

Rコマンダーを用いた統計解析 (17)

担当 繁柵 算男 教授

アシスタント 森 一将

1

心理学演習資料

今日の演習事項

- 分散分析
 - 一元配置分散分析
 - 被験者間／対応のない分散分析(おさらい)
 - 被験者内／対応のある分散分析

2

心理学演習資料

データ概要

- 車の広告キャンペーンの効果測定
 - ある車メーカーが新車Aの広告キャンペーンを全国の主要都市で実施
 - 実施条件は同条件(金額／方法／回数)
 - キャンペーン実施後、アンケートで新車Aへの興味度(買いたいと思うかどうか)を調査
 - 東京、大阪、名古屋、仙台、福岡、札幌で無作為抽出を行い各9人ずつにアンケート
 - 各都市の文化差等の違いで興味度に差が生じているか？

出典: Hoaglin, D., Mosteller, F., and Tukey, J. (1991)より改変

3

心理学演習資料

一元配置分散分析

1. 帰無仮説、対立仮説を設定
2. 統計的仮説検定に用いられる標本統計量を選択
3. 棄却域、有意水準(判断の基準になる確率のしきい値)の設定
4. 実際のデータから標本統計量を計算
5. 検定統計量の値が...
 - 棄却域の値なら、帰無仮説を**棄却**(対立仮説を**採択**)する。
 - そうでなければ、帰無仮説を**保持**する。

4

心理学演習資料

4.実際のデータから標本統計量を計算

2010年1月28日

- 「統計量」→「平均」→「一元配置分散分析」



何を基準に群(グループ)分けをするのか
→「要因」ともいう

5

心理学演習資料

4.実際のデータから標本統計量を計算

2010年1月28日

```
Bartlett test of homogeneity of variances

data: 興味度 by 都市
Bartlett's K-squared = 6.2355, df = 5, p-value = 0.284

> AnovaModel.1 <- aov(興味度 ~ 都市, data=Dataset)
> summary(AnovaModel.1)
          Df Sum Sq Mean Sq F value    Pr(>F)
都市         5 10.9457   2.1891  4.8294 0.001175 **
Residuals   48 21.7581   0.4533
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

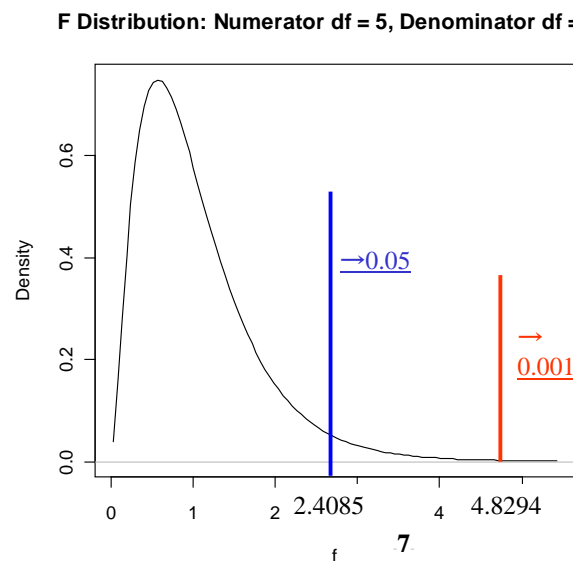
> numSummary(Dataset$興味度, groups=Dataset$都市, statistics=c("mean", "sd"))
      mean      sd n
札幌 12.125 0.200  5
仙台 13.250 0.300  5
大阪 12.500 0.250  5
東京 13.250 0.300  5
福岡 13.750 0.350  5
名古屋 13.250 0.300  5
```

- 計算されたF統計量の値=4.8294
- 自由度(5,48)のF分布モデルの当てはめ
- p値: 有意水準と比較して有意かどうかの判断基準

6

5検定統計量の値が.

2010年1月28日



- 帰無仮説(確率モデル)の元では実際のデータの値が極端なものである
- そのような値が生じるのはまれである
- 帰無仮説は正しくない(棄却する)
⇒対立仮説が正しい

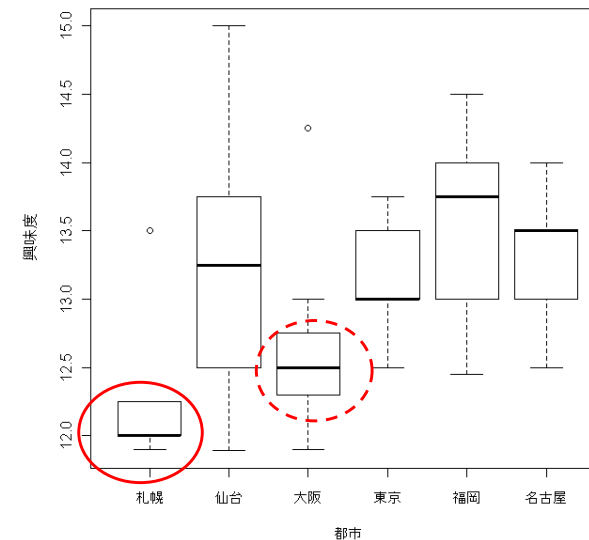
7

心理学演習資料

箱ひげ図(再掲)

2010年1月28日

いくつかの都市の興味度は平均より下がる?



心理学演習資料

1. 帰無仮説、対立仮説を設定
2. 統計的仮説検定に用いられる標本統計量を選択
3. 棄却域、有意水準(判断の基準になる確率のしきい値)の設定
4. 実際のデータから標本統計量を計算
5. 検定統計量の値が...
 - 棄却域の値なら、帰無仮説を**棄却**(対立仮説を**採択**)する。
 - そうでなければ、帰無仮説を**保持**する。

練習: 結果を見ながら対応する手順を考えてみる

- いくつかの方法がある
- 先行研究などを調べて適宜判断する
- TukeyのHSD法
 - Rguiより以下の命令を実行
 - TukeyHSD(aov(興味度~都市,data=Dataset))
 - ↑ ↑ ↑
 - 目的変数 要因 データセット名

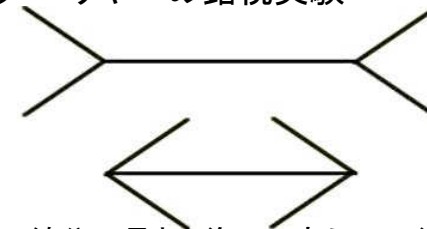
```
> TukeyHSD(aov(興味度~都市,data=Dataset))
Tukey multiple comparisons of means
95% family-wise confidence level
```

Fit: aov(formula = 興味度 ~ 都市, data = Dataset)

調整済みのp値
簡便法で判断

\$都市	diff	lwr	upr	p adj
仙台-札幌	1.04444444	0.10249449	1.98640444	0.0217928
大阪-札幌	0.41111111	-0.53084884	1.35307111	0.7860914
東京-札幌	0.99444444	0.05248449	1.93640444	0.0329734
福岡-札幌	1.28333333	0.34137338	2.22529333	0.0024794
名古屋-札幌	1.10666667	0.16470671	2.04862666	0.0127365
大阪-仙台	-0.63333333	-1.57529329	0.30862666	0.3599444
東京-仙台	-0.05000000	-0.99195995	0.89196000	0.9999856
福岡-仙台	0.23888889	-0.70307106	1.18084888	0.3739595
名古屋-仙台	0.06222222	-0.87973773	1.00418222	0.9999573
東京-大阪	0.58333333	-0.35862662	1.52529333	0.4520106
福岡-大阪	0.87222222	-0.06973773	1.81418222	0.0842029
名古屋-大阪	0.69555556	-0.24640440	1.63751555	0.2606185
福岡-東京	0.28888889	-0.65307106	1.23084888	0.9420131
名古屋-東京	0.11222222	-0.82973773	1.05418222	0.9992262
名古屋-福岡	-0.17666667	-1.11862662	0.76529333	0.9933038

- ミュラー・リヤーの錯視実験



上下の線分の長さを徐々に変えていく
差があると判断した場合の線分の長さの差は何に
起因するか？

一人の被験者に対し複数(羽の角度を変える)実験

データ

2010年1月28日

- 被験者No × 差が生じた角度
- Rで取り入れ可能なデータにすること

被験者	r30	r60	r90	r120
1	42	39	36	34
2	22	17	15	8
3	35	32	25	25
4	34	30	22	20
5	40	33	28	23
6	34	30	23	28

心理学演習資料

演習

2010年1月28日

- 前頁の表データをすべて列(タテ)形式のデータに直せ

心理学演習資料

データ

2010年1月28日

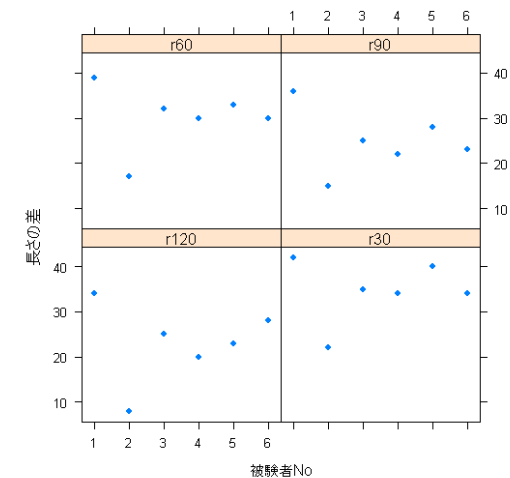
被験者No	角度	長さの差
1	r30	42
2	r30	22
3	r30	35
4	r30	34
5	r30	40
6	r30	34
1	r60	39
2	r60	17
3	r60	32
4	r60	30
5	r60	33
6	r60	30
1	r90	36
2	r90	15
3	r90	25
4	r90	22
5	r90	28
6	r90	23
1	r120	34
2	r120	8
3	r120	25
4	r120	20
5	r120	23
6	r120	28

心理学演習資料

簡単に状態を確認する

2010年1月28日

- グラフー条件付散布図



心理学演習資料

被験者内分散分析

2010年1月28日

- 同じ被験者が異なる条件で実験を行った結果のデータ
- 「対応のある」平均値の差の検定
- 3群以上の場合

被験者	r30	r60	r90	r120
1	42	39	36	34
2	22	17	15	8
3	35	32	25	25
4	34	30	22	20
5	40	33	28	23
6	34	30	23	28

心理学演習資料

(球面性の仮定)

2010年1月28日

- 対応のあるすべての水準においてデータの分散が等しい
- 対応のあるt検定における等分散の仮定に相当
- Rのデフォルト関数ではまだ実装されていない
- 必要な場合はAnova君(筑波大)等の関数をインポートすること

心理学演習資料

一元配置分散分析(被験者内)

2010年1月28日

1. 帰無仮説、対立仮説を設定
2. 統計的仮説検定に用いられる標本統計量を選択
3. 棄却域、有意水準(判断の基準になる確率のしきい値)の設定
4. 実際のデータから標本統計量を計算
5. 検定統計量の値が...
 - 棄却域の値なら、帰無仮説を**棄却**(対立仮説を**採択**)する。
 - そうでなければ、帰無仮説を**保持**する。

心理学演習資料

Rの操作

2010年1月28日

- コマンドラインから以下の分析を実行
summary(aov(長さの差~被験者No+角度))

```
R Console
> summary(aov(長さの差~被験者No+角度))
Anova table:
          Df Sum Sq Mean Sq F value Pr(>F)
被験者No    5  401.467  80.293  2.7420 0.01910 .
角度         1  1150.85  1150.85  39.0400 0.00000 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R version 2.10.1 (2010-04-17)
```

心理学演習資料

- 同じ被験者に対して条件を変えて実験を行う場合
- 必要な統計手法を確認
- 必要に応じてRの関数をインポート
- (可能ならば)SASやSPSSを用いてもよい

Rコマンダーを用いた統計解析 (19)

担当 繁柵 算男 教授

アシスタント 森 一将

1

心理学演習資料

前回の補足

- 指数表示＝極端に大きい／小さい数を表現するとき利用

$$1.169E-05 = 1.169 \times 10^{-5} = 1.169 \times \frac{1}{10^5}$$

指数
10のカタにのる数字

指数がマイナス
分母に移動

=0.0000169

$$2.356E+05 = 2.356 \times 10^5 = 2.356 \times 10000 = 235600$$

指数
10のカタにのる数字

指数がプラス
そのままケル

心理学演習資料

今日の演習事項

- 分散分析
 - 二元配置分散分析
 - 被験者間計画／対応のない(2回目)
心理実験でよく使う手法なので2回演習します
 - 被験者内計画／対応のある
 - 混合計画

3

心理学演習資料

データ概要

- ミネラルウォーターの味(おいしさ)の調査
- 30人の被験者に対し、以下のミネラルウォーターを無作為に割り当て、おいしさを評定してもらった
- ミネラルウォーターの温度:2水準
 - 冷蔵庫で冷たく冷やす:カテゴリ1
 - 常温のまま出す:カテゴリ2
- ミネラルウォーターの種類:3水準
 - アメリカ産(アラスカ):カテゴリ1
 - スイス産(アルプス):カテゴリ2
 - 日本産(富士山):カテゴリ3

(出典:南山大 神谷先生HP) 心理学演習資料

データ

2010年1月28日

- Excelを用いて以下の評価データをRに読み込める形に直す

1:冷蔵庫で冷やす			2:常温		
1:アメリカ	2:スイス	3:日本	1:アメリカ	2:スイス	3:日本
6	10	11	5	7	12
4	8	12	4	6	8
5	10	12	2	5	5
3	8	10	2	4	6
2	9	10	2	3	4

心理学演習資料

ヒント: 前回の事例

2010年1月28日

要因 データ(従属変数)

勉強方法	勉強時間	テストの点
テキスト	少ない	84
テキスト	少ない	105
テキスト	少ない	93
テキスト	多い	140
テキスト	多い	121
テキスト	多い	150
テキスト	普通	137
テキスト	普通	143
テキスト	普通	127
演習	少ない	121
演習	少ない	151
演習	少ない	141
演習	普通	276
演習	普通	225
演習	普通	260
演習	多い	245
演習	多い	208
演習	多い	267

水準
(今回はコード化)

心理学演習資料

データ入力解答例

2010年1月28日

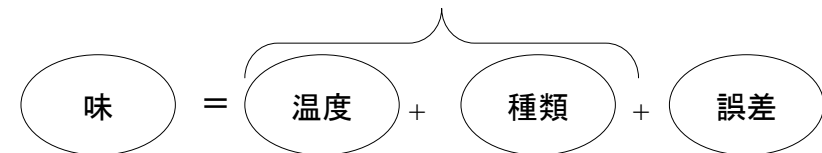
温度	飲料種類	味
1	1	6
1	1	4
1	1	5
1	1	3
1	1	2
1	2	10
1	2	8
1	2	10
1	2	8
1	2	9
1	3	11
1	3	12
1	3	12
1	3	10
1	3	10
2	1	5
2	1	4
2	1	2
2	1	2
2	1	2
2	2	7
2	2	6
2	2	5
2	2	4
2	2	3
2	3	12
2	3	8
2	3	5
2	3	6
2	3	4

心理学演習資料

二元配置分散分析

2010年1月28日

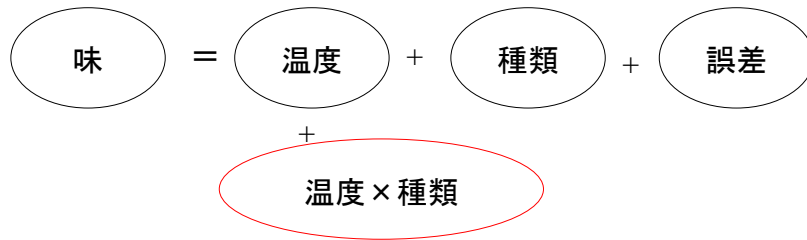
- 今回の事例 「主効果」という



味に影響を与えるのは
温度と種別の2つの要因
2つの要因が対象(味)に影響を与えている

心理学演習資料

- 二元配置分散分析で出てくる独特の要因



2つの要因の水準がある特定の組み合わせのときのみ発生する作用
 (例)
 アラスカの水 + 冷蔵庫で冷やす
 → アラスカの水は他の2種類よりも硬度が高く、冷やすと特においしく感じる

1. 帰無仮説、対立仮説を設定
2. 統計的仮説検定に用いられる標本統計量を選択
3. 棄却域、有意水準 (判断の基準になる確率のしきい値) の設定
4. 実際のデータから標本統計量を計算
5. 検定統計量の値が...
 - 棄却域の値なら、帰無仮説を**棄却** (対立仮説を**採択**) する。
 - そうでなければ、帰無仮説を**保持**する。

1. 帰無仮説、対立仮説を設定

- 帰無仮説

主効果

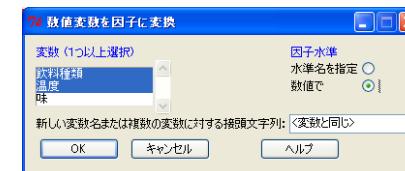
- 温度
 - 「冷蔵庫」「常温」群の味の母平均は等しい
- 種類
 - 「アメリカ産」「スイス産」「日本産」群の味の母平均は等しい

交互作用

- 温度 × 種類
 - 温度と種類の特定の組み合わせにより味の母平均が変わる特定の効果はない

分散分析の前処理

- Rに特有の前処理がある
- 各水準を数値(コード)化した場合のみ実行
- Rコマンド
 「アクティブデータセット内の変数の管理」
 → 「数値変数を因子に変換」
- 要因に対応する変数を選択し、「数値で」



4.実際のデータから標本統計量を計算

2010年1月28日

- コマンドラインより以下のコマンドを実行

```
azi<-aov(味~温度+飲料種類+温度*飲料種類
,data=Dataset)
summary(azi)
```

心理学演習資料

5検定統計量の値が...

2010年1月28日

```
> summary(azi)
          Df Sum Sq Mean Sq F value    Pr(>F)
温度      1  67.500   67.500  21.3158  0.00011 ***
飲料種類  2 155.000   77.500  24.4737 1.608e-06 ***
温度:飲料種類  2  15.000    7.500   2.3684  0.11515
Residuals 24  76.000    3.167
```

心理学演習資料

5検定統計量の値が...

2010年1月28日

- 5%水準で棄却された仮説
 - 温度(棄却)
 - 「冷蔵庫」「常温」群の味の母平均は等しい
 - 種類(棄却)
 - 「アメリカ産」「スイス産」「日本産」群の味の母平均は等しい
 - 温度×種類(保持)
 - 温度と種類の特定の組み合わせにより味の母平均が変わる特定の効果はない

心理学演習資料

どのように結論を出すか？

2010年1月28日

- 温度の違いに関わらず味の評価が等しいとはいえない
- 種類の違いに関わらず味の評価が等しいとはいえない
 - じゃあ、具体的にどの水準間に差が出たか？
 - 多重比較

心理学演習資料

多重比較

2010年1月28日

- TukeyHSD(azi)\$飲料種類

```
> TukeyHSD(azi)$飲料種類
      diff      lwr      upr      p adj
2-1   3.5  1.51260291  5.487397  5.440519e-04
3-1   5.5  3.51260291  7.487397  1.111275e-06
3-2   2.0  0.01260291  3.987397  4.835626e-02
```

どこが有意(平均に差があるとみなす)?

心理学演習資料

多重比較

2010年1月28日

- TukeyHSD(azi)\$温度

```
> TukeyHSD(azi)$温度
      diff      lwr      upr      p adj
2-1   -3 -4.341093 -1.658907  0.0001100036
```

要因の水準が2つ

→実は分散分析表の検定結果と多重比較の結果はほぼ等しくなる

```
> summary(azi)
      Df Sum Sq Mean Sq F value    Pr(>F)
温度      1  67.500   67.500  21.3158  0.00011 ***
飲料種類  2 155.000   77.500  24.4737  1.608e-06 ***
温度:飲料種類  2  15.000    7.500   2.3684  0.11515
Residuals 24  76.000    3.167
```

心理学演習資料

まとめ

2010年1月28日

- 二元配置分散分析
 - 目的変数に影響を与える要因が2つあるときの平均値の検定
- 交互作用の出現
- 主効果の帰無仮説が棄却されたとき
 - 多重検定の実施
- 心理実験では多く使われるため手順と結果の読み方を覚えることが望ましい

心理学演習資料