

多重比較と多重検定

粕谷英一

(九大・理・生物)

印象

多重比較と多重検定

何か補正しなければ文句を言われる

補正すれば有意になりにくい

多重比較と多重検定

誤りの海

多重比較と多重検定

補正が必要だが補正しない
不適切な方法で補正する
補正をすべきでないのに補正する

多重検定と多重比較をめぐる現状

前提

なぜ補正するのか

よくある誤り

補正すべきなのに補正しない

不適当な方法で補正する

補正すべきでないのに補正する

Fisher

Bonferroni法

最小有意差法 (Fisher's LSD, protected LSD)

Tukey

HSD (Honestly Significant Difference)、

T-procedure

Fisher

分散分析

尤度

Tukey

探索的データ解析

ジャックナイフ

統計的方法

検定 & モデル選択

検定

モデル選択

(予測の最適化)

検定とは

検定 & モデル選択

帰無仮説 差なし、相関なし
対立仮説

帰無仮説が正しい可能性はきわめて小さい



帰無仮説を否定（棄却）

確率モデル（帰無仮説）を、ある基準で捨てることにより、反対のモデルを採用

検定

帰無仮説が正しい可能性が小さいと、捨てる

有意水準

有意確率

危険率

帰無仮説が正しいのに、捨ててしまう

第 1 種の誤り

type1 error rate

第 2 種の誤り

type2 error rate

帰無仮説が誤りなのに、捨てない

1-検出力

モデル選択

例 . AIC (赤池情報量基準)

モデル1

データ

モデル2

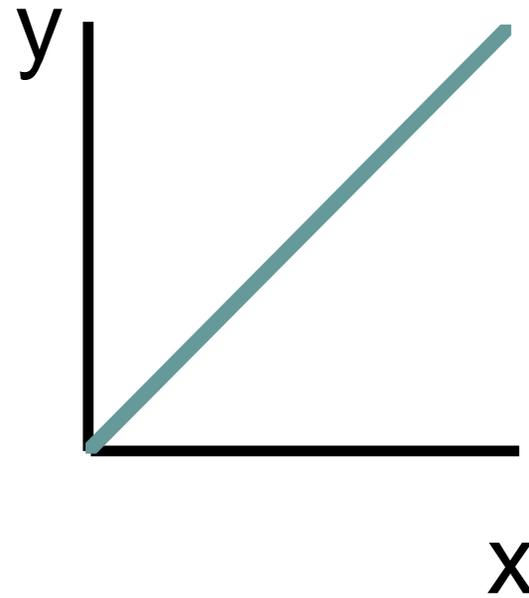
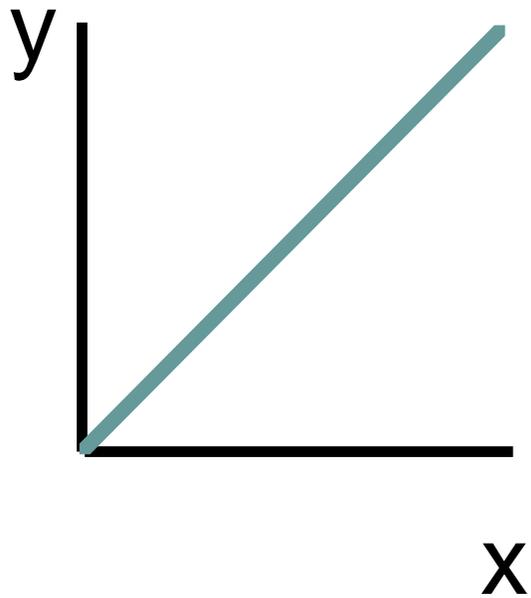
複雑なモデルは適合しやすい

予測の最適化

同じデータ、同じ形式のデータ
について統計的推論はいろいろ
可能

xはyより大きい

xが大きい個体はyも大きい



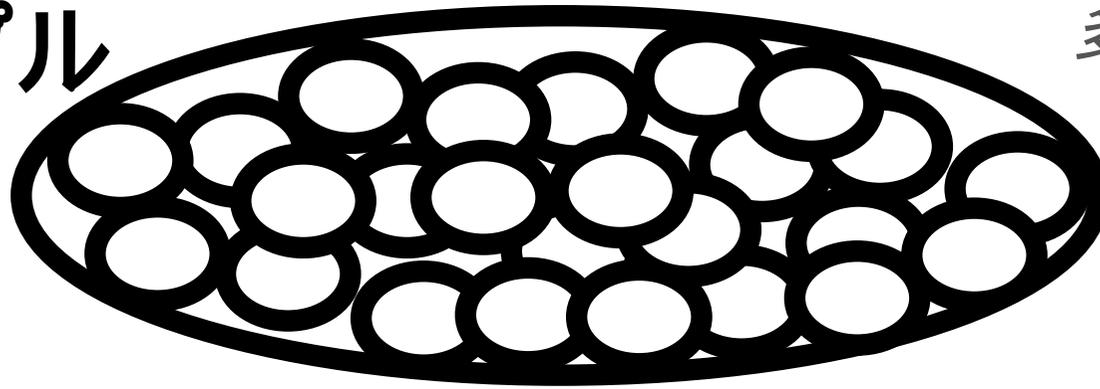
なぜ

多重比較・多重検定するとき

有意水準を調整するのか

サンプル

多重性の例



個々のサブサンプルで検定

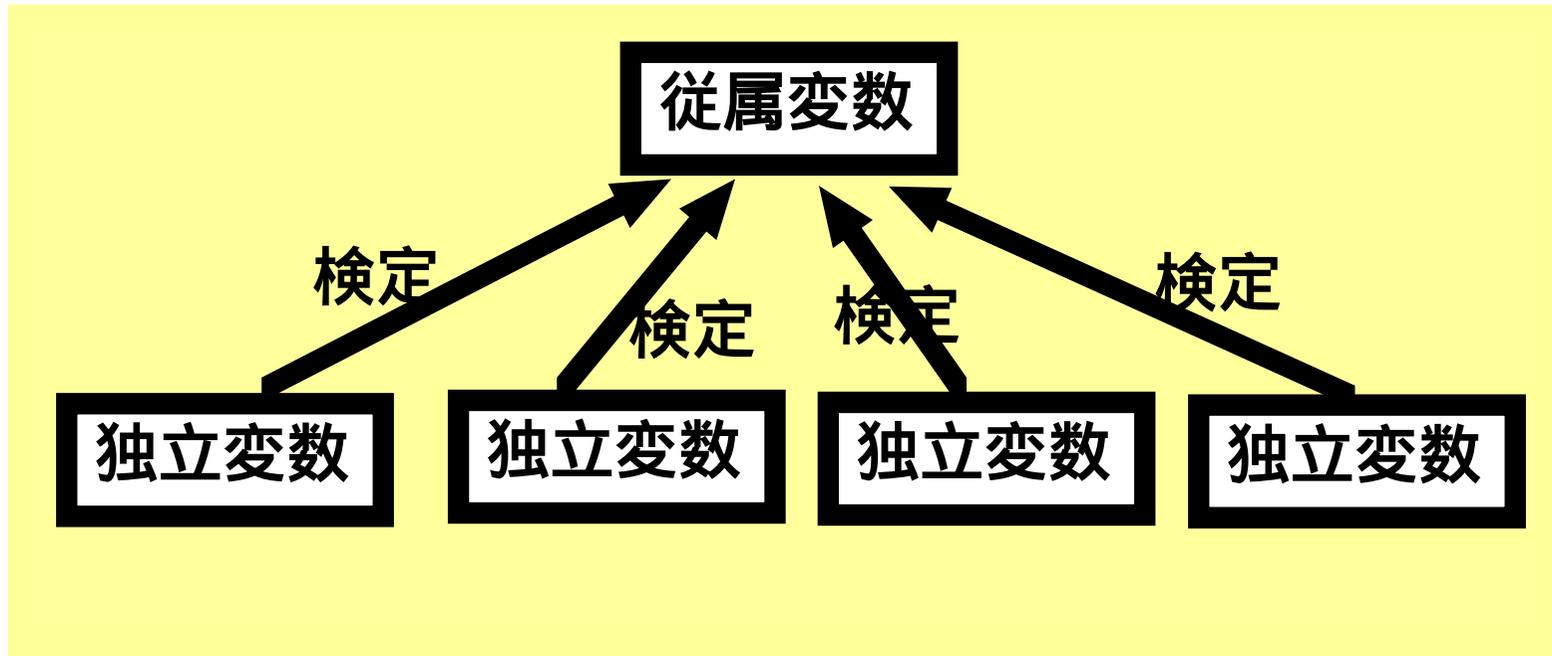


有意なところを使う

例 政治家の血液型はOが多い

たくさんのサブサンプルで検定すれば、どれかは有意

多重性の例



たくさんの独立変数で検定すれば、どれかは有意

Tukey

多重性の例

たくさんの処理

処理1 処理2 処理3 処理4 . . . 処理k

一元配置分散分析

2つの処理間で比べる

$k(k-1)/2$ 個の検定

たくさんの処理間で検定すれば、どれかは有意

多重比較

1 元配置分散分析 - k水準の処理の単純な比較
構造のある仮説を持たない



2つずつ比べて検定 $k(k-1)/2$ 個の検定

どの処理の母平均も同じ（帰無仮説が正しい）でも、

1 検定あたり0.05（5%）の割合で誤って有意と判定

多重比較

検定の個数を考慮に入れるべきだ

1 元配置分散分析 - k水準の処理の単純な比較
構造のある仮説を持たない

2 つずつ比べて検定

$k(k-1)/2$ 個の検定

一連の検定

検定のfamily

キーワード

第1種の誤り

type1 error rate

有意水準

有意確率

危険率

帰無仮説が正しいのに、
捨ててしまう確率

一連の検定 (family) のどこかで 第 1 種の誤りをおかす確率

FWE

キーワード

familywise error rate
experimentwise error rate
とも言う

1 つの検定で 第 1 種の誤りをおかす確率

PCE

キーワード

per-comparison error rate

FWEのコントロール

PCEでなくFWEを5%にすべきだ

FWEのコントロールの考え方は
実はよく使われている

t 検定と分散分析

予備知識

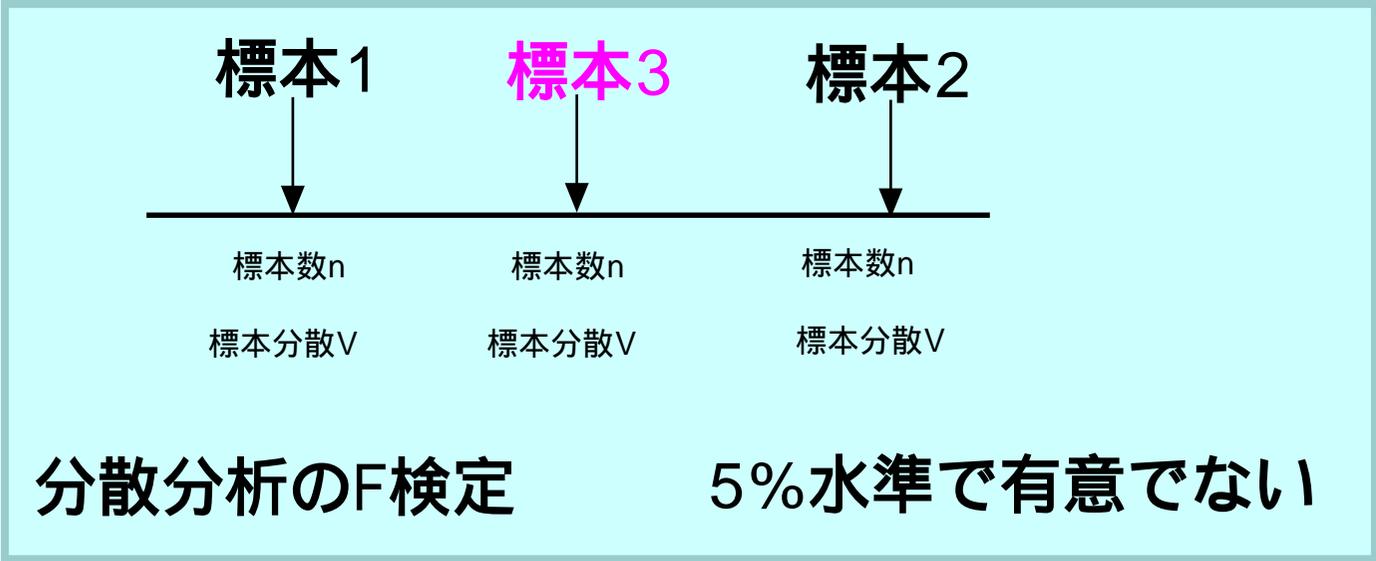
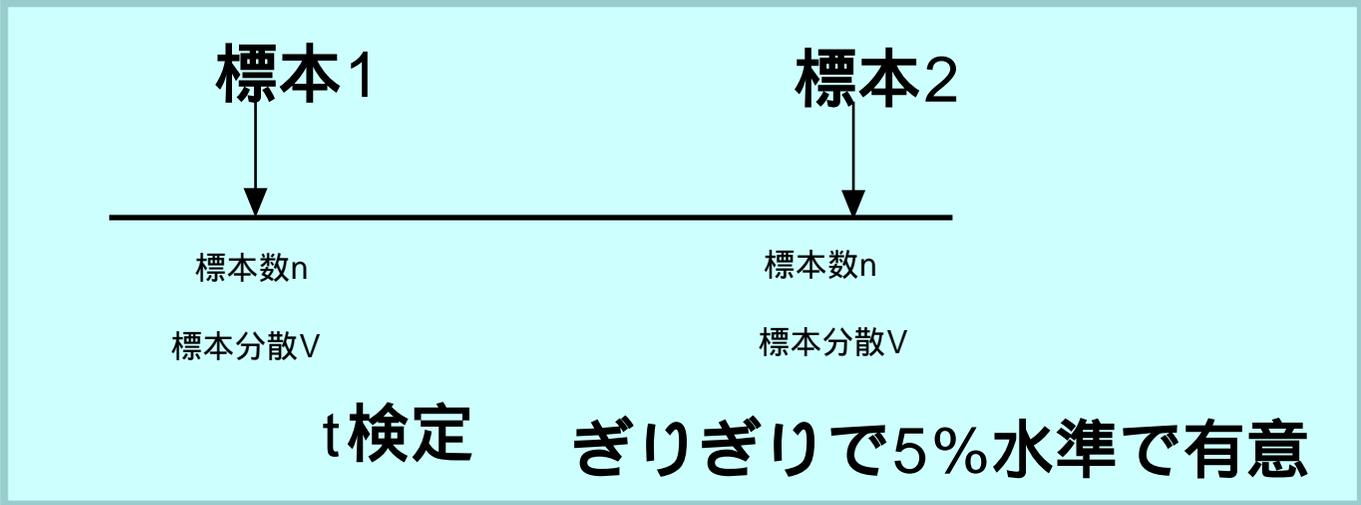
t 検定

分散分析のF検定の分子側自由度が 1 の場合

t分布とF分布

正規分布・等分散

検定統計量 t の2乗 = 検定統計量 F



一元配置分散分析の典型的な全ペア比較の場合

FWEのコントロールではなくPCEのコントロール

| | | |
|------------------------|----------------|---------|
| 分散分析 Kruskal-Wallis | も使わない もおかしい | と言うのと同じ |
|------------------------|----------------|---------|

帰無仮説が複雑

FWEの弱いコントロール

全処理が同じ時

FWEの強いコントロール

すべてのパターン

1 元配置分散分析 構造のある仮説を持たない
2 つずつ比べて検定 $k(k-1)/2$ 個の検定

そんな問題ばかりではない

歴史一時間がかかっている

補正すべきなのに補正しない

一元配置分散分析－処理（水準）間
多重比較 多くの処理（ペア）

サブサンプル 多くのサブサンプル

重回帰、多要因の分散分析
（要因実験） 多くの独立変数

不適当な方法で補正する

FWEがコントロールされない

1 元配置分散分析 構造のある仮説を持たない
2 つずつ比べて検定 $k(k-1)/2$ 個の検定

FisherのLSD (protected LSD)

弱

Newman-Keuls

弱

Duncanの多重範囲検定

不適当な方法で補正する

familyに属する検定の個数

有意かどうかにか

強く影響

そもそも検定か

ある決まった処理（例、コントロール）と比べる

処理1 処理2 処理3 処理4

処理0

Dunnett

最大（最小）の処理と比べる

処理1

処理2

処理3

処理4

処理5

傾向性仮説



仮説

処理1 処理2 処理3 処理4 処理5

isotonic regression 分散分析

Jonckheer Kruskal-Wallis

相関や回帰

傾向性仮説

「全体を通じてどうなっているか」で充分
処理ペア比較は不要

仮説

処理1 処理2 処理3 処理4 処理5

あるペアで差がある 当然差があるペア

傾向性仮説



仮説

処理1 処理2 処理3 処理4 処理5

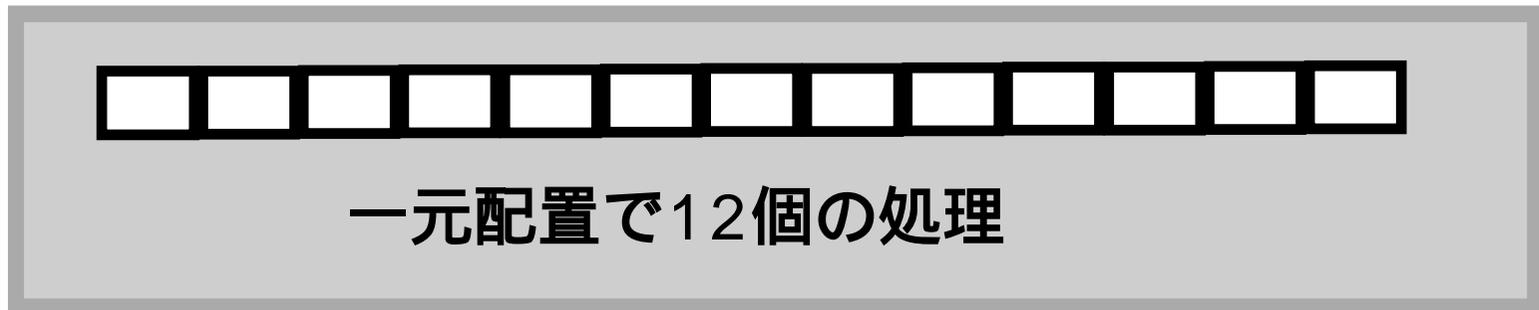
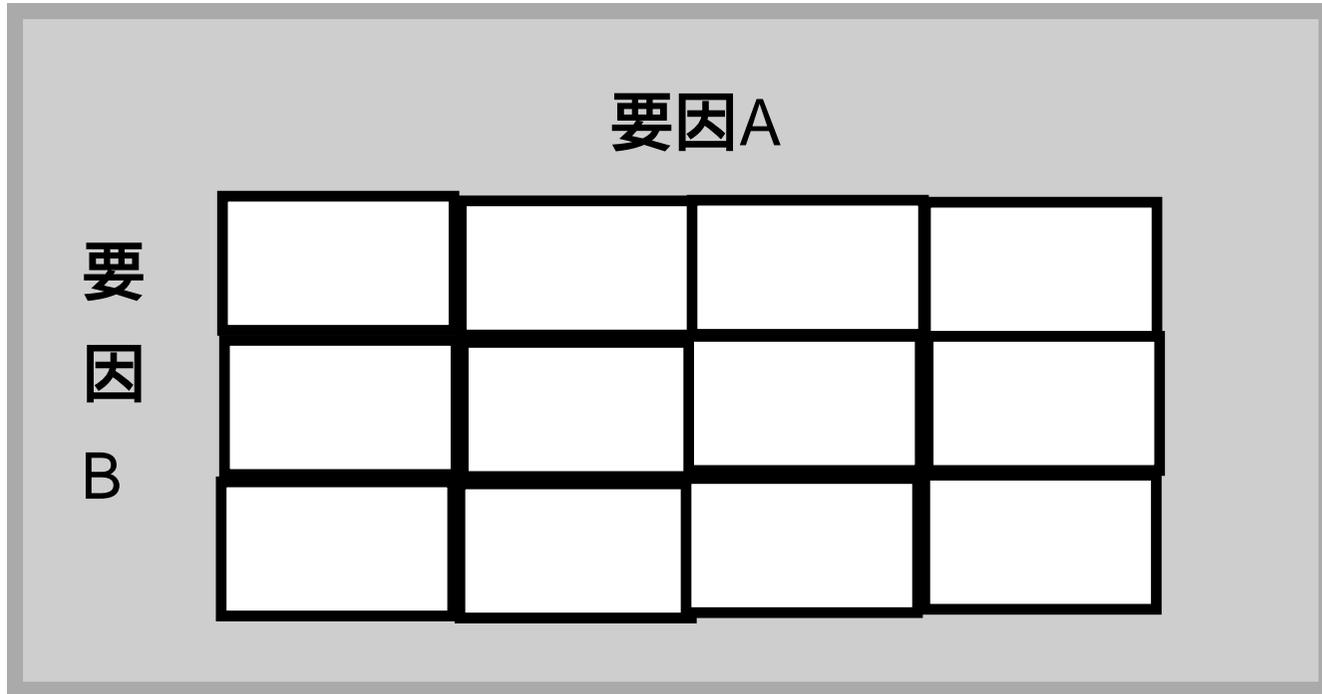
Marcus

傾向性仮説



Williams

2元配置の要因実験



補正すべきでないのに補正する

Bonferroni法

Fisher

c個の検定がfamily
有意水準が

個々の検定は α / c

個々の検定はc倍きびしく

実際によく使われるのは改良型

補正すべきでないのに補正する

調査地の一部

調査地の一部

調査地の一部

調査地の一部

草を食べる
昆虫

体サイズが大きいのでは？

木の葉を食べる

分散分析

補正すべきでないのに補正する

餌による体サイズのちがい
草 木の葉

調査地の一部

検定

調査地の一部

検定

調査地の一部

検定

調査地の一部

検定

有意確率を調査地の数で補正する
(検定の数)

全部 (ほとんど?) 同じ方向に有意差

おかしいーその1

調査地の一部 調査地の一部 調査地の一部 調査地の一部
餌のちがい 餌のちがい 餌のちがい 餌のちがい

分散分析
主効果

ブロック

検定は1つしかない

おかしいーその2

調査地の一部 調査地の一部 調査地の一部 調査地の一部
餌のちがい 餌のちがい 餌のちがい 餌のちがい

「調査地の一部の数」 増えたら
検定の個数

餌のちがいが検出されにくくなる

検定

検定

検定

検定

検定

仮説は共通したものが1つだけ、
検定の個数だけあるわけではない

多重比較・多重検定問題との共通点
は検定がいくつもあることだけ

両方向の結果が混在

| | | | | |
|-----|-----|-----|-----|-----|
| 検定 | 検定 | 検定 | 検定 | 検定 |
| 木の葉 | 木の葉 | 木の葉 | 木の葉 | 木の葉 |
| 草 | 草 | 草 | 草 | 草 |

共通した仮説が1つ：弱い証拠

多重検定では、関係ない

草を食べる 体サイズが大きいのでは？
昆虫
木の葉を食べる

「全体として、同じ場所では草を食べる方が体サイズが大きい」
ことは確かだ（確率論的に）

共通した仮説（例、二元配置分散分析）

「1つ残らずすべての場所で草を食べる方が体サイズが大きい」
ことは確かだ（確率論的に）

多重検定の補正をなさい

共通した仮説

こういう局面で使う方法

ブロック（層）のある検定

分散分析

ノンパラメトリクス

一般化Wilcoxon検定
対応セット検定

一般化線形モデル

有意確率をまとめる方法

p z 多くの方法ーメタ分析

Fisherのcombined probability

得意不得意

結び

似たような検定がいくつもあります

多重比較・多重検定だ、補正だ

病

そもそも多重比較・多重検定問題は、ユーザー
が意志決定する項目が多い

familyを構成する検定個数はそれほど
自明でないことがある

そもそも検定ではないかもしれない