

# 新版 実用 SAS 生物統計ハンドブック

---

---

《SAS®9.4 および R 3.2.0 対応》

監修：東京理科大学 教授 浜田知久馬

執筆：臨床評価研究会 基礎解析分科会

協力：SAS Institute Japan 株式会社



# 監修の言葉

監修させていただいた浜田は SUGIJ のチュートリアルをもう 25 年も続けている。学会で通常、同一人物が 25 年もチュートリアルを行うのは不可能である。それだけの新しいネタがないためである。ところが私は、新鮮なネタ選びにさして苦勞するわけでもなく、チュートリアルを 20 年以上、楽しんで続けることができている。この理由は、SAS が常に進化し続け、新しい統計手法を提供し続けているからである。この間、SPSS、S-PLUS、JMP、R 等の新しい魅力的な統計のソフトウェアが台頭してきたが、私は SAS 一筋である。たいへん高度で強力ではあるが、気紛れで使いこなすことが難しい SAS について、その魅力を一番理解してあげて、人に伝えることができるのは自分であると自負しているからである。

私なりに広報活動をしてきたつもりではあるが、SAS の素晴らしい魅力を多くの人に正確に伝え、使ってもらうにはやはり教科書が必要である。SAS はどんどんバージョンアップするので、タイムリーに教科書を作成するのはなかなか困難である。そんな中、本書は、臨床評価研究会・基礎解析分科会が、臨床試験でよく用いられる統計手法を解説し、SAS でのプログラミングコード、結果の解釈について解説したハンドブックである。製薬企業の第一線の現場で活躍している実務者の目で統計手法を解説しているので、統計学者が書いた本よりわかりやすいのではないかと思う。このハンドブックを利用すれば、多くの生物統計学の手法を SAS で適切に実行することができる。SAS Ver. 9.4 が中心であるが、SAS Ver. 9.3 でも問題ないことが確認されており、まだ当分はこのハンドブックの利用価値は高い。私がこれまで SUGIJ で発表してきたいくつかの SAS マクロについても、このハンドブックに掲載されているのは監修者としてうれしい限りである。

旧版が 2005 年に出たから、すでに 12 年がたち、その間、SAS も様々な進化を遂げた。その中で新版では、様々な SAS の新しい機能の解説を加えている。

---

---

第 1 章の解析事例では、特に経時測定データにおける一般化推定方程式(GEE)、一般化線形混合モデルによる解析として、MIXED、GENMOD、GLIMMIX プロシジャ、多重比較等の大幅な機能拡張がなされた生存時間解析の LIFETEST、PHREG プロシジャ、欠測データの多重補完と解析を行う MI と MIANALYZE プロシジャ等の SAS の最新の機能拡張の利用方法を示している。

また第 2 章の例数設計では、旧版では SAS マクロを示していたが、POWER プロシジャが追加になり、マクロに対応する形式で POWER プロシジャの広範な利用方法を示しており、マクロの内容を理解することで、読者は POWER プロシジャの理解が深まるのではないかと考えられる。

新版の大きな特徴は、R でのプログラム例を新たに追加したことである。SAS と R では、企業と研究者というユーザーの棲み分けがなされているが、本書が、この 2 つのユーザーの架け橋になるのではないかとと思われる。

このハンドブックが SAS の魅力を多くの読者に知ってもらうのに役立ち、SAS の最新の機能を紹介続けるために、時代の要請に応じて再改訂されることを期待してやまない。

平成 28 年 9 月

東京理科大学 教授  
浜田知久馬

# 臨床評価研究会代表より

本書を執筆した臨床評価研究会 (The Association of Clinical Evaluation, 略称 ACE) は、広く医薬品・医療機器の開発に従事する生物統計担当者、データマネジメント担当者、SAS プログラミング担当者、および製造販売後調査担当者らが集う研究会で、日常業務の問題点を深く議論し、解決策を生み出し、かつ実行することを目的としており、現在、約 40 社が参加している。研究会活動は、毎年一回開催する定期研究会および業務分野ごとに課題を設定し活動する分科会から成るが、軸足を分科会活動に置いている。基礎解析分科会はその分科会の一つで、臨床試験や製造販売後調査データの解析に関する諸問題を取り扱う分科会であり、平成 16 年、過去 3 年間の活動成果を集約し、初版ハンドブックを上梓した。その後も、分科会活動は継続しており、今般、この間の活動成果をもとに新版として書き改めた。それが本書である。

医薬品開発における統計学に関する書物は、あまたある。学会やセミナーでも統計的側面を頻繁に取り上げる。つまり、知識を獲得できる機会は多い。しかしながら、知識だけでは日常業務の中で遭遇する現実問題のすべてを解決できない。知識を、成功と失敗に基づく「知恵」に昇華することが必要である。医薬品・医療機器業界にあつては、SAS は統計解析の強力な武器であることから、現場で SAS を利用する際の「知恵」をまとめよう、というのが旧版誕生の舞台裏であり、この情熱が長きにわたり受け継がれ、新版としてパワーアップした。

本書が医薬品および医療機器の開発に従事する方、特に生物統計家および SAS プログラマーの皆さんの一助となり、文字通り座右に備えるハンドブックとして擦り切れるほどに利用して戴ければ、幸いである。またご批評、ご批判をお聞かせ戴ければ、これに勝る喜びはない。

最後に、これまで、臨床評価研究会の活動を暖かく見守り、しかし慧眼をもってご指導、ご助言を戴いている多くのご関係の皆様へ深く感謝申し上げたい。特に、当研究会の前身である医療情報解析研究会を継ぎ、平成 10 年 11 月に当研究会を代表として発足させ、研究会活動の継続に現在もご尽力を頂いている佐野修一顧問へ、深く感謝申し上げます。新版の出版は、佐野顧問が、常に進化

---

---

を考え活動成果を世に問え、と我々を鼓舞し続けて頂いたお蔭である。また、初版ハンドブック巻末にお名前のある旧版執筆者およびその関連の皆様へ、深く感謝申し上げます。皆様のご功績がなければ、本書が誕生することはなかった。

上梓の幸運は、監修の労をおとり頂いた東京理科大学教授 浜田知久馬先生、出版元をお引き受け頂いた株式会社サイエンティスト社との出会いの賜物である。浜田先生の監修がなければ統計学に連なる書物として上梓する価値を生まざ、サイエンティスト社のご支援がなければ「知恵」を継承することはできなかった。浜田先生およびサイエンティスト社に、深く感謝申し上げます。

平成 28 年 12 月

臨床評価研究会代表  
株式会社アスクレップ  
鈴木淳一

# まえがき

新版ハンドブックは、製薬企業や CRO において臨床試験の統計解析業務に携わっている統計解析の実務者およびプログラマーとなる統計解析担当者が業務で直面する問題を容易に解決できるように手引書としてまとめた初版ハンドブックを改訂したものである。

統計解析担当者として業務にあたる場合、統計的な手法とプログラミング技法に関する知識が必要となる。統計的な手法およびプログラミング技法を個々に解説する書は多いが、これら 2 つの側面を同時に解説する書は少ない。さらに、パラメータを入力するだけで必要な解析結果が計算されるソフトウェアも多く存在するが、計算過程を理解せずにこれらのソフトウェアを使用することは、解析結果の解釈を誤ってしまう危険性を含んでいる。

初版と同様にこれらの背景を考慮し、単なるプログラムのサンプルではなく、統計的な手法、さらには結果の考察も含めて、新版ハンドブックには掲載した。最大のポイントは、初版ハンドブックの出版以降に臨床試験の統計解析で使用されるようになった比較的新しい統計手法を追加したことである。

新版ハンドブックでは、統計解析ソフトウェアとして、SAS (SAS Institute Japan 株式会社) の SAS Ver. 9.4 とフリーソフトウェアの R 3.2.0 を使用したことを留意いただきたい。これは、日本における臨床試験や臨床研究の現場では、統計解析ソフトウェアとして SAS が最も多く使用されているためであり、近年は R の使用も増えたためである。

以下に、新版ハンドブックの構成を説明する。

第 1 章に、SAS および R を用いた個々の解析事例として、臨床試験の目的ごとに具体的事例を用いて、SAS および R による解析方法を掲載した。臨床試験の現場で統計解析担当者が、現実として直面することはデータ解析である。やや高度なものも含まれているが、臨床試験でたびたび経験するものなので、有効に活用していただきたい。

第 2 章に、SAS および R を用いた個々の例数設計事例として、第 1 章同様に臨床試験の目的ごとに具体的事例を用いて例数設計実施方法を掲載した。臨

---

---

床試験の計画時には、その試験に必要な例数を設定する。これは、第1章で各種記した臨床試験の主要目的に応じ、統計的な観点から実施することであり、統計解析担当者の業務であることが多い。乱数を利用したシミュレーションによる例数設計の例題もいくつか含めているので、参考にしていただきたい。

新版ハンドブックに掲載されている SAS のプログラムは、前述にもあるように、すべて新版ハンドブック作成当初の最新版である SAS Ver. 9.4 にて作成されている。

第1章および第2章では、プログラムに関して、紙面の都合上、その一部を抜粋した形で掲載されている部分があるが、プログラムのソースコードは臨床評価研究会（ACE）ホームページ（刊行物のご案内：<http://ace-jp.org/book/favor.html>）からダウンロード可能である。

記述の統一性やプログラムの作成形式は、可能な限り例題間で整合性をとるように努めた。特にプログラムに関しては誤りのないように十分な確認を行ったが、万一誤りを発見された方は、臨床評価研究会の基礎解析分科会へご連絡いただければ幸いである。

最後に、新版ハンドブック作成にあたっては、監修者である浜田先生をはじめ、巻末の執筆者リストに掲載された人以外にも多くの人の協力があった。ここに、新版ハンドブック作成を助けていただいたすべての人に感謝申し上げます。また、初版ハンドブックを作成いただいた当時の執筆者の方々にも合わせて感謝を申し上げます。新版ハンドブックが、製薬企業や CRO での臨床試験に関しては勿論のこと、各方面での日頃の統計解析業務に役立つことを願う。

平成 28 年 12 月

臨床評価研究会  
基礎解析分科会一同

# R の記載について

本書では一部の解析手法について SAS のみではなく、統計解析ソフト R による解析プログラムとその解析結果を併記した。解析手法によって、R の解析結果は SAS と同様の結果を示すものもあれば、SAS と異なる結果を示すものもある。異なる主な原因は、R のデフォルトの設定では内部の計算式等が SAS とは異なるためである。SAS と R との解析結果の相違は、可能な限り本書の【R の説明】に記載した。

SAS および R の解析プログラムの併記によって、本書が SAS と R との橋渡しの一助となり、両ソフトウェアのユーザーが増えれば幸いである。

# R の実行環境について

併記したすべての R の解析結果は、バージョン 3.2.0 により解析した結果である。また、本書に記載したいくつかの関数は、実行する際に追加でパッケージを読み込む必要がある。本書で掲載したパッケージは次ページのとおりである。

`install.packages("パッケージ名")`および `library(パッケージ名)`と入力することで指定したパッケージ内の関数を使用することができる。また、`sessionInfo()`と入力すると、使用している R のバージョンやインストールされているパッケージのバージョンが表示され、`packageVersion("パッケージ名")`と入力すると、インストールされている個々のパッケージのバージョンを確認することも可能である。なお、R のプログラムを作成する際、SAS とは異なり大文字と小文字を区別することに注意して頂きたい。

関数等の個別の動作については、R は `help(関数名)`と入力することでオンラインヘルプを閲覧できるため、そちらをご確認いただきたい。

パッケージ	バージョン	章節
pairwiseCI	v0.1-25	1章 1節
gee	v4.13-19	1章 2節
nlme	v3.1-128	1章 3節
lme4	v1.1-12	1章 3節
binom	v1.1-1	1章 5節
MASS	v7.3-45	1章 6節
multcomp	v1.4-6	1章 7節
kSamples	v1.2-4	1章 7節
gmodel	v2.16.2	1章 8節
clinfun	v1.0.12	1章 8節
survival	v2.38-3	1章 9節
Mi	v1.0	1章 10節
lmerTest	v2.0-32	1章 11節
pwr	v1.2-0	2章 1節、2節
EnvStats	v2.1-1	2章 1節
sampleize	v0.2-2	2章 2節

## R のインストール方法について

R のインストール方法については本書には掲載していないが、別途ドキュメントを作成し、SAS や R のプログラムがアップロードされている臨床評価研究会のホームページに合わせて掲載した。必要に応じてご確認いただきたい。

# 目 次

監修の言葉 .....	i
臨床評価研究会代表より .....	iii
まえがき .....	v
R の記載について .....	vii
ダウンロードファイルの案内 .....	ix
第 1 章 SAS/R を用いた個々の解析事例.....	1
第 1 節 データの要約 .....	3
I. 1 変量のデータの要約方法は? .....	3
II. 2 群間の位置母数の差を推定する。(Hodges-Lehmann 推定) .....	10
[参考] データは正規分布? (適合度検定) .....	15
第 2 節 2 群の平均値の比較 .....	20
I. 対応のない 2 群の平均の差の検定として $t$ 検定を行う。 (Student の検定、Aspin-Welch の検定) .....	20
II. 対応のない 2 群の平均の差の検定として Wilcoxon の 順位和検定を行う。 .....	26
III. 対応のある 2 群の平均の比較を行う。(対応のある $t$ 検定、符号検定 および Wilcoxon の符号付き順位和検定) .....	31
IV. 経時測定データにおける群間の違いを一般化推定方程式 (Generalized Estimating Equations, GEE) で推定する。 .....	36
第 3 節 多群の平均値の比較 .....	43
I. 一因子について多群間での違いを検討する。(一元配置分散分析) .....	43
II. 多群間での平均値の違いを検討する。(Kruskal-Wallis 検定) .....	48
III. 交絡因子を調整して一因子について多群間での違いを検討する。 (共分散分析) .....	52

IV. 経時測定データについて多群間での違いを検討する。.....	65
V. 経時測定されたカウントデータの多群間での違いを検討する。 (一般化線形混合モデル) .....	78
 第4節    相関と回帰.....	87
I. 2変数間の相関を調べる。(Pearsonの積率相関係数、 Spearmanの順位相関係数) .....	87
II. 回帰分析を行う。.....	92
III. 母相関係数の区間推定を行う。.....	97
IV. 薬物動態パラメータを非線形回帰モデルにより推定する。.....	100
V. LD <sub>50</sub> を推定する。.....	108
 第5節    2値データの比較.....	115
I. 2×2分割表での出現比率の比較検定を行う。 ( $\chi^2$ 検定、Fisherの正確な検定) .....	115
II. 1標本の正確な出現率の信頼限界を求める。.....	121
III. 2群の出現率の差の信頼限界を求める。.....	125
IV. 対応のある2つの出現率の比較を行う。(McNemar検定) .....	130
V. 信頼区間算出方法の拡張：1標本の割合の信頼区間。.....	135
VI. 信頼区間算出方法の拡張：2群の割合の差の信頼区間。.....	140
 第6節    カテゴリカルデータの比較.....	146
I. 2×c分割表が得られた場合の群間比較を行う。 (Wilcoxonの順位和検定) .....	146
II. r×c分割表の一般連関性を求める。 ( $\chi^2$ 検定、Fisherの正確な検定) .....	150
III. r×c分割表の線形連関性を検討する。.....	155
IV. ブロックのある多群の母平均をノンパラメトリックな手法で 比較する。(Friedman検定) .....	160
V. 2値で得られる反応変数に対して、共変量(およびその交互作用)を 考慮した(model-basedな解析方法で)薬剤の効果を検討する。.....	164

VI. 順序カテゴリカルデータとして得られる反応変数に対して、共変量を考慮した (model-based な解析方法で) 薬剤の効果を検討する。.....	171
第7節 多重比較 .....	178
I. 多群データについて、Dunnett の多重比較を行う。 .....	178
II. 多群データについて、Tukey の多重比較を行う。 .....	184
III. 多群データについて、Williams の多重比較を行う。 .....	192
IV. ノンパラメトリック版の Dunnett の多重比較を行う。 (Steel の検定) .....	198
V. ノンパラメトリック版の Tukey の多重比較を行う。 (Steel-Dwass の検定) .....	203
VI. ノンパラメトリック版の Williams の多重比較を行う。 (Shirley-Williams 検定) .....	208
第8節 傾向性の検定 .....	215
I. 用量-反応関係 (dose-response relationship) の検出を行う。 (回帰分析) .....	215
II. ノンパラメトリックに用量-反応関係の検出を行う。 (Jonckheere - Terpstra の検定) .....	221
III. 用量-反応関係の検出を行う。(Cochran-Armitage 検定) .....	224
IV. 用量-反応関係の検出を行う。(対比を用いた検討) .....	230
V. 背景因子で層別された k 個の層があるとき、層を併合して 用量反応関係があることを検討する。(Mantel-extension 法) .....	239
第9節 生存時間解析 .....	248
I. 生存時間 (生存率) を Kaplan-Meier 法でノンパラメトリックに 推定する。 .....	248
II. 生存時間データでの 2 群の比較を行う。 .....	260
III. 生存時間データで多群の比較を行う。 .....	266
IV. 性別、年齢、治療等が生存時間へ影響を与えているかどうかを Cox 回帰により検討する。(因子により生存率に差があるかの検討) .....	271

V. 共変量と多重性を同時調整する。.....	279
VI. 生存関数の異なる層を含むデータを生存時間解析する。 (層別生存時間解析) .....	284
第10節 欠測のあるデータの解析.....	293
I. 連続量に対する多重補完法 .....	293
II. 欠測メカニズムに対する感度分析.....	305
第11節 同等性の検定.....	312
I. クロスオーバーデザインの解析を行う。(生物学的同等性試験) .....	312
第2章 個々の例数設計事例.....	321
第0節 例数設計 .....	323
第1節 1群の試験.....	325
I. 母比率の検定における例数を求めるには? .....	325
II. 母比率を推定するための例数を求めるには? .....	331
第2節 2群の比較試験.....	334
I. 計量値の優越性試験の例数を求めるには? .....	334
II. 計量値の非劣性試験の例数を求めるには? .....	341
III. 比率の優越性試験の例数を求めるには? .....	346
IV. 順序尺度データの比率の優越性試験の例数を求めるには? .....	350
V. 比率の非劣性および同等性試験の例数を求めるには? .....	355
第3節 用量反応試験 .....	365
I. 計量値の用量反応試験の例数を求めるには? (対比による検定) .....	365
II. 2値データ(比率)の用量反応試験の例数を求めるには? .....	371
III. 2値データ(比率)の用量反応試験の例数を求めるには? (シミュレーション、検出力曲線) .....	376

第4節	生物学的同等性試験 .....	382
I.	生物学的同等性試験において例数を求めるには? .....	382
第5節	生存時間解析 .....	388
I.	生存時間解析において例数を求めるには? .....	388
第6節	医薬品の製造販売後調査の症例数設計 .....	395
I.	医薬品の製造販売後調査において例数を求めるには? .....	395
II.	医薬品の製造販売後調査において例数を求めるには? (2例以上検出) .....	399
索引	.....	405
執筆者リスト	.....	407



---

---

# 第 1 章

## SAS/Rを用いた個々の解析事例

☆本章の構成は①具体的事例②統計手法の説明③データセットの構造④SAS プログラム⑤SAS アウトプット⑥結果⑦考察／補足／備考⑧R プログラム⑨R アウトプット⑩参考文献となっている。

- ① 具体的事例—臨床試験データを中心に、解析する目的を理解しやすい記載とした。他書を引用した場合は引用先を明示した。紙面の都合上、データは省略記載の場合がある。
- ② 統計手法の説明—使用した統計手法に関して概説した。また、SAS 独自の検定統計量を用いている場合もあるため、検定統計量を明示した。
- ③ データセットの構造—プロシジャを実行する際のデータ構造を示した。
- ④ SAS プログラム—SAS のプロシジャには多くのオプションが用意されており、設定によって出力結果が大きく異なるため、使用した理由等について解説を加えた。また、解析の出力結果を編集できるように、できる限りデータセットに出力する形式をとった。
- ⑤ SAS アウトプット—④に記載されている SAS のプログラムを実行した結果を表示した（必要に応じ一部抜粋を用いている）。
- ⑥ 結果—事例を解析した結果を記載した。p 値の表示桁数については、原則として SAS の出力桁数をそのまま用いた。そのため、適切な表示桁になっていない箇所があるのでご注意願いたい。
- ⑦ 考察／補足／備考—結果の解釈の留意点等について解説した。また、他のプロシジャでも同様な解析が可能な場合は SAS コードを例示し、簡単な解説を加えた。
- ⑧ R プログラム—すでに存在するパッケージに入った関数を用いて、SAS と同様の解析を行ったプログラムを掲載した。
- ⑨ R アウトプット—⑧に記載されている R のプログラムを実行した結果を表示した（必要に応じ一部抜粋を用いている）。
- ⑩ 参考文献—本題の記述の際に引用・参照した文献・書籍を記載した。



## 第1節 データの要約

### I. 1変量のデータの要約方法は？

#### 【具体的事例】

被験者集団における薬剤投与前の臨床検査値（ここでは赤血球数を事例として用いている）に対して、基準値からの逸脱、医学的異常値がなかったか吟味したい。

#### 【統計手法（要約方法）の説明】

臨床試験におけるデータの要約は次のような観点から実施されている。

- 1) 得られた全データに対して、計画時に設計した変数の属性からの逸脱がないか、欠側の有無、度数について把握する。
- 2) 被験者特性やベースラインに対して、全体あるいは意味のある群ごとに基本統計量を算出し、標本の特性値を概観する。
- 3) 図表示、分布型の指標を用い、正規性の有無および外れ値の存在を確認する。
- 4) 計量値に対して散布図を作成し、また、相関係数を計算して、変数間の関連の強さを把握する。
- 5) 臨床検査値に対して、基準値からの逸脱、医学的異常値を検出しその取扱いを事前に吟味する。

データの要約方法には、図表示や代表値表現などがある。図表示はデータ全体の傾向を視覚的に示すことができ、データを直感的に理解しやすいというメリットがある。図表示には、ヒストグラム (histogram)、幹葉図 (stem and leaf plot) あるいは箱ひげ図 (box-whisker plot) などがある。代表値表現では、データ全体を数個の数値で表現することになるが、代表値として何を用いるかはその目的やデータの性質に依存する。代表値には、位置の指標として平均値 (mean) や中央値 (median)、ばらつきの指標として範囲 (range)、標準偏差 (standard deviation)、四分位範囲 (interquartile range)、形の指標として歪度 (skewness) や尖度 (kurtosis) などがある。

ここでは、図表示の具体的方法、代表値の具体的な算出方法についての説明は省略する。

### 【データセットの構造】

OBS	NO	RBC
1	1	392
2	2	585
3	3	410
4	4	414
5	5	384
6	6	423
:	:	:

### 【SASプログラム】

SAS では、UNIVARIATE プロシジャを用いてデータの要約を行うことができる。

```
proc univariate data=RBC plot ①normal ②pctldef=5;
  id NO;
  var RBC;
  ③output out=NEW n=N mean=MEAN std=STD median=MEDIAN qrange=QRANGE;
run;
```

①オプションで normal を指定することにより正規性の検定が可能。

Shapiro-Wilk 検定（標本数 2,000 以下の場合のみ）、Kolmogorov-Smirnov 検定、Anderson-Darling 検定および Cramer-von Mises 検定の 4 種の検定結果が出力される。

②pctldef オプションにより、パーセント点の計算アルゴリズムの指定が可能（デフォルトでは pctldef=5）。

③UNIVARIATE プロシジャでは OUTPUT ステートメントを指定することで統計量を SAS データセットに出力することができる。ここでは、例数、平均値、中央値、標準偏差および四分位範囲を出力している。なお、その出力可能な統計量については参考文献等を参照。

### 【SASアウトプット】

```
① UNIVARIATE プロシジャ
   変数 : RBC
```

## モーメント

標本数	50	重み変数の合計	50
平均値	440.86	合計	22043
標準偏差	40.0576167	分散	1604.61265
歪度	1.3351574	尖度	2.93122657
無修正平方和	9796503	修正平方和	78626.02
変動係数	9.08624431	平均の標準誤差	5.66500248

## 基本統計量

位置		ばらつき	
平均値	440.8600	標準偏差	40.05762
中央値	433.5000	分散	1605
最頻値	392.0000	範囲	215.00000
		四分位範囲	42.00000

NOTE: 9 個の最頻値があります (度数 : 2)。表では最頻値のなかで最も小さな値を表示します。

位置の検定 :  $\mu = 0$ 

検定	---統計量---	-----p 値-----
Student の t 統計量	t 77.82168	Pr >  t  <.0001
符号検定	M 25	Pr >=  M  <.0001
符号付順位検定	S 637.5	Pr >=  S  <.0001

## 正規性の検定 ②

検定	---統計量---	-----p 値-----
Shapiro-Wilk	W 0.913383	Pr < W 0.0014
Kolmogorov-Smirnov	D 0.119089	Pr > D 0.0762
Cramer-von Mises	W-Sq 0.182741	Pr > W-Sq 0.0086
Anderson-Darling	A-Sq 1.11832	Pr > A-Sq 0.0059

## パーセント点 (定義 5)

パーセント点	推定値
100% 最大値	585.0
99%	585.0
95%	501.0
90%	499.0
75% Q3	456.0
50% 中央値	433.5
25% Q1	414.0
10%	397.5
5%	392.0
1%	370.0
0% 最小値	370.0

UNIVARIATE プロシジャ  
変数 : RBC

極値のオブザベーション





$$y = (x_j + x_{j+1})/2 \quad (g = 0 \text{ のとき})$$

$$y = x_{j+1} \quad (g > 0 \text{ のとき})$$

で定義される。

定義の詳細については、参考文献を参照。

### 【Rの説明】

summary 関数で要約統計量をまとめて算出できる。また、t.test 関数で  $t$  検定、shapiro.test 関数で正規性の検定が行える。

### 【Rプログラム】

```
d010101.what<-list(NO=0, RBC=0)
tmp1<-scan(file="./d010101.csv",what=d010101.what, sep="")
d010101<-data.frame(tmp1)
summary(d010101[, "RBC"])
t.test(d010101[, "RBC"])
shapiro.test(d010101[, "RBC"])
```

### 【Rアウトプット】

```
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 370.0  414.5   433.5   440.9  455.8   585.0

One Sample t-test

data:  d010101[, "RBC"]
t = 77.822, df = 49, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 429.4758 452.2442
sample estimates:
mean of x
 440.86

Shapiro-Wilk normality test

data:  d010101[, "RBC"]
W = 0.91338, p-value = 0.001371
```

**【参考文献】**

- SAS プロシジャ リファレンス Version 6, First Edition