



Figure 38. 更新後の MEPHAS におけるデータ入力ページの一例

Figure 38 は更新後の MEPHAS でのデータ入力を受け付ける Web ページの一例である。現行の MEPHAS と同様、使用者からのデータ送信には CGI は使用しているため、デザインのみ他のページに合わせたものとしている。



Figure 39. 更新後の MEPHAS におけるデータを入力した際に空欄があった場合に表示されるエラーメッセージ

Figure 39 はデータを入力した際に空欄があった場合に表示されるエラーメッセージである。中央部分のフォントサイズが大きくなっている箇

所に何がエラーの原因となったのかを表示している。

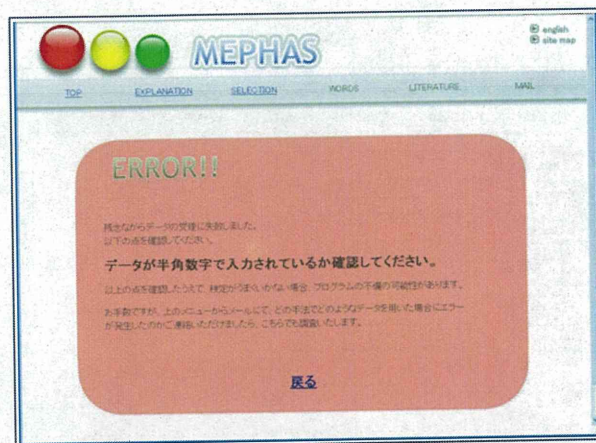


Figure 40. 更新後の MEPHAS における入力情報の誤りについてのエラーメッセージ
半角数字の入力が必要な箇所へ別の文字情報や記号を入力した際に表示される。

同じく Figure 40 は入力されたデータ中に半角数字が予定される箇所に別の文字情報や記号などが入力された場合のエラーメッセージを示している。先の Figure 39 と比較すると、中央部のフォントが大きくなった箇所の文章が異なっていることが分かる。

上記も含めて以下のエラーメッセージのレパートリーを作成した。

- 1). データの入力量が多すぎる場合
- 2). 入力に空欄がある場合
- 3). 入力が半角数字で行われていない場合
- 4). 信頼区間など入力範囲があるもので範囲外の入力の場合
- 5). 群数など整数値が期待されるもので整数値でない場合
- 6). サンプルサイズと入力データが合致しない場合

それぞれにおいて、『データ入力受付プログラム』内でのデータチェック結果に応じて、エラー

に該当するときには、Figure 39, 40 に示したようなエラーメッセージが表示される。

MEPHAS site map

TOP EXPLANATION SELECTION WORDS LITERATURE MAIL

menu

二元配置分散分析

手法の解説・使用法・in English

各要因の水準数を選択してください。

要因A 2
要因B 2

各データを下のテキストボックスに入力してください。

• データは各要因のラベルをつけて入力してください。ラベルには半角数字を使用します。
詳しくは使用法の解説を参考にしてください。
• データと各ラベルの区切りにはコンマかスペースを用いてください。
• 各データは改行で区別してください。

24.8	1	1
23.1	1	1
24.1	1	1
28.1	2	1
22.8	1	2
25	2	1
28.8	2	1
27.2	2	2
27.1	2	2
28.1	2	2
28.5	2	2

Figure 41. 更新後のMEPHASにおける二元配置分散分析のデータ入力画面

二元配置分散分析におけるデータ入力画面の更新については、Figure 34 に示した多数のテキストボックスへの入力を、Figure 41 に示すように一つにまとめることで簡便に済むよう、入力データに区分のためのラベルを含める形式へ変更した。またこのようなラベルを用いた入力形式はMEPHAS 中では新規であるため、入力のための説明事項を追加した。

第1群			第2群		
	データ	順位		データ	順位
1	5.70	15.00	1	4.72	6.00
2	3.82	1.00	2	5.77	16.00
3	4.15	2.00	3	5.02	7.00
4	6.13	19.00	4	5.63	13.00
5	5.81	17.00	5	5.49	12.00
6	7.66	30.00	6	4.39	4.00
7	5.48	11.00	7	6.68	25.00
8	4.48	5.00	8	5.64	14.00
9	4.26	3.00	9	5.26	9.00
10	6.70	28.00	10	6.47	23.00
11	6.69	26.00	11	6.32	21.00
12	6.14	20.00	12	5.09	8.00
13	6.86	29.00	13	5.36	10.00
14	6.71	27.00	14	6.64	24.00
15	6.00	18.00	15	6.45	22.00
合計	251.00		合計	214.00	
U1	94.00		U2	131.00	

Figure 42. 解析結果の出力ページにおける入力データの確認画面

Figure 35 に示した複数行、複数列からなる入力データにおける表示位置のズレに関しては、Excel や Google スプレッドシートのようなセル形式の

データ入力法を活用することも検討した。しかしながら、Web ページ上で動作するスクリプトは使用者の Web ブラウザの環境に依存して動作しない場合があることを考慮し、スクリプトの使用は機能に直接関わらない箇所で動作しない場合の代替が確保できる箇所のみに限定することが妥当と判断し、スクリプトが必要となるセル形式でのデータ入力は設けないことにした。そこで、入力画面の改善ではないが、Figure 42 に示すように解析結果の出力表示において、その統計解析計算プログラムで計算されたデータを確認できるよう、確認画面を作成した。

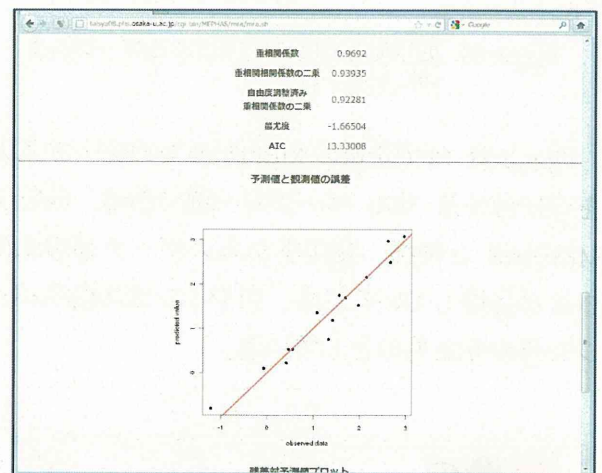


Figure 43. 更新後のMEPHASにおける重回帰分析実施結果のグラフ表示

また解析結果におけるグラフ表示については、Figure 36 に示した Web ブラウザの環境による表示エラーを避けるため、該当する統計解析計算プログラム（現状は重回帰分析と主成分分析の計算プログラム）においてグラフ画像を作成し、Figure 43 に示すように Web ページ上に表示させるように変更した。またそれに合わせて、Figure 44 に示すようにデータ入力画面においてグラフ画像の有無について設定できるオプションを追加した。

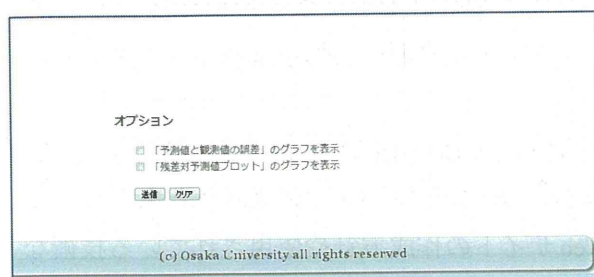


Figure 44. 更新後の MEPHAS における重回帰分析のグラフ表示に関するオプション設定

MEPHAS の Web サイトとしてのデザインやコンテンツの見直しでの検討結果を受けて、多重比較法の一つである Shirley-Williams の方法を追加するために必要となるプログラム作成等を行った。Shirley-Williams の方法について、以下概説する。

Shirley-Williams の方法は、例えば薬剤投与量が 0mg/kg (コントロール群), 10mg/kg, 50mg/kg, ... などのように、多群間に序列があるような数値データについてコントロール群との対比較を複数回行う際に、その序列も考慮に入れた多重比較を行うための手法である。同様の解析手法として Williams の方法(こちらは MEPHAS にて使用できる)があり、Shirley-Williams の方法との違いはパラメトリックな手法かどうか、つまり各群の数値データに正規分布を仮定するかしないかの違い(ちなみに Williams の方法はパラメトリックな手法)である。Shirley-Williams の方法では、他のノンパラメトリックな手法と同様、間隔尺度以上の数値データを順序尺度へと変換する(数値データに順位をつける)ことにより、数値データの分布形の条件から解放されている。この Shirley-Williams の方法も含めて、ノンパラメトリックな多重比較法を収載している既存の統計解析ソフトウェアはあまり多くは存在しておらず、MEPHAS の利用の主目的の一つとなっている。

上記の Shirley-Williams の方法を計算できるようにプログラムを作成した。新たに作成したプログラムであるため、複数のサンプルセットを乱数を用いて作成し、Shirley-Williams の方法を収録している既存の統計解析ソフトウェアの一つである Excel 統計[7] と計算結果の比較を行い、違いの無いことを確認することで、プログラムが適切に作成されていることの確認とした。

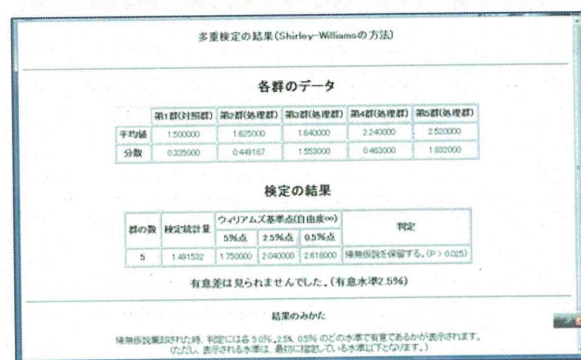


Figure 45. Shirley-Williams の方法による計算結果

Figure 45 は Shirley-Williams の方法による計算結果を示している。

また、分割表の検定の一つであるカイ二乗検定を追加するために必要となるプログラム作成等を行った。現行の MEPHAS が作成された当初は Fisher の正確確率検定が 2×2 の分割表のみならず、 $2 \times m$ の分割表 (2 個の水準と m 個の水準を持つ表) に対しても準備されていた (現在は入力データに依存して生じる過大な計算時間のため停止させている) ため、カイ二乗検定は省略されていた。今回は Figure 46 に示すように分割表の検定を選択するフローチャート中の 2×2 以外の分割表の計算において利用 (2×2 の分割表のみ Fisher の正確確率検定を使用可能) できるよう、作成することにした。

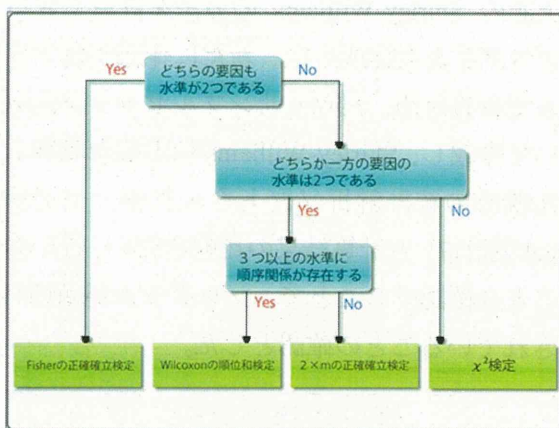


Figure 46. 更新後の MEPHAS における『分割表の検定』のフローチャート

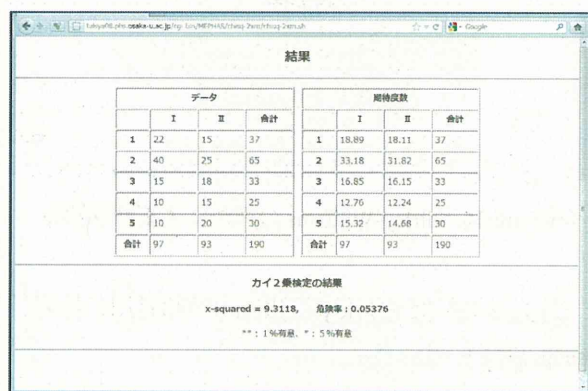


Figure 47. カイ二乗検定による計算結果

Figure 47 は作成したカイ二乗検定による計算結果を示している。

D. 考察

現在、SAS[8] を始めとして S-PLUS[9] や SPSS[10] など様々な統計解析ソフトウェアが販売されており、また S 言語のクローンとしてスタートした R は無償で非常に多くの手法（統計解析手法、データマイニング手法など）を利用可能である。MEPHAS は手法の多様性や高機能化を目指すのではなく、医薬学分野において利用頻度の高い、基本的な統計解析手法を簡便に利用できるよう提供し、かつ統計初学者の学習に資するような Web サイトを目指していくことにより、それら統計解析ソフトウェアとの使い分けができるものと考えている。従って本研究は、ユーザビリティ

の向上や統計解析手法の理解促進につながるような改善を念頭に実施を進めることとした。

現行の MEPHAS は 2000 年代前半、特に 2000 年から 2001 年にかけて作成されたものであり、Web サイトの作成方針や使用されている技術などに関して、現在の WWW や通信環境の状況、Web サイト構築の方向性などに合致していない点が多々見受けられたため、ユーザビリティ向上の一環としてデザインの再検討も含めて、Web サイトの見直しを行った。現在ではあまり推奨されず実際にほとんど使われなくなったフレームによる Web サイトのデザインを更新し、上部に Web サイト全体のメニュー、左側にコンテンツ内のメニューを表示する形式で統一した。また現在の一般的なネットワーク環境において数枚の画像程度ではさほど通信容量へのストレスにはならないと考え、メニュー表示等に関してもより画像を利用し、現行の MEPHAS における容量の軽減重視の方針から使用者の見易さをより重視した Web サイトへの修正を実施した。これらのことにより Web サイトを使用する際に、使用者の行いたいことや知りたいことに、より行き着き易くなったことが期待できる。

『手法の説明』というコンテンツに関しては現行の MEPHAS では文字情報がほとんどであり、また背景色と相まって非常に読みづらいものとなっている。そこで、先のデザインの統一と同様、使用者の見易さにも配慮した本文へと修正を行った。また、説明の内容についての更新に際しては、現行の MEPHAS において用いられている事例や図表を可能な限り利用することで省力化に努めると共に、それらの無い説明箇所については事例や図表等を取り入れることを検討し更新した。事例の紹介や図表による説明は解析手法自体の理解を促進すると共に、その手法により解析されるデータをイメージし易くなるため、使用者の

持つデータに対して適切な解析手法を選び易くする一助となることが期待できる。

『手法の選択』というコンテンツに関して、現行の MEPHAS では、まず使用者の解析したいデータを4つのカテゴリーへと分類し、該当するフローチャートへ誘導していた点を修正し、全てフローチャートの形式へと統一した。また現行の MEPHAS で用いられているフローチャートでは、先述の容量の軽減重視の方針から使用者の見易さ使い易さを重視せず、一般的な Web ブラウザーの一画面中に収まるように作成されていたため、複数のフローチャートにおいて開始点すら揃っていないこと、YES の回答しか行き先が指定されていない設問があること、現在の Web サイトでは計算できないとの結果に行き着く箇所が多いことなどがユーザビリティを下げる要因として挙げられたため、それらに関して修正したフローチャートを作成した。それらフローチャートの修正に伴い、追加が必要になった手法に関しては利用できるようプログラムを作成した。これらのフローチャートの修正は使用者の使い易さの向上につながることを期待できる。

なお、本研究事業における更新では解析手法を選択するためのフローチャートなど解析前の利用し易さの向上を中心に検討、実施してきたため、今後は MEPHAS による解析結果を利用した説明など、解析後の活用、つまり解析手法を利用して得られた結果を適切に用いる点についての説明に関しても充実を図る必要がある。

現行の MEPHAS において該当する語句のリンクにより別ウインドウへ表示することで提供されている語句の説明について、それら説明が利用し易くなるよう検討し更新を行った。語句の説明については、再確認としての使用頻度が高く、より簡便かつより即時的に内容を確認できることがユーザビリティの向上につながると判断し、

ポップアップにより語句の説明を表示できるよう更新した。当初 JavaScript の利用を検討し導入したが試験公開時には CSS による表示に変更した。語句の説明についても解析手法の説明と同様、本研究事業終了後も引き続き順次追加することで内容の充実を図る必要があるだろう。

なお追加実施事項として英語版 Web サイトの構築も、Web サイト更新に並行して進めたが、プログラムの計算結果も含め、文章の英訳中であり、まだ成果としては出ていないため結果からは割愛した。

現行の MEPHAS のプログラム構成を見直し、使用者の使い易さの向上につながるよう、修正点を検討した。その結果、CGI を介して使用者からデータを受け取るためのプログラム内でのデータチェックにおいて間違いを見つけた際の対応（エラーメッセージの内容も含む）がほぼ同一であり、使用者に間違いのあることを伝えるのみとなっており、使用者がその間違いの修正しづらいことが、MEPHAS 上の統計解析手法を利用することの妨げとなり得ると考えられた。そこで、間違いの内容に応じた対応ができるようプログラムを修正することで、使用者の使い易さの向上が図れると考え、プログラムの修正を行い、間違いの内容に応じたエラーメッセージが表示されるようにした。今後はデータ入力時のみならず、統計解析手法の実行時に生じるエラーについてもエラーメッセージの表示など使用者側へエラー内容について情報提供できるように修正することも必要と考えている。

現行の MEPHAS における各プログラムとデータ入力画面、解析結果表示画面を見直し、使用者の使い易さの向上につながるよう、修正点を検討した。その結果、以下の点を修正の対象とした。

- 1). データ入力画面における入力必要箇所の大量発生（例：二元配置分散分析）
- 2). データ入力画面における行列形式データの表示位置のズレとそれによる解析結果への不安（例：Dunnett の方法などの多重比較法）
- 3). 解析結果におけるグラフ表示の Web ブラウザ環境における表示エラー（例：重回帰分析）
- 4). 試験公開以降のメンテナンス等を考慮した統計解析計算プログラムの更新

1)から 3)のそれぞれについて、結果の箇所为例を示したように修正を行った。また 4)に関しては、今後の発展性等も念頭に、R 言語、R 実行環境への統計解析計算プログラムの更新を実施した。今後は、大阪大学内からのアクセスのみに限定した試験公開中にも修正の検討を進め、例えば使用者が選択した統計解析手法が適切でない場合はその理由を明示し適切な手法へと誘導する、適切であっても同様の解析手法がある場合にはそれを紹介するなどの機能を追加することで、統計解析の適切な利用、あるいは学習へとつながるような機能の追加を検討する。

また、手法選択のフローチャートの修正に伴い、追加が必要となった多重比較法、及び分割表の検定を実施するためのプログラムの作成を行った。追加された多重比較法は、群間に序列がある場合のコントロールとの対比較を行うパラメトリックな手法の Williams の方法と対となる、ノンパラメトリックな手法の Shirley-Williams の方法である。また分割表の検定はカイ二乗検定であり、Fisher の正確確率検定では過大な計算時間が発生してしまう 2×2 以外の分割表を対象とする手法として利用することを念頭に追加を行った。作成したプログラムは既存の統計解析ソフトウェアにて行った解析結果と比較し、違いが無いことを確認した。本事業終了後も手法を選択する上で必

要となった統計解析手法を中心に、手法の追加は引き続き検討する予定である。

今後は、解析手法の実行前に適切な手法を選択できるよう補助するだけでなく、何らかの手法により解析した際に、そのデータに応じて他の適切と思われる手法、あるいは同様の解析を可能な手法を紹介するなど、手法の計算結果を出した後の情報提供なども行えるようになれば、統計学への理解促進という点に関してより発展性があると考えている。

E. 結論

本研究事業では MEPHAS のコンテンツや Web サイトのデザインを追加、修正することによりユーザビリティの向上を図る開発研究と、MEPHAS における統計解析手法の実行時に動作しているプログラム群の修正と解析手法を追加するためのプログラムの作成を主に行う開発研究を実施し、一定の成果を得た。得られた成果は下記の通りである。

- a. 現行の MEPHAS における改善点の抽出とそれに基づく Web サイト修正の方向性の決定、及びデザインの作成
- b. a で決定した Web サイト修正内容に沿った Web ページ更新の作業実施と試験公開に向けた準備（若干遅れたが、予定通りアクセス可能な範囲を限定した試験公開まで完了）
- c. 手法の説明における事例や図表等の追加による更新
- d. 手法の選択のためのフローチャートの整理と改訂版の作成
- e. MEPHAS 内の用語の説明法の改訂と用語説明の一覧ページの作成
- f. 入力データの誤りへの対応に向けたプログラムの修正
- g. ユーザビリティ向上を目指した入力、解析結果表示画面の更新とプログラムの修正

- h. メンテナンスや更新等を考慮した統計解析
計算プログラムの更新
- i. 統計解析手法を追加するためのプログラム
等の作成

引用文献など

- [1]. MEPHAS,
<http://www.gen-info.osaka-u.ac.jp/MEPHAS/>
- [2]. R Development Core Team. R: A language and
environment for statistical computing. R
Foundation for Statistical Computing, Vienna,
Austria., <http://www.R-project.org/>, last
accessed: 2012/05/17
- [3]. The Apache Software Foundation,
<http://www.apache.org/>, last accessed:
2012/05/17
- [4]. W3C XHTML™ 1.0 The Extensible HyperText
Markup Language (Second Edition) A
Reformulation of HTML 4 in XML 1.0,
<http://www.w3.org/TR/xhtml1/>, last accessed:
2012/05/25
- [5]. W3C HTML 4.01 Specification,
<http://www.w3.org/TR/html401/>, last accessed:
2012/05/25
- [6]. 情報処理振興事業協会セキュリティセンタ
ー, 「スクリプトウイルスに関する報告書」
2001 年 2 月 ,
[http://www.ipa.go.jp/security/fy12/contents/virus
/report/script-rep.PDF](http://www.ipa.go.jp/security/fy12/contents/virus/report/script-rep.PDF), last accessed: 2012/05/17
- [7]. Excel 統計, <http://software.ssri.co.jp/ex2010/>
- [8]. SAS,
[http://www.sas.com/offices/asiapacific/japan/plat
form/sas9/index.html](http://www.sas.com/offices/asiapacific/japan/platform/sas9/index.html)
- [9]. S-PLUS, <http://www.msi.co.jp/splus/>
- [10]. SPSS, <http://www.spss.co.jp/>

F. 健康危険情報

特になし。

(分担研究報告書には記入せずに、総括研究報告
書にまとめて記入)

G. 研究発表

1. 論文発表

特になし。

2. 学会発表

福井大介¹, 田雨時², 岡本晃典², 川下理日人^{2,3},
後藤直久³, 安永照雄³, 高木達也^{2,3} (¹阪大薬,
²阪大院薬, ³阪大微研), 医薬学データ用統計解
析プログラム、MEPHAS の更新について,
29P-0115, 日本薬学会 第131年会, 静岡, 2011年3
月 (東日本大震災のため開催中止となったが、
組織委員会より要旨等も配布済みのため年会は
成立したものとするとのお知らせあり)

福井大介¹, 田雨時², 岡本晃典², 川下理日人^{2,3},
後藤直久³, 安永照雄³, 高木達也^{2,3} (¹阪大薬,
²阪大院薬, ³阪大微研), 医薬学データ用統計解
析プログラム、MEPHASの更新について(その2),
30P2-am115, 日本薬学会 第132年会, 札幌, 2012
年3月

H. 知的財産権の出願・登録状況

1. 特許取得

特になし。

2. 実用新案登録

特になし。

3. その他

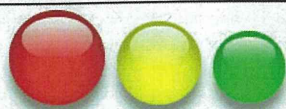
特になし。

(参考資料)

更新後の MEPHAS の Web ページキャプチャー画像（作成した Web サイトの一部のみ抜粋）



図 1. MEPHAS 新規サイトのトップページ



MEPHAS

[English](#)
[site map](#)

[TOP](#)

[EXPLANATION](#)

[SELECTION](#)

[WORDS](#)

[LITERATURE](#)

[MAIL](#)

menu

[分散分析](#)

[2 標本の差の検定](#)

[分割表の検定](#)

[母平均の多重比較](#)

[多変量解析など](#)

Explanation

分散分析

実験の結果得られたデータ（測定値）は、いくら同じ条件で同じ結果を出そうとしても必ず実験誤差を伴います。実験結果の解析をする際にばらつきを伴って得られる実験データから妥当な結論を得るための解析方法の一つとして使われるのが分散分析です。

[分散分析](#)

平均値の差の検定

平均値の差の検定には正規分布を前提としたパラメトリックな解析方法と正規分布を前提としないノンパラメトリックな方法があります。詳しくは上の“平均値の差の検定”のところをクリックしてください。平均値の差の検定はさらに以下のような種類に分けることができます。

[2 標本の差の検定](#)

[分割表の検定](#)

[母平均の多重比較](#)

多変量解析など

ある対象から得られた多数の変量を要約する手法を多変量解析と呼びます。

[スピアマンとケンドールの順位相関係数](#)

[重回帰分析](#)

[ロジスティック回帰分析](#)

[数量化Ⅱ類](#)

[主成分分析](#)

を行うことができます。各々の詳しい説明へは、上の“多変量解析など”のところか、それぞれの手法をクリックしてみてください。

[表紙へ戻る](#)

(c) Osaka University all rights reserved

図 2. MEPHAS 新規サイトにおける手法の説明コンテンツのトップページ





MEPHAS

[English](#)
[Site map](#)

[TOP](#)
[EXPLANATION](#)
[SELECTION](#)
[WORDS](#)
[LITERATURE](#)
[MAIL](#)

menu

分散分析

2標本の差の検定

分割表の検定

母平均の多重比較

多変量解析など

2 標本の差の検定

平均値の差の検定のうち、特に2つの集団の間に差があるかないかを調べるために用いる方法です。さまざまな種類の検定方法があります。

- 「パラメトリック検定」であるか「ノンパラメトリック検定」であるか
- 「対応のあるデータ」か「対応のないデータ」か

でそれぞれの手法を分類しています。

☒ 「パラメトリック検定」か「ノンパラメトリック検定」か

パラメトリック検定

パラメトリック検定は母集団パラメーター（平均、分散、標準偏差など）について仮説を設ける検定方法です。一般に母集団の正規性や等分散性を仮定します。

「処理解群と非処理解群の間で差が生じているか」や、「被験薬を投与した集団と対照薬を投与した集団間で差が生じているか」を調べる場合などのように2つの集団の比較を行うような場合、どちらの群もデータが正規分布に従うとすると2つの集団の母集団の母平均の差に関する検定を考えると比較しやすくなります。この検定を行うために用いるのがt検定です。本サイトには以下の3種類の検定があります。

- ☐ Studentのt検定
- ☐ Welchのt検定
- ☐ 対応のあるt検定

ノンパラメトリック検定

ノンパラメトリック検定はパラメトリック検定に対して一切の仮説を設けない検定方法です。データ数が少ない場合にも適用でき、名義尺度や順序尺度のデータも使用できます。

ノンパラメトリック検定である順位和検定は、測定値を大きさの順番に置きなおしその順位の差を考える検定です。本サイトでは以下の2種類の検定を行うことができます。

- ☐ Wilcoxonの順位和検定（マン・ホイットニーのU検定）
- ☐ Wilcoxonの符号付き順位和検定

☒ 「対応のあるデータ」か「対応のないデータ」か

2標本の差の検定では、さらにデータに対応がある場合、対応がない場合で2つに分けることができます。

対応のある場合

対応のある場合とは、2つの集団が「術後と術前」や、「薬の投与前と投与後」のように同一対象から異なる2地点での観測値の「ペア」が得られる場合、または異なる母集団から同じ条件を持つものとして選択するような場合のことを言います。

本サイトでは以下の2種類の検定を行うことができます。

- ☐ 対応のあるt検定
- ☐ Wilcoxonの符号付順位和検定

詳しくは[こちら](#)。

対応のない場合

対応のない場合とは、上記の「対応のある場合」の条件を満たさないものです。2つの集団の母集団がたがいに独立である場合のことでもあります。

本サイトでは以下の3種類の検定を行うことができます。

- ☐ Studentのt検定
- ☐ Welchのt検定
- ☐ Wilcoxonの順位和検定

詳しくは[こちら](#)。

[戻る](#)

(c) Osaka University all rights reserved

図 3. MEPHAS 新規サイトにおける 2 標本の差の検定に関する説明コンテンツのページ

[English](#)
[site map](#)

[TOP](#)
[EXPLANATION](#)
[SELECTION](#)
[WORDS](#)
[LITERATURE](#)
[MAIL](#)

menu

分散分析

2 標本の差の検定

分割表の検定

母平均の多重比較

多変量解析など

2 標本の差の検定

平均値の差の検定のうち、特に2つの集団の間に差があるかないかを調べるために用いる方法です。

対応のある場合

対応のある場合とは、2つの集団が「術後と術前」や、「薬の投与前と投与後」のように同一対象から異なる2地点での観測値の「ペア」が得られる場合、または異なる母集団から同じ条件を持つものとして選択するような場合のことを言います。

同一個体でなくても同じ条件下、同じ系統の場合もデータを「ペア」として扱うことができるためデータに対応があるといえます。データを「ペア」としているため、2群のデータ数は等しくなります。

このような対応のあるデータに用いるのが以下の検定です。

方法1 →

方法2 →

- ☒ 対応のある t 検定
- ☒ Wilcoxon の符号付き順位和検定

対応のある t 検定は、データに正規分布が仮定できる場合に用いる、パラメトリック検定です。

Wilcoxon の符号付き順位和検定は、2群が正規分布をなすかどうか分からない場合に用いるノンパラメトリック検定です。

データに正規分布が仮定できる場合

データに正規分布が仮定できる場合、対応のある t 検定を用います。

対応のある t 検定は、2つのデータの差を計算してその差を元にして検定を行います。Student の t 検定や Welch の t 検定より優位差が見られやすくなります。

前提

帰無仮説：「母平均に差はない」
対立仮説：「母平均に差がある」

有意水準を α として両側検定を行います。

$P > \alpha$ の時、帰無仮説を採択し、「母代表値に差があるとは言えない」
 $P \leq \alpha$ の時、帰無仮説を棄却し、「母代表値に差がある」とします。

データに正規分布が仮定できない場合

データに正規分布が仮定できない場合、Wilcoxon の符号付き順位和検定を用います。

前提

帰無仮説：「母代表値に差はない」
対立仮説：「母代表値に差がある」

有意水準を α として両側検定を行います。

$P > \alpha$ の時、帰無仮説を採択し、「母代表値に差があるとは言えない」
 $P \leq \alpha$ の時、帰無仮説を棄却し、「母代表値に差がある」とします。

[戻る](#)

(c) Osaka University all rights reserved

図 4. MEPHAS 新規サイトにおける 2 標本の差の検定内、対応のある場合に関する説明コンテンツのページ

[english](#)
[site map](#)

[TOP](#)
[EXPLANATION](#)
[SELECTION](#)
[WORDS](#)
[LITERATURE](#)
[MAIL](#)

menu

分散分析

2 標本の差の検定

分割表の検定

母平均の多重比較

多変量解析など

分散分析

ある実験をいくつかの条件で行ったとします。一つの条件に対し繰り返し実験を行ったとき、同じ条件であっても、そのなかで平均値からのずれが多かれ少なかれ生じます。これを実験誤差と呼びます。

また条件を変えて同様の実験を繰り返し行ったとき、条件の違いからほかの条件の場合の平均値との差が出た場合、その差を生みだした **要因** を **因子** と呼びます。

データの持つばらつきは分散として得られるのですが、分散の大小は実験誤差と因子に左右されるといえます。分散が大きければ散らばりも大きく、分散が小さければ散らばりも小さいと解釈できます。分散分析とは、データの持つばらつきが因子によるものよりも実験誤差によるもののほうが大きいかを検定し、因子によるばらつきのほうが大きければ **母平均** に差があるとする検定です。

一元配置分散分析

ある一つの **因子** を取り上げて、その因子について条件を変えることによりデータに違いが生じるかどうかを検定する方法です。前提条件として各データはほぼ **正規分布** に従い、かつ等分散であるものとします。各群のデータ数は同じでなくてもかまいません。

一元配置分散分析の例は以下のようなものです。

ある化学合成において温度を 200、230、250℃ の 3 種類に設定して実験を行う。また各温度の設定で複数回の実験を行いデータをとるものとする。

[一元配置分散分析を行う。](#)

二元配置分散分析

2 つの **因子** を取り上げて、それらの因子についての条件を変えることによってデータに違いが生じるかを検定する方法です。

2 元配置分散分析では、2 つの因子それぞれ単独に生じるばらつきのほかにそれら 2 つの因子が組み合わせることによって生じるばらつきも考慮に入れなければなりません。

それぞれの因子単独で生じるばらつきは「主効果」、2 つの因子の組み合わせによって生じるばらつきは「相互効果」と呼ばれます。

具体的な 2 元配置分散分析の例として、以下のようなケースが考えられ、それぞれ検定の仕方が少し異なります。

- [各水準の繰り返し数が等しく、1 である場合。](#)
- [各水準の繰り返し数が等しく、2 以上である場合。](#)
- [各水準の繰り返し数が等しくないが、周辺度数に比例する場合。](#)
- [各水準の繰り返し数が等しくなく周辺度数に比例しない場合。](#)

[二元配置分散分析を行う](#)

戻る

(c) Osaka University all rights reserved

図 5. MEPHAS 新規サイトにおける分散分析の説明コンテンツのページ

English

Site map

TOP

EXPLANATION

SELECTION

WORDS

LITERATURE

MAIL

menu

分散分析

2標本の差の検定

分散表の検定

母平均の多重比較

多変量解析など

多変量解析

3つ以上の変数を同時に扱う統計手法のことを多変量解析と呼びます。

本項では、多変量解析を中心に紹介していきます。

重回帰分析

重回帰分析は、いくつかのデータ（説明変数）から目的のもの（目的変数）を予測する回帰式を求める、または説明変数が有意な影響を与えるかどうかを知るために用いられます。回帰式は以下のような形で表すことができます。この式を「重回帰式」と言います。

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_jx_j + \dots + b_px_p \quad (j = 1, \dots, p)$$

y : yの推定値

b_0 : 定数

b_j : 偏回帰係数

例えば、性別、年齢、体重などの要因（説明変数）から寿命（目的変数）を推定する回帰式を求めたいとき、重回帰分析によりそれぞれの説明変数が目的変数に与える影響の大きさ（偏回帰係数）を求め、上記のような回帰式を導きます。また、求められた偏回帰係数の信頼度からその目的変数（性別や年齢）が目的変数（寿命）に対して有意な影響を与えているかどうかを知ることができます。

重回帰分析において目的変数が正規分布をとる連続変数であることが原則です。

結果

当サイトのプログラムで重回帰分析を行うと以下の値が得られます。

t値

求めた偏回帰係数や定数項が0でないかどうかを検定するために用いる検定統計量です。これをもとにt検定を行い、「偏回帰係数（または定数項）は0である」という帰無仮説が棄却されるかどうかを見ます。帰無仮説が棄却された場合、その偏回帰係数は、0でないと言えるので、目的変数（y）に何らかの影響を与えていることになります。また、帰無仮説を採用する場合、その偏回帰係数は0でない、つまり目的変数（y）に影響を与えていない可能性があるといえます。

決定係数

目的変数の実測値の分散は、求められた回帰式の予測値の分散と、残差の分散の和で表わされます。回帰方程式では、この残差の分散ができるだけ小さくなるように回帰式を定めています。そこで、回帰式の当てはまりの良さを表す指標として用いられるのが、決定係数です。決定係数は、0以上1以下の値をとり、1に近ければ近いほど予測精度は良いということです。一般に決定係数が0.8以上であれば、かなり良い精度で予測されているといえます。

重相関係数

実測される目的変数（y）の値と重回帰分析の結果として予測される目的変数の間の相関係数のことです。この値が大きければそれだけ回帰方程式から得られる予測値が、目的変数の実測値に近づくことになります。もしも重回帰分析による予測が完全であればこの値は1になります。

自由変調整済み重相関係数

説明変数を増やしていくと、重相関係数は増大していき、説明変数（x）の偏数が「データセット数 - 1」になると重相関係数は1になり、データセット数が独立変数の数に比べて小さい時には良い指標であるとは見えなくなってしまう。そのために重相関係数をデータセットの数と自由度で修正します。この値が身になった時は予測した重回帰式には意味がないということになります。

回帰の分散分析表

ここでは、得られた回帰式の正当性を検定しています。重回帰分析の場合、結果の正当性は上記にある決定係数で判断されます。そこで、重回帰分析の正当性を評価するために、決定係数が十分に大きいといえるかを判断するために分散分析を行っています。ここで決定係数が大きいということが言えれば、回帰式は正当であるといえます。分散分析に関する詳しい解説は、こちらをご覧ください。

信頼限界

設定した有意水準のもとで母数（ここでは偏回帰係数のこと）が存在する区間の上限と下限のことです。有意水準を5%とした時、95%の確率でこの区間に偏回帰係数が存在するということになります。

AIC

赤池の情報量基準。モデルの最大対数尤度と独立変数（説明変数）の数により解析的評価から求められたもの。この値が小さいほどそのモデルは予測が良いといえます。

信頼限界

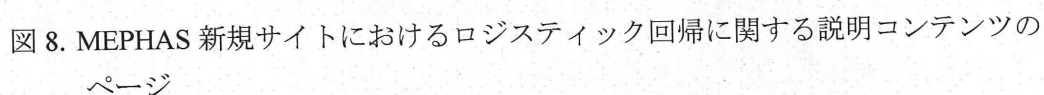
設定した有意水準のもとで母数（ここでは偏回帰係数のこと）が存在する区間の上限と下限のことです。

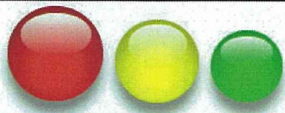
戻る

(c) Osaka University all rights reserved

図 7. MEPHAS 新規サイトにおける重回帰分析に関する説明コンテンツのページ

- 30 -





[english](#)
[site map](#)

TOP
EXPLANATION
SELECTION
WORDS
LITERATURE
MAIL

chart menu

2 群間の差の検定のチャートへ

分割表の検定のチャートへ

分散分析・多重比較のチャートへ

多変量解析などのチャートへ

Selection

正しい手法を選択するためのチャートです。

下のチャートの黄色の部分をクリックするとさらに詳しいチャートが表示されます。

赤字をクリックするとその説明が表示されます。

変数の数はいくつですか？

2 群

3 群以上

データは 2 群に分けられている

データは 3 群以上に分かれている

分割表のデータである

2 群間の差の検定のチャートへ

多重比較・分散分析のチャートへ

分割表の検定のチャートへ

順位相関係数

多変量解析などのチャートへ

[戻る](#)

(c) Osaka University all rights reserved

図 9. MEPHAS 新規サイトにおける手法の選択補助コンテンツのページ

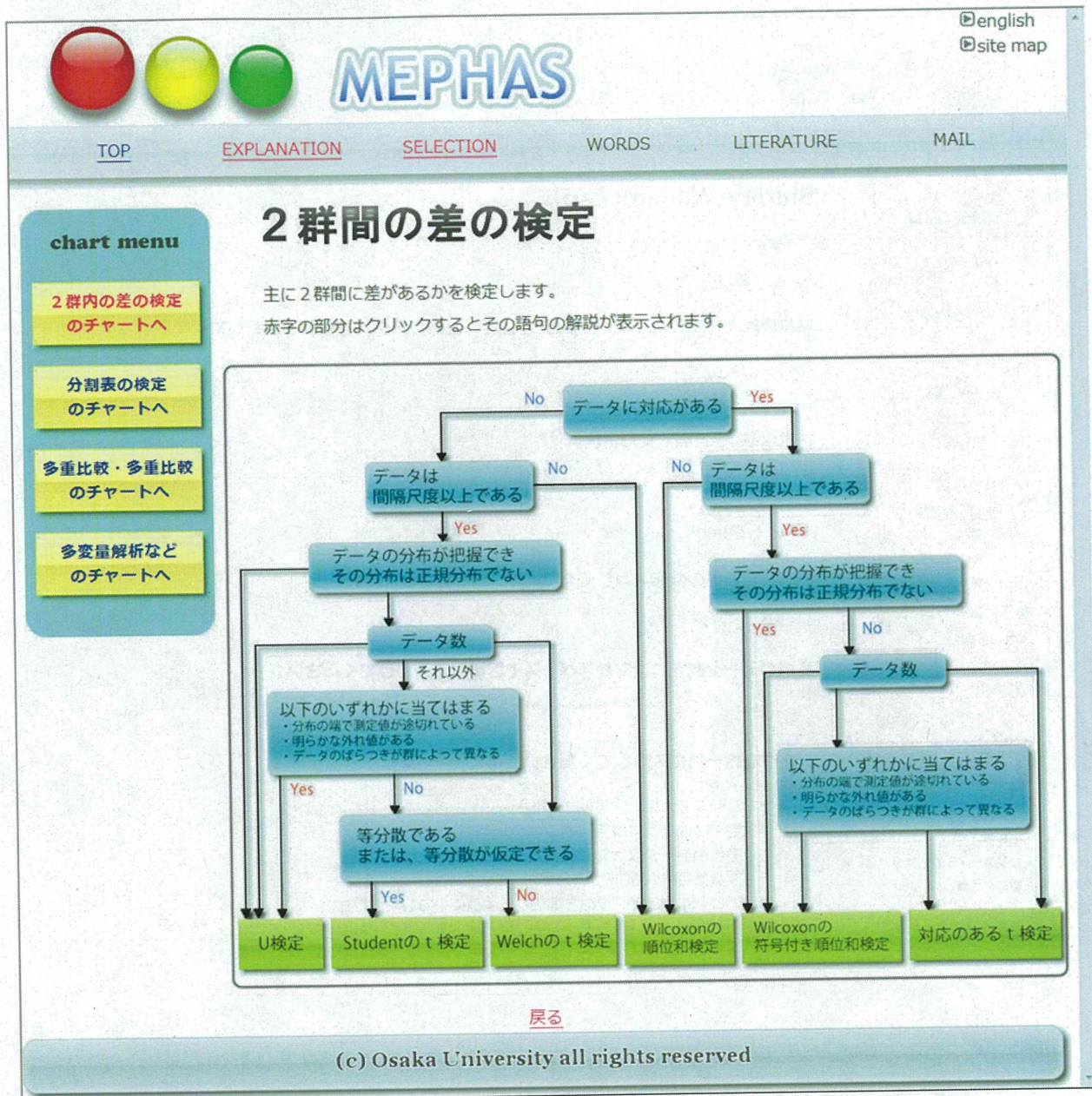


図 10. MEPHAS 新規サイトにおける手法の選択補助内、2 群間の差の検定に関する
選択補助のフローチャートのページ



[english](#)
[site map](#)

[TOP](#)
[EXPLANATION](#)
[SELECTION](#)
[WORDS](#)
[LITERATURE](#)
[MAIL](#)

menu

分散分析

[一元配置分散分析](#)

[二元配置分散分析](#)

2 標本の差の検定

[Studentのt検定](#)

[Welchのt検定](#)

[Wilcoxonの順位和検定](#)

[対応のあるt検定](#)

[Wilcoxonの符号付順位和検定](#)

分割表の検定

[Fisherの正確確率検定](#)

[カイ2乗検定\(2x2\)](#)

[カイ2乗検定\(2xm\)](#)

[カイ2乗検定\(mxn\)](#)

母平均の多重比較検定

[Tukeyの方法](#)

[Dunnettの方法](#)

[Williamsの方法](#)

[Steel-Dwassの方法](#)

[Steelの方法](#)

[Shirley-Williamsの方法](#)

多変量解析など

[順位相関係数](#)

[重回帰分析](#)

[ロジスティック回帰分析](#)

[数量化二類](#)

[主成分分析](#)

Shirley-Williamsの方法

手法の説明・使用法・in English

帰無仮説を「 $\mu_1 = \mu_2 = \dots = \mu_a$ 」として、証明したい対立仮説を選択してください。

☐ $\mu_1 < \mu_2 < \dots < \mu_a$
☐ $\mu_1 > \mu_2 > \dots > \mu_a$

有意水準 α を定めてください。

☐ $\alpha = 0.050$
☐ $\alpha = 0.025$
☐ $\alpha = 0.005$

検定したい群の数を入力してください。

群数

各群のデータ数をコンマかスペースで区切って入力してください。

各群のデータ数

データを縦一列に入力してください。

- ただし、以下の点に注意してください。
- 群と群の間はコンマかスペースで区切ってください。
- それぞれのデータ数は異なっていてもかまいません。
- データ数は各群1000まで入力可能です。
- コントロール群はデータの左端になるように入力してください。
- 結果のページでデータの左から第2群、第3群...sと呼んでいます。
- データ入力はエクセルからの「コピー＆ペースト」でも可能です。

(c) Osaka University all rights reserved

図 11. MEPHAS 新規サイトにおける Shirley-Williams の方法のデータ入力ページ

多重検定の結果 (Shirley-Williamsの方法)

各群のデータ

	第1群(対照群)	第2群(処理群)	第3群(処理群)	第4群(処理群)
平均値	4.846485	5.847370	5.999199	7.345175
分散	1.550319	1.260604	1.106185	0.660575

検定の結果

群の数	検定統計量	ウィリアムズ基準点(自由度 ∞)			判定
		5%点	2.5%点	0.5%点	
4	4.152207	1.739000	2.032000	2.615000	有意水準0.5%で有意
3	2.053485	1.716000	2.015000	2.607000	有意水準2.5%で有意
2	1.954751	1.645000	1.960000	2.576000	帰無仮説を保留する。(P > 0.025)

第3群以上で有意差が認められます。(有意水準2.5%)

結果のみかた

帰無仮説棄却された時、判定には各 5.0%、2.5%、0.5% のどの水準で有意であるかが表示されます。
(ただし、表示される水準は、最初に指定している水準以下となります。)

帰無仮説が最初に指定した有意水準で保留された場合、その時点で検定を終了します。

図 12. MEPHAS 新規サイトにおける Shirley-Williams の方法の計算結果の表示ページ

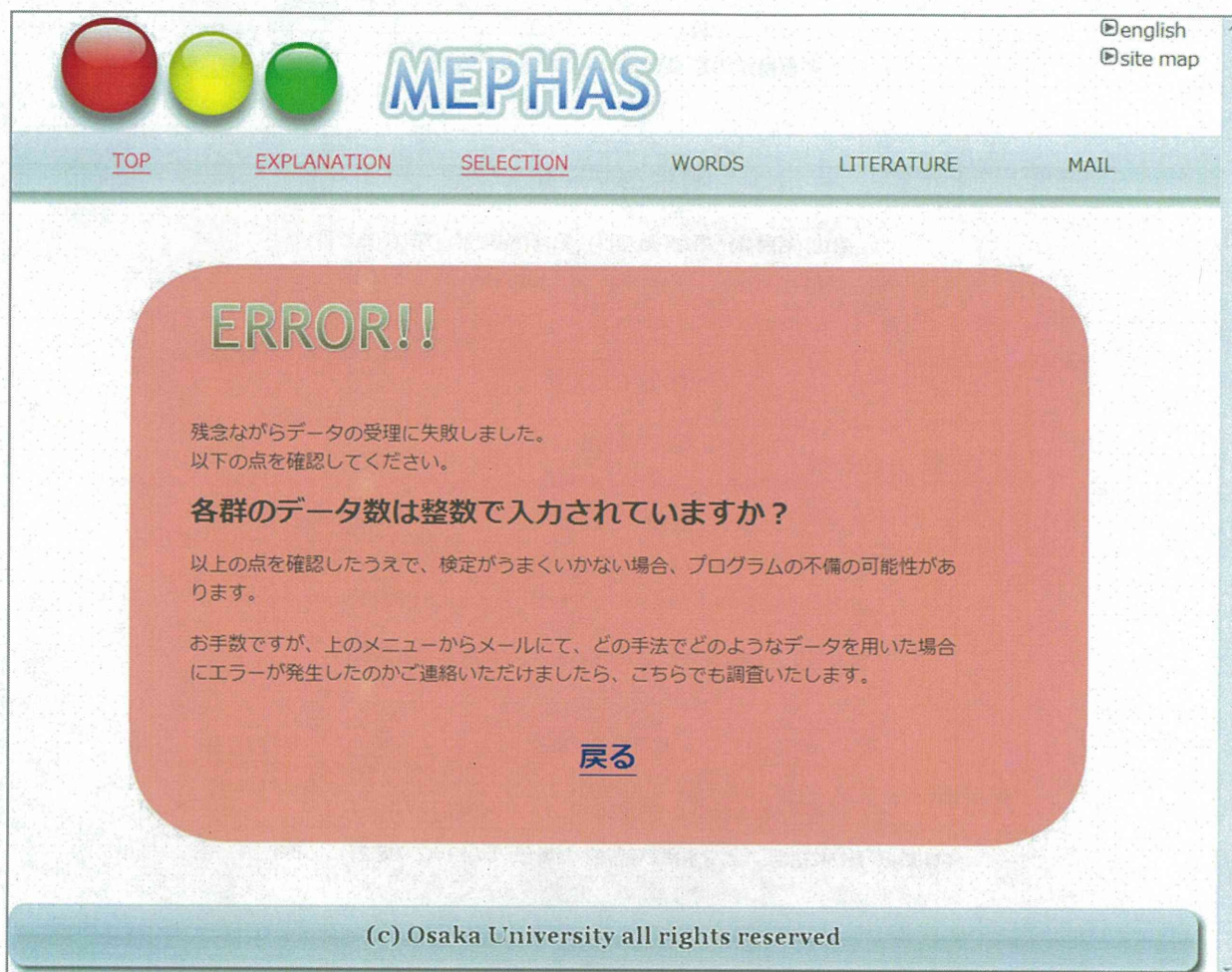


図 13. MEPHAS 新規サイトにおけるエラーメッセージ(データ数が整数で入力されていない場合) のページ