と観察されなかった層の分割表を求める方法を紹介した．ケースコントロール研究については、ケースやコントロールを選択する確率、すなわち選択確率をパラメータとしてもちいる方法が考えやすい．例えば、以下のパラメータを設定する．

**ケース選択の確率：**アウトカムあり（ $D = 1$ ）のとき

- ・  $S_{A_1}$ ：曝露あり（ $E = 1$ ）の対象者が選択される確率
- ・  $S_{A_0}$ ：曝露なし（ $E = 0$ ）の対象者が選択される確率

**コントロール選択の確率：**アウトカムなし（ $D = 0$ ）のとき

- ・  $S_{B_1}$ ：曝露あり（ $E = 1$ ）の対象者が選択される確率
- ・  $S_{B_0}$ ：曝露なし（ $E = 0$ ）の対象者が選択される確率

以上の選択確率により、選択前の集団での分割表のデータ ( $A_1$ 、 $A_0$ 、 $B_1$ 、 $B_0$ ) と、ケースコントロール研究によって選択された集団での分割表のデータ ( $A_{11}$ 、 $A_{01}$ 、 $B_{11}$ 、 $B_{01}$ ) を、以下のように結びつけることができる (分割表のレイアウトについては Table 3- 12 参照) .

$$A_1 = \frac{A_{11}}{S_{A_1}}, A_0 = \frac{A_{01}}{S_{A_0}}, B_1 = \frac{B_{11}}{S_{B_1}}, B_0 = \frac{B_{01}}{S_{B_0}}.$$

したがって、選択前の集団でのオッズ比、すなわち選択バイアスの影響を補正したオッズ比は

$$OR = \frac{A_1 B_0}{A_0 B_1} = \frac{A_{11}/S_{A_1} \cdot B_{01}/S_{B_0}}{A_{01}/S_{A_0} \cdot B_{11}/S_{B_1}} = \frac{A_{11} B_{01}}{A_{01} B_{11}} \left( \frac{S_{A_1}/S_{B_1}}{S_{A_0}/S_{B_0}} \right)^{-1}$$

となる．選択後の集団におけるオッズ比をあらためて  $OR^* = (A_{11} B_{01}) / (A_{01} B_{11})$  と表せば

$$OR = OR^* \left( \frac{S_{A_1}/S_{B_1}}{S_{A_0}/S_{B_0}} \right)^{-1}$$

という関係が、選択前後のオッズ比についてなりたつ．これもコホート研究におけるリスク比の式 3-9 に対し、 $N_1 \rightarrow B_1$ 、 $N_0 \rightarrow B_0$  という置き換えを行った式とみなすことができる．また、この場合も、ふたつのオッズ比  $OR^*$  と  $OR$  の比、つまり系統誤差に相当する量  $ROR$  をあらためてパラメータとして定め、感度分析を行うことができる．すなわち

$$ROR = \frac{S_{A_1}/S_{B_1}}{S_{A_0}/S_{B_0}} \text{ とおくと、 } OR = OR^*/ROR.$$

また、オッズ比についても、以下の条件が満たされればバイアスが生じないこと、すなわち選択の前後でオッズ比の値が変化しないことがわかる．

$$ROR = \frac{S_{A_1}/S_{B_1}}{S_{A_0}/S_{B_0}} = 1$$

この条件が成立する場合の一例として、例えば以下の条件 (十分条件) が挙げられる．

$$S_{A_1} = S_{A_0} \text{ かつ } S_{B_1} = S_{B_0}$$

ケースコントロール研究において、この条件は、ケースとコントロールで選択確率は異なるものの、ケースとコントロールそれぞれについては、曝露の有無によらず選択確率が等しいことを意味している．なお、ケースを全例選択する場合は  $S_{A_1} = S_{A_0} = 1$  であるから 1 つめの条件は自動的に成り立っている．さらに、コントロール選択が曝露の有無によらなければ 2 つめの条件もなりたつから、以上の条件は、コントロール選択の妥当性をしめす条件でもある．

また、ケースコントロール研究の場合も、選択前の集団の人数が既知の場合は、4 つの選択確率が取りうる値に式 3-8 と同様な制約があるかもしれない．個々の研究ごとに、その状況に応じた注意が必要である．

## 3.8 人-時間を分母とした発生率を含む効果指標の場合

### 3.8.1 発生率の差、及び比

アウトカムの発生数を分子、また人-月や人-年等、人-時間を分母とした発生率をもちいた効果指標がある。発生率の差や比、すなわち発生率差や発生率比である。集団ごとの人-時間の合計を曝露ありの集団で  $T_1$ 、また曝露なしの集団で  $T_0$  と表すと、これらの効果指標の推定値は以下のようなになる<sup>18</sup>。

$$\text{発生率差: } \widehat{ID} = \frac{A_1}{T_1} - \frac{A_0}{T_0}, \text{Var}(\widehat{ID}) = \frac{A_1}{T_1^2} + \frac{A_0}{T_0^2},$$

$$\text{発生率比: } \widehat{IR} = \frac{A_1/T_1}{A_0/T_0}, \text{Var}[\ln(\widehat{IR})] = \frac{1}{A_1} + \frac{1}{A_0}.$$

以下でも、混乱が生じない限り推定値をしめす記号  $\hat{\phantom{x}}$  を省略する。これらの式の形から、リスク差やリスク比に関する関係式の多くを

$$N_1 \rightarrow T_1, N_0 \rightarrow T_0$$

という置き換えにより、発生率差や発生率比に対してもちいることができる。以下では、このような観点から、発生率差や発生率比に対する感度分析の方法を概観する。

### 3.8.2 交絡

リスク差やリスク比については、未知の交絡要因に関する量的な想定をパラメータの値として与え、それらをもちいて交絡要因の有無別に層別した分割表を求めることが、感度分析の大まかな方針だった。ただし、発生率差や発生率比の場合、分割表の分子にあたる部分は人数のままだが、分母にあたる部分が人数から人-時間に変わる (Table 3-17)。

Table 3-17 交絡要因の有無で層別した分割表：発生率

アウトカム	合計（元のデータ）		未知の交絡要因			
			$C = 1$		$C = 0$	
	曝露		曝露		曝露	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	$A_1$	$A_0$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
$D = 0$	$B_1$	$B_0$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
合計 人-時間	$T_1$	$T_0$	$T_{11}$	$T_{01}$	$T_{10}$	$T_{00}$

発生率比を例にとると、以下のようにパラメータを設定することが考えられる。

- $IR_{DC}$  : 未知の交絡要因 ( $C$ ) とぜん息 ( $D$ ) の関連性をしめす発生率比 (交絡要因あり / 交絡要因なし)。ただし、この発生率比は、曝露の有無によらず等しいと仮定する。
- $P_{C1}$  : 曝露あり ( $E = 1$ ) の集団で、未知の交絡要因あり ( $C = 1$ ) の対象者が占める人-時間の割合。

- $P_{C0}$  : 曝露なし ( $E = 0$ ) の集団で、未知の交絡要因あり ( $C = 1$ ) の対象者が占める人-時間の割合.

以上のようにパラメータを定めると、例えば交絡要因ありの層のデータは

$$A_{11} = \frac{IR_{DC} P_{C1} A_1}{IR_{DC} P_{C1} + 1 - P_{C1}}, A_{01} = \frac{IR_{DC} P_{C0} A_0}{IR_{DC} P_{C0} + 1 - P_{C0}}, T_{11} = P_{C1} T_1, T_{01} = P_{C0} T_0$$

となる. したがって、交絡の影響を補正した発生率比  $IR$  は

$$IR = \frac{A_{11}/T_{11}}{A_{01}/T_{01}} = \frac{A_1/T_1}{A_0/T_0} \cdot \frac{IR_{DC} P_{C0} + 1 - P_{C0}}{IR_{DC} P_{C1} + 1 - P_{C1}}.$$

さらに、元のデータから計算した未補正の発生率比  $IR^*$  との関係は

$$IR = IR^* / \frac{IR_{DC} P_{C1} + 1 - P_{C1}}{IR_{DC} P_{C0} + 1 - P_{C0}}$$

となる. これらは、いずれもリスク比の場合と同様である (3.1.6、及び参考 3- 2). 対応する式を比較すると、いずれも  $N_1 \rightarrow T_1$ 、 $N_0 \rightarrow T_0$  という置き換え、及びこれにともなう置き換え  $RR_{DC} \rightarrow IR_{DC}$  によって得られる関係であることがわかる.

さらに、交絡要因の影響が曝露の有無により異なることが想定される場合についても、曝露の有無により異なる発生率比を定めると、層ごとに異なる発生率比が得られる. それらを Mantel-Haenszel の方法をもちいて併合したり、標準化にもちいることができる点も、リスク比の場合と同様である.

以上では発生率比を例にとったが、発生率差についても同様である. すなわち、 $N_1 \rightarrow T_1$ 、 $N_0 \rightarrow T_0$  という置き換え、及びこれにともなう置き換え  $RD \rightarrow ID$  により、リスク差の場合と同様な関係が発生率差についても得られる.

### 3.8.3 曝露 人-時間 の誤分類

発生率の分母である曝露 人-時間 の誤分類については、リスク差やリスク比の場合とやや異なり、例えば以下のパラメータを定めて感度分析を行う.

- Se : 曝露 人-時間 の分類に関する感度
- Sp : 曝露 人-時間 の分類に関する特異度

このようにパラメータを定義すると、例えば曝露ありのときの人-時間  $T_1$  のうち、正しく曝露ありの人-時間  $T_1^*$  の一部として観察されるのは Se  $T_1$ 、また誤って曝露なしの人-時間  $T_0^*$  の一部として観察されるのは  $(1 - \text{Se}) T_1$  となる. これらは、本来なら曝露ありに分類されるべき人-時間なので、この間に観察されるアウトカム数の期待値は、それぞれの人-時間に、曝露ありの場合の発生率  $A_1/T_1$  をかけることで得られ

$$\text{Se } T_1 \cdot \frac{A_1}{T_1} = \text{Se } A_1, (1 - \text{Se}) T_1 \cdot \frac{A_1}{T_1} = (1 - \text{Se}) A_1$$

となる. 前者は、正しく曝露ありのときのアウトカム  $A_1^*$  の一部として観察され、後者は誤って曝露なしの場合のアウトカム  $A_0^*$  の一部として観察される.

同様な関係が曝露なしのときの人-時間  $T_0$  についてもなりたつので、まとめると、誤分類後の分割表のデータ ( $A_1^*$ 、 $A_0^*$ 、 $T_1^*$ 、 $T_0^*$ ) と誤分類前の分割表のデータ ( $A_1$ 、 $A_0$ 、 $T_1$ 、 $T_0$ ) との間に、以下がなりたつことがわかる。

$$A_1^* = \text{Se } A_1 + (1 - \text{Sp})A_0, \quad A_0^* = (1 - \text{Se}) A_1 + \text{Sp } A_0,$$

$$T_1^* = \text{Se } T_1 + (1 - \text{Sp})T_0, \quad T_0^* = (1 - \text{Se}) T_1 + \text{Sp } T_0.$$

これらを解くことで、誤分類前の分割表のデータが得られるが、それは、リスク差やリスク比について導いた同様な式に  $N_1 \rightarrow T_1$ 、 $N_0 \rightarrow T_0$  という置き換えを行っただけの内容である。例えば、曝露の誤分類の影響を補正した発生率比  $IR$  は、式 3-4 に対してこの置き換えを行うことで得られ

$$IR = \frac{A_1/T_1}{A_0/T_0} = \frac{\text{Sp}A_1^* - (1 - \text{Sp})A_0^*}{\text{Se}A_0^* - (1 - \text{Se})A_1^*} \bigg/ \frac{\text{Sp}T_1^* - (1 - \text{Sp})T_0^*}{\text{Se}T_0^* - (1 - \text{Se})T_1^*}$$

となる。

以上をふりかえってみると、人-時間 ( $T_1$ 、 $T_0$ ) に対する誤分類の感度と特異度が、そのままアウトカム ( $A_1$ 、 $A_0$ ) の誤分類にも反映されている。このため、結果的に非差別的誤分類を想定した内容となっているが、これは偶然ではない。曝露によってアウトカムの発生が速くなったり、遅くなったりする場合、その影響によって人-時間もまた変化する。つまり、人-時間もまた曝露の結果 (アウトカム) の一部だとみなすことができる。この点は、アウトカムの誤分類を補正する際にも問題となる。

### 3.8.4 アウトカムの誤分類

アウトカムの誤分類についても、リスク差やリスク比の場合とやや異なり、例えば以下のパラメータを定めて感度分析を行う。

- $\text{Se}$  : アウトカム分類の感度
- $\text{Fr}$  : 単位人-時間あたりに、誤ってアウトカムありと分類される (偽陽性の) 人数

ここでの偽陽性は、本当はアウトカムがないのに、単位人-時間あたり  $\text{Fr}$  人がアウトカムありに分類されるという意味である。つまり  $\text{Fr}$  は誤分類の発生率 (rate) を表している。定義から、 $\text{Fr}$  の単位は人/(人-時間) = 1/時間 である。したがって、曝露人-時間が  $T$  のときに、誤ってアウトカムありに分類される対象者数は  $\text{Fr } T$  となる。

さて、曝露ありの場合を考えると、アウトカムありに分類される対象者数の合計  $A_1^*$  は、真にアウトカムありの  $A_1$  人のうち正しくありと分類される  $\text{Se } A_1$  人、及び実際にはアウトカムなしなのに曝露人-時間  $T_1$  のうちに誤ってありと分類される  $\text{Fr } T_1$  人の和となる。すなわち

$$A_1^* = \text{Se } A_1 + \text{Fr } T_1.$$

これを  $A_1$  について解いて  $T_1$  で割ると

$$\frac{A_1}{T_1} = \frac{A_1^* - \text{Fr } T_1}{\text{Se}T_1}$$

という関係が得られる。

なお、アウトカムが起こっていないのに、ある時点で誤ってアウトカムありに分類されると、その時点で観察が打ち切られる。あるいはデータ上、そのようなあつかいがなされる。このため、アウトカムに誤分類が生じると人-時間にも影響が及ぶ。すなわち、曝露ありの場合の人-時間  $T_1$  がそのまま観察されることはなく、アウトカムの誤分類の影響を受けた不正確な人-時間が観察される。これを  $T_1^*$  で表すと、厳密には  $T_1^* \neq T_1$  となるはずである。

しかし、アウトカムや誤分類の発生が稀な場合、人-時間への影響はわずかであろう。このように  $T_1^* \approx T_1$  と考えてよい場合、上式は

$$\frac{A_1}{T_1} \approx \frac{A_1^* - Fr T_1^*}{Se T_1^*}$$

と近似できる。この関係をもちいると、誤分類後の分割表のデータ ( $A^*$ 、 $A_0^*$ 、 $T_1^*$ 、 $T_0^*$ ) とパラメータの値から、誤分類前の分割表のデータ ( $A_1$ 、 $A_0$ 、 $T_1$ 、 $T_0$ ) を求めることができる。例えば、曝露の誤分類の影響を補正した発生率比  $IR$  は

$$IR = \frac{A_1/T_1}{A_0/T_0} \approx \frac{A_1^* - Fr T_1^*}{A_0^* - Fr T_0^*} \bigg/ \frac{T_1^*}{T_0^*}$$

となる。以上では、非差別的誤分類を想定したが、曝露の有無別に  $Se$  と  $Fr$  が異なる差別的誤分類を想定すると、発生率比の近似式として以下が得られる。

$$IR = \frac{A_1/T_1}{A_0/T_0} \approx \frac{A_1^* - Fr_1 T_1^*}{A_0^* - Fr_0 T_0^*} \bigg/ \frac{Se_1 T_1^*}{Se_0 T_0^*}.$$

なお、アウトカムや誤分類の発生が稀でない場合は、ハザード比等に対して感度分析を行う方法がいくつか提案されている。やや発展的な内容であるため本書では割愛したが、関心がある読者のために参考文献等を紹介した (参考 3- 16)。

### 3.8.5 選択バイアス

発生率を含む効果指標についても、選択確率による方法と観察されなかった層の分割表を求める方法をもちいることができる。

選択確率をもちいる方法については、例えば以下のようにパラメータを定めることが考えられる。

- ・  $S_{A_1}$  : 曝露あり ( $E = 1$ )、アウトカムあり ( $D = 1$ ) の対象者が選択される確率
- ・  $S_{A_0}$  : 曝露なし ( $E = 0$ )、アウトカムあり ( $D = 1$ ) の対象者が選択される確率
- ・  $S_{T_1}$  : 曝露ありの人-時間が選択される確率
- ・  $S_{T_0}$  : 曝露なしの人-時間が選択される確率

このようにパラメータを設定すると、コホート研究におけるリスク比に関する式 3-9 に対し、 $N_1 \rightarrow T_1$ 、 $N_0 \rightarrow T_0$  という置き換えを行った下式が得られる。

$$IR = \frac{A_1/T_1}{A_0/T_0} = \frac{A_{11}/T_{11}}{A_{01}/T_{01}} \left( \frac{S_{A_1}/S_{T_1}}{S_{A_0}/S_{T_0}} \right)^{-1} = IR^* \left( \frac{S_{A_1}/S_{T_1}}{S_{A_0}/S_{T_0}} \right)^{-1}.$$

この場合も、ふたつの発生率比 ( $IR$  と  $IR^*$ ) の比、すなわち系統誤差に相当する量  $RIR$  をあらためてパラメータとして定め、感度分析を行うことができる。



$$RIR = \frac{S_{A_1}/S_{T_1}}{S_{A_0}/S_{T_0}} = \frac{IR^*}{IR}.$$

ところで、曝露ありの場合の人-時間  $T_1$  のうち、ある研究で  $S_{T_1} T_1$  だけ選択されるとすると、この人-時間において期待されるアウトカム数は

$$S_{T_1} T_1 \cdot \frac{A_1}{T_1} = S_{T_1} A_1$$

である。一方、曝露ありで、かつアウトカムありの対象者  $A_1$  人のうち、この研究で選択されるのは  $S_{A_1} A_1$  人である。したがって両者が一致すれば、すなわち  $S_{T_1} = S_{A_1}$  ならば曝露ありの場合の発生率にバイアスが生じない。同様に、 $S_{T_0} = S_{A_0}$  ならば曝露なしの場合の発生率にもバイアスが生じないので、「 $S_{T_1} = S_{A_1}$  かつ  $S_{T_0} = S_{A_0}$ 」ならば、発生率比にバイアスが生じないことが確認できる。

### 3.9 多次元感度分析

これまで個々のバイアスについて見てきたように、1つのバイアスを補正するために、通常、複数のパラメータを設定する。さらに、各パラメータに対して値を1つ定め、その値が正しいと仮定して、バイアスを補正するのだった。しかし、あらためて指摘するまでもなく、パラメータの値を1つに定めることは、そう簡単なことではない。「この値のはずだ」と明確に想定できることは、むしろ稀かもしれない。

以下では、計算方法そのものは単純な感度分析とおなじだが、複数のパラメータに対して複数の値を代入することで、バイアスの影響をより多角的に検討する方法を概観する。計算の内容は、これまで述べてきたものとおなじだが、複数の、つまり多次元のパラメータを同時にあつかうという意味で、こうした方法を**多次元感度分析**と呼ぶ。多次元感度分析により、個々のパラメータの値をこの範囲で変化させると、効果指標の値がこの範囲で変化する、といった傾向を大まかに把握することができる。

なお、多次元感度分析については、これまで述べてきた、交絡、誤分類、及び選択バイアスについて網羅的に述べるのではなく、交絡とアウトカムの誤分類を例にとってその概略のみ紹介する。

#### 3.9.1 交絡

ふたたび、WTCHRの数値例を検討しよう。交絡に対し、以下の3つのパラメータの値を定め、単純な感度分析を行った(3.1)。

- $RR_{DC} = 2$  :  
未知の交絡要因(C)とぜん息(D)の関連性をしめすリスク比(交絡要因あり/交絡要因なし)。ただし、このリスク比は、曝露の有無によらず等しいと仮定。
- $P_{C1} = 0.9$  :  
曝露あり( $E = 1$ )の集団で未知の交絡要因あり( $C = 1$ )の対象者が占める割合。
- $P_{C0} = 0.5$  :  
曝露なし( $E = 0$ )の集団で未知の交絡要因あり( $C = 1$ )の対象者が占める割合。

これらのパラメータに対し、以下でしめした複数の値を代入してみよう。

- $RR_{DC} = 2, 4, 6, 8$
- $P_{C1} = 0.1, 0.3, 0.5, 0.7, 0.9$
- $P_{C0} = 0.1, 0.3, 0.5, 0.7, 0.9$

これらの値を式3-2に代入すると、値の組み合わせごとに、交絡の影響を補正したリスク比RRが得られる。その結果を表にまとめた(Table 3-18)。

この結果から、大まかな傾向として以下が読み取れる。

- 交絡要因ありの対象者の割合が曝露の有無によらず等しい場合、つまり  $P_{C1} = P_{C0}$  となる各表の対角線上では、補正後のリスク比は未補正のリスク比(1.73)と等しい。

すなわち、アウトカムに影響を及ぼす要因があったとしても、その分布が曝露の有無によらず等しければ交絡が生じないことが確認できる。

- ・ 交絡要因ありの対象者の割合が、曝露なしの群で大きい場合、つまり  $P_{C1} < P_{C0}$  となる各表の対角線の上側では、補正後のリスク比が未補正（1.73）のリスク比より大きい。また、この逆もなりたつ。また、交絡要因がアウトカム及ぼす影響（ $RR_{DC}$ ）が大きいほど、こうした変化が大きい。

Table 3- 18 多次元感度分析の例：交絡の影響を補正したリスク比

$RR_{DC}$	$P_{C1}$	$P_{C0}$				
		0.1	0.3	0.5	0.7	0.9
2	0.1	1.73	2.05	2.36	2.68	2.99
	0.3	1.47	1.73	2.00	2.27	2.53
	0.5	1.27	1.50	1.73	1.96	2.19
	0.7	1.12	1.32	1.53	1.73	1.94
	0.9	1.00	1.19	1.37	1.55	1.73
4	0.1	1.73	2.53	3.33	4.13	4.93
	0.3	1.19	1.73	2.28	2.83	3.37
	0.5	0.90	1.32	1.73	2.15	2.56
	0.7	0.73	1.06	1.40	1.73	2.07
	0.9	0.61	0.89	1.17	1.45	1.73
6	0.1	1.73	2.89	4.04	5.20	6.35
	0.3	1.04	1.73	2.43	3.12	3.81
	0.5	0.74	1.24	1.73	2.23	2.72
	0.7	0.58	0.96	1.35	1.73	2.12
	0.9	0.47	0.79	1.10	1.42	1.73
8	0.1	1.73	3.16	4.59	6.01	7.44
	0.3	0.95	1.73	2.51	3.30	4.08
	0.5	0.65	1.19	1.73	2.27	2.81
	0.7	0.50	0.91	1.32	1.73	2.14
	0.9	0.40	0.74	1.07	1.40	1.73

単純な感度分析で例示したのは、この表の1箇所、つまり一番上の  $OR_{DC} = 2$  の表で、行が  $P_{C1} = 0.9$ 、列が  $P_{C0} = 0.5$  となる箇所だった（3.1.5）。すでに見た通り、補正後のリスク比は1.37である。この周辺をざっと眺めてみると、リスク比が1未満となる場合は見あたらない。しかし交絡要因がアウトカム及ぼす影響（ $RR_{DC}$ ）がさらに大きく4以上となり、さらに曝露なしの群における交絡要因ありの対象者の割合（ $P_{C0}$ ）が30%以下となると、太い青字でしめたように、リスク比が1未満となる場合があることがわかる。

はたして、ぜん息のリスクを4倍以上も高める要因が存在するのか、またそのような要因の分布が3倍以上（例えば30%と90%等）も異なることが現実にあるのか、解析者の見識や判断に基づき、さらに考察を深めることができるだろう。

このように、多次元感度分析では、解析者があらかじめ想定を置くことなく、どのような場合にどの程度のバイアスが生じうるか、大まかな状況を把握することができる。以上で

例示したように、効果指標の値がある値を超えるのはどのような場合か、といった具体的な目標があるとさらに考えやすい。

### 3.9.2 アウトカムの誤分類

アウトカムの誤分類については、WTCHR の数値例に対して以下のパラメータを定め、単純な感度分析を行った (3.4.3)。

- $Se_1 = 0.85$  : 曝露ありの場合の、アウトカム分類の感度
- $Sp_1 = 1.00$  : 曝露ありの場合の、アウトカム分類の特異度
- $Se_0 = 0.90$  : 曝露なしの場合の、アウトカム分類の感度
- $Sp_0 = 0.98$  : 曝露なしの場合の、アウトカム分類の特異度

これらのパラメータに対し、以下でしめした複数の値を代入してみよう。

- $Se_1 = 0.85, 0.90$
- $Sp_1 = 0.98, 0.99, 1.00$
- $Se_0 = 0.85, 0.90$
- $Sp_0 = 0.98, 0.99, 1.00$

これらの値を式 3-6 に代入すると、アウトカムの誤分類の影響を補正したリスク比が得られる。その結果を表にまとめた (Table 3- 19)。

Table 3- 19 多次元感度分析の例：アウトカムの誤分類の影響を補正したリスク比

$Se_1$	$Sp_1$	$Se_0$ $Sp_0$	0.85			0.90		
			0.98	0.99	1.00	0.98	0.99	1.00
0.85	0.98		3.53	1.60	1.05	3.74	1.70	1.11
	0.99		4.70	2.14	1.39	4.98	2.26	1.48
	1.00		5.84	2.65	1.73	6.19	2.81	1.83
0.90	0.98		3.33	1.51	0.99	3.53	1.60	1.04
	0.99		4.44	2.02	1.32	4.70	2.14	1.39
	1.00		5.52	2.51	1.64	5.85	2.66	1.73

この結果から、大まかな傾向として以下が読み取れる。

- 曝露の有無によらず感度が等しく ( $Se_1 = Se_0 = 0.85$ 、または  $Se_1 = Se_0 = 0.90$ )、かつ特異度も等しい ( $Sp_1 = Sp_0 = 0.98$ 、 $Sp_1 = Sp_0 = 0.99$ 、または  $Sp_1 = Sp_0 = 1.00$ ) 非差別的誤分類の場合、補正後のリスク比は、未補正のリスク比 (1.73) 以上の値となる。言い換えれば、補正後のリスク比が真の値だと仮定すると、非差別的誤分類により、未補正のリスク比は関連性がないことをしめす方向 (より 1 に近い方向) に偏っていることが確認できる。また、特別な場合として、曝露の有無によらず特異度が 1 の場合 ( $Sp_1 = Sp_0 = 1.00$ ) は、補正の前後でリスク比の値が変わらず、バイアスが生じないことが確認できる (3.4.5)。

- ・ 上の表でも下の表でも、曝露ありの場合の感度 ( $Se_1$ ) と特異度 ( $Sp_1$ ) を固定して (おなじ行に注目して)、曝露なしの場合の感度 ( $Se_0$ ) が 5%異なる場合の (ただし特異度  $Sp_0$  は等しい場合の) リスク比を左右で比較すると、その違いはわずかであることが確認できる。一方、曝露の有無別に特異度が 1%異なる場合は、より大きく補正值が変化する。また、曝露ありの群の特異度の方が大きく、 $Sp_1 > Sp_0$  となる各表の対角線の下側では補正值が 1.73 より大きく、逆に曝露なしの群の特異度の方が大きく、 $Sp_1 < Sp_0$  となる各表の対角線の上側では補正值が 1.73 より小さい。

単純な感度分析で例示したのは、この表の 1 箇所、つまり右上の表 ( $Se_1 = 0.85$ 、 $Se_0 = 0.90$  の表) で、 $Sp_1 = 1.00$ 、 $Sp_0 = 0.98$  となる左下角の箇所だった (3.4.3)。すでに見たように、補正後のリスク比は 6.19 である。この事例では、感度についての想定はあまり影響がなく、特異度の値の組み合わせがこの事例とおなじ場合 ( $Sp_1 = 1.00$ 、 $Sp_0 = 0.98$  となる場合) を比較すると、リスク比の補正值がいずれも 5.5 を超えている。このため、曝露ありの群の特異度の方がわずかながら高いという想定が確かなら、おそらく真のリスク比は 2.0 より大きいのではないかと判断することができる。

このように、多次元感度分析では、解析者の想定がやや漠然としていても、どのような場合にどの程度のバイアスが懸念されるか、大まかに状況を把握することができる。

### 3.9.3 多次元感度分析の有用性と限界

バイアスに対する解析者の想定が正しいとは限らないし、その想定が漠然としていることもある。こうした場合、ただ 1 つの想定にもとづく単純な感度分析には限界がある。検討したいバイアスについて、いったん単純な感度分析の方法が確立できたら、おなじ計算を繰り返すだけで、容易に多次元感度分析を行うことができる。こうした手軽さが、多次元感度分析の有用性である。

一方、多次元感度分析には限界もある。例えば、多次元感度分析では、個々のパラメータに代入する複数の値をおなじ重みであつかっている。これは実質的に、各パラメータに代入するいずれの値も同様に確からしいと想定しているようなものである。複数のパラメータのあつかいについても同様で、それらの値は独立であり、どのパラメータがどの値をとることも同様に確からしいと、暗黙のうちに想定しているかのようである。しかし、こうした想定は、はたして現実的だろうか。

例えば、あるパラメータについて、この値の方が他の値より尤もらしいと考えられることがある。つまり、あるパラメータの値が不確かだと言っても、あらゆる値が同様に確からしいわけではないかもしれない。また、あるパラメータの値が大きければ、もう一方のパラメータの値も大きいと考えられる場合がある。つまり両者には相関があると想定される場合である。さらに、あるパラメータについては値を想定しやすいが、別のパラメータについてはもっと漠然としている場合もあるだろう。この場合は、パラメータごとに解析者が抱いている不確かさの程度が異なっている。

あるバイアスに対して解析者がいなく、こうした想定を表現しようとする、第 2 部で紹介する確率的感度分析の方が、むしろ自然だと感じられるかもしれない。

## 第 1 部のまとめ

この第 1 部では「ある観察研究の、ある効果指標の推定値におけるバイアス」と対象を絞って感度分析を行う方法のうち、おもに単純な感度分析について紹介した。

まず、「1 実践的な感度分析に向けて」では、その観察研究の標的とみなせる理想的な試験（ターゲット試験）を想定し、現実の研究との差をさまざまな角度から比較・検討する方法を紹介した（1.1）。また、こうした考え方をヒントに、観察研究の段階（介入前、介入時、介入後）ごとに、どのようなバイアスが問題となるか、その分類と、対応する感度分析の方法について概観した（1.2）。さらに、こうした感度分析の背後でしばしば想定されている推定値の内部構造、とくに系統誤差と偶然誤差の関係について大まかに要約した（1.3）。また、本書であつかっている感度分析の方法の概略をしめした（1.4）。

「2 本書であつかう効果指標と数値例」では、効果指標の定義（2.1.1）にくわえ、こうした感度分析でしばしば行われる問題の単純化と、その意義について述べた。その際、こうした単純化が、計算の簡略化だけでなく解釈の容易さを意図したものであることにふれ、こうした感度分析の目的が、より厳密なバイアス補正を行い、より正確な推定値を得ることではないことを指摘した（2.1.2）。また、いっけん単純な方法が意外にも広範な応用をもつことを紹介した（2.1.3）。数値例としては、以降でもちいるレジストリー研究の事例を簡単に紹介した（2.2）。

第 1 部の大部分を占める「3 単純な感度分析とその応用」では、まずコホート研究の場合について、単純な感度分析の方法を交絡（3.1）、誤分類（3.2、3.3、3.4、3.5）、選択バイアス（3.6）の順に紹介した。また、ケースコントロール研究（3.7）や人時間を分母とした発生率を含む効果指標（3.8）に対する感度分析についても概説した。さらに、単純な感度分析の、直接的な応用として多次元感度分析について、その概略を紹介した（3.9）。

すでに指摘したように、第 1 部で紹介した方法では、以下の 2 種類の不確かさを明確に、また量的にあつかっているとは言い難い（1.4、3.9.3）。

- ・ 系統誤差の方向や大きさに対して、解析者自身が抱いている不確かさ
- ・ 偶然誤差による、効果指標の推定値がもつ不確かさ

こうした不確かさを柔軟に、また量的にあつかえる方法として、第 2 部では確率的感度分析とその応用について紹介する。

また、第 1 部では、これらの方法を運用するための準備や、得られた情報を感度分析においてどうもちいるか、また結果の報告をどう行うか等についてはごく限られた内容しかあつかっていない。こうした、さらに実践的な内容についても第 2 部で紹介する。

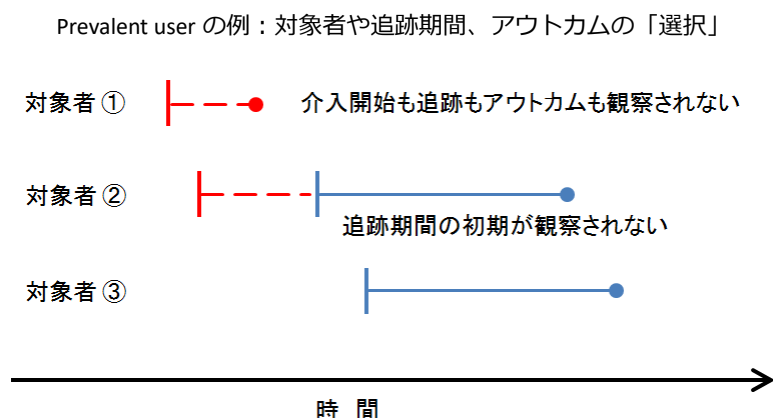


## 参 考 [本文中の対応箇所を括弧内に表示]

### 参考 1- 1 new user ではなく、prevalent user を試験に含めた場合 [本文：1.1.3]

その研究の目的が、あるアウトカムの発生状況を、新薬と既存薬とで比較することだったとしよう。ただし研究者の意図としては、いずれの薬剤についても **new user** のみを選択し、両者のアウトカムを比較したい。ところが新薬群については **new user** のみを選択することができるのだが、既存薬群については、データを過去に遡って調べることに限界があり、結果的に **prevalent user** も選択される可能性があるとしよう。

この例を下図を見ながら考えてみよう。いずれも既存薬群に含まれるべき対象者だとして、データベースの情報から、どのようなことが確認できるだろうか。なお、図中の青い部分がデータベースから確認できる情報であり、赤い部分はデータベースには含まれていない過去の情報だとする。



この図で対象者①は実際には **new user** だが、データベースに情報が含まれていないためこの研究からは除外される。

一方、対象者②はこの研究では **new user** としてあつかわれる可能性がある。しかし赤い点線でしめされているように、追跡期間の初期の情報がデータベースに含まれていない。このため、追跡期間の一部がデータ上は除外される。さらにこの期間に生じたアウトカムも把握できないため、アウトカムの一部も除外される可能性がある。

対象者③は実際に **new user** であり、この研究でも **new user** としてあつかわれる。

さて、対象者①のように対象者ごと除外されたり、対象者②のように追跡期間やアウトカムの一部が除外されることは、この事例では既存薬群に限って生じる。言い換えれば、この研究における患者やデータの選択や除外は、比較したい介入と関連している（介入 → 選択・除外）。

また、この選択や除外は、おそらくアウトカムとも関連している。というのは、既存薬を過去に使用してアウトカムが良好だった対象者ほど、現在もそれを使用している可能性が高いからである（アウトカムが良好な健康状態 → 選択・除外）。結局のところ、この選択や除外は、介入ともアウトカムとも関連している。つまり、この事例は「適格な対象者の一部や、一部の対象者の介入開始直後の追跡期間、もしくはアウトカムの一部を除外することが、介入ともアウトカムとも関連している」場合であり、構造上、選択バイアスが生じる懸念がある（1.2.1 Table 1- 1）。

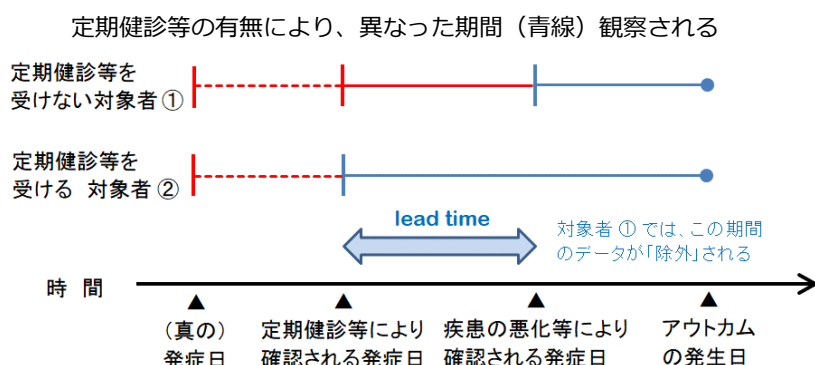
同様な構造で生じるバイアスとして **lead time bias** や **immortal time bias** があるので概観しておこう。

### Lead Time Bias

例えば、定期健診等を積極的に行う場合と、そうでない場合とで、ある疾患の発症からアウトカム発生までの時間を比較したいのだとしよう。既存のデータベース等をもちいてこの比較を行うと、定期健診等を積極的に行う対象者ほど、疾患の発症が早期に発見され、見かけ上、発見からアウトカムが生じるまでの時間が長い、というデータが得られるかもしれない（下図参照：対象者②）。逆に、定期健診等を積極的に行わない対象者では、疾患が発見された時点で既に病態が悪化しており、発見からアウトカム発生までの時間が短く観察されるかもしれない（下図参照：対象者①）。言い換えれば、後者の対象者では、積極的に定期健診等を行っていたとしたら確認できたはずの疾患の発症日から、実際に発症が確認された日までの期間（下図の **lead time** にあたる期間）が観察されないため、データとしては除外される可能性がある。また、アウトカムによっては、この期間に発生していたとしても、定期健診等の頻度が低いせいで確認できないかもしれない。つまり、この期間のアウトカムについても、一部除外される可能性がある。

こうした期間やアウトカムの除外は、定期健診等を積極的に行わない対象者でより多く生じだろう（定期健診等への積極性 → 選択・除外）。またこの除外は、健診に対する対象者の意識や生活習慣、経済状況等、アウトカムとも関連している可能性がある（アウトカムと関連する要因 → 選択・除外）。したがってこの事例でも「介入開始直後の追跡期間、もしくはアウトカムの一部」を除外することが介入ともアウトカムとも関連している可能性があり、構造上、選択バイアスが生じる懸念がある。

Lead time bias の例：発症からアウトカム発生までの時間が等しい対象者が



### Immortal Time Bias

ある方法で介入や曝露を定義すると、追跡期間のなかに、関心があるアウトカムがまったく発生しない期間が生じることがある。アウトカムが死亡だとすると、対象者が死亡しない、つまり不死 (immortal) である期間が生じることがある。

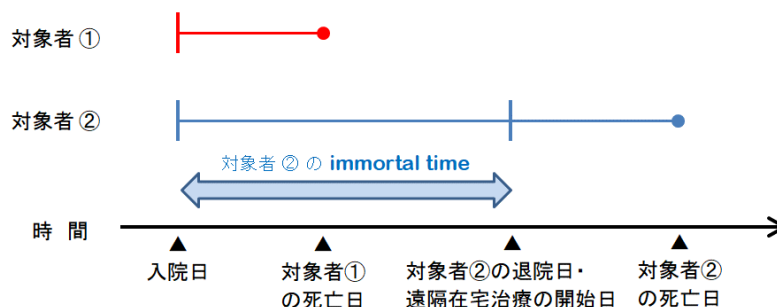
例えば、慢性心不全で入院した患者の予後と、通常治療を行う場合と、遠隔在宅治療を行う場合とで比較したいのだとしよう。このとき、遠隔在宅治療を受けたことがデータから確認できる患者を「遠隔在宅治療群」と定義すると、入院後、退院に至ることなく死亡した患者は、この定義により除外される（下図参照：対象者①）。一方、「遠隔在宅治療群」の患者が、入院開始から遠隔在宅治療の開始までに死亡することは定義上起こりえない（下図参照：対象者②）。このような患者では、入院開始から遠隔在宅治療の開始までの期間が不死である時間 (immortal time) となる。

さて、ここで「もし入院時に、ランダムに治療方針を割り付けていたとしたら」と考えてみよう。すると「遠隔在宅治療群」においても、入院中の死亡は生じえただろう。ところがこの観察研究では、これに該当する患者は少なくとも「遠隔在宅治療群」からは除外され、一方、定義しだいで「通常治療群」として選択されるかもしれない。この問題は、ランダム化の時期（もしくは追跡開始に相当する時点の定義）を少し遅らせ「入院後、一定の日数が経過した時点」に変更しても生じうることに注意したい。

この事例でも、一方の群に限って「介入開始直後の追跡期間、もしくはアウトカムの一部」が除外、もしくは選択されることになり（通常治療か遠隔在宅治療か → 選択・除外）、しかもこの除外や選択は、予後が悪い患者で生じやすいだろうから、アウトカムの発生と無関係とは考えにくい（アウトカムと関連す

る要因 → 選択・除外)．つまりこの事例でも、これまでと同様なメカニズムで「介入開始直後の追跡期間、もしくはアウトカムの一部」が選択・除外され、選択バイアスが生じる懸念がある．

Immortal time bias の例：ランダム化により遠隔在宅治療に割り付けられた可能性のある対象者が、入院中の死亡により「除外」される



ところで、以上のようなバイアスの名称を数多く知っていることが重要なわけではない．こうしたバイアスが発生する理由や構造を、個々の試験についてよく見極めることが重要である．その際、immortal time bias の事例に対して行ったように「もし理想的なランダム化比較試験（ターゲット試験）を行ったとしたら」と考えると、どのような情報が欠落しがちか、またはその懸念があるか、具体的な見通しが得られることがある．

## 参考 1- 2 意図した介入からの逸脱、及び介入後の欠測によるバイアス [本文：1.2.2、3.6.1]

### 意図した介入からの逸脱によるバイアス

ROBINS-I では、割り当てられた、あるいは開始された介入の効果を intention-to-treat 効果（以下、ITT 効果）、また介入が開始され、かつそれが遵守された場合の効果を per-protocol 効果（以下、PP 効果）と呼び区別している<sup>3, 4</sup>．

ITT 効果に関心がある場合「意図した介入からの逸脱」が通常の治療経過のなかで自然に生じうるものなら、バイアスの懸念は低いかもしれない．例えば、有害事象の発現により介入が中断したとしても、その後も追跡が継続され、必要なデータが得られていれば、割り当てられた、あるいは開始された介入の効果を推定することができる．

一方、PP 効果に関心がある場合「意図した介入からの逸脱」はそのままバイアスの原因となる．このバイアスは、同様な効果（例えば、割り付けられた治療を逸脱なく完了した場合の効果）に関心がある場合は、ランダム化比較試験でも生じうる．

ある要因の値が時間とともに変化し、かつその値が、その後のアウトカムの発生だけでなく、その後の介入や治療の変更（意図した介入からの逸脱）にも影響を及ぼす場合、いわゆる時間依存性交絡（time-varying confounding）が生じる．

時間依存性交絡を、回帰モデル等をもちいた一般的な解析で調整することは適切ではない．例えば、関心がある効果指標がハザード比の場合、アウトカムに影響を及ぼす要因の値の変化や、介入や曝露の変化を時間依存性共変量として考慮した Cox 回帰モデルをもちいても、こうしたバイアスは制御できない<sup>11</sup>．

また「意図した介入からの逸脱」を「打ち切り」とみなして解析することも考えられる．すなわち、開始された介入が初めて変更された時点で（例えば、単剤の効果を調べたい研究でアウトカムの発生に影響を及ぼす他剤の併用が開始された時点で）観察が打ち切られたと考え、それ以降のデータを除外する．ただし、こうしたデータの除外は「追跡が開始された対象者の、その後の情報が欠測することにより生じるバ

バイアス」を生み出す懸念がある（1.2.1 Table 1-1）．このため「打ち切り」によって生じる選択バイアスの影響を、適切に考慮した解析が必要になる．

時間依存性交絡も考慮したうえで、意図した治療を遵守した場合の介入効果（PP 効果）を推定するには、G-method と総称される以下の方法をもちいることができる．

- Parametric g-formula <sup>12</sup>
- Inverse probability weighting of marginal structural models <sup>13</sup>
- G-estimation of nested structural models <sup>14, 15</sup>

ただし、こうした方法によって無条件に PP 効果が推定できるわけではない．必要なデータはどのようなものか、またそれらが満たすべき条件はなにか、事前に検討すべきことがある．その基礎となる概念や、これらの方法を紹介した、以下の教科書の出版が予定されている．

Hernán MA, Robins JM. Causal Inference. Boca Roton: Chapman & Hall/CRC, 2019.

また、その草稿の一部が Web 上で公開されている．

<https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/> [accessed March 16, 2019].

### 介入後の欠測によるバイアス

ROBINS-I でも述べられているように、介入を受けた対象者の、追跡後の情報が欠測することで、選択バイアスが生じる懸念がある（1.2.1 Table 1-1）．追跡不能によるアウトカムの欠測が構造的に選択バイアスを生むメカニズムについては、入門編でも簡単に紹介した（例えば、付録 1）．こうした欠測の発生もまた「中間事象」の一種であり、ランダム化比較試験でも生じうる（1.2.2）．

ROBINS-I でも解説されているように、1) 欠測データの割合が比較したい群間で同様であり、かつ 2) 欠測の原因が比較したい群間で同様なら、効果指標へのバイアスはさほど深刻なものではないはずである．しかし、2つの条件の一方が成立しない場合は注意が必要となる．なお、こうした欠測データの解析にも G-method をもちいることができる．先述した文献で紹介されているので、あわせて参照して頂きたい．

## 参考 1- 3 臨床試験等で行われる統計解析における「感度分析」 [ 本文 : 1.3.3 ]

臨床試験等で行われる統計解析では系統誤差が十分に制御されていることを、暗黙のうちに仮定していることが少なくない．このため実質的に、以下の 2 点を問題にしていることが多い．

- 偶然誤差に対して想定しているモデル（sampling model）は適切か？  
単純な分割表に対する解析でも、その背景には、データの生成過程で発生する偶然誤差に対する想定がある．例えば 2 項分布、ポアソン分布、正規分布等、何らかの確率モデルに従ってデータが発生することを想定している．
- 真の値に対して想定しているモデル（structural model）は適切か？  
関心があるアウトカムと、介入や曝露、また共変量との関係を、数学的なモデルをもちいて表すことがある．例えばロジスティックモデル、比例ハザードモデル等をもちいて、要因相互の関係をモデル化する．

臨床試験等で行われる感度分析の多くは、以上で述べた「モデル」の妥当性を問題にする．例えば、偶然誤差に対する想定をゆるめ、分布に依存しない解析を試みる（sampling model に対する検討）．また、交互作用項や別の共変量をモデルに含めるなどして結果を比較する（structural model に対する検討）．こうした「モデル」の誤特定によって生じるバイアスを specification bias と呼ぶことがある <sup>16</sup>．

なお、ICH E9「臨床試験のための統計的原則」では安定性（ロバストネス）について以下の記載がある．

安定性（ロバストネス）とは、データ、仮定及び解析方針についての様々な制限に対して全体の結論がどの程度変わり易いか（原文：the sensitivity of the overall conclusions）に関連した概念である。安定性は、異なる仮定又は異なる解析方針に基づいて解析を行った場合でも、試験治療の効果と試験の主要な結論は大きく影響されないことを意味している。

この記載は「全体の結論」や「主要な結論」に対する感度について述べており、効果指標の推定値に対する感度（試験治療の効果）のみに言及した内容ではないかもしれない。しかし「異なる仮定又は解析方針」という箇所に、上述した意味での「モデル」の妥当性に関する検討も含意されていると考えられる。

なお、本書であつかっているような感度分析では、むしろ系統誤差の影響が大きいと見え、上述した「モデル」の誤特定によるバイアスを直接的な対象とはしない。また、データの誤入力等も、現実には起こりうるが対象としない。ただし、実際の観察研究では、上述した意味での「感度分析」をまず十分に行い、さらに本書で述べているような感度分析も行い、双方の結果をともに参照しながら最終的な判断を行うことが薦められる。

## 参考 2- 1 WTCHR に対する研究の概略 [ 本文 : 2.2、3.1.5 ]

本文中で参照した 2 つの文献<sup>19,20</sup>の概略を紹介する。

**Wheeler K, McKelvey W, Thorpe L, et al. Asthma diagnosed after 11 September 2001 among rescue and recovery workers: findings from the World Trade Center Health Registry. Environ Health Perspect 2007; 115(11):1584-1590.**

**目的** WTC の現地で作業やボランティアに従事した人たちを対象として、あらたに診断されたぜん息のリスクを記述すること、また現地で生じた様々な曝露とぜん息との関連性を明らかにすることが、この研究の目的であった。

**方法** 研究には World Trade Center Health Registry (WTCHR) に含まれる 2003 年から 2004 年までのインタビュー・データをもちいた。まず、9 月 11 日までにぜん息を発症していなかった対象者を定め、あらたに診断されたぜん息のリスクと、マスクや防塵マスクの使用を含む作業履歴との関連性を推定した。

**結果** 926 人 (3.6%) の対象者が、あらたにぜん息と診断されたと報告した。到着が早いほど、また作業時間が長いほど、高いリスクが認められ、それぞれが独立したリスク要因であった。同様に、粉じんへの曝露と、堆積物上での直接的作業も、他とは独立したリスク要因であった。9 月 11 日に到着した作業者たちについて、マスクや防塵マスクの使用開始が遅れるほどぜん息のリスクが高いという関連性が認められ、調整済オッズ比は、1 日遅れた場合 1.63 (95%信頼区間は 1.03 - 2.56)、また 16 - 40 週遅れた場合は 3.44 (95%信頼区間は 1.43 - 8.25) であった。  
なお、本文中で参照されている粉じんへの曝露の影響をしめす調整済オッズ比は 1.28 (95%信頼区間は 1.09 - 1.50) であり、調整にもちいた要因は性、年齢、年齢の 2 乗、2001 年 9 月 11 日におけるニューヨーク市居住の有無、作業中の所属機関、堆積物上での作業の有無等であった。

**結論** 自己報告により、あらたにぜん息と診断された割合がこの研究集団では高く、現地での曝露の程度と関連していた。防塵マスク等の使用が適切であったか否かについては、この研究では識別できないが、少なくともマスクや防塵マスクの使用に防護効果があることがわかった。この結果は、呼吸器への曝露を他の方法で制御できない場合に、適切な防護用具を適切かつ迅速に配給し、その使用を促すことの必要性を明確にしめている。



Jurek AM, Maldonado G. Quantitative bias analysis in an asthma study of rescue-recovery workers and volunteers from the 9/11 World Trade Center attacks. *Annals of Epidemiology* 2016; 26(11): 794-801.

**目的** バイアス分析について学ぶとき、疫学者は、誤差が発生するのは逆の順序で複数のバイアスを補正しなさいと教えられる。ある試験で誤差が発生する順序を理解するには、その研究で疫学データが生成した過程を注意ぶかく検討する必要がある。本稿では、人間が引き起こした災害に対する疫学研究ならではのデータの生成過程について記述する。

**方法** 2001年9月11日以降に救助や復旧に携わった作業者やボランティアについて、事故当日の粉じんへの曝露と、自己報告による、あらたに診断されたぜん息とを関連付けている研究に対し、そのデータが生成された過程を記述し、バイアス分析を行った。欠測、アウトカムの誤分類、また研究への不参加がオッズ比の推定値に及ぼした影響を補正した。

**結果** 系統誤差に対するある想定のもとで、3つのバイアスをすべて補正したオッズ比は1.33から3.84という範囲だった。ほとんどの補正值は、この研究で観察されたオッズ比1.77より大きく、また95%信頼区間(1.55, 2.01)からも外れていた。

**結論** 人間が引き起こした災害では、他の領域の疫学ではあまり観察されない、いくつかの状況が生じる。結果の解釈をさらに有意義なものとするためにデータ収集の技術的な側面と、バイアスパラメータに関する情報収集を重視し、さらに積極的なアプローチをとることが今後の災害に対する疫学研究にとって有益であろう。

### 参考 3- 1 交絡の補正：リスク差（効果指標の修飾がある場合）[ 本文：3.1.3 ]

アウトカム	合計		交絡要因			
			$C = 1$		$C = 0$	
	曝露		曝露		曝露	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	$A_1$	$A_0$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
$D = 0$	$B_1$	$B_0$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
合計	$N_1$	$N_0$	$N_{11}$	$N_{01}$	$N_{10}$	$N_{00}$

リスク差が、交絡要因の有無別に異なると考えられる場合は、交絡要因とアウトカムの関連性をしめすリスク差が曝露の有無別に異なると想定し、以下のパラメータを定める。

$RD_{DC1}$ ：曝露ありの場合の、交絡要因とアウトカムの関連をしめすリスク差  
(交絡要因あり－交絡要因なし)。

$RD_{DC0}$ ：曝露なしの場合の、交絡要因とアウトカムの関連をしめすリスク差  
(交絡要因あり－交絡要因なし)。

$P_{C1}$ ： $E = 1$  の場合の  $C = 1$  の割合、すなわち  $P_{C1} = N_{11}/N_1$ 。

$P_{C0}$ ： $E = 0$  の場合の  $C = 1$  の割合、すなわち  $P_{C0} = N_{01}/N_0$ 。

以上の関係から、交絡要因ありの層の度数は

$$A_{11} = P_{C1}\{A_1 + (1 - P_{C1})RD_{DC1}N_1\}, A_{01} = P_{C0}\{A_0 + (1 - P_{C0})RD_{DC1}N_0\}, N_{11} = P_{C1}N_1, N_{01} = P_{C0}N_0.$$

よって、この層のリスク差は

$$RD_1 = \frac{A_{11}}{N_{11}} - \frac{A_{01}}{N_{01}} = \frac{A_1}{N_1} - \frac{A_0}{N_0} - (P_{C1} - P_{C0})RD_{DC1}.$$

一方、交絡要因なしの層の度数は



$$A_{10} = (1 - P_{C1})(A_1 - P_{C1}RD_{DC0}N_1), A_{00} = (1 - P_{C0})(A_0 - P_{C0}RD_{DC0}N_0),$$

$$N_{10} = (1 - P_{C1})N_1, N_{00} = (1 - P_{C0})N_0.$$

よって、この層のリスク差は

$$RD_0 = \frac{A_{10}}{N_{10}} - \frac{A_{00}}{N_{00}} = \frac{A_1}{N_1} - \frac{A_0}{N_0} - (P_{C1} - P_{C0})RD_{DC0}.$$

層別した分割表から、Mantel-Haenszel の方法を用いて層を併合したリスク差を求めることができる。  
また、標準集団を適当に定め、交絡要因の分布を標準集団の分布にそろえた場合のリスク差、すなわち標準化リスク差を求めることもできる。

### 参考 3- 2 交絡の補正：リスク比 [ 本文：3.1.4、3.1.5、3.7.2、3.8.2 ]

アウトカム	合計		交絡要因			
			C = 1		C = 0	
	曝露		曝露		曝露	
	E = 1	E = 0	E = 1	E = 0	E = 1	E = 0
D = 1	A <sub>1</sub>	A <sub>0</sub>	A <sub>11</sub>	A <sub>01</sub>	A <sub>10</sub>	A <sub>00</sub>
D = 0	B <sub>1</sub>	B <sub>0</sub>	B <sub>11</sub>	B <sub>01</sub>	B <sub>10</sub>	B <sub>00</sub>
合計	N <sub>1</sub>	N <sub>0</sub>	N <sub>11</sub>	N <sub>01</sub>	N <sub>10</sub>	N <sub>00</sub>

リスク比については、上表のように度数を表し、以下のパラメータを定める。

$RR_{DC}$ ：交絡要因とアウトカムの関連性をしめすリスク比（交絡要因あり / 交絡要因なし）。ただし、このリスク比は曝露の有無によらず等しいと仮定する。すなわち

$$RR_{DC} = \frac{A_{11}/N_{11}}{A_{10}/N_{10}} = \frac{A_{01}/N_{01}}{A_{00}/N_{00}}, \text{ よって } RR_{DC} = \frac{A_{11}/N_{11}}{(A_1 - A_{11})/(N_1 - N_{11})} = \frac{A_{01}/N_{01}}{(A_0 - A_{01})/(N_0 - N_{01})}.$$

$P_{C1}$ ：E = 1 の場合の C = 1 の割合、すなわち  $P_{C1} = N_{11}/N_1$ 。よって  $N_{11} = P_{C1}N_1$ 。

$P_{C0}$ ：E = 0 の場合の C = 1 の割合、すなわち  $P_{C0} = N_{01}/N_0$ 。よって  $N_{01} = P_{C0}N_0$ 。

以上の関係から、 $N_{11}$  と  $N_{01}$  を消去して整理すると

$$(1 - P_{C1})A_{11} = RR_{DC}(A_1 - A_{11})P_{C1}, (1 - P_{C0})A_{01} = RR_{DC}(A_0 - A_{01})P_{C0}.$$

よって

$$A_{11} = \frac{RR_{DC} P_{C1} A_1}{RR_{DC} P_{C1} + 1 - P_{C1}}, A_{01} = \frac{RR_{DC} P_{C0} A_0}{RR_{DC} P_{C0} + 1 - P_{C0}}.$$

以上から、交絡要因ありの層の度数  $A_{11}$ 、 $A_{01}$ 、 $N_{11}$ 、 $N_{01}$  を元のデータ ( $A_1$ 、 $A_0$ 、 $N_1$ 、 $N_0$ ) とパラメータ ( $RR_{DC}$ 、 $P_{C1}$ 、 $P_{C0}$ ) によって表すことができた。したがって、交絡の影響を補正したリスク比  $RR$  は

$$RR = \frac{A_{11}/N_{11}}{A_{01}/N_{01}} = \frac{A_1/N_1}{A_0/N_0} \cdot \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}P_{C1} + 1 - P_{C1}}.$$

また、元のデータから計算した、未補正のリスク比  $RR^*$  との関係は

$$RR = RR^* \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}P_{C1} + 1 - P_{C1}} = RR^* / \frac{RR_{DC}P_{C1} + 1 - P_{C1}}{RR_{DC}P_{C0} + 1 - P_{C0}}.$$

以上では、交絡要因ありの層のリスク比を求めたが、なしの層で同様な計算を行ってもおなじリスク比が得られる。その原因は、 $RR_{DC}$  の値が曝露の有無によらず等しいと仮定したことにある。

交絡要因の影響が曝露の有無により異なることが想定される場合は、曝露の有無により異なるリスク比、例えば曝露ありの場合は  $RR_{DC1}$ 、曝露なしの場合は  $RR_{DC0}$  等と定める。すると、上式の該当箇所が層ごとに異なる値を取り、層ごとのリスク比は交絡要因の有無別に以下ようになる。

$$RR_1 = \frac{A_1/N_1}{A_0/N_0} \cdot \frac{RR_{DC1}P_{C0} + 1 - P_{C0}}{RR_{DC1}P_{C1} + 1 - P_{C1}}, \quad RR_0 = \frac{A_1/N_1}{A_0/N_0} \cdot \frac{RR_{DC0}P_{C0} + 1 - P_{C0}}{RR_{DC0}P_{C1} + 1 - P_{C1}}.$$

ただし、 $RR_1$  は交絡要因ありの層のリスク比であり、 $RR_0$  は交絡要因なしの層のリスク比である。

層別した分割表から、Mantel-Haenszel の方法をもちいて層を併合したリスク比を求めることができる。また、標準集団を適当に定め、交絡要因の分布を標準集団の分布にそろえた場合のリスク比、すなわち標準化リスク比を求めることもできる。

### 参考 3- 3 交絡の補正：オッズ比（コホート研究の場合） [ 本文：3.1.4、3.7.2 ]

アウトカム	合計		交絡要因			
			$C = 1$		$C = 0$	
	曝露		曝露		曝露	
	$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
$D = 1$	$A_1$	$A_0$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
$D = 0$	$B_1$	$B_0$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
合計	$N_1$	$N_0$	$N_{11}$	$N_{01}$	$N_{10}$	$N_{00}$

コホート研究におけるオッズ比については、以下のようにパラメータを定めることが考えられる。

$OR_{DC}$ ：交絡要因とアウトカムの関連性をしめすオッズ比（交絡要因あり/ 交絡要因なし）。ただし、このオッズ比は曝露の有無によらず等しいと仮定する。すなわち

$$OR_{DC} = \frac{A_{11}/B_{11}}{A_{10}/B_{10}} = \frac{A_{01}/B_{01}}{A_{00}/B_{00}}, \quad \text{よって}$$

$$OR_{DC} = \frac{A_{11}/B_{11}}{(A_1 - A_{11})/(B_1 - B_{11})} = \frac{A_{01}/B_{01}}{(A_0 - A_{01})/(B_0 - B_{01})}.$$

$P_{C1}$ ： $E = 1$  の場合の  $C = 1$  の割合、すなわち  $P_{C1} = N_{11}/N_1 = (A_{11} + B_{11})/N_1$ 。

$P_{C0}$ ： $E = 0$  の場合の  $C = 1$  の割合、すなわち  $P_{C0} = N_{01}/N_0 = (A_{01} + B_{01})/N_0$ 。

このようにパラメータを設定すると、例えば  $A_{11}$  に関する 2 次方程式が導かれ、層ごとの度数やオッズ比はかなり複雑な式になる。ただし、アウトカムが稀な場合は合計欄の度数 ( $N$ ) とアウトカムなしの場合の度数 ( $B$ ) が近似的に等しいため、以下がなりたつ。

$$P_{C1} = N_{11}/N_1 \cong B_{11}/B_1, \quad \text{及び} \quad P_{C0} = N_{01}/N_0 \cong B_{01}/B_0.$$

この場合はリスク比とよく似た関係が得られる。すなわち、交絡要因ありの層の度数は

$$A_{11} \cong \frac{OR_{DC}P_{C1}A_1}{OR_{DC}P_{C1} + 1 - P_{C1}}, \quad A_{01} \cong \frac{OR_{DC}P_{C0}A_0}{OR_{DC}P_{C0} + 1 - P_{C0}}, \quad B_{11} \cong P_{C1}B_1, \quad B_{01} \cong P_{C0}B_0.$$

したがって交絡の影響を補正したオッズ比は

$$OR_{DE} = \frac{A_{11}B_{01}}{A_{01}B_{11}} \cong \frac{A_1B_0}{A_0B_1} \cdot \frac{OR_{DC}P_{C0} + 1 - P_{C0}}{OR_{DC}P_{C1} + 1 - P_{C1}}.$$

また、元のデータから計算した未補正のオッズ比  $OR^*$  との関係は

$$OR_{DE} \cong OR^* \frac{OR_{DC}P_{C0} + 1 - P_{C0}}{OR_{DC}P_{C1} + 1 - P_{C1}} = OR^* / \frac{OR_{DC}P_{C1} + 1 - P_{C1}}{OR_{DC}P_{C0} + 1 - P_{C0}}.$$

以上では、交絡要因ありの層のオッズ比を求めたが、交絡要因なしの層で同様な計算を行ってもおなじオッズ比が得られる。その原因は  $OR_{DC}$  の値が曝露群によらず等しいと仮定したことにある。

交絡要因の影響が曝露の有無により異なることが想定される場合は、曝露の有無により異なるオッズ比、例えば曝露ありの場合は  $OR_{DC1}$ 、曝露なしの場合は  $OR_{DC0}$  と定める。すると上式の該当箇所が層ごとに異なる値を取り、層ごとのオッズ比は交絡要因の有無別に以下ようになる。

$$OR_{DE1} = \frac{A_1 B_0}{A_0 B_1} \cdot \frac{OR_{DC1} P_{C0} + 1 - P_{C0}}{OR_{DC1} P_{C1} + 1 - P_{C1}}, \quad OR_{DE0} = \frac{A_1 B_0}{A_0 B_1} \cdot \frac{OR_{DC0} P_{C0} + 1 - P_{C0}}{OR_{DC0} P_{C1} + 1 - P_{C1}}.$$

この場合も、層別した分割表から、Mantel-Haenszel の方法をもちいて層を併合したオッズ比を求めることができる。また、標準集団を適当に定め、交絡要因の分布を標準集団の分布にそろえた場合のオッズ比、すなわち標準化オッズ比を求めることもできる。

なお、コホート研究における以上の関係は、アウトカムが稀な場合になりたつ近似式であった。しかし、ケースコントロール研究では、コントロール選択が適切なら、アウトカムが稀でない場合にもこれらの関係がなりたつ (3.7.2)。またパラメータの設定をかせ

$P_{C1}$  :  $E = 1$  かつ  $D = 0$  の場合の  $C = 1$  の割合、すなわち  $P_{C1} = B_{11}/B_1$

$P_{C0}$  :  $E = 0$  かつ  $D = 0$  の場合の  $C = 1$  の割合、すなわち  $P_{C0} = B_{01}/B_0$

などとすれば、やはり同じ関係が導かれる。コホート研究なのにアウトカムなし ( $D = 0$ ) の場合の割合を想定するのは不自然に思えるかもしれないが、パラメータの意味に注意しながら運用すればよいだろう。

## 参考 3- 4 交絡の補正：回帰モデルによる感度分析、E-value と感度分析

### [ 本文 : 2.1.3、3.1.6 ]

以下では、既知の交絡要因の影響を調整した効果指標がすでに得られている場合について考えよう。つまり交絡調整がある程度はできているはずだが、それでも未知・未測定の影響によるバイアスが懸念され、その影響を評価したい。こうした状況でどんな感度分析が可能か、2つの方法を紹介する。

#### 回帰モデルによる感度分析

以下では、おもに Linx ら (1998) <sup>21</sup> を参照して、回帰モデルをもちいた感度分析について紹介する。

曝露 ( $E$ ) の有無、データが得られている交絡要因 ( $Z$ )、及び未知・未測定の影響要因 ( $C$ ) の有無が、アウトカムの発生 ( $D = 1$ ) に及ぼす影響を、対数線形モデル

$$\ln[\Pr(D = 1|E, C, Z)] = \alpha + \beta E + \gamma C + \theta' Z$$

をもちいて表すことができたとして、このモデルには、データが得られていない未知・未測定の影響要因も含まれている。しかし、それが既知だとしたら知ることができる真のリスク比  $RR$  を

$$RR = \Pr(D = 1|E = 1, C, Z) / \Pr(D = 1|E = 0, C, Z) = \exp(\beta)$$

という関係を通じて、つまり回帰係数  $\beta$  によって表している。また回帰係数  $\gamma$  は、未知・未測定の影響要因によって何倍アウトカムが発生するかという関連性、すなわち  $RR_{DC}$  を

$$RR_{DC} = \Pr(D = 1|E, C = 1, Z) / \Pr(D = 1|E, C = 0, Z) = \exp(\gamma)$$

という関係を通じて表している。なお、このモデルには交互作用項が含まれていないため、曝露の有無によらず未知・未測定の影響要因の影響が等しいことを想定している。

さて、実際には交絡要因  $C$  のデータは得られないので、この項を除いた回帰モデル

$$\ln[\Pr(D = 1|E, Z)] = \alpha^* + \beta^* E + \theta^* Z$$

をもちいて、バイアスを含むリスク比  $RR^* = \exp(\beta^*)$  が推定される。

このとき、真の回帰係数  $\beta$  とバイアスを含む回帰係数  $\beta^*$  の間に以下の関係がなりたつ<sup>21</sup>。

$$\beta = \beta^* - \ln \frac{\exp(\gamma)P_{C1} + 1 - P_{C1}}{\exp(\gamma)P_{C0} + 1 - P_{C0}}$$

ただし、これまでと同様に  $P_{C1}$  は曝露あり ( $E = 1$ ) の群に含まれる交絡要因あり ( $C = 1$ ) の対象者の割合、また  $P_{C0}$  は曝露あり ( $E = 0$ ) の群に含まれる交絡要因あり ( $C = 1$ ) の対象者の割合である。こうして、真のリスク比  $RR = \exp(\beta)$  と未補正のリスク比  $RR^* = \exp(\beta^*)$  の間に、以下がなりたつ。

$$RR = RR^* / \frac{RR_{DC} P_{C1} + 1 - P_{C1}}{RR_{DC} P_{C0} + 1 - P_{C0}}$$

これは、本文中でしめした式 3-2 (3.1.4、及び 3.1.6) と一致している。

以上は対数線形モデルについて成立する関係だが、アウトカムが稀な場合は、ロジスティックモデルや比例ハザードモデルについても同様な関係が近似的になりたつことが知られている<sup>21</sup>。

### E-value と感度分析

以下では、おもに Ding ら (2016)<sup>22</sup> と VanderWeele ら (2017)<sup>23</sup> を参照し、層別解析や回帰モデル等により、既知の交絡要因について調整された推定値に対して感度分析を行う方法を紹介する。この方法は、本文中でしめした式 3-2 (3.1.4、及び 3.1.6) を少し変形して得られる以下の関係をさらに一般化した内容となっている。

$$RR = RR^* / \frac{RR_{DC} P_{C1} + 1 - P_{C1}}{RR_{DC} P_{C0} + 1 - P_{C0}}$$

この式をさらに変形すると、以下が得られる。

$$RR = RR^* / \frac{(RR_{DC} \cdot \frac{P_{C1}}{P_{C0}}) + \frac{1 - P_{C1}}{P_{C0}}}{(RR_{DC} + \frac{P_{C1}}{P_{C0}} - 1) + \frac{1 - P_{C1}}{P_{C0}}}$$

ここで、 $RR_{DC} \geq 1$  かつ  $P_{C1}/P_{C0} \geq 1$  と仮定しても一般性を失わない。まず、 $RR_{DC} \leq 1$  のときは、リスク比  $RR_{DC}$  の定義を変更し、比のとりかたを「交絡要因あり/交絡要因なし」から「交絡要因なし/交絡要因あり」にかえれば  $RR_{DC} \geq 1$  となる。また、 $P_{C1}/P_{C0} \leq 1$  のときは、交絡要因「あり」の割合ではなく「なし」の割合をあらためて  $P_{C1}$ 、 $P_{C0}$  等と定義すれば  $P_{C1}/P_{C0} \geq 1$  となる。これらの条件を仮定すると、以下の不等式が成り立つ ( $b \geq a \geq 0, c \geq 0$  のとき  $(a+c)/(b+c) \geq a/b$  が成り立つこと等から)。

$$RR \geq RR^* / \frac{RR_{DC} \cdot \frac{P_{C1}}{P_{C0}}}{RR_{DC} + \frac{P_{C1}}{P_{C0}} - 1}$$

さらに交絡要因に関する割合の比を  $RR_{CE} = P_{C1}/P_{C0}$  とリスク比のように表すと、この不等式は

$$RR \geq RR^* / \frac{RR_{DC} \cdot RR_{CE}}{RR_{DC} + RR_{CE} - 1}$$

と、さらに簡単な式になる。これは (真の値)  $\geq$  (推定値)/(系統誤差) という形をしており、バイアスの程度を不等式で表したものとなっている。

以上では、 $RR^*$  は未調整のリスク比であったが、既知の交絡要因  $C$  についてはデータが得られており、層ごとの (あるいは値  $C = c$  ごとの) 推定値が得られていたとする。このとき、関連する定義を以下のように変更することで、さらに広範な応用が可能になる<sup>22</sup>。

- $RR$  : 既知の交絡要因の値が  $C = c$  のときの、真のリスク比
- $RR^*$  : 既知の交絡要因の値が  $C = c$  のときの、観察されたリスク比
- $RR_{UE}$  : 既知の交絡要因の値が  $C = c$  のときに、未知の交絡要因の値が  $U = k$  となる割合が曝露の有無別にどれほど異なるかという比の最大値  $\max_k \{P(U = k|E = 1, C = c)/P(U = k|E = 0, C = c)\}$ 。すなわち、未知の交絡要因の分布が曝露の有無により最大で何倍 ( $\geq 1$ ) 異なるか。
- $RR_{DU}$  : 既知の交絡要因が  $C = c$  のとき、曝露なしの場合に未知の交絡要因によりアウトカムの発生割合がどれほど異なるかという比  $\max_k P(D = 1|E = 0, C = c, U = k)/\min_k P(D = 1|E = 0, C = c, U = k)$

と、曝露ありの場合の同様な比の、いずれか大きいほうの値。すなわち、未知の交絡要因により、アウトカムの発生割合が最大で何倍 ( $\geq 1$ ) 異なるか。

このように定めると、以下が成り立つ。

$$RR \geq RR^* / \frac{RR_{UE} \cdot RR_{DU}}{RR_{UE} + RR_{DU} - 1}$$

この不等式は、厳密には既知の交絡要因の値 ( $C = c$ ) ごとに成立するものだが、これらの値をつうじての、つまり集団全体での関係にも一般化することができる<sup>22</sup>。

上式で  $x = RR_{UE} = RR_{DU}$  とおくと  $x$  に関する 2 次不等式が得られるが、これを  $RR^* > RR$  という条件下で、すなわち真のリスク比より大きなリスク比が観察された場合を想定して解くと、以下が得られる。

$$RR_{UE} = RR_{DU} \geq \frac{RR^* + \sqrt{RR^*(RR^* - RR)}}{RR}$$

最後の項で  $RR = 1$  とした場合の値、すなわち  $RR^* + \sqrt{RR^*(RR^* - 1)}$  が E-value である。例えば、ある研究で観察されたリスク比が  $RR^* = 3.9$  であったとすると、E-value は  $3.9 + \sqrt{3.9(3.9 - 1)} = 7.26$  となる。すると、以上の不等式から  $RR_{UE} = RR_{DU} \geq 7.26$  となるから、真のリスク比が  $RR = 1$  のときに  $RR^* = 3.9$  というリスク比が観察されるのは、アウトカムの発生割合を 7.26 倍以上増加させる要因が、一方の群に 7.26 倍以上多い場合であることがわかる。すでに複数の要因について調整したうえで、さらにこうした要因の存在が懸念されるようなら、真のリスク比が 1 である可能性も否定できない。しかし、そんな要因が存在するとは考えにくい、ということなら真のリスク比が 1 である可能性は稀だと判断できる。

以上の不等式は  $RR = 1$  の場合に限らずもちいることができる。例えば、観察されたリスク比 ( $RR^*$ ) は 1.2 だが、真のリスク比 ( $RR$ ) は 0.8 程度ではないか、という想定があったとする。このとき、上式の右辺は  $\{1.2 + \sqrt{1.2(1.2 - 0.8)}\} / 0.8 = 2.37$  となる。したがって、こうしたバイアスが生じるのは、アウトカムの発生割合を 2.37 倍以上増加させる要因が、一方の群に 2.37 倍以上多く含まれている場合である。

さらに、はじめに示した不等式

$$RR \geq RR^* / \frac{RR_{UE} \cdot RR_{DU}}{RR_{UE} + RR_{DU} - 1}$$

から、右辺の分母の値がおなじなら、真のリスク比と観察されたリスク比との関係が変わらないことがわかる。例えば、この例では、 $RR_{UE} = RR_{DU} = 2.37$  の場合に、真のリスク比が 1.5 倍になって観察された ( $RR^* / RR = 1.2 / 0.8 = 1.5$ )。したがって

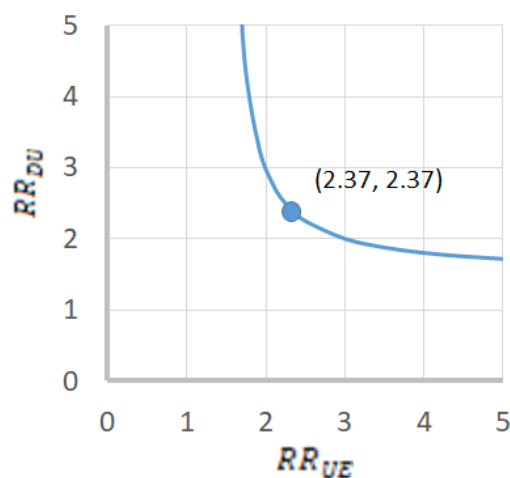
$$\frac{RR_{UE} \cdot RR_{DU}}{RR_{UE} + RR_{DU} - 1} = 1.5$$

を満たす値の組 ( $RR_{UE}, RR_{DU}$ ) は、すべておなじ大きさのバイアスをもたらす。したがって、

( $RR_{UE}, RR_{DU}$ ) = (2.37, 2.37) の場合に限らず、例えば

(2.1, 2.75)、(4.0, 1.8) 等、上式をみたす値の組み合わせを求め、それらの値がどれほど現実的か、あるいは妥当かを検討することができる。参考までに、この場合の値の組をプロットした (上図参照)。

なお、この方法を信頼区間の上限や下限に対してもちいることができる。また、リスク比以外の効果指標にも応用することができる<sup>22,23</sup>。



### 参考 3- 5 交絡の補正：別のパラメータをもちいる場合 [ 本文：3.1.6 ]

交絡については、パラメータの設定をかえることで、様々な補正式が得られることが知られている<sup>16</sup>。例えばリスク比について、以下の 3 つのパラメータをもちいた補正が考えられる。

- ・  $RR_{DC}$  : 交絡要因  $C$  とアウトカム  $D$  の関連性をしめすリスク比。

- ・  $O_{C1}$  : 曝露ありのとき、交絡要因ありの対象者が、なしの対象者の何倍多いかという比（オッズ）、すなわち  $O_{C1} = P_{C1}/(1 - P_{C1})$ .
- ・  $O_{C0}$  : 曝露なしのとき、交絡要因ありの対象者が、なしの対象者の何倍多いかという比（オッズ）、すなわち  $O_{C0} = P_{C0}/(1 - P_{C0})$ .

以上の定義から、補正後のリスク比  $RR$  と、未補正のリスク比  $RR^*$  との関係（式 3-2 : 3.1.4、及び 3.1.6）を、以下のように変形することができる。

$$RR = RR^* \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}P_{C1} + 1 - P_{C1}} = RR^* \frac{(RR_{DC}O_{C0} + 1)(O_{C1} + 1)}{(RR_{DC}O_{C1} + 1)(O_{C0} + 1)}.$$

したがって、以上 3 つのパラメータをもちいても交絡の影響を補正することができる。

また、以下の 3 つのパラメータをもちいた補正も考えられる。

- ・  $RR_{DC}$  : 交絡要因  $C$  アウトカム  $D$  の関連性をしめすリスク比。
- ・  $P_{C0}$  : 曝露なしのとき、交絡要因ありの対象者が含まれる割合。
- ・  $RR_{CE}$  : 交絡要因ありの対象者が含まれる割合の比（曝露あり/曝露なし）、すなわち  $RR_{CE} = P_{C1}/P_{C0}$ 。

以上の定義から  $P_{C1} = RR_{CE}P_{C0}$  となるので、これを式 3-2（3.1.4、及び 3.1.6）に代入すると

$$RR = RR^* \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}P_{C1} + 1 - P_{C1}} = RR^* \frac{RR_{DC}P_{C0} + 1 - P_{C0}}{RR_{DC}RR_{CE}P_{C0} + 1 - RR_{CE}P_{C0}}.$$

したがって、以上 3 つのパラメータをもちいても交絡の影響を補正することができる。

なお、どのモデルをもちいるかは、バイアスに対する解析者の想定や、計算の手間等にも依存する。例えば、確率的感度分析では複数のパラメータに相関をもたせたい場合がある。本文中でしめたように、交絡要因ありの対象者が占める割合、つまり曝露なしのとき  $P_{C0}$ 、ありのとき  $P_{C1}$  という 2 つのパラメータについて、一方の値が大きければ他方も大きいのではないかと解析者が考えていたとする。つまり、これら 2 つのパラメータは独立でなく、なんらかの相関があるのではないかと、という想定である。こうした想定をシミュレーションによって表現する際、これらの値を、相関をもつ 2 つの乱数をもちいて生成することができる。例えば互いに独立な乱数を 3 つ生成し、それらを 2 つずつ合成した 2 つの乱数を構成することで、2 つの乱数に相関をもたせることができる<sup>8, 16</sup>。また  $P_{C0}$  と  $P_{C1}$  ではなく、 $P_{C0}$  と  $RR_{CE}$  をパラメータとして定め、これらの値を 2 つの独立な乱数をもちいて生成してもよい。というのは、 $P_{C0}$  と  $RR_{CE}$  を独立に定めたとしても、上述した  $P_{C1} = RR_{CE}P_{C0}$  という関係により  $P_{C1}$  と  $P_{C0}$  に相関が生じるからである。手間だけを考えれば、独立な乱数を 2 つ生成させる後者の計算のほうが簡単かもしれない。また前者のほうが、解析者が抱いている想定を表現しやすい場合もあるかもしれない。本文中でも指摘したように、感度分析の方法には自由度があるので、解析者の意図や目的にふさわしい方法を選ぶとよい（3.1.6）。

### 参考 3- 6 差異的で独立な誤分類、非差異的で従属な誤分類 [ 本文 : 3.2.1 ]

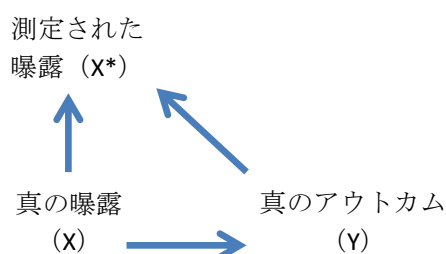
本書では誤分類の性質をしめす differential を「差異的」と訳したが、ほかにも様々な訳語がもちいられている。例えば、差別的、特異的、差がない、差異がない、偏りがない、等々である。また differential misclassification を特異誤分類、差異誤分類等と「differential」を名詞のように訳している例もある。これらはいずれも、ある要因の測定値の正確さが、べつの要因の真の値によって異なる、差がある、偏っている、といった状況を表している。例えば、入門編の付録 1 でしめたように、主治医が「薬剤  $X$  への曝露により有害事象  $Y$  というアウトカムが生じやすいはずだ」という先入観をもっている状況では、薬剤への曝露（ $X$  の真の値）によって、観察されるアウトカム（ $Y$  の測定値  $Y^*$ ）の正確さが異なる可能性がある。また、認知症患者に対して聞き取り調査を行ったところ曝露について誤分類が生じた、という状況では、認知症というアウトカムの有無（ $Y$  の真の値）により、この調査から得られた曝露の有無（ $X$  の測定値  $X^*$ ）の正確さが異なっている可能性がある。どちらも、一方の測定値の正確さが他方の真の値により異なる場合であり、差異的な誤分類の例となっていた。



ところで誤分類には、差異性の有無と従属性の有無により4つ場合が考えられる。以下では差異的で独立な誤分類と、非差異的で従属な誤分類の例をしめした。ただし説明のため、やや極端な事例を取り上げている。また要因相互の関係を図でしめす際「測定誤差」にあたる要因を明示せず、説明を簡略化している。

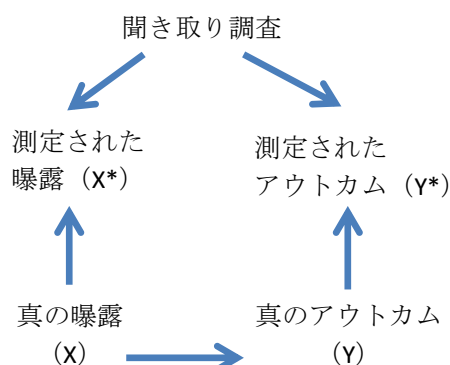
### 差異的で独立な誤分類

例えば、アウトカムが記憶障害の有無で、ある薬剤使用の有無を聞き取り調査によって調べる研究を行ったとしよう。この場合、記憶障害ありの人は、薬剤使用についても記憶が曖昧であり、誤った回答をする可能性が高いと考えられる。すると、アウトカムの値（記憶障害の有無）という、曝露とは別な要因（の真の値）によって、曝露の値（薬剤使用の有無）の正確さが異なることになる。したがって、これは差異的な誤分類にあたる。ただし、アウトカム（記憶障害）については、医療記録等から正確に調べることができたとしよう。また、曝露状況（薬剤使用）について、他の要因（例えば交絡要因として解析にもちいる要因）の情報を参照することなく聞き取りが行われていたとすれば、曝露（薬剤使用）に関する誤分類は他の要因と独立だと考えてよいだろう。以上から、この例での曝露（薬剤使用）の誤分類は、差異的だが独立だと考えられる。以上の状況を参考までに図でしめした。



### 非差異的で従属な誤分類

今度は、曝露（薬剤使用）だけでなく、アウトカムも聞き取り調査によって調べる場合を考えよう。聞き取り調査である以上、曝露にもアウトカムにも誤分類が生じることは避けがたい。また、おなじ回答者から聞き取りを行うと、アウトカムのことをよく記憶している人ほど曝露についても正確に記憶している、といった形で両者の誤分類に関連があるかもしれない。つまりこの例では、曝露とアウトカムの誤分類は、互いに従属であると考えられる。ただし、先ほどの記憶障害の例とは異なり、この研究で関心がある曝露やアウトカムが、回答者の記憶の正確さに直接的な影響を及ぼすものでなければ、一方の値によって他方の値の正確さが異なるとは考えにくい。つまり、曝露やアウトカムの誤分類は非差異的であろう。まとめると、この例では、曝露にもアウトカムにも誤分類が疑われ、いずれも非差異的だが従属である（独立でない）と考えられる。以上の状況を参考までに図でしめした。



参考 3- 7 曝露の誤分類の補正：差別的誤分類 [ 本文：3.3.2、3.3.4、3.7.3 ]

アウトカム	誤分類			
	なし		あり (元のデータ)	
	曝露		曝露	
	$E = 1$	$E = 0$	$E^* = 1$	$E^* = 0$
$D = 1$	$A_1$	$A_0$	$A_1^*$	$A_0^*$
$D = 0$	$B_1$	$B_0$	$B_1^*$	$B_0^*$

上表のように度数を表し、以下のパラメータを定める．

$Se_A$ ：アウトカムありの場合の、曝露分類の感度

$Sp_A$ ：アウトカムありの場合の、曝露分類の特異度

$Se_B$ ：アウトカムなしの場合の、曝露分類の感度

$Sp_B$ ：アウトカムなしの場合の、曝露分類の特異度

本文で述べた非差別的誤分類の場合と同様に、誤分類がない（真の）分割表のデータと、誤分類がある（元の）分割表のデータとの間には以下の関係がある（3.3.2）．

$$A_1^* = Se_A A_1 + (1 - Sp_A)A_0, A_0^* = (1 - Se_A) A_1 + Sp_A A_0,$$

$$B_1^* = Se_B B_1 + (1 - Sp_B)B_0, B_0^* = (1 - Se_B) B_1 + Sp_B B_0.$$

これらを解くと、誤分類がない分割表のデータを以下のように求めることができる．

$$A_1 = \frac{Sp_A A_1^* - (1 - Sp_A)A_0^*}{Se_A + Sp_A - 1}, B_1 = \frac{Sp_B B_1^* - (1 - Sp_B)B_0^*}{Se_B + Sp_B - 1},$$

$$A_0 = \frac{Se_A A_0^* - (1 - Se_A)A_1^*}{Se_A + Sp_A - 1}, B_0 = \frac{Se_B B_0^* - (1 - Se_B)B_1^*}{Se_B + Sp_B - 1}.$$

さらに、これらを合計することで曝露の有無別の合計人数  $N_1$ 、及び  $N_0$  も求めることができる．

また、陽性的中率と陰性的中率を以下のように定めて補正式を求めることもできる．

$PVP_A$ ：アウトカムありの場合の、曝露分類に関する陽性的中率

$PVN_A$ ：アウトカムありの場合の、曝露分類に関する陰性的中率

$PVP_B$ ：アウトカムなしの場合の、曝露分類に関する陽性的中率

$PVN_B$ ：アウトカムなしの場合の、曝露分類に関する陰性的中率

この場合、誤分類がある分割表のデータ ( $A_1^*$ 、 $A_0^*$ 、 $B_1^*$ 、 $B_0^*$ ) から考えると、以上の的中率のもとでは、誤分類のない分割表の（真の）度数 ( $A_1$ 、 $A_0$ 、 $B_1$ 、 $B_0$ ) は以下のようなる．

$$A_1 = PVP_A A_1^* + (1 - PVN_A)A_0^*, A_0 = PVN_A A_0^* + (1 - PVP_A)A_1^*$$

$$B_1 = PVP_B B_1^* + (1 - PVN_B)B_0^*, B_0 = PVN_B B_0^* + (1 - PVP_B)B_1^*$$

これらからも、曝露の差別的誤分類の影響を補正した効果指標を求めることができる．

参考 3- 8 非差別的で独立な曝露の誤分類 [ 本文：3.2.2、3.3.3、3.3.5 ]

曝露もアウトカムも 2 値の場合、非差別的で独立な誤分類により効果指標の推定値に、関連性を弱める方向のバイアス (bias toward the null) が生じる (3.2.2)．以下では、曝露の誤分類を例にとるが、同様なこ

とがアウトカムの誤分類についてもなりたつ。また、本文中で述べたように、この性質は曝露の水準が 3 以上の場合には成り立たない。

さて、式 3-4 (3.3.2) を導く際、他の要因やアウトカムの有無によらず曝露の誤分類の程度が等しいこと、つまり非差異的で独立な誤分類を想定した。本文中でしめした関係 (3.3.2) をもちいて、未補正のリスク差  $RD^*$  を表すと以下が得られる。

$$RD^* = \frac{A_1^*}{N_1^*} - \frac{A_0^*}{N_0^*} = \frac{A_1^*}{A_1^* + B_1^*} - \frac{A_0^*}{A_0^* + B_0^*} = \frac{Se A_1 + (1 - Sp) A_0}{Se N_1 + (1 - Sp) N_0} - \frac{(1 - Se) A_1 + Sp A_0}{(1 - Se) N_1 + Sp N_0}$$

$$= \frac{(Se + Sp - 1) N_1 N_0}{\{Se N_1 + (1 - Sp) N_0\} \{(1 - Se) N_1 + Sp N_0\}} \left( \frac{A_1}{N_1} - \frac{A_0}{N_0} \right).$$

最後の式の最終項は補正後のリスク差、もしくは誤分類を受ける前の真のリスク差  $RD$  なので、この関係は以下のように表せる。

$$RD^* = K \cdot RD, \text{ ただし } K = \frac{(Se + Sp - 1) N_1 N_0}{\{Se N_1 + (1 - Sp) N_0\} \{(1 - Se) N_1 + Sp N_0\}}$$

なお、 $K$  の分母が正であり、かつ分子より大きいことを容易に示すことができるので  $K < 1$  である ( $Se = Sp = 1$  のときに等号がなりたつが、ここでは  $0 < Se < 1$  かつ  $0 < Sp < 1$  とする)。また、 $K$  の符号が  $Se + Sp - 1$  の正負 (感度+特異度と 1 との大小) に依存することもわかる。さらにこの式から、真のリスク差  $RD$  が 0 なら、誤分類後のリスク差  $RD^*$  も 0 になることがわかる。つまり、非差異的で独立な誤分類により、帰無仮説の下でのバイアス (bias under the null) が生じないことが確認できる。

さて、真のリスク差  $RD$  が 0 でない場合、すなわち対立仮説が正しい場合のバイアス (bias under the alternative) については以下が導かれる。

#### 1) 感度 + 特異度 > 1 の場合

$0 < K < 1$  となる。このため、真のリスク差  $RD$  が 0 より大きければ  $0 < RD^* < RD$  がなりたつ。また真のリスク差が 0 より小さければ  $RD < RD^* < 0$  がなりたつ。いずれの場合も、誤分類後のリスク差  $RD^*$  は、真のリスク差  $RD$  より 0 に近い値をとる。言い換えれば、誤分類後のリスク差  $RD^*$  は、真のリスク差  $RD$  よりも「曝露による影響がない」という方向に偏っている。この場合が「非差異的な誤分類により、差がないという方向に偏る (帰無仮説の方向に偏る)」等と言われている内容である。

#### 2) 感度 + 特異度 = 1 の場合

$K = 0$  となり  $RD^* = 0$  となる。この場合、真のリスク差  $RD$  とは無関係に、誤分類後のリスク差  $RD^*$  は常に 0 となる。このようなことが起こるのは、感度か特異度が 0.5 以下となる場合なので、誤分類の程度がかなり深刻な場合である。まるで、コインやサイコロで、ランダムに分類を決めたかのようにであり、通常こうしたことは起こらないと考えられる。

#### 3) 感度 + 特異度 < 1 の場合

$K < 0$  となり、関連性の向きが逆転する。つまり、 $RD > 0$  なら  $RD^* < 0$  となり、 $RD < 0$  なら  $RD^* > 0$  となる。このようなことが起こるのは、2) で述べた「ランダムに分類を決めたかのような」場合より、さらに誤分類の程度が深刻な場合である。通常、ここまで深刻な誤分類を想定することはなく、結果的に、1) の場合 (感度と特異度の和が 1 を超える場合) を想定することが多い。

以上から、曝露もアウトカムも 2 値で、かつそれらの誤分類が非差異的で独立な場合は、誤分類後のリスク差が、真のリスク差より 0 に近い値を取る (差がないという方向に偏っている) と考えることができる。

なお、同様のことがリスク比やオッズ比についてもなりたつ。すなわち、感度 + 特異度 > 1 という条件下では、誤分類後のリスク比やオッズ比は真の値より 1 に近い値を取る。

参考 3-9 アウトカムの誤分類の補正：差別的誤分類 [ 本文：3.4.2、3.7.4 ]

誤分類	アウトカム	曝露	
		$E = 1$	$E = 0$
あり (元のデータ)	$D^* = 1$	$A_1^*$	$A_0^*$
	$D^* = 0$	$B_1^*$	$B_0^*$
なし	$D = 1$	$A_1$	$A_0$
	$D = 0$	$B_1$	$B_0$

上表のように度数を表し、以下のパラメータを定める．

$Se_1$ ：曝露ありの場合の、アウトカム分類の感度

$Sp_1$ ：曝露ありの場合の、アウトカム分類の特異度

$Se_0$ ：曝露なしの場合の、アウトカム分類の感度

$Sp_0$ ：曝露なしの場合の、アウトカム分類の特異度

本文でも述べたように、誤分類がない（真の）分割表のデータと、誤分類がある（元の）分割表のデータとの間には以下の関係がある（3.4.2）．

$$A_1^* = Se_1 A_1 + (1 - Sp_1) B_1, \quad A_0^* = Se_0 A_0 + (1 - Sp_0) B_0,$$

$$B_1^* = (1 - Se_1) A_1 + Sp_1 B_1, \quad B_0^* = (1 - Se_0) A_0 + Sp_0 B_0.$$

これらを  $A_1 + B_1 = A_1^* + B_1^* = N_1$ 、及び  $A_0 + B_0 = A_0^* + B_0^* = N_0$  に留意して解くと、誤分類がない場合の度数を以下のように求めることができる

$$A_1 = \frac{A_1^* - (1 - Sp_1) N_1^*}{Se_1 + Sp_1 - 1}, \quad B_1 = \frac{B_1^* - (1 - Se_1) N_1^*}{Se_1 + Sp_1 - 1}, \quad N_1 = N_1^*,$$

$$A_0 = \frac{A_0^* - (1 - Sp_0) N_0^*}{Se_0 + Sp_0 - 1}, \quad B_0 = \frac{B_0^* - (1 - Se_0) N_0^*}{Se_0 + Sp_0 - 1}, \quad N_0 = N_0^*.$$

また、陽性的中率と陰性的中率を以下のように定めて補正式を求めることもできる．

$PVP_1$ ：曝露ありの場合の、アウトカム分類の陽性的中率

$PVN_1$ ：曝露ありの場合の、アウトカム分類の陰性的中率

$PVP_0$ ：曝露なしの場合の、アウトカム分類の陽性的中率

$PVN_0$ ：曝露なしの場合の、アウトカム分類の陰性的中率

この場合、誤分類がある分割表のデータ（ $A_1^*$ 、 $A_0^*$ 、 $B_1^*$ 、 $B_0^*$ ）から考えると、以上の的中率のもとでは、誤分類のない分割表の（真の）度数（ $A_1$ 、 $A_0$ 、 $B_1$ 、 $B_0$ ）は以下ようになる．

$$A_1 = PVP_1 A_1^* + (1 - PVN_1) B_1^*, \quad A_0 = PVP_0 A_0^* + (1 - PVN_0) B_0^*,$$

$$B_1 = (1 - PVP_1) A_1^* + PVN_1 B_1^*, \quad B_0 = (1 - PVP_0) A_0^* + PVN_0 B_0^*.$$

これらからも、アウトカムの誤分類の影響を補正した効果指標を求めることができる．

参考 3- 10 交絡要因の誤分類の補正：差別的誤分類 [ 本文：3.5.2、3.5.3 ]

誤分類	アウトカム	合計		交絡要因			
				$C = 1$		$C = 0$	
		曝露		曝露		曝露	
		$E = 1$	$E = 0$	$E = 1$	$E = 0$	$E = 1$	$E = 0$
なし	$D = 1$	$A_1$	$A_0$	$A_{11}$	$A_{01}$	$A_{10}$	$A_{00}$
	$D = 0$	$B_1$	$B_0$	$B_{11}$	$B_{01}$	$B_{10}$	$B_{00}$
あり				$C^* = 1$		$C^* = 0$	
				曝露		曝露	
				$E = 1$	$E = 0$	$E = 1$	$E = 0$
	$D = 1$			$A_{11}^*$	$A_{01}^*$	$A_{10}^*$	$A_{00}^*$
	$D = 0$			$B_{11}^*$	$B_{01}^*$	$B_{10}^*$	$B_{00}^*$

上表のように度数を表し、以下のパラメータを定める。

$Se_A$  アウトカムありの場合の、交絡要因の分類に関する感度

$Sp_A$ ：アウトカムありの場合の、交絡要因の分類に関する特異度

$Se_B$ ：アウトカムなしの場合の、交絡要因の分類に関する感度

$Sp_B$ ：アウトカムなしの場合の、交絡要因の分類に関する特異度

このとき、誤分類がない分割表の（真の）度数から考えると、以上の感度と特異度のもとでは、誤分類がある分割表の（元の）度数は以下ようになる。まず、交絡要因ありの層（ $C^* = 1$ ）について

$$A_{11}^* = Se_A A_{11} + (1 - Sp_A) A_{10}, \quad A_{01}^* = Se_A A_{01} + (1 - Sp_A) A_{00}, \\ B_{11}^* = Se_B B_{11} + (1 - Sp_B) B_{10}, \quad B_{01}^* = Se_B B_{01} + (1 - Sp_B) B_{00}.$$

これらを解くと、誤分類がない交絡要因ありの層（ $C = 1$ ）の度数を以下のように求めることができる

$$A_{11} = \frac{A_{11}^* - (1 - Sp_A) A_{10}}{Se_A + Sp_A - 1}, \quad A_{01} = \frac{A_{01}^* - (1 - Sp_A) A_{00}}{Se_A + Sp_A - 1}, \\ B_{11} = \frac{B_{11}^* - (1 - Sp_B) B_{10}}{Se_B + Sp_B - 1}, \quad B_{01} = \frac{B_{01}^* - (1 - Sp_B) B_{00}}{Se_B + Sp_B - 1}.$$

同様に、交絡要因なしの層の度数を求めると、以下が得られる。

$$A_{10} = \frac{Se_A A_{11} - A_{11}^*}{Se_A + Sp_A - 1}, \quad A_{00} = \frac{Se_A A_{01} - A_{01}^*}{Se_A + Sp_A - 1}, \quad B_{10} = \frac{Se_B B_{11} - B_{11}^*}{Se_B + Sp_B - 1}, \quad B_{00} = \frac{Se_B B_{01} - B_{01}^*}{Se_B + Sp_B - 1}.$$

以上から交絡要因の層ごとの効果指標や、2つの層を Mantel-Haenszel の方法で併合した効果指標を求めることができる。また、曝露ありの集団、曝露なしの集団、あるいは全体集団の、いずれかを標準集団として標準化した効果指標を求めることもできる。

参考 3- 11 交絡要因に対する独立で非差別的誤分類の影響 [ 本文：3.5.5 ]

本文中でしめした数値例では、交絡要因について未調整のリスク比が 1.73、誤分類がある交絡要因で調整したリスク比が 1.37、また誤分類がない、真の交絡要因をもちいて調整したリスク比が 1.33 であり、以下の大小関係があった（3.5.3）。

$$\left( \begin{array}{c} \text{真の交絡要因} \\ \text{で調整したリスク比} \end{array} \right) < \left( \begin{array}{c} \text{誤分類ありの交絡要因} \\ \text{で調整したリスク比} \end{array} \right) < (\text{未調整のリスク比})$$

このことは、疫学研究では一般的といえる多くの状況下で、2 値の交絡要因についてなりたつ性質であることが知られている<sup>31</sup>。その条件とは、例えば交絡要因ありのときに発生割合が高い（あるいは低い）、といった関係が曝露の有無によって逆転しないこと、すなわち、交絡要因の有無と曝露の有無による質的な交互作用がないことである。この条件下で、2 値の交絡要因が他の要因とは独立に非差別的誤分類を受けている場合、以下のいずれかがなりたつ。

$$\left( \begin{array}{c} \text{真の交絡要因で} \\ \text{調整した効果指標} \end{array} \right) < \left( \begin{array}{c} \text{誤分類ありの交絡要因} \\ \text{で調整した効果指標} \end{array} \right) < (\text{未調整の効果指標}),$$

$$(\text{未調整の効果指標}) < \left( \begin{array}{c} \text{誤分類ありの交絡要因} \\ \text{で調整した効果指標} \end{array} \right) < \left( \begin{array}{c} \text{真の交絡要因で} \\ \text{調整した効果指標} \end{array} \right).$$

この関係から、誤分類ありの交絡要因で調整した効果指標の推定値の方が、未調整の推定値より真の値に近いことが期待できる。このため、仮に誤分類を受けていたとしても、交絡要因をもちいて調整を行う方が、まったく行わない（未調整）よりましであると考えられている。

なお、交絡要因に対する誤分類が差別的な場合、また従属な場合、このような性質はない。このため、誤分類の影響で効果指標の値がどの方向に偏るかを予測することは難しい。

### 参考 3- 12 選択バイアスの補正：観察されなかった層の分割表を求める方法

#### [ 本文 : 3.6.7 ]

アウトカム	対象者選択				合計	
	あり (C = 1)		なし (C = 0)			
	曝露		曝露		曝露	
	E = 1	E = 0	E = 1	E = 0	E = 1	E = 0
D = 1	A <sub>11</sub>	A <sub>01</sub>	A <sub>10</sub>	A <sub>00</sub>	A <sub>1</sub>	A <sub>0</sub>
D = 0	B <sub>11</sub>	B <sub>01</sub>	B <sub>10</sub>	B <sub>00</sub>	B <sub>1</sub>	B <sub>0</sub>
合計	N <sub>11</sub>	N <sub>01</sub>	N <sub>10</sub>	N <sub>00</sub>	N <sub>1</sub>	N <sub>0</sub>

上表のように度数を表し、以下のパラメータを定める。

- $O_{C0}$  : 対象者選択なし (C = 0) の集団での曝露オッズ。すなわち、この集団で曝露ありの対象者が曝露なしの対象者の何倍多いかという比  $N_{10}/N_{00}$ 。
- $IP_{C0}$  : 対象者選択なし (C = 0) の集団での、曝露なしの場合のアウトカムの発生割合  $A_{00}/N_{00}$ 。
- $RR_{DE0}$  : 対象者選択なし (C = 0) の集団での、曝露とアウトカムの関連性をしめすリスク比  $(A_{10}/N_{10})/(A_{00}/N_{00})$ 。

ここで、対象者選択なしの集団の合計人数  $N_{10} + N_{00}$  は、選択前の集団の人数  $N$  から、実際に研究に参加した対象者の人数  $N_{11} + N_{01} = N_{+1}$  を引いたものなので

$$N_{10} + N_{00} = N_{+0} = N - N_{+1}.$$

したがって 1 つめのパラメータの定義から

$$N_{10} = \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1}, \quad N_{00} = \frac{N - N_{+1}}{O_{C0} + 1}.$$

さらに、残りのパラメータの定義から



$$A_{10} = \frac{IP_{C0} \cdot RR_{DE0} \cdot O_{C0}(N - N_{+1})}{O_{C0} + 1}, A_{00} = \frac{IP_{C0}(N - N_{+1})}{O_{C0} + 1}.$$

以上は対象者選択なしの集団、つまり実際には観察されなかった集団における分割表のデータである。よって対象者選択ありの集団となしの集団の合計、つまり選択前の集団における分割表のデータは

$$A_1 = A_{11} + \frac{IP_{C0} \cdot RR_{DE0} \cdot O_{C0}(N - N_{+1})}{O_{C0} + 1}, A_0 = A_{01} + \frac{IP_{C0}(N - N_{+1})}{O_{C0} + 1},$$

$$N_1 = N_{11} + \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1}, N_0 = N_{01} + \frac{N - N_{+1}}{O_{C0} + 1}$$

等となる。したがって選択バイアスを補正したリスク比は

$$RR = \frac{\left\{ A_{11} + \frac{IP_{C0} \cdot RR_{DE0} \cdot O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\} / \left\{ N_{11} + \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\}}{\left\{ A_{01} + \frac{IP_{C0}(N - N_{+1})}{O_{C0} + 1} \right\} / \left\{ N_{01} + \frac{N - N_{+1}}{O_{C0} + 1} \right\}}.$$

### 参考 3- 13 対象者選択によるバイアスが生じない場合（十分条件） [ 本文 : 3.6.7 ]

すでにデータが得られている対象者選択あり（ $C = 1$ ）の集団における、曝露なしの場合のアウトカムの発生割合  $A_{01}/N_{01}$  を  $IP_{C1}$  で、またこの集団におけるリスク比  $(A_{11}/N_{11})/(A_{01}/N_{01})$  を  $RR^*$  で表そう。すると、参考 3- 12 で求めた選択バイアスを補正したリスク比は以下のように変形できる。

$$RR = \frac{\left\{ IP_{C1} \cdot RR^* \cdot N_{11} + IP_{C0} \cdot RR_{DE0} \cdot \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\} / \left\{ N_{11} + \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\}}{\left\{ IP_{C1} \cdot N_{01} + \frac{IP_{C0}(N - N_{+1})}{O_{C0} + 1} \right\} / \left\{ N_{01} + \frac{N - N_{+1}}{O_{C0} + 1} \right\}}$$

$$= RR^* \cdot \frac{\left\{ N_{11} + \frac{IP_{C0} \cdot RR_{DE0} \cdot O_{C0}(N - N_{+1})}{IP_{C1} \cdot RR^* \cdot O_{C0} + 1} \right\} / \left\{ N_{11} + \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\}}{\left\{ N_{01} + \frac{IP_{C0} \cdot N - N_{+1}}{IP_{C1} \cdot O_{C0} + 1} \right\} / \left\{ N_{01} + \frac{N - N_{+1}}{O_{C0} + 1} \right\}}$$

したがって、 $IP_{C0} = IP_{C1}$  かつ  $RR_{DE0} = RR^*$  がなりたつとき

$$RR = RR^* \cdot \frac{\left\{ N_{11} + 1 \cdot \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\} / \left\{ N_{11} + \frac{O_{C0}(N - N_{+1})}{O_{C0} + 1} \right\}}{\left\{ N_{01} + 1 \cdot \frac{N - N_{+1}}{O_{C0} + 1} \right\} / \left\{ N_{01} + \frac{N - N_{+1}}{O_{C0} + 1} \right\}} = RR^*,$$

すなわち、 $RR = RR^*$  がなりたつことがわかる。

なお、条件「 $IP_{C0} = IP_{C1}$  かつ  $RR_{DE0} = RR^*$ 」は十分条件であり、対象者選択の有無を層としたとき、曝露ありの場合も曝露なしの場合も、アウトカムの発生割合が層によらず等しいことを意味している。

### 参考 3- 14 別のパラメータによる選択バイアスの補正 [ 本文 : 3.6.4、3.6.6、3.6.9 ]

選択確率をもちいる場合、以下のように定めることもできる。

- ・  $S_{A_1}$  : 曝露あり（ $E = 1$ ）、アウトカムあり（ $D = 1$ ）の対象者が選択される確率
- ・  $S_{A_0}$  : 曝露なし（ $E = 0$ ）、アウトカムあり（ $D = 1$ ）の対象者が選択される確率
- ・  $S_{B_1}$  : 曝露あり（ $E = 1$ ）、アウトカムなし（ $D = 0$ ）の対象者が選択される確率
- ・  $S_{B_0}$  : 曝露なし（ $E = 0$ ）、アウトカムなし（ $D = 0$ ）の対象者が選択される確率

このようにパラメータを定めると

$$A_1 = \frac{A_{11}}{S_{A_1}}, A_0 = \frac{A_{01}}{S_{A_0}}, B_1 = \frac{B_{11}}{S_{B_1}}, B_0 = \frac{B_{01}}{S_{B_0}}$$

から

$$RR = \frac{A_1/N_1}{A_0/N_0} = \frac{A_1/(A_1+B_1)}{A_0/(A_0+B_0)} = \frac{(A_{11}/S_{A_1})/(A_{11}/S_{A_1} + B_{11}/S_{B_1})}{(A_{01}/S_{A_0})/(A_{01}/S_{A_0} + B_{01}/S_{B_0})} = \frac{A_{11}/(A_{11} + B_{11} \cdot \frac{S_{A_1}}{S_{B_1}})}{A_{01}/(A_{01} + B_{01} \cdot \frac{S_{A_0}}{S_{B_0}})}.$$

したがって「 $S_{A_1} = S_{B_1}$  かつ  $S_{A_0} = S_{B_0}$ 」ならば、つまり曝露ありの場合もなしの場合も、アウトカムの有無によらず選択確率が等しければ、リスク比にバイアスが生じないことがわかる（十分条件）。

なお、ここで

$$S_{A_1} = S_{B_1} \Leftrightarrow S_{A_1} = S_{B_1} = S_{N_1} \Leftrightarrow S_{A_1} = S_{N_1}, \text{ 及び } S_{A_0} = S_{B_0} \Leftrightarrow S_{A_0} = S_{B_0} = S_{N_0} \Leftrightarrow S_{A_0} = S_{N_0}$$

という関係がなりたつ。したがって、例えば「 $S_{A_1} = S_{B_1}$  かつ  $S_{A_0} = S_{B_0}$ 」ならば「 $S_{A_1} = S_{N_1}$  かつ  $S_{A_0} = S_{N_0}$ 」、すなわち

$$\frac{S_{A_1}/S_{N_1}}{S_{A_0}/S_{N_0}} = 1$$

もなりたつ。本文中でしめしたように、これはリスク比に選択バイアスが生じない条件に他ならない

（3.6.6 式 3-9）。逆も成り立つので、選択バイアスが生じない条件「 $S_{A_1} = S_{B_1}$  かつ  $S_{A_0} = S_{B_0}$ 」と「 $S_{A_1} = S_{N_1}$  かつ  $S_{A_0} = S_{N_0}$ 」は実質的におなじ内容である。すなわち、アウトカムなしの対象者について選択確率を定めても、集団全体について選択確率を定めても本質的な違いはない（3.6.4）。

また、観察されなかった層の分割表を求める際、以下のようにパラメータを設定することも考えられる。ただし本文中と同様に、選択前の集団の大きさ（ $N = N_1 + N_0$ ）は既知とする。

- $O_{CD1}$  : 対象者選択なし（ $C = 0$ ）の場合の、アウトカムありの集団（ $D = 1$ ）における曝露オッズ。  
すなわち、この集団で、曝露ありの対象者が、曝露なしの対象者の何倍多いかという比  $A_{10}/A_{00}$ 。
- $O_{CD0}$  : 対象者選択なし（ $C = 0$ ）の場合の、アウトカムなしの集団（ $D = 0$ ）における曝露オッズ。  
すなわち、この集団で、曝露ありの対象者が、曝露なしの対象者の何倍多いかという比  $B_{10}/B_{00}$ 。
- $IP_{C0}$  : 対象者選択なし（ $C = 0$ ）の集団での、曝露なしの場合のアウトカムの発生割合  $A_{00}/N_{00}$ 。

このようにパラメータを定めると

$$A_{10} = O_{CD1}A_{00}, B_{10} = O_{CD0}B_{00}, A_{00} = IP_{C0}N_{00}.$$

これらと  $N_{10} = A_{10} + B_{10}$ 、 $N_{00} = A_{00} + B_{00}$  から

$$\begin{aligned} N_{10} &= A_{10} + B_{10} = O_{CD1}A_{00} + O_{CD0}B_{00} = O_{CD1}A_{00} + O_{CD0}(N_{00} - A_{00}) \\ &= O_{CD1}IP_{C0}N_{00} + O_{CD0}(N_{00} - IP_{C0}N_{00}) = \{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1}\}N_{00}. \end{aligned}$$

さらに、対象者選択なしの集団の合計人数（ $N_{10} + N_{00}$ ）は、対象者選択の源泉となった集団の人数  $N$  から、実際に研究に参加した対象者の人数  $N_{11} + N_{01} = N_{+1}$  を引いたものなので

$$N_{10} + N_{00} = N_{+0} = N - N_{+1}.$$

この左辺に上で求めた  $N_{10}$  を代入して  $N_{00}$  について解くと

$$N_{00} = \frac{N - N_{+1}}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1} \text{ が得られ、よって } N_{10} = \frac{\{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1}\}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}.$$

したがって

$$\begin{aligned} A_{00} &= \frac{IP_{C0}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}, A_{10} = \frac{IP_{C0}O_{CD1}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}, \\ B_{00} &= \frac{(1 - IP_{C0})(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}, B_{10} = \frac{(1 - IP_{C0})O_{CD0}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}. \end{aligned}$$

以上から、選択前の集団における分割表の度数を以下のように求めることができる。

$$\begin{aligned} A_1 &= A_{11} + \frac{IP_{C0}O_{CD1}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}, A_0 = A_{01} + \frac{IP_{C0}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}, \\ B_1 &= B_{11} + \frac{(1 - IP_{C0})O_{CD0}(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}, B_0 = B_{01} + \frac{(1 - IP_{C0})(N - N_{+1})}{(1 - IP_{C0})O_{CD0} + IP_{C0}O_{CD1} + 1}. \end{aligned}$$

あるいは、以下のようにパラメータを設定することもできる。

- $S_{N_0}$  : 曝露なし ( $E = 0$ ) の場合に、対象者選択あり ( $C = 1$ ) となる割合  $N_{01}/N_0$ 。これは曝露なしの集団における選択確率でもある。例えば、対象者に研究への参加を呼びかけたとき、曝露なしの集団で参加に応じる確率に相当する。
- $RR_{CE0}$  : 曝露なし ( $E = 0$ ) の場合の、患者選択とアウトカムの関連性をしめすリスク比  $(A_{01}/N_{01})/(A_{00}/N_{00})$ 。
- $RR_{CE1}$  : 曝露あり ( $E = 1$ ) の場合の、患者選択とアウトカムの関連性をしめすリスク比  $(A_{11}/N_{11})/(A_{10}/N_{10})$ 。

このようにパラメータを定めると、選択前の集団における分割表の度数は以下のようになる。

$$A_1 = A_{11} \left\{ 1 + \frac{1}{RR_{CE1} \cdot N_{11}} \left( N - \frac{N_{01}}{S_{N_0}} - N_{11} \right) \right\}, A_0 = A_{01} \left( 1 + \frac{1 - S_{N_0}}{RR_{CE0} \cdot S_{N_0}} \right), N_1 = N - \frac{N_{01}}{S_{N_0}}, N_0 = \frac{N_{01}}{S_{N_0}}.$$

いずれも、数式の見かけが異なるだけで同種の計算である。しかし、解析者が抱いている選択バイアスに対する量的な想定を表現する際、考えやすさやあつかいやすさの面で違いがあるかもしれない。すでに得られている情報も考慮して、パラメータを設定するとよい。

### 参考 3- 15 アウトカムの誤分類の補正 : ケースコントロール研究 [ 本文 : 3.7.5 ]

誤分類	アウトカム	曝露	
		$E = 1$	$E = 0$
あり (選択後の集団)	$D^* = 1$	$A_1^*$	$A_0^*$
	$D^* = 0$	$B_1^*$	$B_0^*$
なし (選択前の集団)	$D = 1$	$A_1$	$A_0$
	$D = 0$	$B_1$	$B_0$

上表のように度数を表し、また対象者選択の源泉となった集団（選択前の集団）において、以下のパラメータを定める。

$Se_1$  : 曝露ありの場合の、アウトカム分類の感度

$Sp_1$  : 曝露ありの場合の、アウトカム分類の特異度

$Se_0$  : 曝露なしの場合の、アウトカム分類の感度

$Sp_0$  : 曝露なしの場合の、アウトカム分類の特異度

まず、選択前の集団における誤分類のない度数 ( $A_1$ 、 $A_0$ 、 $B_1$ 、 $B_0$ ) に注目しよう。これらの度数が、以上の感度と特異度のもとで誤分類を受けると、実際にケースコントロール研究を行う場合にも、不正確なアウトカムを参照しながらケースとコントロールを選択することになる。すなわち、分類上アウトカムありの対象者は（真のアウトカムがなしの対象者も含め）全例ケースとして選択する。また、分類上アウトカムなしの対象者から（真にアウトカムありの対象者も含め）確率  $S$  ( $0 < S_N < 1$ ) でコントロールを選択する。こうして得られた、ただし誤りを含んだケースやコントロールの人数が、上表の度数 ( $A_1^*$ 、 $A_0^*$ 、 $B_1^*$ 、 $B_0^*$ ) だとしてよう。

さて、選択前の集団ですでに誤分類を受けたアウトカム、すなわち分類上のアウトカムが「あり」だった対象者がケースとなり、また「なし」だった対象者がコントロールとなるのだから、それぞれの候補者数は以下ようになる（参考 3-9）。

ケースの候補者 : 曝露あり  $Se_1 A_1 + (1 - Sp_1) B_1$ 、曝露なし  $Se_0 A_0 + (1 - Sp_0) B_0$ 。

コントロールの候補者 : 曝露あり  $(1 - Se_1)A_1 + Sp_1B_1$ 、曝露なし  $(1 - Se_0)A_0 + Sp_0B_0$ .

これらのうち、ケースについては候補者が全例選択され、またコントロールについては確率  $S$  で候補者の一部が選択されるので、このケースコントロール研究によって選択される対象者数 ( $A_1^*$ 、 $A_0^*$ 、 $B_1^*$ 、 $B_0^*$ ) は以下になる。

ケース : 曝露あり  $A_1^* = Se_1A_1 + (1 - Sp_1)B_1$ 、曝露なし  $A_0^* = Se_0A_0 + (1 - Sp_0)B_0$ .

コントロール : 曝露あり  $B_1^* = S \{(1 - Se_1)A_1 + Sp_1B_1\}$ 、曝露なし  $B_0^* = S \{(1 - Se_0)A_0 + Sp_0B_0\}$ .

これらを解くと、誤分類がない、選択前の集団における度数を以下のように求めることができる。

$$A_1 = \frac{S \cdot Sp_1 A_1^* - (1 - Sp_1) B_1^*}{S (Se_1 + Sp_1 - 1)}, A_0 = \frac{S \cdot Sp_0 A_0^* - (1 - Sp_0) B_0^*}{S (Se_0 + Sp_0 - 1)},$$

$$B_1 = \frac{Se_1 B_1^* - S (1 - Se_1) A_1^*}{S (Se_1 + Sp_1 - 1)}, B_0 = \frac{Se_0 B_0^* - S (1 - Se_0) A_0^*}{S (Se_0 + Sp_0 - 1)}.$$

以上から、選択前の集団での感度と特異度、及びコントロール選択の確率をもちいて、差別的で独立なアウトカムの誤分類の影響を補正したオッズ比  $OR = (A_1 B_0) / (A_0 B_1)$  が得られる。

さらに以上の関係から、このケースコントロール研究で選択された集団（選択後の集団）における感度と特異度も求めることができる。例えば曝露ありの集団に注目すると、選択後の集団に含まれる真にアウトカムありの対象者数は  $Se_1 A_1 + S (1 - Se_1) A_1$  である。そのうち正しくアウトカムありと分類された、すなわちケースとして選択された対象者数は  $Se_1 A_1$  である。したがって、選択後の集団における感度は、曝露ありの集団では

$$\frac{Se_1 A_1}{Se_1 A_1 + S (1 - Se_1) A_1} = \frac{Se_1}{Se_1 + S (1 - Se_1)}$$

となる。同様に、選択後の集団における特異度は、曝露ありの集団では

$$\frac{S \cdot Sp_1 B_1}{S \cdot Sp_1 B_1 + (1 - Sp_1) B_1} = \frac{S \cdot Sp_1}{1 - Sp_1 + S \cdot Sp_1}$$

となる。添え字の **1** を **0** に置き換えると、曝露なしの集団における同様な関係が得られる。本文中では非差別的誤分類を想定し ( $Se_1 = Se_0 = Se$ 、かつ  $Sp_1 = Sp_0 = Sp$ )、また選択後の集団における感度と特異度をそれぞれ  $Se_c$  及び  $Sp_c$  で表し、以上の関係をしめした (3.7.5)。

### 参考 3- 16 アウトカムの誤分類：アウトカム発生までの時間やハザード等の補正

#### [ 本文：3.8.4 ]

時点  $t$  でのアウトカム発生に関するハザード関数  $\lambda(t)$  や累積ハザード関数  $\Lambda(t) = \int_0^t \lambda(s) ds$ 、またアウトカム発生までの時間  $T$  の累積分布関数  $F(t) = \Pr(T \leq t) = 1 - e^{-\Lambda(t)}$  を推定する際、アウトカムの誤分類の影響を補正する方法がある。時間の経過とともに発生するアウトカムの誤分類がどのようなメカニズムによって生じるのか、その状況や想定に応じてさまざまな方法が提案されている。

#### 時間 $T$ を測定結果（連続量）とみなし、その測定誤差を補正する方法

以下では、おもに Oh ら (2018) <sup>38</sup> を参照して、アウトカム発生までの時間  $T$  の測定誤差を補正する方法について紹介する。

アウトカム発生までの時間について加速モデルを考え、曝露  $E$  と時間  $T$  の関係が  $T = \exp(\alpha_0 + \alpha_1 E + \sigma \epsilon)$  というモデルで表されると仮定する。ただし、 $\alpha_0$  と  $\alpha_1$  は回帰係数であり、 $\sigma$  は形状パラメータ、また  $\epsilon$  は極値分布に従う誤差項である。さらに、真の時間  $T$  と、誤分類を受けた時間  $T^*$  とのあいだに  $T^* = T \cdot \exp(v)$ 、すなわち  $\ln T^* = \ln T + v$  という関係を想定する。ただし、 $v$  は、例えば平均 0、分散  $\sigma_v^2$  の分布に従う。この場合、誤差の程度を定めているのは  $v$  の分散  $\sigma_v^2$  であり、この値が大きいほど真の時間  $T$  とは異なる時間、すなわち誤差を含む時間  $T^*$  が観察される。このため、 $T^*$  をもちいて求めたパラメータの推定値 ( $\alpha_0$ 、 $\alpha_1$ 、及び  $\sigma$ ) にバイアスが生じる。また、おなじデータに対して Cox 比例ハザードモデルをもちいて推定した対数ハザード比  $\beta$  と、加速モデルに含まれるパラメータとの間には  $\beta = -\alpha_1/\sigma$  という関係がある。このため、加速モデルのパラメータの値からハザード比を推定することができる。以上の枠組みのもとで、例えば分散  $\sigma_v^2$  の値を定め、以下の手順で誤分類の影響を補正した対数ハザード比  $\beta$  を求めることができる<sup>38</sup>。なお、この分散の値が、解析者が想定する測定誤差の程度を表しており、この感度分析におけるパラメータである。

- 正の数  $\lambda$  の値を適当に定め、誤差項にあたる量  $\omega$  の値を、正規分布  $N(0, \lambda\sigma_v^2)$  からランダムに選ぶ。
- 得られた  $\omega$  を、データとして得られている  $\ln T^*$  に加算する。すなわち  $\ln T = \ln T^* + \omega$  と、測定誤差をさらに上乗せした時間  $T$  を求める。この加算により、 $\ln T$  の測定誤差の分散は  $(1 + \lambda)\sigma_v^2$  となる。
- $\ln T$  に対して加速モデルをあてはめ、得られた推定値から対数ハザード比  $\beta_\lambda$  を求める。
- 以上の計算を  $\lambda$  の値ごとに  $k$  回繰り返し、対数ハザード比  $\beta_\lambda$  を  $k$  個求める。さらに  $\lambda$  の値を変えて同様な計算をくりかえす。これらの計算により、イメージとしては横軸に  $\lambda$ 、縦軸に  $\beta_\lambda$  の値をプロットしたグラフが得られる。
- 以上のグラフに曲線（例えば 2 次曲線）をあてはめ、 $\lambda = -1$  とした場合の  $\beta_\lambda$  の値を求める。ここで、 $\lambda$  は正の値なので、0 以下の値については  $\beta_\lambda$  の値が存在しない。すなわち、このあてはめは外挿にあたる。また以上の手順から、 $\lambda = -1$  という値は、 $\ln T$  の測定誤差の分散  $(1 + \lambda)\sigma_v^2$  が 0 になる場合と対応していることがわかる。つまりこの外挿により、測定誤差の影響を補正した対数ハザード比の推定値が得られる。

シミュレーションと外挿によって測定誤差の影響を補正するこうした方法を **Simulation and Extrapolation (SIMEX)** 法と呼ぶが、測定誤差の補正にひろくもちいられている方法の 1 つである<sup>39</sup>。

### アウトカムの発生を Current Status Data と関連付ける方法

以下では、おもに McKeown ら (2010)<sup>40</sup> を参照して、代替となるアウトカムの発生を観察することで真のアウトカムの発生を判定する際に生じる誤差と、それを補正する方法について紹介する。

時間の経過とともに発生する真のアウトカムを直接調べるかわりに、例えば複数の臨床検査値等を参照して、それらがある基準を満たせばアウトカムありとみなす、といった形でアウトカムの有無を判定することがある。この場合、実際に観察されるのは、真のアウトカム  $D$  ではなく、その代替となるアウトカム  $\Delta$  である。以下ではアウトカムありの場合を 1 で、なしの場合を 0 で表す。

この際、判定にもちいた検査の精度（感度と特異度）によっては、誤分類が生じる。例えば、ある検査の感度は  $Se = \Pr(\Delta = 1 | D = 1)$ 、また特異度は  $Sp = \Pr(\Delta = 0 | D = 0)$  等と表すことができる。

また、アウトカムが発生した時間  $T$  についても、それを直接観察することはできない。あらかじめ計画された観測時間（monitoring time） $C$  を通じ、直前の検査から陽性となった検査までのある時点でアウトカ

ムが生じた、といった間接的な情報が得られるのみである。なお以下では、観測時間  $C$  が  $T$  と独立であることを仮定する。臨床試験のように事前に観測時間を定め、アウトカムやそれと関連する変数の観察を前向きに行う場合には自然な仮定だが、観察研究の場合はその限りではない。以下では、いったんこれを仮定して先へ進もう。

こうして、各時点での状況をしめすデータ（current status data）、すなわち時間  $C$  ごとに得られる代替アウトカム  $\Delta$  の値から、真のアウトカムの有無とその発生時間、すなわち時間  $T$  と  $D$  の関係を推定することが問題となる。例えば、ある検査の  $i$  番目の観測時間を  $C_i$ 、そのときの代替アウトカムを  $\Delta_i$  で表すと、観測時間  $C_i = c_i$  にアウトカムありと判定される（ $\Delta_i = 1$  となる）確率  $\Pr(\Delta_i = 1 | c_i)$  は、この時点における真のアウトカムの有無（ $d_i = 1$  または  $d_i = 0$ ）や、この検査の感度と特異度（Se と Sp）をもちいて以下のように表すことができる。

$$\begin{aligned}\Pr(\Delta_i = 1 | c_i) &= \Pr(\Delta_i = 1 | d_i = 1, c_i) \Pr(d_i = 1 | c_i) + \Pr(\Delta_i = 1 | d_i = 0, c_i) \Pr(d_i = 0 | c_i) \\ &= \text{Se} \cdot \Pr(d_i = 1 | c_i) + (1 - \text{Sp}) \{1 - \Pr(d_i = 1 | c_i)\} = (\text{Se} + \text{Sp} - 1)F(c_i) + 1 - \text{Sp}\end{aligned}$$

ただし、 $F(c_i) = \Pr(T \leq c_i)$  は、時点  $c_i$  までに真のアウトカムが発生する確率である。同様に

$$\Pr(\Delta_i = 0 | c_i) = \text{Sp} - (\text{Se} + \text{Sp} - 1)F(c_i)$$

がなりたつ。これらから、時点  $c_i$  での尤度  $L_i$  が、観察された代替アウトカムの有無（ $\delta_i = 1$  または  $\delta_i = 0$ ）をもちいて以下のように表される。

$$L_i = \{(\text{Se} + \text{Sp} - 1)F(c_i) + 1 - \text{Sp}\}^{\delta_i} \{\text{Sp} - (\text{Se} + \text{Sp} - 1)F(c_i)\}^{1-\delta_i}.$$

このような尤度を、時点を通じて、また対象者を通じて掛け合わせることで、アウトカム発生までの時間  $T$  の累積分布関数  $F(t)$  を推定することができる<sup>40</sup>。また以上では、感度と特異度（Se と Sp）を一定とみなしているが、これらが時間とともに変化する場合もあつかうことができる。

このような考え方をベースに、 $F(t)$  をノンパラメトリックに推定することができるし、パラメトリックな、あるいはセミパラメトリックなモデルをもちいる方法もある<sup>41</sup>。さらに、複数の検査結果を同時に考慮する場合や、打ち切りも考慮する場合、時間  $T$  が観測時間  $C$  と独立とは限らない場合等、さまざまな応用がある<sup>42</sup>。

#### アウトカムの誤分類を、競合イベントと関連付ける方法

コホート研究では、他のアウトカムが先に生じたせいで、関心があるアウトカムが観察されないことがある。これらのアウトカムが互いに識別しにくい場合は、さらに誤分類も生じる。例えば、がんによる死亡が、別の死因による死亡とみなされ、分類上、がんによる死亡は「なし」に分類される。こうした場合、以下の3つの要素を組み合わせて考えることが多い。

- 1) 関心があるアウトカムの発生状況を表すハザード関数や累積ハザード関数、または累積分布関数。
- 2) 関心があるアウトカム以外のアウトカムの発生状況を表すハザード関数や累積ハザード関数、または累積分布関数。
- 3) 一方のアウトカムが他方のアウトカムとして、誤分類される確率（感度や特異度に相当する確率）。

最後の「誤分類される確率」は、理論的には時点ごとに異なる値を取る関数だが、あつかいを簡単にするため、時間を通じて一定とみなすこともある。基本的な考え方としては、関心があるアウトカムとそれ以



外のアウトカムがそれぞれに発生し（1 及び 2）、その後、ある確率で正しく、もしくは誤って分類される（3）。こうした考え方に基づき、誤分類の影響を補正したハザード比や累積ハザード関数、累積分布関数を推定する方法が提案されている<sup>43、44、45、46</sup>。

## 第 1 部の参考文献

1. Rothman KJ, Greenland S, Lash TL. Validity in epidemiologic studies. Chapter 9. In: Rothman KJ, Greenland S, Lash TL (eds). *Modern Epidemiology*. 3rd edn. Philadelphia, PA: Lippincott Williams&Wilkins, 2008.
2. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology* 2016; 183(8):758–764.
3. Sterne JAC et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *British Medical Journal* 2016; 355: i4919.
4. Sterne JAC, Higgins JPT, Elbers RG, Reeves BC and the development group for ROBINS-I. Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I): detailed guidance, updated 12 October 2016. Available from [https://www.bristol.ac.uk/media-library/sites/social-community-medicine/images/centres/cresyda/ROBINS-I\\_detailed\\_guidance.pdf](https://www.bristol.ac.uk/media-library/sites/social-community-medicine/images/centres/cresyda/ROBINS-I_detailed_guidance.pdf) [accessed March 16, 2019].
5. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; 15: 615–625.
6. Higgins J, Altman D, Sterne J. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 (Updated March 2011).
7. Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in randomized trials-A conversation between trialists and epidemiologists. *Epidemiology* 2017; 28: 54–9.
8. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York: Springer, 2009.
9. Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *International Journal of Epidemiology* 2014; 43(6): 1969–85.
10. Thomas L, Peterson ED. The value of statistical analysis plans in observational research: Defining high-quality research from the start. *Journal of American Medical Association* 2012; 308: 773–774.
11. Greenland S. Introduction to regression modeling. Chapter 21. In: Rothman KJ, Greenland S, Lash TL (eds). *Modern Epidemiology*. 3rd edn. Philadelphia, PA: Lippincott Williams&Wilkins, 2008.
12. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period--application to control of the healthy worker survivor effect. *Mathematical Modelling*. 1986; 7:1393–1512.
13. Robins JM. Marginal structural models. In: 1997 Proceedings of the Section on Bayesian Statistical Science, Alexandria, VA: American Statistical Association, 1998; 1–10.
14. Robins JM. The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology: A Focus on AIDS* (L. Sechrest, H. Freeman and A. Mulley, eds.) 1989; 113–159.
15. Robins JM, Blevins D, Ritter G, Wulfsohn M. G-estimation of the effect of prophylaxis therapy for pneumocystic carinii pneumonia on the survival of AIDS patients. *Epidemiology*. 1992; 3 319–336.

16. Greenland S, Lash TL. Bias analysis. Chapter 19. In: Rothman KJ, Greenland S, Lash TL (eds). *Modern Epidemiology*. 3rd edn. Philadelphia, PA: Lippincott Williams&Wilkins, 2008.
17. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use: ICH Harmonised Tripartite Guideline. Statistical Principles for Clinical Trials E9. Available from <http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html> [accessed March 16, 2019].
18. Greenland S, Rothman KJ. Introduction to categorical statistics. Chapter 14. In: Rothman KJ, Greenland S, Lash TL (eds). *Modern Epidemiology*. 3rd edn. Philadelphia, PA: Lippincott Williams&Wilkins, 2008.
19. Wheeler K, McKelvey W, Thorpe L, et al. Asthma diagnosed after 11 September 2001 among rescue and recovery workers: findings from the World Trade Center Health Registry. *Environ Health Perspect* 2007; 115(11): 1584-1590.
20. Jurek AM, Maldonado G. Quantitative bias analysis in an asthma study of rescue-recovery workers and volunteers from the 9/11 World Trade Center attacks. *Annals of Epidemiology* 2016; 26(11): 794-801.
21. Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* 1998; 54: 948–963.
22. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. *Epidemiology*. 2016; 27: 368-77.
23. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of Internal Medicine* 2017; 167: 268-274.
24. Dosemeci M, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect toward the null value? *American Journal of Epidemiol* 1990; 132(4): 746–748.
25. Weinberg CR, Umbach DM, Greenland S. When will non-differential misclassification of an exposure preserve the direction of a trend? *American Journal of Epidemiology* 1994; 140: 565-71.
26. Jurek AM, Greenland S, Maldonado G. How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? *International Journal of Epidemiology* 2008; 37: 382–5.
27. NYC Health. World Trade Center Health Registry Wave 1 Survey (2003–2004) Public Use Data Set Codebook. <http://sites.bu.edu/tec/files/2016/01/wave1-codebook.pdf> [accessed March 16, 2019].
28. Hernan MA, Sauer BC, Hernandez-Diaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *Journal of Clinical Epidemiology* 2016; 79: 70-5.
29. Funk MJ, Landi SN. Misclassification in administrative claims data: quantifying the impact on treatment effect estimates. *Current Epidemiology Reports* 2014; 1(4): 175–85.
30. Greenland S. The effect of misclassification in the presense of covariate. *Amedical Journal of Epidemiology* 1980; 112: 564-569.
31. Ogburn EL, VanderWeele TJ. On the nondifferential misclassification of a binary confounder. *Epidemiology* 2012; 23: 433-439.
32. Olsen J, Melbye M, Olsen SF et al. The Danish National Birth Cohort—its background, structure and aim. *Scandinavian Journal of Public Health* 2001; 29: 300–7.

33. Nohr EA, Liew Z. How to investigate and adjust for selection bias in cohort studies. *Acta Obstetrica et Gynecologica Scandinavica*. 2018 Apr; 97(4): 407-416.
34. Farfel M, DiGrande L, Brackbill R, Prann A, Cone J, Friedman S, et al. An overview of 9/11 experiences and respiratory and mental health conditions among World Trade Center Health Registry enrollees. *Journal of Urban Health* 2008; 85(6): 880–909.
35. Murphy J, Brackbill RM, Thalji L, Dolan M, Pulliam P, Walker DJ. Measuring and maximizing coverage in the World Trade Center Health Registry. *Statistics in Medicine* 2007; 26(8): 1688–1701.
36. Greenland S, Salvan A, Wegman D H, Hallock M F, Smith T J. A case-control study of cancer mortality at a transformer assembly facility. *Int Arch Occup Environ Health* 1994; 66: 49-54.
37. 竹内文乃, 野間 久史. 観察研究におけるバイアスの感度解析. *統計数理* 2014; 62(1): 77-92.
38. Oh E, Shepherd BE, Lumley T, Shaw PA. Considerations for analysis of time-to-event outcomes measured with error: Bias and correction with SIMEX. *Statistics in Medicine* 2018; 37(8): 1276-1289.
39. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu C. Measurement error in nonlinear models. Boca Roton, FL: Cphapman and Hall, 2006.
40. McKeown K, Jewell NP. Misclassification of current status data. *Lifetime Data Analysis* 2010; 16, 215–30.
41. Sal Y Rosas VG, Hughes JP. Nonparametric and semiparametric analysis of current status data subject to outcome misclassification. *Statistical Communications in Infectious Diseases* 2011; 364.
42. McKeown KM. Topics in Current Status Data. 2011; PhD thesis, University of California, Berkeley.
43. Van Rompaye B, Goetghebeur E, Jaffar S. Design and testing for clinical trials faced with misclassified causes of death. *Biostatistics* 2010; 11(3): 546–58.
44. Van Rompaye B, Jaffar S, Goetghebeur E. Estimation with Cox models: Cause-specific survival analysis with misclassified cause of failure. *Epidemiology* 2012; 23(2): 194–202.
45. Ha J, Tsodikov A. Isotonic estimation of survival under a misattribution of cause of death. *Lifetime Data Analysis*. 2012; 18(1):58–79.
46. Bakoyannis G, Yiannoutsos CT. Impact of and Correction for Outcome Misclassification in Cumulative Incidence Estimation. *PLoS One* 2015; 109: e0137454.

執筆者・タスクフォースメンバー

日本製薬工業協会 医薬品評価委員会 データサイエンス部会

2016 年度タスクフォース 2

○ 兼山 達也	大日本住友製薬株式会社
○ 小森 哲志	バイエル薬品株式会社
塩境 一仁	第一三共株式会社
中水流 嘉臣	ファイザー株式会社

担当副部長

小山 暢之	第一三共株式会社
-------	----------

○：タスクフォースリーダー兼推進委員

レビューア

第 1 部

データサイエンス部会

山本 英晴	中外製薬株式会社
酒井 弘憲	エーザイ株式会社
菅波 秀規	興和株式会社
土屋 悟	大日本住友製薬株式会社
奥山 ことば	MSD 株式会社

第 2 部

データサイエンス部会

山本 英晴	中外製薬株式会社
酒井 弘憲	エーザイ株式会社
松下 泰之	第一三共株式会社
土屋 悟	大日本住友製薬株式会社
奥山 ことば	MSD 株式会社

ESATJ (Excel Tool)

データサイエンス部会

淡路 直人	中外製薬株式会社
土綿 慎一	ファイザー株式会社
松下 泰之	第一三共株式会社