
プロジェクト実習 I

ヒューマンインタフェース

対象：情報工学課程 2 回生

担当教員：西崎友規子

2024 年 9.20 改訂 (2024 年度用)

§1. はじめに

ヒューマンインタフェース (Human Interface) とは、人間 (human) と機器 (computer 等) との間を繋ぐもの、または、そこで生ずる課題全般を指す用語です。同じような意味の言葉として、ヒューマン-コンピュータ・インタラクション (Human Computer Interaction)、ユーザインタフェース (User Interface)、マン-マシン・インタフェース (Man Machine Interface)、ヒューマン-マシン・インタフェース (Human Machine Interface) などがあります。本実習では、ヒューマンインタフェースの省略表現として「インタフェース」という用語を用います。

さて、私たち人間にとって、良いインタフェースとはどのようなものでしょうか。まず最低限のレベルとして、操作対象となる機器を「使うことができる」ようにインタフェースは提供されなければなりません。どんなに優秀な機能があっても、その機能を使う手段がユーザにインタフェースとして提供されない限り、その機能に意味があるとはいえません。次に、単に「使うことができる」だけでなく、「使いやすい」ことは対象となる機器を利用する上で大切です。この「使いやすさ=ユーザビリティ (usability)」がインタフェースの良し悪しを決定づける主要な要素となります。

1-1. 実習の目的

本実習は、より良いインタフェースを実現するために必要な手法を学ぶことを目的とします。

- 目的 1) 人間の認知特性の特徴を明らかにするための実験を体験し、認知特性の測定方法や分析方法の一端を学びます。
- 目的 2) インタフェースの開発手順を学びます。

1-2. 実習の準備

本実習では、下記の 2 つのソフトウェアを使用します。いずれも、無料でダウンロード可能ですので、自宅の PC で作業を行う場合は、各自でインストールするようにしてください

- 統計解析用のプログラム言語 R (R コマンドー)
- Windows 上で動作する総合開発環境 Microsoft Visual Studio 2022 で利用できる Microsoft Visual C#
(Visual Studio 2022 for Mac のサポートは対応しません)

また、実験データをグループ内で交換したり、持ち帰るために、USB メモリなどを準備してください。

§2. 認知特性

認知（機能）特性とは、感覚器（視覚、聴覚、触覚など）から入る様々な情報を処理、記憶、整理して、それをもとに理解、判断、行動などに繋げる、知的活動全般の基盤となる機能です。

認知特性を知ることは、より良いインタフェースを設計・デザインする上で重要です。人間全般が持つ共通の認知特性の傾向を把握することにより、人間の自然なふるまいに対して、多くの人がストレスなく快適に使用できるインタフェースを提供することに繋がります。また、認知特性には個人差が大きく生じるものもありますが、事前に対象となるユーザの特性を掴んでいれば、そのユーザ群にとって適切なインタフェースを設計・デザインすることが可能になります。

§3. 人間中心設計

ユーザビリティの高いシステムを開発するには、ユーザの特性を理解し、ユーザの視点に立って、ユーザがそのシステムのメンタルモデルⁱを構築しやすく、あるいは、ユーザの抱いているメンタルモデルⁱと実際のシステムとのギャップが少ないように設計することが重要です。このような製品設計は、人間中心設計（Human Centered Design (HCD)）と呼ばれます。人間中心設計のための基本原則として、国際標準化機構 ISO 9241-210 では以下の6項目が挙げられています。

1. 設計は、ユーザ、タスク、環境の明確な理解に基づくべきである
2. ユーザは、設計から開発までの間、関与すべきである
3. 設計は、ユーザを中心に据えた評価によって駆動され、改良されるべきである
4. これらの過程は反復的に行われるべきである
5. 設計はユーザエクスペリエンス全体を指向すべきである
6. 設計チームとして多様なスキルや視点を持った人を参加させるべきである

§4. インタフェース開発手順

人間中心設計の観点からインタフェースを設計・開発するには、人間の情報処理モデルを前提におくとその工程が考えやすくなります。すなわち、目標を設定し、ユーザやヒューマンファクタ（人間とシステムの間で考慮すべき要因）を分析、システムを設計、プロトタイプ（試作品）を開発、ユーザビリティを評価、その結果をまた仕様（開発に必要な詳細な設計）にフィードバックさせるというものです（図1）。

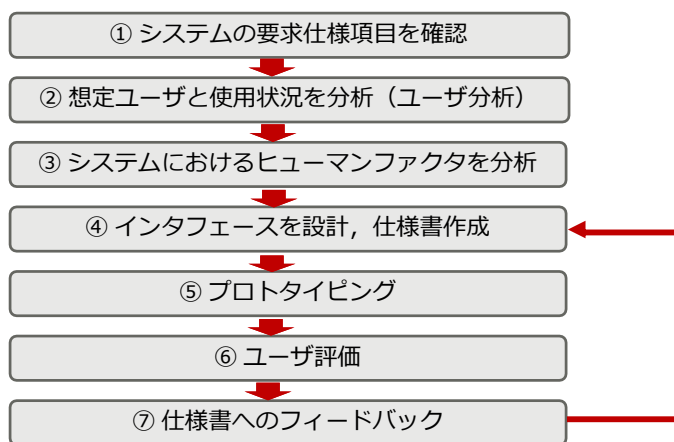


図1 標準的なインタフェース設計工程

ⁱ メンタルモデル：ユーザが頭の中に描いている像のこと

本実習の前半では、ユーザ分析（②）の重要性を学び、後半では、ヒューマンファクタ分析→設計→プロトタイピング→ユーザ評価（①～⑦）の一連のプロセスを体験します。

① システムの要求仕様項目を確認

作ろうとしているシステムは、何をするものなのか、どのくらいの性能を達成する必要があるのか、入力と出力はそれぞれ何かなどの使用項目を決定、確認します。

② 想定ユーザと使用状況を分析

システムを使用すると思われるユーザは、どのような人なのか、年齢や性別、能力、特性などを分析します（＝ユーザ分析）。能力や特性については、想定ユーザに近いモデルユーザに協力を得て、実験を行なって測定したり、観察したり、意見をもらったりします。また、使用状況や環境についても分析が必要です。

③ システムにおけるヒューマンファクタを分析

入力部、出力部、メニューなど、システムのインタフェース部分について、ヒューマンファクタ（人間とシステムの間で考慮すべき要因）を列挙し、要件などを分析します。

④ インタフェースを設計

システム本体の設計と合わせて、①、②および人間の特性を考慮し、インタフェースを設計します。結果は、インタフェース仕様書としてまとめます。インタフェース仕様書とは、システムのインタフェース部分に関する設計項目の詳細を設計思想とともに記述したものです。

⑤ プロトタイピング

4の設計と仕様書に基づいて、プロトタイプを開発します。実際に使用しないと決められない値は、パラメータ化しておきます。

⑥ ユーザ評価

作成されたプロトタイプについて、インタフェースが必要な要件を満たしているか、操作上の問題はないか、ユーザビリティはどうか、実際の製品のユーザもしくはモデルユーザに使用して評価を行います。評価手法については、§5で詳しく解説します。

⑦ 仕様へのフィードバック

評価した結果を分析し、もとの仕様に反映させます。修正すべき項目があれば修正し、パラメータの値も再設定します。

§5. インタフェース設計原則

より良いインタフェース設計のための設計原則として、Shneiderman の 8 原則を紹介します。これらの原則は経験から導かれた一般論であり、実際の設計に適用する場合には、盲目的に適用しないように注意が必要です。特に、設計に利用するには具体性に欠ける原則があったり、場合によっては、複数の原則が互いに矛盾することもあります。したがって、環境に応じて、解釈し、具体的に検討し、適切に拡張しなければならないことを念頭に置くといいでしょう。

1. 一貫性を持たせる

メニュー、ヘルプ、メッセージなどにおいて、用語、部品の表示位置、部品の形、部品のサイズ、並びの意味、色の意味、形の意味などに一貫性を持たせるべきです。

2. 熟練ユーザには近道を用意する

システムに慣れたユーザに対して、あるいは頻繁に使う機能には、すばやくその機能を利用できるように近道を用意するべきです（例えば、短縮形、特殊キー、隠しコマンド、マクロ機能など）。逆に、利用頻度が低かったり、誤操作による影響が大きい場合には、操作を複雑にすると良いでしょう（例えば、2 つ以上のシーケンスによる操作や、複数のボタンを同時に押すなど）。

3. ユーザに有益なフィードバックを提供する

全ての操作に対して、インタフェースは常に何らかのフィードバックを与えるべきです。頻繁に行われる操作には簡潔な対応を、頻度が少ない操作には多くの有益な情報をユーザにフィードバックします。

4. ユーザに段階的な達成感を与える

操作の流れにも起承転結が必要です。一連の操作を完遂したときにユーザにそれを知らせるフィードバックは、一つのことをやり遂げたという満足感、安心感を与え、不測の事態を起こす可能性を少なくするとともに、次の動作への準備を促します。

5. エラー処理を簡単にする

できる限りユーザが致命的なエラーを起こさないように設計することが前提であるが、もしエラーが起きてしまったら、システムがその原因を見つけ出し、単純でわかりやすいエラーの処理方法を提供するように心がけるべきです（例えば、コマンド全部を再入力するのではなく、間違えた箇所だけを訂正すればよいようにするなど）。

6. 逆操作を許す

操作は可能な限り可逆にすべきです。操作を誤ったとしても、それを取り消せることを知っていれば、ユーザは不安にはならず、システムを利用することに対する気軽さを誘うことができます。

7. ユーザに主体的な制御権を与える

ユーザは、自分がシステムに使われているのではなく、自分がシステムを制御しているという意識でいたいものです。やるべきことをシステムが決めるのではなく、ユーザが主体的に選択しているという意識にさせるべきです。

8. ユーザの短期記憶の負担を少なくする

人間の短期記憶には限度があり、古くから「 7 ± 2 」の情報しか記憶できないと言われています（最近では「 4 ± 1 」の方が一般的です）。そのため、ユーザが操作途中に覚えておかなければならないことは、極力少なくすべきでしょう（例えば、表示を簡潔にする、ウィンドウ間の移動を少なくするなど）。

§6. ユーザ評価法

ユーザ評価は、可能な限り設計の早い段階から行い、問題点を抽出してその解決を心がけることが、より良いインタフェースを目指すうえで重要です。評価手法は、大まかに、「分析的手法」と「実験的手法」の2つに分けることができます。以下に両者の相違点をまとめます（表2）。

表2 分析的手法と実験的手法

	分析的手法	実験的手法
特徴	ユーザが作業を実行する過程をシミュレートすることにより評価	プロトタイプあるいは製品を実際のユーザに使ってもらうことにより評価
利点	実際のインタフェースやユーザを必要としないため、初期の設計段階でも評価が可能である。一般の評価者（被験者）雇用の手間がかからないが、熟練した評価者が必要である場合が多い。	実際のシステムや製品を使って課題解決場面を見るため、「ある操作は実行時間が想像以上にかかっている」など、分析的手法では発見しづらい問題点を見つけることが可能。実際のユーザに近い被験者を選ぶことで、設計者が気づかなかったユーザ側の問題を発見することができる。
代表的な手法	チェックリスト法 ヒューリスティック法 認知的ウォークスルー	客観評価（パフォーマンス評価） 主観評価 発話プロトコル分析

6-1. 分析的手法

6-1-1. チェックリスト法

ユーザビリティを評価するための質問項目を人間工学や認知工学の基準に基づいて作成したものであり、開発担当者によって評価されます。評価項目は、例えば「見やすくなっているか」、「重要な情報は強調されているか」などであり、問題領域、改善範囲ならびにインタフェースの良い点などを明確にし、標準的で系統だったインタフェースの評価を可能にします。問題点としては、チェックリストを作成する段階でどのような質問事項を取り入れるかは各種のガイドラインに沿っており、チェックリストで浮かび上がってくる製品の使いにくさの程度はリスト作成者の能力に左右されることです。

6-1-2. ヒューリスティック法

数人のユーザ・インタフェース専門家が経験則に基づいて、仕様やプロトタイプを評価します。各評価者は個別に仕様やプロトタイプを評価し、その後、評価者間で意見の交換が行われ問題点をまとめたりリストを作成します。ヒューリスティックにおける評価項目として Nielsen は次の 10 項目を挙げています（表3）。この方法は、初期段階の評価手法としては極めて有力ですが、評価者に知識と経験がなければあまり効果を発揮できない、といわれています。

表3 Nielsen のユーザビリティ 10 項目

1	システムの状態を視認できるようにする
2	実環境にあったシステムを構築する（専門用語は避け、ユーザが普段使う言葉を使用する）
3	ユーザにコントロールの主導権と自由度を与える
4	操作と表示に一貫性を持たせる
5	フィードバックを与え、エラーの発生を事前に防止する
6	記憶の負担を最小限にし、見ただけでわかるようにする
7	柔軟性と効率性を持たせる（ショートカットなど）
8	余分な情報を提示しない最小限で美しいデザインにする
9	ユーザがエラーを認識し、回復できるようにする
10	ヘルプやマニュアルを用意する

6-1-3. 認知的ウォークスルー

人間の情報処理モデルに倣いながら、ユーザになったつもりで、ユーザがどのように機器を操作するか、その過程をレビューする方法です。担当設計者、担当外設計者、他の設計チームのメンバーなど複数の人が参加することが常であり、必ずしも、ユーザ評価の専門家の参加は必要ではありません。この手法の欠点は、不足している機能や不要な機能を評価できないことです。

6-2. 実験的手法

6-2-1. 客観評価（パフォーマンス評価実験）

インタフェースの良し悪しを客観的かつ定量的に測るための一般的な手法です。実験参加者に、ある課題（タスク）を与え、開発したシステム・機器、あるいはプロトタイプを使ってタスクを行わせます。このタスク達成のパフォーマンス（性能）により、評価を行います。定量的に測定可能なパフォーマンスには、例えば、以下（a, b, c）があります。また、定量的に測定した値を正確に評価するには統計分析が必要であり、そのためには統計的に意味のある実験参加者数やタスク数が必要となります。

- a: 作業時間（タスク達成に要した時間）
- b: 作業の正確さ（エラー率を用いる場合もある）
- c: 習得時間（操作方法を習得するまでに要した時間）

なお、客観評価においては、実験における偏りをなくすために、実験参加者の特性のばらつき（年齢、性別、学歴、職歴、その他の個人特性）、および、タスクの遂行順（カウンターバランス）に注意する必要があります。

6-2-2. 主観評価

実験参加者の主観的な評価内容を指標とする方法で、質問紙調査（いわゆる、アンケート）、インタビューがあります。主観的な指標を定量的に扱うことができるのは、質問紙調査による回答です。質問紙調査は、容易にユーザの実態を掴むことができると考えられがちですが、有用な質問紙を作成しなければ間違った結論を導くことにもなるため、その作成が最も重要な鍵となります。質問文作成の注意点（表4）に十分に気を配る必要があります。

表4 質問文作成の注意点

・	回答者が理解できない、専門用語や難しい言葉を使っていないか
・	回答者が誤解するような表現、曖昧な表現はないか
・	2つの事例を一度に聞いていないか
・	誘導質問になっていないか
・	社会的に望ましい回答をするなどのバイアスのかかった回答をする質問はないか
・	選びたいと思う選択肢がないという可能性はないか
・	質問の順番は適切か

通常、「リッカート法」と呼ばれる、選択肢が5～9個から成る方法が用いられます（表5）。例えば、選択肢が5個から成る方法は、5件法と呼びます。選択肢が少ない場合は粗い情報しか得られませんが、素早く回答できるため、多くの質問項目を設けることができます。逆に、選択肢が多いほど微妙な差異を拾うことができますが、回答者の選択に迷いが生じたり、時間がかかるという問題があります。主観評価は、一般に、実験参加者の主観の基準があいまいであり、結果の再現性が乏しいという欠点もあります。

また、実施方法としては、最近ではアンケート用紙ではなくオンラインアンケートシステムが普及し、回答者の利便性だけでなく、集計の負担も軽減できるようになっています。

表5 リッカート法の例

質問	このシステムを継続して使ってみたいと思いますか？	
回答選択肢	1	そう思わない
	2	どちらかというと思わない
	3	どちらともいえない
	4	どちらかといえばそう思う
	5	そう思う

6-2-3. 発話プロトコル法

ユーザが機器を操作しながら、そのとき頭の中に浮かんだことをその場で口に出させて、ユーザの操作に対する認知過程がどのようなになっているかを引き出す手法です。プロトタイプの評価には特に効果的で少ない実験参加者数でも豊富なデータを得ることができますが、参加者がうまく発話できるようになるには訓練が必要です。また、得られたデータの解析には技術と時間を要します。

§7. 本実習の進め方

本実習の目的は、より良いインタフェースを実現するための設計手順を学ぶことであり、以下の2種類のフェーズに分けて、全部で3つの実験を実施します。

#1 (1日目, 2日目) 認知課題実験

インタフェース設計におけるユーザ分析の一環として、人間の特性を考慮する必要性を2つの認知課題実験を通して理解します。各自で行った実験データを持ち寄り、統計ソフト R を用いて統計分析の手法を学びます。(6-2. 参照)

#2 (3日目, 4日目, 5日目) インタフェース実験

インタフェース設計手順を習得するために、与えられたインタフェース(銀行 ATM 機)を題材とし、ヒューマンファクタ分析→設計→プロトタイプング→ユーザ評価の一連のプロセスを体験します。(6-3. 参照)

7-1. 全体の流れ

1日目 認知課題実験(1) (9/30) *詳細は 7-2. を参照

- (1) 実習全体に関する説明を受けます。
- (2) 認知課題実験(1)の内容を講義により外観します。
- (3) 認知課題実験(1)を行います。
- (4) 各自のデータ、および2班分(14~15名)のデータを Excel に集計します。
- (5) 統計ソフト R に Excel データを読み込みます。
- (6) 統計的分析(t検定, 分散分析)について、講義により概観します(付録 A 参照)。
- (7) 認知課題実験(1)の結果を、R を用いて分析します。【t検定×5, 分散分析×2】
- (8) 分析結果について考察し、レポートを作成します。

提出物：レポート1 *詳細は 7-5. を参照

認知課題実験(1)の内容(実験目的, 方法, 結果と分析, 考察)
→Redmine に提出すること 10/7 (月) 12:45 〆切

2日目 認知課題実験(2) (10/7) *詳細は 7-3. を参照

- (1) インタフェース開発手順を講義により概観します。
- (2) 認知課題実験(2)の内容を講義により外観します。
- (3) 認知課題実験(2)(3種類の課題)を行います。
- (4) 各自のデータ、および2班分(14~16名)のデータを Excel に集計します。
- (5) 統計ソフト R に Excel データを読み込みます。
- (6) 認知課題実験(2)の結果を、R を用いて分析します。【t検定×6 (3実験×2分析)】
- (7) 分析結果について考察し、レポートを作成します。

提出物：レポート2 *詳細は 7-5. を参照

認知課題実験(2)の内容(実験目的, 方法, 結果と分析, 考察)
→Redmine に提出すること 10/14 (月) 12:45 〆切

3日目 インタフェース実験 (10/14) *詳細は 7-4. を参照

- (1) インタフェースの分析的評価手法を講義により概観します。
- (2) 与えられた ATM インタフェース (プロトタイプ(ATM-A)) を分析的に評価し, 問題点を抽出します。まず, 各自で問題点を抽出し, その後, 班員間で問題点を突き合わせます。
- (3) 上記をもとに, 各自で要求仕様を整理し, 設計を行います。設計の根拠を含め, 各自でどのように設計 (改修) するのかをメモしながら進めます。
- (4) インタフェース設計に用いるアプリケーション (Microsoft Visual Studio C#) の基本的操作について実習します (付録 B 参照)。
- (5) 決定した仕様に基づいて, インタフェースをコンピュータ上に実装します (ATM-X を作る)。

4日目 インタフェース実験 (10/21)

- (1) 決定した仕様に基づいて, インタフェースをコンピュータ上に実装します。
- (2) インタフェースの実験的評価手法を講義により概観します。
- (3) 実験的評価を行うにあたり, 評価方法 (客観評価, 主観評価) の内容を各班で決めて, TA・教員に確認を得ます。

提出物：レポート 3 *詳細は 6-4. を参照

プロトタイプ ATM-A の分析的評価, 要求仕様, 設計の内容 (ATM-X の写真や図)
→Redmine に提出すること 10/28 (月) 12:45 〆切

5日目 インタフェース実験 (10/28)

- (1) 実装した各自のインタフェースと, 新たに追加する 1 つのプロトタイプ (ATM-F), および ATM-A の計 3 つのインタフェースについて実験を実施します。
- (2) 統計的分析 (t 検定, 分散分析) について, 講義により概観します
- (3) 2 班分 (14~15 名) のデータを集計し, R を用いて分析 (分散分析) を行います。
- (4) 実験結果をまとめて考察し, 班の中で共有します。

提出物：実験報告書

授業中に作成して提出

7-2. 認知課題実験（1）

<目的>

私たちの見る「もの」の形や大きさは、外界の刺激対象の形や大きさの忠実な模写ではありません。私たちを取り巻く客観的世界と、それに対応する認知の世界の間には、多かれ少なかれなんらかの不一致が生じており、この不一致は、主として刺激布置と、主体的要因（経験、構え、欲求などの個人的条件）によって規定されています。

刺激配置による不一致の典型的な例として、幾何学的錯視があります。幾何学的錯視はいわゆる誤差的現象ではなく、刺激図形の幾何学的性質によって規則的に変化します。そこで生じる見えの歪みは、決して錯視図にのみ顕著に現れるのではなく、日常場面でも同様の現象はしばしば生じています。

そこで、ミュラーリヤー錯視図形を例として、刺激条件と認知との間の法則性を理解するとともに、認知特性に関する実験方法、分析方法を学びます

<方法>

(1) 刺激図形

ミューラーリヤーの錯視図形(図)を用います。標準刺激の主線は10cm, 矢羽の長さは3cmに固定されていますが、鋭角は5種類(60°, 120°, 180°, 240°, 300°)とします

標準刺激の折り目を内側に曲げて比較刺激を差し込み、スライドさせて使用します。標準刺激を差し替えて、5種類の条件を順次変化させます。

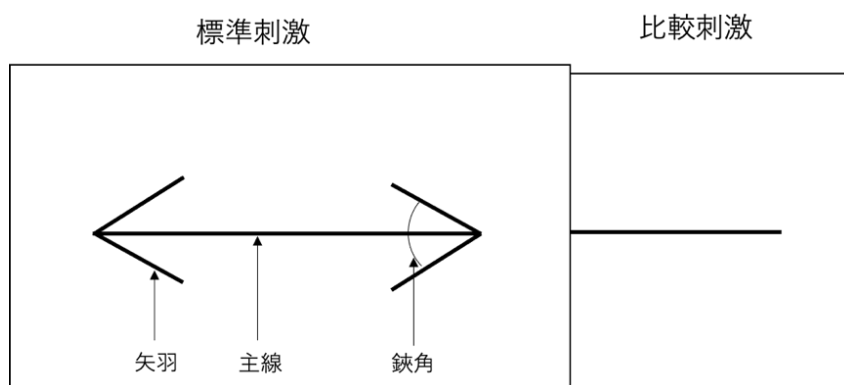


図 2 実験に使用する錯視図形

(2) 手続き

比較刺激を調整し、標準刺激の主線と等しい長さに見える、比較刺激の直線の長さ（主観的等価点 point of subjective equality: PSE）を求めます。PSE と主線の客観的な長さ（10cm）の差分が、錯視量 (I) となります

比較刺激が最も短く見える地点（数 mm しか見えない地点）から調節を開始する「上昇系列（A）」と、最も長く見える地点（比較刺激用紙の最も端の地点）から開始する「下降系列（D）」の2つの条件を設定します。それぞれ4回ずつ、計8回の実験を行い、8つのPSEを求めます。最小メモリ（1mm）の1/10の値を目分量で測りPSEとします。

5種類の刺激条件（欠角）の実施順が偏らないように、上昇系列（A）と下降系列（D）の試行順は指定の通りとします（＊moodle 掲示の Excel ファイル（認知課題実験（1）結果分析）を使用して、生まれ月に応じた各シートを利用すること）。

<結果の分析方法>

- (1) 個人毎に、5つの刺激条件それぞれについて、上昇系列(A)、下降系列(D)を別々にして、PSEの合計と平均値、錯視量(I)を求めます。
- (2) 上記(1)データを2班分(14~15名)集めます。
- (3) 2班分のデータについて、5つの刺激条件毎に、上昇系列(A)と下降系列(D)の錯視量(I)の平均値のどちらが有意に値が大きいか統計分析(t検定)を行って明らかにします。
Rでt検定を行うために必要な形式にExcelでデータ整理し、それぞれ新しいシート名をつけます(英字のみ)。上昇系列(A)と下降系列(D)のデータは2条件を横に並べるようにします(表6)。
*t検定については、付録A2.を参照。

表6 t検定のためのデータの並べ方

60°	
A	D
-0.4	-1.2
-2.8	-3.2
-1.2	-2.8
-2.0	2.0
-1.9	-1.0
-2.5	-1.8
-1.9	-2.0
-1.0	-2.1
-2.4	-2.7
-1.2	-2.1
-1.0	-0.9
-2.1	-2.0
-1.8	-2.0
-0.5	-3.0
-1.0	-0.9
-0.7	-2.7

- (4) 2班分のデータについて、上昇系列(A)、下降系列(D)毎に、5つの刺激条件間の錯視量(I)の平均値に有意な差があるか否か統計分析(一元配置(一要因)分散分析)を行って明らかにします。
Rを用いて分散分析を行うために必要な形式に、Excelでデータ整理し、それぞれ新しいシート名をつけます(英字のみ)。5種類の角度条件は縦に並べます(表7参照のこと)。*分散分析については、付録A3.を参照。5条件間の平均値の差異を視覚的に提示するために、グラフを作成すると良いでしょう。

表7 一元配置分散分析のためのデータの並べ方

“数値”として認識されるので、
“文字列”にする。あるいは。
「60A」など英数字にすると良い。

angle	I
60	-1.2
60	-0.9
60	-0.8
120	-0.9
120	-0.5
120	-0.3
180	0
180	0
180	-0.2
240	1.1
240	1.7
240	0.8
300	1.9
300	2.6
300	2.2

7-3. 認知課題実験（2）

<目的>

画面に提示された図（アイコン）に対応するキーをできるだけ早く押すという 3 種類の認知課題について、実験を行います。複数の参加者のデータを取得し統計解析を行うことによって、インタフェース設計の基礎としてのユーザ理解の重要性を学びます。

<方法>

3 つの実験では、それぞれ、図とキーとの対応に迷いやすいものと、そうでないものが含まれています。正しいキーを押すことができたとしても、反応が遅くなるものと、そうでないものがあるかもしれません。迷いにくいものや素早い反応ができる図（アイコン）やキーは、ユーザに負担をかけない良い対応と評価することができるでしょう。人間の認知的な特性を考えると、どのような時に間違いやすかったり、遅れがちだったりするのでしょうか。反応が遅くなったという自覚がなくとも、データを取得して統計的な分析を行うと、図（アイコン）とキーの対応によって、無意識に生じている反応の差異が正確にわかります。

- (1) 3 つの認知課題（task1, task2, task3）の実行ファイル（.exe）を moodle からダウンロードします。デスクトップ上など、わかりやすい場所に本実習用のフォルダ（フォルダ名は英字）を作っておくとい良いでしょう。
- (2) 実行ファイルをクリックすると、Visual Studio でプログラムが実行されます。全員、3 つの認知課題を実施します。ただし、実施順は、班員間で同じにならないように、カウンターバランスⁱⁱをとるようにしてください。
- (3) 各課題の最初に、その課題の説明が書かれているので、よく読んで指示の通り進めます。いずれの課題も、最初に練習課題があり、終了すると本試行に移ります。

<結果の分析方法>

- (1) 各課題は終了すると、実行ファイルと同じフォルダに、課題ごとのログファイル（.txt）が出力されます。集計し直すために、まず Excel を開きます。メニュー [ファイル] → [開く] でログファイルを指定し、「開く」を選択します。「選択したデータは区切り文字で区切られています」という画面を次に進み、「フィールドの区切り文字を指定してください」というダイアログで「カンマ」を選択して完了します。表 8 のように Excel で編集できる形式に表示できたら、「ファイル名を指定して保存」でファイル形式を Excel 有効ブック（.xls など）で保存しておきます。

表 8 ログファイルの例

Count	ReactTime	color	Text	Match	T/F
1	0.5660499	green	Go	1	1
2	0.3889566	green	Go	1	1
3	0.5861383	red	Stop	1	1
4	0.3907544	green	Stop	0	0
5	0.4329367	red	Go	0	1
6	0.6599866	green	Stop	0	1
7	0.4411837	red	Go	0	1
8	0.4048147	red	Go	0	1
9	0.4511053	green	Go	1	0
10	0.5098285	red	Go	0	1

ⁱⁱ カウンターバランス：刺激の提示順や課題の実施順等のバランスをとり、実験結果に関与する可能性のある変数の効果を除去すること

(2) ログファイルのタイトル行の意味は、表 9 を参照してください。

表 9 ログファイルのタイトル行の説明

	全試行	正答した試行のみ
Task1	一致条件 vs. 不一致条件	一致条件 vs. 不一致条件
Task2	一致条件 vs. 不一致条件	一致条件 vs. 不一致条件
Task3	遠隔条件 vs. 近接条件	遠隔条件 vs. 近接条件

(3) 各課題それぞれ 2 条件の反応時間 (ReactTime) について、どちらが有意に値が大きい (時間を長く要している) か統計分析 (t 検定) を行って明らかにします。いずれも、全試行分を対象にした分析と、正答した試行のみ対象とした分析を行います (表 10)。*t 検定については、付録 A2.を参照。

表 10 統計分析を行う条件

Task1	Task2	Task3
Count: 試行回数 ReactTime: 反応時間 [100ns=0.0001ms] color: 刺激の色 (red, green) Text: 刺激の文字 (Go, Stop) Match: 刺激の色と文字の一致 (一致:1 不一致:0) T/F: 反応の正誤 (正:1 誤:0)	Count: 試行回数 ReactTime: 反応時間 [100ns=0.0001ms] direction: 刺激の向き (right, left) Text: 刺激の文字 (Next, Back) Match: 刺激の向きと文字の一致 (一致:1 不一致:0) T/F: 反応の正誤 (正:1 誤:0)	Count: 試行回数 ReactTime: 反応時間 [100ns=0.0001ms] Text: 刺激の数字列 Space: スペースの有無 (有:1 無:0) T/F: 反応の正誤 (正:1 誤:0)

(4) 全ての課題は、まず個人ごとに各条件での平均値を算出します。その後、2 班分のデータを集めて集計します。

(5) R で t 検定を行うために必要な形式に Excel でデータ整理し、それぞれ新しいシート名をつけます (英字のみ)。反応時間 (ReactTime) のデータは 2 条件を横に並べるようにします (表 11)。正答した試行のみ対象にする分析は必要なデータのみ使い、別の列にまとめます。シートを分ける場合は、後でわかりやすい名前にします。

表 11 t 検定のためのデータの並べ方
一致条件 不一致条件

RT_1	RT_0
0.4740692	0.3861639
0.4905802	0.3515014
0.5710242	0.4959399
0.4527296	0.5534622
0.4851226	0.5337731
0.3644231	0.5183284
0.4268746	0.3855127
0.3398046	0.3841327
0.4199558	0.4134894
0.4181048	0.5654502
0.5233244	0.5337731
0.5438584	0.5183284
0.4348166	0.3855127
0.3937553	0.5937685
0.4892879	0.4584327

条件毎に列を分け、縦にデータを並べる

7-4. インタフェース実験

銀行 ATM のインタフェースを対象とし、ヒューマンファクタ分析→設計→プロトタイピング→ユーザ評価の一連のヒューマンインタフェース設計プロセスを体験します。最終的に、下記の課題を実行可能なユーザビリティに優れたインタフェースを実装し、評価を行います。

【課題】

- ・残高紹介
- ・振込

振込先：KIT 銀行・松ヶ崎支店・普通預金 ＊＊＊＊＊＊
振込金額：1 万 3 千円
暗証番号：＊＊＊＊

*口座番号（7 桁）と暗証番号（4 桁）は、授業内で周知します。

7-4-1. 分析的評価

与えられた ATM プロトタイプ (ATM-A) について、ヒューリスティック評価を行い、問題点を抽出します。

- (1) まず最初に、ATM-A を一度実施し、ログを取得します。取得したログデータは統計分析（5 日目）で使いますので取得できているか、確認しておいてください（ログファイルは、ATM-A ファイルと同じフォルダに保存されます）。
- (2) ログ取得後に、各自で分析的評価を行います。評価の基準となるガイドラインは、Nielsen の 10 項目（6-1-2. 参照）とします。
- (3) 各自で問題点を抽出し、その後、それを班員間で突き合わせて、問題点とその重要度をまとめます。抽出された問題点に対する考察は各自で行い、「自分が抽出できなかった問題」についても必ず考察してください。

7-4-2. 設計と実装

プロトタイプ (ATM-A) を基に、要求仕様に不足する条件を決定・追加し、再設計、実装します (ATM-X とする)。

- ・要求仕様は、その理由とともに箇条書きなどで整理してまとめます。
- ・ATM-A から大幅な変更がある場合は、画面設計図や状態遷移図を作成してください。大幅変更がない場合は、特に必要とはしません。
- ・再設計したインタフェースデザインについては、図（写真）とともに、設計上の留意点などもレポートに記述してください。（プログラミング環境の利用方法は付録 B. 参照）

7-4-3. 実験的評価

実装したインタフェースを利用して、以下の手順で評価実験を行います。

- (1) 各自が実装したインタフェース (ATM-X)、元のプロトタイプ (ATM-A)、新たに配付する 1 つのプロトタイプ (ATM-F) の計 3 種類のインタフェースを比較対象として、客観的評価、主観的評価を行います。
- (2) 評価に用いる指標（客観的評価、主観的評価（アンケート））内容は、班で検討し統一してください。
- (3) 各自が実装したインタフェースの評価は、班全員に行ってもらいます。実施順は、班員間で同じにならないように、カウンターバランスをとってください。ATM-F の評価は全員が同じものを行うため、データのみ交換し合って、班員分集計します。先に実施した ATM-A の客観的評価データも交換し合い、主観的評価は新たにデータ取得してください。
- (4) 各自が実装したインタフェース (ATM-X)、プロトタイプ ATM-A、ATM-F の計 3 種類のインタフェースの客観的評価、主観的評価のデータは、R を用いて分散分析（一元配置）を行います。デー

タの並べ方は、表 12 を参照してください。

*分散分析については付録 A3.を参照。ログを Excel から R にインポートする方法は認知課題実験と同様。

- (5) 統計的分析の結果に基づき、各自が実装したインタフェースについて考察（分析結果の理由・原因）、改善点の提案を行います。

表 12 一元配置分散分析のためのデータの並べ方

ATMtype	RT_1
A	0.56390
A	0.47409
A	0.36140
A	0.62905
A	0.45788
A	0.48293
<hr/>	
X	0.42243
X	0.25796
X	0.48879
X	0.36140
X	0.22359
X	0.67895
<hr/>	
F	0.42557
F	0.56671
F	0.42956
F	0.61915
F	0.98765
F	0.59455

条件（例：ATMのタイプ）分かりやすいラベルをつけて一列に並べデータと対応させる

要因（例：反応時間）1つなので（一元配置）、全データを一列に並べる

■客観的評価の指標

下記を用いるのが良いですが、具体的に、“どこからどこまでの操作時間”か、“何のエラー率”かを詳細に設定してください。

- 操作時間
- エラー率

■主観的評価の指標

質問紙（アンケート）による5件法、あるいは7件法を用います。各インタフェースの主観的満足度、快適度等を測定するために、詳細な項目を設定します。主観的評価指標の作成に関しては、6-2-2.（表4、表5）、付録Cを参照してください。

7-5. 提出物について

提出すべきものは、レポート3種類、最終日の実験報告書の計4点です。

レポートはいずれも、指定の期日までに、Redmine に提出してください。最終日の実験報告書は、当日回収します。期日に間に合わない提出物は大幅に減点します。レポートに不足箇所があるまま提出した場合は、その不足箇所の分量に比例した分だけ、期日遅れと同様に減点します。また、再提出の期日遅れも同様に扱いますので、注意してください。

レポート作成時には、「プロジェクト実習履修の手引き」を熟読の上、「ヒューマンインタフェース報告書チェックリスト」を使って自身で確認し、書式等に誤りがないようにしてください。

各提出物について、最低限含めるべき事項は下記のとおりです。

■レポート1（認知課題実験（1））＊PDF

- ・実験目的
- ・実験方法
- ・実験結果（個人のPSE合計、平均値、1、2班分のデータに関する基準統計量（平均値、標準偏差）、統計分析の結果）
- ・考察

■レポート1（認知課題実験（2））＊PDF

- ・実験目的
- ・実験方法
- ・3つの実験に関する結果（2班分のデータに関する基準統計量（平均値、標準偏差）、統計分析の結果）
- ・考察

■レポート3（プロトタイプATM-Aの分析的評価と設計）＊PDF

- ・プロトタイプATM-Aの分析的評価（個人の評価、班員の評価）
- ・要求仕様（理由）
- ・設計の内容（設計上の留意点、自分で実装したATM-Xインタフェースのスクリーンショット、あるいは図）

■実験報告書（インタフェース実験の結果まとめ）＊配布用紙に記入

<参考文献>

木下富雄編：教材心理学〔新装版〕，ナカニシヤ出版（2020）

ヒューマンインタフェース設計

Nielsen, J. : Usability Engineering, Academic Press (1993) [ユーザビリティエンジニアリング 原論第2版, 東京電機大学出版局, (2002)]

シュナイダーマン：ユーザーインタフェースの設計，日経BP出版センター（1995）

スミス他：利用者インタフェース・ソフトウェア設計ガイドライン，インタフェース・ソフトウェア研究会（1986）

ニューマン他：インタラクティブシステムデザイン，ピアソン（1999）

北原義典：イラストで学ぶヒューマンインタフェース，講談社（2019）

志堂寺知則：ヒューマンインタフェース，コロナ社（2019）

R, 統計的分析

南風原：心理統計学の基礎 — 総合的理解のために，有斐閣アルマ（2002）

山田剛史他：よくわかる心理統計，ミネルヴァ書房（2016）

緒賀郷志：Rによる心理・調査データ解析，東京図書（2010）

付録 A. 統計的分析

A1. 統計的分析とは

例えば、システム A とシステム B という 2 つの異なるシステムの操作時間を比較してどちらが良いかを確定するためには、2 つのシステムを利用する可能性のあるすべてのユーザに使ってもらって操作時間を収集し、それぞれのシステムの操作時間の差異を比較する必要があります。ところが、一般にそのように全ユーザを集めてくることは不可能です。そこで、ユーザの一部を選び出し、選ばれた参加者（被験者）にシステムを使ってもらい、その結果を利用して両者の操作時間を比較することになります。このように、母集団（全ユーザ）から標本（選ばれた被験者）を抽出して、その結果を元に母集団の傾向を確率的に推測する手法を、統計的分析（統計的仮説検定）といいます。

本実習では、ヒューマンインタフェースの評価で頻繁に用いられる統計的分析手法である、t 検定（平均値の差の検定）と分散分析（ANOVA）について説明します。

A2. 帰無仮説と対立仮説、有意水準

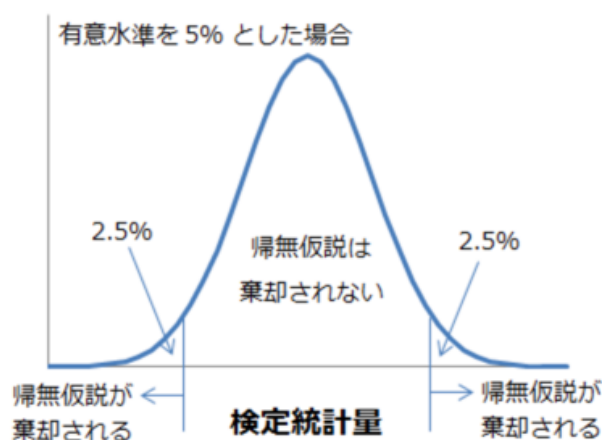
“ある仮説”が正しいかどうかの判断のために立てられる仮説を“帰無仮説”と呼びます。例えば、先の 2 つの異なるシステム間では「操作時間に差がない」という帰無仮説をおき、得られたデータと比較し、その仮説が成立する可能性が極めて小さいことを説明することで、システム間に差が存在することを説明します。仮説が正しくないと判断することを“棄却”といいます。また、帰無仮説に対立している証明したい仮説を“対立仮説”（例：操作時間に差があるとする仮説）といいます。

帰無仮説を棄却するための基準となる確率は、“有意水準”と呼ばれ、この数値は検定を行う前に決めておく必要があります。ヒューマンインタフェースの評価においては、有意水準 5%（95%の確率で帰無仮説が棄却されない）が用いられることが多いです。

検定統計量：帰無仮説が正しいと仮定したときに、観測した事象よりも稀なことが起こる確率を計算するための値です。“統計量”とよばれる場合もあります。

p 値：帰無仮説が正しいとした仮定とき、観測した事象よりも稀なことが起こる確率のことです。

p 値が有意水準より小さい時は、帰無仮説を棄却し、対立仮説を採択します。つまり、対立仮説が正しいと結論づけられます。



A3. t 検定（平均値の差の検定）

先の2つのシステム間の操作時間を比較するような場合，“t 検定”を用います。t 検定では、母集団の分散が等しい場合に用いるのが適切であり、2 標本のデータに対応があるかないかで方法が異なります。

$$\text{検定統計量} = \frac{\text{群 1 標本平均} - \text{群 2 標本平均}}{\text{標準誤差}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- 対応のあるデータ（標本間で対になったデータ）
例：ある薬を投与された患者の「投与前の血圧」と「投与後の血圧」
ある授業を受けた学生の「授業前のテスト得点」と「授業後のテスト得点」
- 対応のないデータ
例：「A 高校の学生の平均身長」と「B 高校の学生の平均身長」
あるテストについての「男子の平均点」と「女子の平均点」

検定統計量（t 値）が算出できたら、t 分布表ⁱⁱⁱから、その自由度^{iv}では有意差があると判定できるかの判定を行います（R などの統計ソフトでは、p 値を算出してくれるので、自分で t 分布表を確認する必要はありません）。2 標本の個々の平均値と標準偏差を求め、t 検定の結果を行い、その結果を論文に記す場合は、以下のような流れと書き方になります。

(1) 2 標本の個々の平均値と標準偏差を求める

例： システム A の操作時間の平均値は 4.34(標準偏差 1.09)、システム B の操作時間の平均値は 3.87 (標準偏差 0.88)

(2) 対応のない t 検定を実施

例： 自由度:11, t 値: 2.84, p 値: <.05 (5%より小さい)

(3) 検定の結果を解釈

例： 2 標本の間に有意な差が見られた。＝システム A はシステム B よりも、有意に時間を要する。

＝論文への記述方法例＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝

「システム A の操作時間の平均値は 4.34(標準偏差 1.09)、システム B の操作時間の平均値は 3.87 (標準偏差 0.88) であった。対応のない t 検定の結果、システム A とシステム B の操作時間は有意な差が見られた (t(11)=2.84, p<.05)。つまり、システム B はシステム A よりも、操作時間を必要としないシステムであることがわかった。」

＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝＝

ⁱⁱⁱ t 分布表：t 分布は自由度によって分布の形が変わるため、標準正規分布表とは異なる t 分布の値を一覧とする表が必要となります。

^{iv} 自由度：ある変数において自由な値をとることのできるデータの数。例えば、 n 個のデータ x_1, x_2, \dots, x_n があるとき、これらはどれも自由な値を取りうるので自由度は n です。ここで、もし平均値 $\bar{x} = a$ であるとき、平均値が変わらないようにするためには $n - 1$ 個の x_i は自由な値を取りうるが、 n 個目の x_n は自由な値を取ることはできません。その場合、自由度は $n - 1$ となります。一般に、 n 個のデータの間で k 個の条件があるとき、自由度は $n - k$ となります。

A3. 分散分析 (ANOVA)

3 標本以上の平均値の差の比較をする際には、“分散分析”が用いられます。t 検定は 2 標本の平均値の差の比較を行う点が分散分析と異なります。分散分析はデータの分散をもとにした分析方法です。分散分析では、標本ごとのデータのばらつきを元に、F 分布^vを用いて検定を行います。分散分析では、帰無仮説を「3 標本間の平均値に差がない」と設定し、対立仮説を「3 標本間の平均値に差がある」とします。3 標本間のいずれかに差があるということだけがわかりますが、どの標本とどの標本に差があるかは分散分析の結果だけではわかりません。標本間の差を明らかにするためには、さらに多重比較(post hoc test)が必要となります。

- 一元配置分散分析：1 つの因子からなるデータを分析する方法で、因子に含まれる水準間の平均値の差を見ることができます。例えば、ある学校の 1 組、2 組、3 組の数学のテストのデータがある場合、一元配置分散分析を用いて、1 組、2 組、3 組の数学のテストの平均点に差があるかどうかを検定できます。
- 二元配置分散分析：2 つの因子からなるデータを分析する方法で、各因子における水準間の平均値の差を見ることができます。また、2 つの要因が組み合わさることで現れる相乗効果の有無の確認もできます。例えば、薬 A、B、C をそれぞれ 10mg、20mg 投与した場合の効果についてのデータがある場合、二元配置分散分析を用いて薬の種類によって得られる平均値に差があるか、あるいは薬の投与量によって得られる平均値に差があるかどうかを検定できます。

【多重比較法】

水準間の平均値の差を比較する検定手法です。多重比較法を用いると、どの水準間に差があるかを調べることができます。水準間の平均値の差を調べる方法として、t 検定を利用することも考えられますが、t 検定で 2 群ごとの比較を単純に繰り返してしまうと、検定の多重性の問題^{vi}が生じます。多重比較法は、このような問題を回避するために、全体の有意水準をあらかじめ設定している有意水準になるよう調整を行う検定方法です。複数の方法が存在しますが、正規性と等分散性を仮定する多重比較法では、Tukey 法 (HSD 検定) が最もよく利用されています。

検定統計量 (F 値) が算出できたら、F 分布表から、その自由度では有意差があると判定できるかの判定を行います (R などの統計ソフトでは、p 値を算出してくれるので、自分で F 分布表を確認する必要はありません)。3 標本の個々の平均値と標準偏差を求め、分散分析を行い、その結果を論文に記す場合は、以下のような流れと書き方になります。

(1) 4 標本の個々の平均値と標準偏差を求める

例： システム A の操作時間の平均値は 4.34 (標準偏差 1.09)、システム X の操作時間の平均値は 2.95 (標準偏差 1.02)、システム F の操作時間の平均値は 3.87 (標準偏差 0.88) 一元配置分散分析を実施

例： 自由度: 3, 28, F 値: 8.24, p 値: <.05 (5%より小さい)

(2) 多重比較を実施

例： A と X, A と F, X と F 3 種類の組み合わせで実施検定の結果を解釈

例： システムの効果は有意であった。多重比較の結果、A と X の間以外の組み合わせで有意な差が認められた。

＝論文への記述方法例＝＝＝

「システム A の操作時間の平均値は 4.34 (標準偏差 1.09)、システム X の操作時間の平均値は 2.95 (標準偏差 1.02)、システム F の操作時間の平均値は 3.87 (標準偏差 0.88) であった。一元配置分散分析の結果、

^v F 分布：自由度が k_1, k_2 のカイ二乗分布 $\chi^2_1 \sim \chi^2(k_1), \chi^2_2 \sim \chi^2(k_2)$ が互いに独立である場合に、 $F = \frac{\chi^2_1/k_1}{\chi^2_2/k_2}$ から算出される F が従う確率分布のことです。このとき F は自由度 (k_1, k_2) の F 分布に従います。F 分布は t 分布と同様、自由度によって形が異なります。ただし、t 分布と異なり 2 つの自由度から分布の形が決まります。

^{vi} 多重性の問題：例えば、3 グループのうちどのグループの間に差があるのかを検討するために有意水準 5% で t 検定を 3 回繰り返すとして、それぞれの検定で、本当は母集団では差がないのに偶然差がでてしまう確率は 5% で、差がでない確率は 95% です。ここで、3 回の検定でどれも差がない (有意でない) という結果になる確率は $0.95 \times 0.95 \times 0.95 = 0.86$ となります。3 回の検定で少なくとも 1 回は有意になる確率は $1 - 0.86 = 0.14$ となり、検定全体では最初に設定した有意水準の 5% を超えてしまい、検定を繰り返すほど、どこかで偶然有意差が出てしまうという問題が生じます。

システムの効果は有意であった ($F(3,28)=8.24, p<.05$). Tukey の多重比較の結果, システム A と X, システム X と F, システム A と F の間に有意差があり, システム A が有意に操作時間が短いことがわかった。」

=====

A5. R の使い方

- (1) R を起動します.
 - (2) library(Rcmdr) と入力し, R コマンダーを開きます.
 - (3) Excel ファイルからデータを読み込みます.
- [データ] → [データのインポート] → [エクセルファイルから]

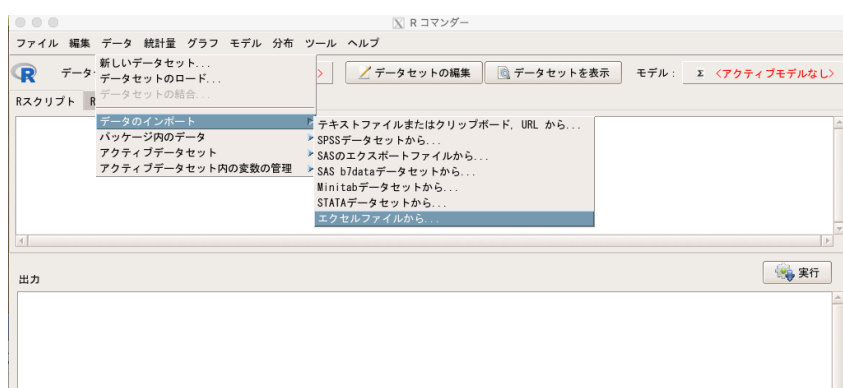


図 A-1 エクセルファイル読み込み方法

ファイルからのデータ

A5-1. 対応のある t 検定

[統計量] → [平均] → [対応のある t 検定]

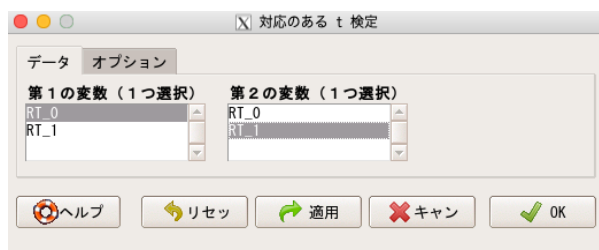


図 A-2 対応のある t 検定

```

Paired t-test
data: match and unmatch
t = -3.954, df = 31, p-value = 0.0004151
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.15661772 -0.05002743
sample estimates:
mean of the differences
 -0.1033226
    
```

検定統計量 (t値) 自由度 p値

図 A-3 R による t 検定結果の見方

A5-2. 一元配置分散分析

[統計量] → [平均] → [1 元配置分散分析]

→ 多重比較を行うために「2 組ずつの平均の比較 (多重比較)」をチェックする

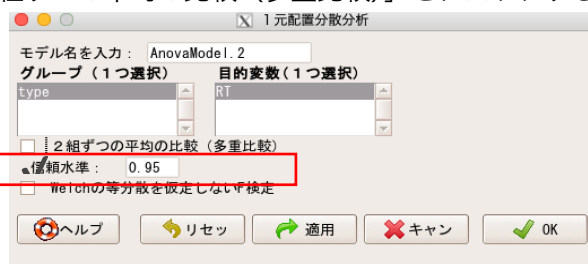


図 A-4 一元配置分散分析

```

> AnovaModel.6 <- aov(RT_1 ~ ATM.type, data=Dataset)
> summary(AnovaModel.6)

```

	DT	Sum Sq	Mean Sq	F value	Pr(>F)
ATM.type	2	64104	32052	16.19	0.0000557 ***
Residuals	21	41573	1980		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

> with(Dataset, numSummary(RT_1, groups=ATM.type, statistics=c("mean", "sd")))

```

	mean	sd	data:n
A	159.1225	38.53818	8
B	241.2175	44.72190	8
C	283.6250	49.53480	8

検定統計量 (F値) p値 有意水準0.1%で有意という意味 (* = 5%, **=1%, ***=0.1%)

自由度

各条件の平均値と標準偏差

図 A-5 R による一元配置分散分析結果の見方 (水準数が3つの例)

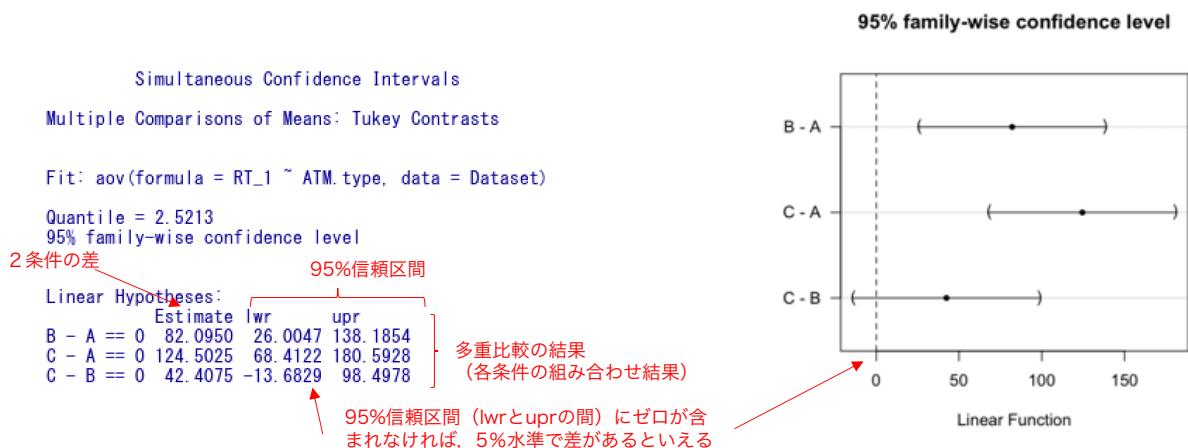


図 A-6 R による一元配置分散分析, 多重比較 (Tukey 法) 結果の見方 (水準数が3つの例)

付録 B. プログラミング環境の利用方法

本実習では、インタフェースの開発環境として、Windows 上で動作する統合開発環境 Microsoft Visual Studio で利用できる Microsoft Visual C# を利用します。ここでは、この統合開発環境の使い方について簡単に述べます。

B1. 利用手順

プログラミング環境の利用手順は以下になります。

- (1) コンピュータの電源を入れ、Windows を起動します。
- (2) 指定されたログオン名およびパスワードにより ログオンします。
- (3) Visual C# を起動し、プログラミングを行います。
- (4) 作業終了後は、コンピュータをシャットダウンし 電源を切ります。

B2. 本実験プロトタイプにおける特殊な設定

本実験で提供するプロトタイプは、実験に利用するためにいくつかの特殊な手続きを用いています。以下、そのそれぞれについて述べるので、参考にしてください。

B2-1. 新規フォームの継承による作成

本実験では、プロトタイプの各画面をひとつの既存フォームから継承する形式で作成されています。これは、

- ・ 操作ログ取得のための関数の実装
- ・ インタフェース間で利用できる (static 型) 共用変数の実装
- ・ フォームサイズの固定

の 3 つを実現するためです。与えられたプロトタイプにフォームを追加する場合は、次の作業を行います。

- (1) メニューの [プロジェクト]→[Windows フォームの追加]を選択
- (2) ダイアログから [継承されたフォーム] を選択 (通常、左列の上から 2 番目にある)
- (3) [追加 (A)] ボタンを押す
- (4) [継承ピッカー] ダイアログから、フォーム「Form_orig」を選択 (通常、列の最上段にある)
- (5) [OK] ボタンを押す

作成されたフォームはすべて「無題ページ」というタイトルになっているので、必要があれば各フォームの Text プロパティを修正すること。

B2-2. 次ウィンドウへの移動

次ウィンドウへの移動は、以下の 3 ステップからなります。ここで、FormNew は移動先のフォームの名称 (クラス名) をさします。

```
FormNew newForm = new FormNew( );
newform, show( );
this.Dispose( );
```

ただし、最初の画面 (つまり、実験開始ボタンのある画面) だけは、このようになっていません。これは、最初の画面を Dispose するとアプリケーションそのものが終了してしまうためです。アプリケーションの終了は、終了のためのボタンを押した処理で this.Close(); 関数を呼び出すことで行われます。それ以外のところで Close 処理を行ってはなりません^{vii}。

^{vii} アプリケーションの右上にある [x] ボタンを押してもアプリケーションは正常終了しますが、実験においてはこのボタンを押させる操作を要求しないこと。

B2-3. ログの取得

ログの取得は、以下の関数を、ログを取得したいインタフェースを操作した際に呼び出される処理の冒頭に記述することで行われる。

```
this.log(this, sender, e);
```

ログは、アプリケーションと同じ場所（実装中であれば、ATM_A\helloworld\bin\Debug）に出力されます。ファイル名は、log に長桁の数値のついたものであり、テキストファイルとして保存されます。各行には、フォームのクラス名、フォームの Text プロパティの値、操作したインタフェースの種別、操作したインタフェースの Text プロパティの値 6、操作を行った時間（アプリケーション起動時からの経過時間を 100 ナノ秒単位で表記）がこの順にコンマ区切りで保存される。なお、第 1 行はそれぞれの情報の意味を表すラベルです。

ログは、本実験プロトタイプを正しく終了させた場合にのみ出力されます。デバッグなど途中で強制終了させた場合には出力されません。出力されたログファイルは、認知課題（task1, task2, task3）のログファイルと同様に、Excel で読み込むことができます。

付録 C. 主観評価指標（アンケート）の例

操作性：

- ☐ 操作形式は一貫しているか
- ☐ 操作しやすいと感じられるか
- ☐ 関連するボタンはグループ化されているか
- ☐ 重要なボタンは強調されているか
- ☐ 操作の完了をユーザに知らせているか
- ☐ 次の操作がすぐわかるか
- ☐ 操作ミスをすぐ回復できるか

理解性：

- ☐ メッセージとボタン名は対応しているか
- ☐ メッセージは簡潔明瞭か
- ☐ ボタンの名称はシンプルか
- ☐ ボタンの機能はわかりやすいか
- ☐ わかりやすいラベルがついているか
- ☐ いろんなユーザに対して機能はわかりやすいか
- ☐ 普遍的に学習されたものとの関連性は強い
- ☐ 用語には一般的な言葉を使用しているか

識別性：

- ☐ 関連の深い機能は近くにあるか
- ☐ 使用できるボタンと使用できないボタンの区別ができていますか
- ☐ 副次的な要素は主要素と関連してはっきりと区別されているか

記憶性：

- ☐ 表示形式は一貫しているか
- ☐ 操作手順を覚えやすいか
- ☐ 適切なデフォルト値を用意しているか
- ☐ 視覚的イメージは統一されているか
- ☐ イメージは躍動的で生き生きしているか
- ☐ デザインは重要な箇所に焦点を当て注意を払っているか

親密性：

- ☐ オブジェクトはユーザにとって親しみのあるものか
- ☐ オブジェクトはユーザの仕事あるいは日常の環境に共通してあるものか
- ☐ 現実世界について知っていることを操作に適用できるか

可視性：

- ☐ 表示部の背景と文字とのコントラストは十分か
- ☐ 表示文字の形と太さは十分か
- ☐ 視覚に異常のあるユーザ（近視、遠視、乱視、色盲）でもオブジェクトを識別できるか
- ☐ 表示されるあらゆるオブジェクトの大きさは適切か
- ☐ すべてのオブジェクト、すべての線分、すべての点が必要か

誘目性：

- ☐ 各オブジェクトは視覚的に均衡がとれており安定感があるか
- ☐ パターン、明るさは調和のとれた組み合わせになっているか
- ☐ あいまいさをなくすように付加的なキュー（ラベル、アイコン、文章）を使っているか