【数式を使わない】
医療機器における統計的手法と
サンプルサイズ決定方法

村山 浩一 著

目次

第 1 章　はじめに …… 3
1.　統計的手法の目的 …… 5
2.　ISO 13485 における統計的手法の要求事項 …… 5
2.1.　設計・開発の検証（7.3.6）…… 5
2.2.　設計・開発のバリデーション（7.3.7）…… 6
2.3.　製造およびサービスの提供（7.5）…… 7
3.　統計的手法の適用範囲 …… 8
4.　製品要求事項の保証 …… 9
5.　統計手法の基本的な分類 …… 10
5.1.　記述統計 …… 10
5.2.　推測統計 …… 10
5.3.　何を検定したいのか …… 11
6.　サンプルサイズに関する誤解 …… 11
7.　統計的手法の重要なポイント …… 12
8.　サンプルサイズ決定の基本原則 …… 13
9.　まとめ …… 14

第 2 章　統計的手法の基礎 …… 15
1.　データの構造 …… 17
2.　計量値と計数値の分布 …… 18
3.　計量値の分布：正規分布 …… 18
3.1.　正規分布の定義と重要性 …… 18
3.2.　正規分布の定義と特徴 …… 18
3.3.　正規分布の重要性 …… 19
3.4.　正規分布の性質 …… 19
4.　計量値の分布 :t 分布 …… 20
4.1.　t 分布の定義と特徴 …… 20
4.2.　t 分布の重要性 …… 20
4.3.　t 分布の性質 …… 21
4.4.　t 分布と正規分布の関係 …… 21
5.　計数値の分布：二項分布 …… 21
5.1.　二項分布の実用的な応用例 …… 22
6.　計数値の分布：ポアソン分布 …… 23
6.1.　ポアソン分布の実用的な応用例 …… 24
7.　分布の違いによる確率密度関数式 …… 24
8.　サンプルサイズ決定の基本原則 …… 26
9.　統計的検定に関する重要概念 …… 27

9.1.　有意水準（α）…… 27
9.2.　検出力（1- β）…… 27
9.3.　P 値　…… 27
10.　誤差に関する重要概念 …… 28
10.1.　第 1 種の過誤（生産者危険）…… 28
10.2.　第 2 種の過誤（消費者危険）…… 28
11.　サンプリングに関する概念 …… 28
11.1.　信頼区間 …… 28
11.2.　サンプルサイズ …… 28
12.　実践的な応用 …… 29

第 3 章　統計的推定とは　…… 31
1.　統計的推測の実例と重要性　…… 33
2.　推定の種類　…… 34
2.1.　点推定　…… 34
2.2.　区間推定　…… 34
2.3.　信頼区間と信頼度　…… 34

第 4 章　第 1 種の過誤と第 2 種の過誤　…… 37
1.　第 1 種の過誤と第 2 種の過誤の定義　…… 39
2.　第 1 種の過誤：生産者危険　…… 40
3.　第 2 種の過誤：消費者危険　…… 41
4.「消費者危険」と「生産者危険」の関係 …… 42
5.　OC 曲線の概念と特性 …… 43

第 5 章　サンプルサイズの決定　…… 45
1.　基本的な考え方　…… 47
2.　統計的検定における基本的な考え方　…… 48
3.　サンプルサイズ決定の重要性　…… 49
4.　なぜサンプルサイズを決定しないといけないか　…… 52
5.　サンプルサイズ検討に必要な情報　…… 53
6.　サンプル評価のステップ　…… 53

第 6 章　帰無仮説とは　…… 55
1.　背理法による仮説検定の基本概念　…… 57
2.　第 1 種の過誤と第 2 種の過誤　…… 57
3.　帰無仮説の概念と特徴　…… 58
4.　帰無仮説の棄却と解釈　…… 58

第 7 章　有意水準とは　…… 61

1. 検定の基本概念 …… 63
2. 検定と仮説の関係 …… 63
3. 有意水準の定義 …… 64
4. P 値の概念と解釈…… 65
5. P 値と有意水準の実践的な関係…… 66
6. 有意水準と第 1 種の過誤の関係 …… 67
7. 医療機器開発における有意水準の重要性 …… 67

第 8 章　検出力とは …… 69
1. 検出力（Statistical Power）とは …… 71
2. 検出力の特性 1…… 72
3. 検出力の特性 2…… 72
4. 第一種の誤りと第二種の誤り …… 73
5. 検出力の定義 …… 74
6. コイン投げの具体例による検出力の説明 …… 75
7. 棄却域の設定と検出力の関係 …… 76
8. 連続型確率分布における検出力 …… 77
9. 検出力の重要性 …… 78
10. 消費者危険が受容可能な設計を行う …… 79
11. 実践的な検出力の設定 …… 80
12. 検出力を高める方法 …… 81
12.1. 二項検定による α エラーの具体例 …… 81
13. サンプルサイズの考え方 …… 84
13.1. 区間推定の場合のサンプルサイズの考え方 …… 84
13.2. 検定の場合のサンプルサイズの考え方：正規分布、片側検定の場合…… 84
13.3. 検定の場合のサンプルサイズの考え方：正規分布、両側検定…… 85

第 9 章　サンプルサイズ、検出力の計算式 …… 87
1. 一つの母平均の検定におけるサンプルサイズの検討（母分散既知の場合）…… 89
2. 一つの母平均の検定におけるサンプルサイズの検討（母分散未知の場合）…… 90
3. 一つの母分散の検定におけるサンプルサイズの検討 …… 91
4. 母不良率の検定におけるサンプルサイズの検討 …… 91
5. 二つの母平均の差の検定におけるサンプルサイズの決定 …… 92
6. 二つの母不良率の差の検定 …… 92
7. 二つの母分散の比の検定におけるサンプルサイズの検討 …… 93

8. 対応がある場合の母平均の差に関するサンプルサイズの検討　93
9. 一つの母平均の区間推定に対するサンプルサイズの検討（母分散既知）…… 94
10. 一つの母平均の区間推定におけるサンプルサイズの検討（母分散未知）…… 94
11. 一つの母分散の区間推定におけるサンプルサイズの検討 …… 95
12. 一つの母不良率の推定におけるサンプルサイズの決定 …… 95
13. 二つの母平均の差の区間推定におけるサンプルサイズの決定　95
14. 二つの母分散の比の区間推定におけるサンプルサイズの決定　96

第 10 章　例題 …… 97
1. 例題 1：一つの母平均の検定におけるサンプルサイズの検討（母分散既知）…… 99
2. 例題 2：一つの母平均の区間推定に対するサンプルサイズの検討（母分散既知）…… 100
3. 例題 3：一つの母平均の検定におけるサンプルサイズの検討（母分散未知）…… 100
4. 例題 4：一つの母平均の区間推定におけるサンプルサイズの検討（母分散未知）…… 101
5. 統計的手法の実際 …… 102
5.1. 製品、プロセスのリスク評価 …… 102
5.2. 計数値（Attribute）試験に適用するサンプルサイズ計算 … 102
5.3. 計数値（Attribute）試験に適用するサンプルサイズ計算 … 102
5.4. 計量値（Variable）試験に適用するサンプルサイズ計算…… 102
5.5. 平均値の信頼区間（分散既知）…… 103
5.6. 仮説検定 …… 103
5.7. 平均の仮説検定（分散既知）…… 104
5.8. 平均の仮説検定（分散既知）…… 104
5.9. 平均の仮説検定（分散未知）…… 105
5.10. プロセスバリデーションへの統計的手法の適用 …… 106

はじめに

医療機器の開発・製造において、統計的手法の活用は不可欠なものとなっている。しかし、多くの医療機器設計者にとって統計学は「難しい」「複雑」「数式が多い」というイメージが強く、敬遠されがちな分野でもある。特に医療機器業界では、ISO 13485 をはじめとする品質マネジメントシステムの要求事項に対応するため、統計的手法の理解と適切な活用が強く求められている。

本書は、そうした統計学への心理的なハードルを下げることを目的として執筆した。特に重視したのは、なるべく数式を使わず、図や具体例を用いて直感的に理解できる説明を心がけたことである。統計学の本質的な考え方や、実務で本当に必要な知識に焦点を当てることで、読者に「使える統計学」を提供することを目指している。

また、現代では統計解析のほとんどがコンピュータソフトウェアで行われる。そのため本書では、手計算による複雑な数式の展開は最小限に抑え、代わりに各統計手法の基本的な考え方や、結果の解釈方法に重点を置いている。これにより、読者は統計学の本質的な理解に集中することができ、実務での活用に直結する知識を効率的に習得することが可能となっている。

本書の構成は以下の通りである。

第 1 章では、医療機器における統計的手法の目的と、ISO 13485 における要求事項について解説している。特に設計・開発の検証、バリデーション、製造およびサービスの提供における統計的手法の適用範囲について詳しく説明している。また、統計手法の基本的な分類や、サンプルサイズに関する一般的な誤解についても触れている。

第 2 章から第 4 章では、統計的手法の基礎となる概念について詳しく解説している。データの構造、計量値と計数値の分布、正規分布、t 分布、二項分布、ポアソン分布などの基本的な確率分布について、実用的な応用例を交えながら説明している。また、第 1 種の過誤（生産者危険）と第 2 種の過誤（消費者危険）の概念、OC 曲線の特性についても詳しく解説している。

第 5 章から第 8 章では、サンプルサイズの決定方法について、その理論的背景と実践的な適用方法を説明している。特に、帰無仮説の概念、有意水準の設定、検出力の重要性について、具体例を用いて分かりやすく解説している。また、医療機器開発における有意水準の重要性や、検出力を高めるための実践的な方法についても詳しく説明している。

第 9 章では、様々な統計的検定や推定におけるサンプルサイズの計算式について解説している。一つの母平均の検定から、二つの母平均の差の検定、母不良率の検定、区間推定など、実務で頻繁に遭遇する場面での具体的な計算方法を示している。

第 10 章では、具体的な例題を通じて、これまでに学んだ内容の実践的な適用方法を解説している。特に、製品・プロセスのリスク評価、計数値試験と計量値試験におけるサンプルサイズ計算、平均値の信頼区間、仮説検定など、実務に直結する事例を多く取り上げている。

本書の特徴は、統計学の理論的な側面だけでなく、実務における適用方法に重点を置いていることである。例えば、品質保証、設計検証、プロセスバリデーションなど、医療機器開発の現場で直面する具体的な課題に対して、どのように統計的手法を活用すべきかを、豊富な事例とともに説明している。

また、本書では統計ソフトウェアの使用を前提としているため、複雑な計算式の暗記や手計算は必要ないことを強調している。代わりに、統計的手法の選択基準、入力するパラメータの意味、得られた結果の解釈方法など、実務で本当に必要となる知識に焦点を当てている。

さらに、医療機器業界特有の規制要件や品質保証の観点から、統計的手法の適用方法について詳しく解説している。ISO 13485 の要求事項との関連性を明確にしながら、コンプライアンスの観点からも適切な統計的手法の選択と適用が可能となるよう配慮している。

本書が、医療機器業界で働く技術者にとって、統計的手法を実践的に活用するための良き指針となることを願っている。統計学の本質を理解し、適切に活用することは、より安全で信頼性の高い医療機器の開発・製造につながる。そのための第一歩として、本書を活用していただければ幸いである。

読者が本書を通じて統計的手法への理解を深め、その実践的な活用方法を習得されることを期待している。

2025 年 1 月

株式会社イーコンプライアンス
村山 浩一

第1章

はじめに

1. 統計的手法の目的

医療機器にとって必要なことは、その意図する用途（Intended use）に基づき、適切な使用法において安全かつ有効に使用されることが保証されなければならない。取扱説明書に従った適切な使用法において、安全性と有効性が確保されることが求められる。そのため、医療機器の開発から製造において、完成機器またはユニット、部品などの要求事項を満足させることが保証できること、および医学的に安全で有効であることを確実にする必要がある。

そうした保証のため、臨床評価、設計開発の検証（ベリフィケーション）、設計開発バリデーション（妥当性確認）、製造工程における検査などの活動が行われる。統計的手法は、こうした製品実現のための活動において、それらの結果が統計的に有意であると主張することにより、より確実な方向に導くことを支援することを目的としている。これらの活動を通じて、製品の品質と安全性を確保し、患者や医療従事者に対して信頼性の高い医療機器を提供することが可能となる。また、こうした統計的手法は、その意図する用途に基づき、適切な使用法において、安全かつ有効に使用されることが保証されなければならないという医療機器に対する基本的要求に応えるものである。

2. ISO 13485 における統計的手法の要求事項

2.1. 設計・開発の検証（7.3.6）

> 7.3.6 設計・開発の検証
>
> 設計・開発からのアウトプットが、設計・開発へのインプットの要求事項を満たしていることを確実にするために、計画し文書化した取り決めに従って設計・開発の検証を実施する。
>
> 組織は、方法、許容基準を含む検証計画を文書化する。これには、適切な場合、サンプルサイズの根拠とした統計的手法を含む。
>
> 意図した使用が、医療機器の他の機器への接続またはインターフェースを要求している場合、検証には、接続またはインターフェースした状態で、設計からのアウトプットが、設計へのインプットに合致していることの確認を含む。
>
> 必要な処置を含む、この検証の結果および結論の記録を維持する（4.2.4、4.2.5 参照）。

設計・開発からのアウトプットが設計・開発へのインプットの要求事項を満たしていることを確実にするために、計画し文書化した取り決めに従って設計・開発の検証を実施する必要がある。組織は検証計画において、まず検証の方法と手順を明確にし、具体的な合否判定基準（許容基準）を設定する。そして適切な場合には、サンプルサイズとその根拠となる統計的手法について詳細な説明を含める。これには、サンプルサイズの決定方法、統計的手法の選定理由、そして具体的な計算過程を含む。

また、記録要件として必要な項目、保管期間、アクセス権限等を規定し、使用する測定機器および試験設備についても、その選定理由や校正要件、保守管理方法を明確にする。環境

条件が結果に影響を与える可能性がある場合には、温度、湿度、清浄度等の具体的な要求事項も規定する。

意図した使用が医療機器の他の機器への接続またはインターフェースを要求している場合、検証には接続またはインターフェースした状態での確認が不可欠となる。この際、設計からのアウトプットが設計へのインプットに合致していることを、具体的な検証項目と判定基準に基づいて確認する。特に、相互接続性、互換性、信号のやり取り、データの整合性などの観点から詳細な検証を行う。これらの検証における結果および結論は、判定結果、判定理由、関連データ、検証実施者、判定者等を含めて文書化し、必要な処置についても明確に記録を残す。さらに、これらの記録は規定された期間にわたって適切に維持管理される必要がある。

2.2. 設計・開発のバリデーション（7.3.7）

7.3.7 設計・開発のバリデーション

結果として得られる製品が、規定した適用または意図した使用への要求事項を満たす能力があることを確実にするために、計画し文書化した取り決めに従って、設計･開発のバリデーションを実施する。

組織は、方法、許容基準を含むバリデーション計画を文書化する。これには、適切な場合、サンプルサイズの根拠とした統計的手法を含む。

設計バリデーションは、製品を代表するもので実施する。製品を代表するものは、初回生産品、バッチ、またはこれらに同等なものを含む。バリデーションに用いる製品の選択の根拠を記録する（4.2.5 参照）。

設計・開発のバリデーションの一部として、組織は、適用される規制要求事項に従い、医療機器の臨床評価または性能評価を実施する。臨床評価または性能評価1こ用いる医療機器は、顧客の使用のための出荷とは見なさない。

設計・開発のバリデーションは、結果として得られる製品が規定した適用または意図した使用への要求事項を満たす能力があることを確実にするために実施する。組織はバリデーション計画において、まずバリデーションの目的を明確にし、その目的に応じた具体的な方法と手順を規定する。合否判定基準については、定量的な基準と定性的な基準の両面から設定し、その根拠を明確にする。

サンプルサイズとその根拠となる統計的手法については、想定される使用条件やリスク、技術的な制約等を考慮して決定する。記録要件としては、バリデーション結果、データ解析、判定結果、実施者、判定者等の情報を含め、製品の選択基準とその根拠も詳細に文書化する。使用する測定機器および試験設備については、要求される精度、再現性、安定性等を考慮して選定し、必要な校正や保守管理の要件も規定する。

環境条件については、実際の使用環境を考慮し、温度、湿度、気圧、照度、電磁環境等の条件を具体的に設定する。設計バリデーションは製品を代表するもので実施する必要があり、

初回生産品、バッチ、またはこれらに同等なものを対象とする。これらの製品の選択においては、製造条件、製造設備、作業者等の変動要因を考慮し、その選択の根拠を詳細に記録する。

また、設計・開発のバリデーションの一部として、組織は適用される規制要求事項に従い、医療機器の臨床評価または性能評価を実施する。この評価では、安全性、有効性、使用性等の観点から包括的な検証を行う。ただし、この臨床評価または性能評価に用いる医療機器は、あくまでも評価目的であり、顧客の使用のための出荷とは見なさない。これは、バリデーション用の製品と市販用の製品を明確に区別し、管理する必要があることを意味する。

2.3. 製造およびサービスの提供（7.5）

7.5　製造およびサービスの提供

組織は、次の事項を含むプロセスのバリデーションの手順を文書化する。

a)　プロセスのレビュおよび承認のために定めた基準

b)　設備の承認および要員の資格認定

c)　固有の方法、手順および判断基準の使用

d)　適切な場合、サンプルサイズの根拠となる統計的手法

e)　記録に関する要求事項（4.2.5 参照）

f)　再バリデーションの判断基準を含む、再バリデーション

g)　プロセスに対する変更の承認

製造およびサービスの提供に関して、組織はプロセスのバリデーションの手順を文書化しなければならない。まず、プロセスのレビューおよび承認のために定めた基準として、プロセスパラメータの設定根拠を明確にする。これには、重要品質特性に影響を与えるパラメータの特定、その管理範囲の設定根拠、そして管理方法の詳細を含む。また、各工程における管理項目と管理値の設定根拠を文書化し、それらのモニタリング方法と頻度についても具体的に規定する。

QMSのステップ	統計的手法が利用できる場面	統計的手法の例
製品設計開発 ・製品仕様の検討 ・設計検証 ・設計バリデーション	・大きさ、重さ、強さ、機能の決定 ・要求事項への適合性評価 ・ばらつきの管理	・検定と推定 ・実験計画
製造工程設計開発 ・製造・試験基準の決定 ・製造・試験バリデーション	・稼働条件の最適化 ・要求事項への適合性評価	
製造 ・製造管理 ・品質管理	・規格への品質適合の確認 ・工程管理	・抜取検査 ・工程能力 ・管理図
市販後管理 ・製品の監視測定	・想定外のイベント監視	・QC七つ道具 ・新QC七つ道具

表 1-1 統計の種類

3. 統計的手法の適用範囲

医療機器の品質マネジメントシステム（QMS）において、統計的手法は製品のライフサイクル全体を通じて幅広く活用される。

製品設計開発段階においては、物理的特性である大きさ、重量、強度などの設計値設定や、機能的特性の性能規格設定、製品寿命の予測などに統計的手法が用いられる。具体的には、検定と推定、実験計画法、相関・回帰分析、信頼性解析などの手法を用いて、設計検証試験のデータ解析や測定値のばらつき評価、信頼性評価などが実施される。

製造工程設計開発段階では、工程パラメータの最適化や許容範囲の設定、検査基準の確立といった製造・試験基準の決定に統計的手法が活用される。また、プロセスバリデーションの計画と評価、稼働条件の最適化、要求事項への適合性評価においても、実験計画法やプロセス能力分析、測定システム解析（MSA）、リスク評価手法などが適用される。

製造段階における統計的手法の活用は、主に製造管理と品質管理の二つの側面で行われる。製造管理においては、工程パラメータの監視や異常検知と是正、工程能力の継続的評価が実施される。品質管理では、規格適合性の確認や品質特性の測定と評価、不適合品の管理などが行われる。これらの活動には、抜取検査、工程能力評価、管理図法、ヒストグラム分析などの手法が用いられる。

市販後管理では、製品の監視測定と想定外のイベント監視が重要な活動となる。市場品質データの収集と分析、顧客満足度の評価、製品性能の長期モニタリングといった製品の監視測定活動が行われる。また、不具合データの傾向分析や苦情データの分析、市場安全性情報の評価などの想定外のイベント監視も実施される。これらの活動では、QC 七つ道具（特性要因図、パレート図、ヒストグラム、散布図、管理図、チェックシート、グラフ）や新 QC 七つ道具（連関図法、親和図法、系統図法、マトリックス図法、アローダイアグラム法、PDPC 法、マトリックスデータ解析法）などの手法が活用される。

これらの統計的手法の適用により、客観的なデータに基づく意思決定やプロセスの安定性

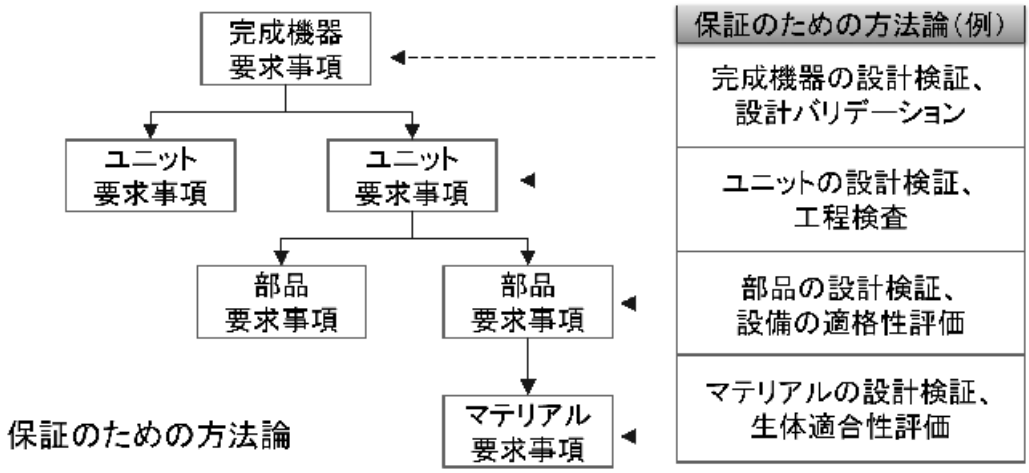

図 1-1 保証のための方法論

と能力の定量的評価が可能となり、品質問題の早期発見と予防、継続的改善活動の推進、リスクマネジメントの強化、製品品質の向上と安定化といった効果が期待できる。

統計的手法の選択と適用に際しては、目的に適した手法の選択、データの収集方法と品質の確保、解析結果の正しい解釈、適切な判断基準の設定、担当者の教育訓練といった点に留意する必要がある。これらの要素を適切に考慮することで、統計的手法は医療機器の品質保証における有効なツールとして機能する。

4. 製品要求事項の保証

医療機器の製品実現プロセスにおいて、製品要求事項が確実に保証されることは不可欠である。この保証を確立するために、検証やバリデーションなどの各段階で統計的手法を用いた有効性の実証を考慮する必要がある。

製品要求事項の保証は、完成機器、ユニット、部品、マテリアルという階層構造に応じて、それぞれのレベルで適切な方法を確立しなければならない。

完成機器レベルでは、設計検証および設計バリデーションが実施される。設計検証では製品仕様への適合性を確認し、設計バリデーションでは意図する用途に対する適合性を評価する。この過程では、性能試験データの統計解析、信頼性評価、臨床データの統計的評価などが行われる。

ユニットレベルでは、ユニットの設計検証や工程検査が行われる。ここでは各ユニットの機能や性能が要求事項を満たしていることを確認するため、機能試験データの解析や工程能力評価などの統計的手法が適用される。

部品レベルでは、部品の設計検証や設備の適格性評価が実施される。個々の部品の寸法精度や強度などの特性評価、製造設備の性能評価において、測定システム解析（MSA）や工程能力指数などの統計的評価が行われる。

マテリアルレベルでは、マテリアルの設計検証や生体適合性評価などが必要となる。原材料の物理的・化学的特性の評価、生体適合性試験データの統計解析、安定性評価などが実施される。

これらの各階層における保証活動では、以下のような統計的手法の活用が重要となる。

要求事項への適合性を評価するための検定や推定、製品特性のばらつきを評価するための工程能力解析、品質の経時的な変化を監視するための管理図法、製品の信頼性を評価するための寿命試験データ解析、生体適合性を確認するための生物学的評価データの統計解析など、目的に応じた適切な統計的手法を選択し活用する必要がある。

さらに、これらの統計的手法の適用においては、適切なサンプリング計画の立案、データの品質確保、解析結果の妥当性評価、判定基準の設定など、統計的な観点からの十分な検討が必要である。また、得られた結果の解釈と、それに基づく意思決定のプロセスも明確に文書化されなければならない。

このように、製品要求事項の保証は、各階層における適切な評価方法の確立と、それらを裏付ける統計的手法の効果的な活用によって達成される。これにより、医療機器の安全性と

有効性に関する客観的な保証が可能となる。

5. 統計手法の基本的な分類

統計的手法は、大きく記述統計と推測統計という二つの基本的な分野に分類される。

5.1. 記述統計

記述統計は、収集したデータを整理・要約し、その特徴を明確に把握するための手法である。具体的には、データを表やグラフなどの視覚的な形式に整理し、平均値、中央値、標準偏差などの統計量を算出することで、データの傾向や特徴を理解する。医療機器の品質管理においては、製造工程の状態把握、製品の品質特性の分布確認、不適合率や苦情率の時系列的な推移の把握、品質問題の原因分析などに活用される。また、工程能力の評価や品質改善活動の効果確認においても重要な役割を果たす。

5.2. 推測統計

推測統計は、母集団の一部であるサンプルのデータから、母集団全体の特性を推測するための手法である。この手法では、サンプルから得られた統計量に基づいて、母集団のパラメータを推定し、その推定結果の信頼性を統計的に評価する。医療機器の品質保証においては、製造工程における抜取検査、設計検証におけるサンプリング試験、プロセスバリデーションなどで活用される。例えば、製品ロットからの抜取検査結果に基づくロット全体の品質評価や、設計検証における試作品の試験結果から量産品の性能を予測する場合などに用いられる。

これらの統計手法は相互に補完的な関係にあり、医療機器の品質マネジメントシステムにおいては、両者を適切に組み合わせて活用することが重要である。記述統計によってデータの特徴を的確に把握し、推測統計によってその結果の一般化や予測の信頼性を確保することで、より効果的な品質保証活動が可能となる。また、これらの統計的手法の適用に際しては、データの性質や目的に応じて適切な手法を選択し、結果の解釈においても統計的な観点からの妥当性を十分に検討する必要がある。

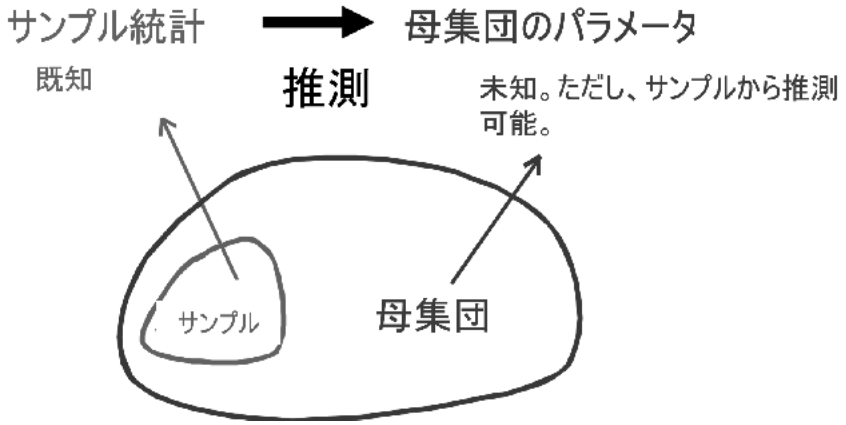

図 1-2 推測統計

5.3. 何を検定したいのか

統計的検定を行う際､まず明確にすべきことは「何を検定したいのか」という目的である。

母集団の分布を知りたい

1) 一つの母平均 / 母不良率の検定を行いたい

例：バッテリーの充電時間を短くしたい

→ 充電時間の平均値の検定 / 推定

例：工程を改良した結果、不良率が減少したかを調べたい

→ 不良率の検定 / 推定

例：臨床試験で開発した医療機器に効果があったかどうかを判定したい

→ 不良率の検定 / 検定

2) 一つの母分散 / 母比率の検定を行いたい

例：成型部品の大きさのばらつきを小さくしたい

→ ばらつきの検定 / 推定

3) 二つの母平均 / 母不良率の差を検定したい

例：二つの機械で製造した部品の大きさが同じかどうかを調べたい

→ 平均値の差の検定

例：開発品と対象品で効果に差があるかどうかを調べたい

→ 不良率の差の検定

4) 二つの母分散の比の検定を行いたい

例：二つの機械で製造した部品の大きさのばらつきが同じといえるかどうかを調べたい → 分散の比の検定

5) 対応がある場合の母平均の差の検定

例：同じ温度ロガーで倉庫の二か所の温度を測定したときに倉庫の場所で差があるかどうかを調べたい → 分散の比の検定

6. サンプルサイズに関する誤解

サンプルサイズの決定に関しては、一般に「大きいサンプルサイズの方が良い」という考え方や、効率性の観点から「なるべく小さいサンプルサイズにしたい」という考え方が見られる。しかしながら、これらの考え方はいずれも適切ではない。サンプルサイズは、統計的な根拠に基づいて「何例」という具体的な数値として決定されなければならない。

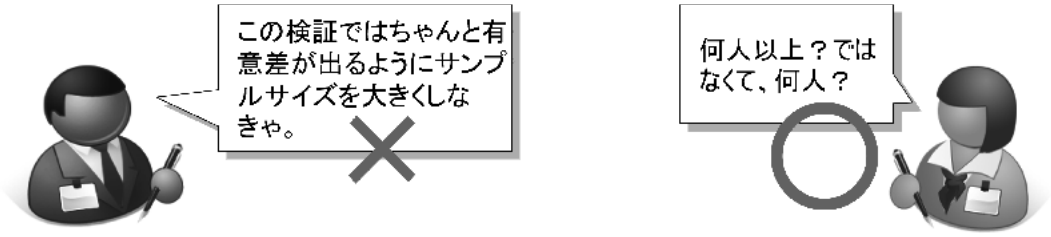

図 1-3 サンプルサイズに関する誤解

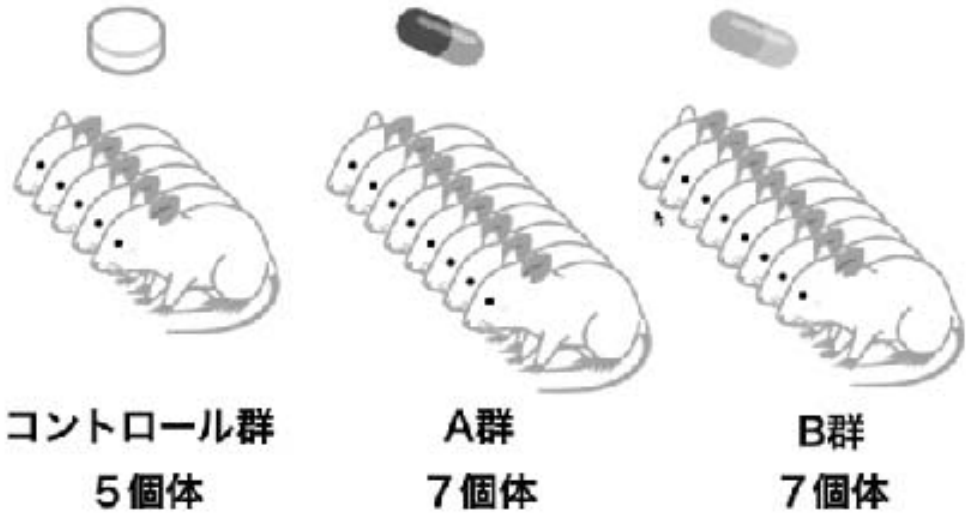

図 1-4 サンプル数とサンプルサイズ　～マウスの投薬実験～

ここで重要なのは、サンプル数（the number of samples）とサンプルサイズ（sample size）という二つの概念の違いを正確に理解することである。サンプル数は標本抽出を実施した回数、すなわち群数や標本数を表す概念である。一方、サンプルサイズは 1 回の標本抽出において調査対象とした個体数、すなわち各標本または各群における個体数（N 数）を表す概念である。

この違いを具体的に説明するため、がん治療薬の効果を評価する実験を例として考えてみる。がんを発症したマウスに対して、コントロール薬（C 群）、新薬 A（A 群）、新薬 B（B 群）を投与し、一定期間後の腫瘍径の母平均の差を比較する実験を想定する。この実験において C 群から 5 個体、A 群から 7 個体、B 群から 7 個体のデータを取得した場合、サンプル数は投与群の数である 3 となる。一方、サンプルサイズは C 群が 5、A 群が 7、B 群が 7 となり、各群で異なる値をとることになる。

このように、サンプル数とサンプルサイズは異なる統計的概念であり、実験計画や統計解析においてはこれらを明確に区別して扱う必要がある。サンプルサイズの決定に際しては、検出したい差の大きさ、望ましい検出力、有意水準、想定される変動などの要因を考慮し、統計的に妥当な数値を設定することが重要である。また、倫理的な観点や実験の実行可能性なども考慮に入れる必要がある。

これらの概念を正しく理解し、適切なサンプルサイズを設定することは、信頼性の高い統計的評価を行う上で極めて重要である。過大なサンプルサイズは資源の無駄遣いとなり、過小なサンプルサイズは信頼性の低い結果をもたらす可能性がある。そのため、目的に応じた適切なサンプルサイズを科学的な根拠に基づいて決定することが求められる。

7. 統計的手法の重要なポイント

統計的手法を適用する際の重要な基本概念として、母集団全体の把握が困難な場合におけ

るサンプリングの必要性を理解しなければならない。最終検査や苦情管理など全数把握が可能な場合とは異なり、大量の部品の受入検査や連続的な製造工程における製品特性の評価など、母集団全体の直接的な測定が現実的でない場合には、適切なサンプリングが不可欠となる。

サンプリングの必要性は、対象となる特性のばらつきの性質によって判断される。均質性が保証された状態、すなわちばらつきが存在しない場合においては、サンプルサイズ 1（N=1）で十分な場合がある。例えば、十分に撹拌された味噌汁のように、測定箇所に関わらず同一の特性を示す状態では、単一のサンプルが母集団を代表する値として扱うことが可能である。これは、一昼夜放置後に上澄み、中層、底層で味が異なってくるような不均質な状態とは本質的に異なる状況である。

サンプリングは確率論を基盤とする手法であり、この特性を深く理解することが重要である。統計的な解析において、サンプルサイズはその結果に重要な影響を与える。一般的に、サンプルサイズが小さい場合には統計的有意差の検出が困難となり、サンプルサイズが大きい場合には有意差が検出されやすく（P 値が小さくなる）なる傾向がある。同様に、群間における差（効果量）が大きい場合にも統計的有意差が検出されやすく（P 値が小さくなる）なる性質がある。

これらの特性を踏まえ、統計的手法の適用においては、以下の点に注意を払う必要がある。

適切なサンプリング方法の選択、サンプルサイズの科学的な設定、データの品質確保、解析結果の適切な解釈、そして結果に基づく意思決定の妥当性評価である。また、統計的有意差の検出は、実用的な意味での重要性とは必ずしも一致しないことにも留意が必要である。

医療機器の品質保証において、これらの統計的手法の特性を正しく理解し、適切に活用することは、信頼性の高い品質評価と意思決定を行う上で極めて重要である。特に、サンプリングを伴う評価においては、その統計的な性質を十分に考慮した上で、適切な計画立案と実行が求められる。

8. サンプルサイズ決定の基本原則

サンプルサイズの決定には、科学的な根拠に基づく四つの重要な要素の考慮が必要である。これらの要素は相互に関連しており、適切なサンプルサイズの決定に不可欠な役割を果たす。

1) まず第一の要素は、群間の効果の差（効果量）である。これは検出したい差の大きさを表し、臨床的または実用的な意味を持つ差として設定される。
2) 第二の要素は、ばらつき（サンプルの標準偏差）である。これは測定値や観察値の変動の程度を表す指標であり、予備実験データや過去の類似研究から推定される。
3) 第三の要素は、有意水準（P 値、信頼度、α エラーを防ぐ力）である。これは第一種の過誤（偽陽性）を制御するための基準値であり、実際には差がないのに差があると判断してしまう確率の許容限界を示す。
4) 第四の要素は、検出力（検定力、パワー、β エラーを防ぐ力）である。これは第二種

の過誤（偽陰性）を制御するための指標であり、実際に存在する差を検出できる確率を表す。

これら四つの要素が具体的な数値として定められれば、必要なサンプルサイズは数学的に一意に決定される。実務において一般的に用いられる基準値として、有意水準 α は 0.05（5%）、検出力（1- β）は 80% が広く採用されている。これらの値は長年の研究実績と実践的な経験に基づいて設定された標準的な基準である。

サンプルサイズの具体的な計算には、統計ソフトウェアや Excel などの計算ツールを利用することが推奨される。複雑な計算式を暗記する必要はなく、信頼性の高いソフトウェアツールを適切に活用することで、効率的かつ正確なサンプルサイズの決定が可能となる。

なお、これらの要素を考慮したサンプルサイズの決定は、研究や試験の計画段階で行われるべきであり、データ収集後に遡って決定することは適切ではない。また、実験や試験の実行可能性、倫理的な配慮、コストなども考慮に入れながら、総合的な判断のもとでサンプルサイズを決定することが重要である。

9. まとめ

本章では、医療機器の品質保証システムにおける統計的手法の重要性と、その中でも特にサンプルサイズ決定に関する基本的な考え方について解説した。

医療機器の安全性と有効性を確保するためには、客観的なデータに基づく評価と意思決定が不可欠である。統計的手法は、この目的を達成するための重要なツールとして位置づけられており、ISO 13485 においても、設計開発の検証、バリデーション、製造工程の管理など、様々な場面での適用が明確に要求されている。

製品要求事項の保証においては、完成機器から個々のマテリアルに至るまで、各階層に応じた適切な評価方法を確立する必要がある。この過程で統計的手法は、データの客観的な分析と評価を可能とし、品質保証活動の信頼性を高める重要な役割を果たす。

特にサンプルサイズの決定は、統計的評価の信頼性を左右する重要な要素である。単純に「大きければ良い」という考え方や、効率性のみを重視して「小さければ良い」とする考え方は適切ではない。サンプルサイズは、効果量、ばらつき、有意水準、検出力という四つの要素を科学的に考慮し、具体的な数値として決定されなければならない。

これらの基本的な概念を正しく理解し、適切に適用することは、医療機器の品質保証において極めて重要である。次章以降では、これらの概念についてより詳細な解説を行い、実務における具体的な適用方法について説明していく。その過程で、統計的手法の実践的な活用法と、それによって得られる品質保証上の利点について、さらに深く理解を深めることができる。

第2章

統計的手法の基礎

1. データの構造

データは平均値（μ）とばらつき（σ）という二つの基本要素から構成される。この関係は数式で表現すると、データ = 平均値 (μ) + ばらつき (σ) として表される。この構造は実際の測定データにおいて普遍的に観察される特徴である。

実際の測定値が理論値や目標値から異なる値を示す現象は、様々なばらつきの要因が複合的に作用した結果として理解できる。例えば、A 薬品を 100.0 グラムとして秤量した場合、実際の測定結果には平均値に加えて、秤量操作そのもののばらつき、使用する天秤のばらつき、その他の環境要因によるばらつきが含まれることになる。

さらに複雑な例として､A 薬品を純水に溶解して C 濃度の溶液を調製する場合を考えると、最終的な濃度には以下の複数のばらつきが影響する。

A 薬品量に関連するばらつき、純水量の測定に関するばらつき、溶液濃度の計測に伴うばらつき、その他の環境要因によるばらつきが合算された結果として、実測値が得られることになる。

ばらつきは、その性質から大きく二つに分類することができる。一つは特定の要因に起因する系統的なばらつきであり、もう一つは偶然的な誤差によるばらつきである。要因として考慮すべき主な項目としては、作業者などの人的要因、測定装置の精度、測定器の特性、温度変化の影響、圧力変化の影響、使用する原材料のロット間差、その他の環境要因（気圧、室温、保管条件など）が挙げられる。

これらのばらつきの評価においては、その大きさと影響度を適切に分析し、管理すべき要因として扱うべきか、あるいは許容可能な誤差として扱うべきかを慎重に検討する必要がある。この判断は、製品の品質特性に対する影響度、管理の実行可能性、コストなどを総合的に考慮して行われる。

特に医療機器の品質管理においては、これらのばらつきが製品の安全性や有効性に与える影響を適切に評価し、必要な管理方法を確立することが重要である。また、各要因の寄与度を定量的に評価し、重要度に応じた管理戦略を構築することで、効率的かつ効果的な品質保証システムを実現することができる。

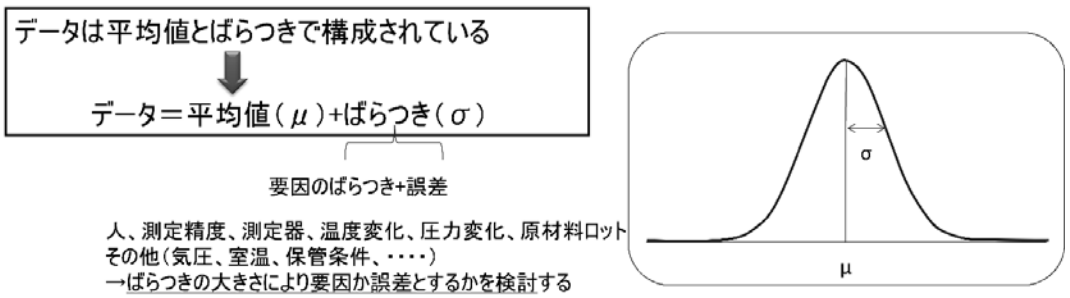

図 2-1 データの構造

執 筆 者

村山 浩一（むらやま こういち）
株式会社イーコンプライアンス 代表取締役

【経歴】

1986 年 4 月　日本ディジタルイクイップメント株式会社（日本 DEC）　ソフトウェアサービス部　入社

- GCP 管理システム、症例データ管理システム企画・開発担当
（現 ClinicalWorks/GCP/CDM）
- 改正 GCP（J-GCP）に対応した標準業務手順書作成コンサルティング
- 製薬業界におけるドキュメント管理システム導入コンサルティング

1999 年 1 月　日本ディジタルイクイップメント株式会社　退社

1999 年 2 月　日本アイ・ビー・エム株式会社　コンサルティング事業部　入社

- NY の TWG（The Wilkerson Group）で製薬業界に特化したコンサルタントとして研修
- 製薬企業におけるプロセス　リエンジニアリング担当
- Computerized System Validation(CSV)、21 CFR Part11 コンサルティング

2001 年 7 月　IBM 認定主幹コンサルタント

- アイビーエム・ビジネスコンサルティングサービス株式会社へ出向
マネージング・コンサルタント

2004 年 7 月　日本アイ・ビー・エム株式会社　退社

2004 年 8 月　株式会社イーコンプライアンス　設立

現在に至る。

【活動】

医薬品業界・医療機器業界を担当し 30 年以上のキャリアをもつ。

医薬品企業･医療機器企業における､コンピュータ化システムの品質保証（CSV､Part11 対応）をはじめ、リスクマネジメント、CAPA（是正処置および予防処置）、QMS 構築支援、FDA 査察対応等のコンサルテーションなどを幅広く展開している。

サイエンス＆テクノロジー株式会社におけるセミナー開催多数。

【主な著書】

- 実践ベンダーオーディット実施の手引き（2008 年）イーコンプライアンス刊
- コンピューターシステムバリデーション・厚労省 ER/ES 指針・21 CFR Part 11「社内監査の手引き」（2009 年）イーコンプライアンス刊
- GAMP5,Annex11, 厚労省 CSV 指針を基礎から解説【超入門シリーズ 1】コンピュータバリデーション（2009 年）イーコンプライアンス刊
- 【厚労省新ガイドライン対応シリーズ】医薬品・医薬部外品製造販売業者における「コンピュータ化システム適正管理ガイドライン」対応準備の手引き（2010 年）イーコンプライアンス刊

著 者 紹 介

- GAMP 5, FDA, ANNEX 11 に対応した【改定版】コンピュータバリデーション実施の手引き（2011 年）イーコンプライアンス刊
- 対応のためのガイドラインサンプル付【改定版】厚労省 ER/ES 指針対応実施の手引き（2011 年）イーコンプライアンス刊
- 【要点をわかりやすく学ぶ】製薬･医療機器企業におけるリスクマネジメント（2015 年）サイエンス＆テクノロジー刊
- 【要点をわかりやすく学ぶ】PIC/S GMP Annex15 適格性評価とバリデーション（2015 年）サイエンス＆テクノロジー刊
- 医療機器設計管理入門（2020 年）イーコンプライアンス刊
- 当局要求をふまえた データインテグリティ手順書作成の要点（2020 年）サイエンス＆テクノロジー刊
- 【改正 GMP 省令対応シリーズ 2】改正 GMP 省令で要求される『医薬品品質システム』と継続的改善（2021 年）サイエンス＆テクノロジー刊
- 【改正 GMP 省令対応シリーズ 3】改正 GMP 省令で要求される『CAPA(是正措置・予防措置)』導入・運用手順（2023 年）サイエンス＆テクノロジー刊

【数式を使わない】医療機器における統計的手法と
サンプルサイズ決定方法

2025年1月31日　第1版　第1刷発行

定価：44,000円（税込）

著　者　村山 浩一
発行人　村山 浩一
発行所　株式会社イーコンプレス
〒630-0244 奈良県生駒市東松ヶ丘1-2　奥田第一ビル102
TEL　050-3733-8134　FAX　03-6745-8626
http://eCompress.co.jp

印刷・製本　株式会社マツモト