

# ご案内

## 【資料について】

ワークショップの資料は教育目的で作成されていますが、著作権等が設定されています。営利目的での転用、第三者への配布（複写）、改変等をされないようお願いいたします。

## 【出席証明書について】

本ワークショップの出席証明書（押印済み）は、受付でお渡しした配付資料に同封しています。氏名および所属をご記入下さい。

なお、日本臨床薬理学会認定 CRC の出席証明書は、受付でお受け取りください。

## 【日本臨床薬理学会認定 CRC 単位について】

単位の必要な方は、受付で出席証明書をお渡ししますので、所属・氏名を記入して下さい。休憩時間あるいはワークショップ終了後に受付にて押印します。

なお、14日・15日の両日参加される方の出席証明書は、15日にお渡します。

## 【記録について】

本ワークショップは、平成20年度厚生労働科学研究費補助金（医療技術実用総合研究事業）を受けて開催しております。記録のために写真撮影やビデオの撮影を行いますことをご了承下さい。

## 【お願い】

参加者のワークショップおよび講演中の写真撮影・ビデオ撮影は、ご遠慮下さい。

## 【2月15日（日）にご参加の方へ】

- ① 受付にて2月15日（日）のワークショップで使用する資料をお渡します。この資料を用いて「ワークショップ1」を行いますので、内容について必ず準備をしておいて下さい。
- ② 本ワークショップではグループワークを行います。座席は、当日配布しますグループ分けに従って下さい。

# プログラム

2月14日(土) 13:30~18:30

総合司会：小手川 勤（大分大学医学部臨床薬理学 准教授）

開会の挨拶 13:30~13:35 中野 重行

## 臨床研究のポイントがわかる！ なぜなぜ ミニレクチャー

司会：大橋 京一（大分大学医学部臨床薬理学 教授）

13:35~14:10 1) なぜ、臨床研究が必要なのか？

谷内 一彦（東北大学大学院医学系研究科機能薬理分野 教授）

14:10~14:45 2) なぜ、「被験者」なのか？

三輪 亮寿（三輪亮寿法律事務所 所長）

14:45~15:20 3) なぜ、倫理原則が必要なのか？

笹栗 俊之（九州大学大学院医学研究院臨床薬理学分野 教授）

15:20~15:55 4) なぜ、併用禁止薬なのか？

藤村 昭夫（自治医科大学医学部臨床薬理学部門 教授）

15:55~16:15 （休憩）

司会：小林 真一（聖マリアンナ医科大学薬理学 教授／日本臨床薬理学会 前理事長）

16:15~16:50 5) なぜ、治験の安全性情報の取り扱いが改善が必要なのか？

依頼者の立場から

大島 裕之（日本製薬工業協会医薬品評価委員会臨床評価部会）

16:50~17:25 6) なぜ、安全性情報は取扱いが難しいのか？ CRC の立場から

井草 千鶴（町田市民病院治験支援室 CRC）

17:25~18:00 7) なぜ、メーカーの監査担当者は厳しいのか？

西村（鈴木）多美子（就実大学薬理学・薬物治療学 教授）

18:00~18:10 8) なぜ、日本人は温泉が好きなのか？

中野 重行（国際医療福祉大学大学院／大分大学医学部創薬育薬医学 教授）

18:10~18:30 総合討論

2月15日(日) 9:00~15:30

司会：小手川 勤（大分大学医学部臨床薬理学 准教授）

## CRCの技をみがく！ワクワクドキドキ ワークショップ

コーディネーター 福岡 敏雄（倉敷中央病院総合診療科／医師教育研修部 部長）

09:00~10:30 ワークショップ1：どうすれば効果的に伝えられるか？

10:30~10:50 （休憩）

10:50~15:00 ワークショップ2：疑問はどのようにして解決するのか？

10:50~11:30 レクチャー

11:30~13:00 グループワーク（昼食）

13:00~15:00 発表とディスカッション

15:00~15:25 総合討論

閉会のあいさつ 15:25~15:30 小林 真一

認定CRCのためのワークショップ  
2009年2月14日(土)&15日(日)、別府

## なぜ、日本人は温泉が好きなのか？

国際医療福祉大学 大学院 (創薬育薬医療)  
大分大学 医学部 創薬育薬医学  
中野 重行

## お話の流れ

1. 別府温泉の紹介をします
2. その健康に関する情報は正しい？
3. 明日のワークショップの課題：  
疑問の解き方を考えてみよう！

## 全国第一位の湧出量を誇る温泉 別府

源泉数：約2800ヶ所 (日本一)

湧出量：約13万7千 KL/日 (日本一)

別府市：人口 約12万人

## 油屋 熊八

1863年(文久3年) - 1935年(昭和10年)

- ・ 亀の井旅館(現在の 亀の井ホテル)を創業(1909年)
- ・ 亀の井自動車(現在の 亀の井バス)を開業(1928年)  
日本初の女性バスガイドによる定期観光バス  
別府駅前にフロンズ像(2007年)
- ・ 山は富士、海は瀬戸内、湯は別府  
標柱を富士山頂付近に建てた(1925年)  
全国各地に建てた
- ・ 全国大賞大会(1931年~)  
温泉マーク(別府温泉のシンボルマーク)
- ・ 生きているだけで丸儲け

## お話の流れ

1. 別府温泉の紹介をします
2. その健康に関する情報は正しい？
3. 明日のワークショップの課題：  
疑問の解き方を考えてみよう！

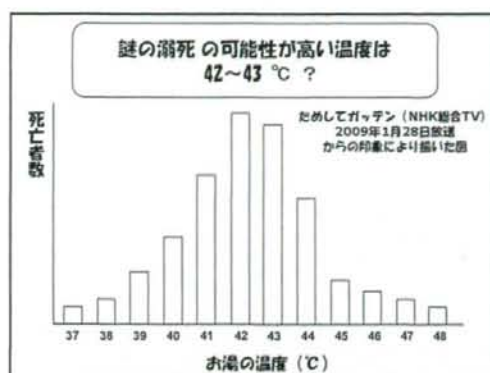
## ためしてガッテン

お風呂大変身！超快感・安全入浴術  
NHK総合テレビ 2009年1月28日放送

入浴事故(死亡)：1万4千人(年間)

死因：謎の溺死(肺にほとんど水が入って  
いない状態で、湯船で亡くなる)

謎の溺死：入浴事故の大半 推定1万人以上



**お話の流れ**

1. 別府温泉の紹介をします
2. その健康に関する情報は正しい?
- 3. 明日のワークショップの課題:  
疑問の解き方を考えてみよう!**

**温泉の効用**

皆様にご利用いただいている「神皇の村天然温泉」の効用は下記のとおりです。  
健康づくりのため、大いにご活用下さい。

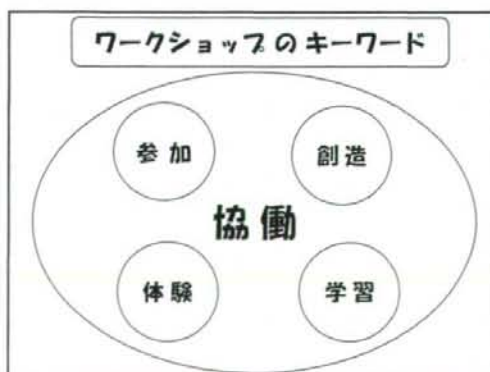
P H 値	8.3
泉 質	単純温泉 (弱アルカリ性硫酸性高温泉)
効 用	神経痛、筋肉痛、関節痛、五十肩、運動麻痺、関節のこわばり、うちみ、くじき、慢性消化器病、痔疾、冷え性、病後回復期、疲労回復、健康増進
禁忌症	急性疾患 (特に熱のある場合)、運動性の筋痛、悪性腫瘍、重い心臓病、出血性疾患、高度の貧血、呼吸不全、腎不全、その他一般に病勢進行中の疾患、妊婦中 (特に胎動と末胎)
分析所	大分公衆衛生センター

明日のWSへ向けた課題: 考えてみよう!

**A. 温泉は「効用」に書いてある疾患や症状に本当に効くのだろうか?**

**B. 筋緊張性頭痛の多くの患者が、温泉につかることで症状が改善するとすれば、なぜなのだろうか?**

このような疑問は、どうすれば解決できる?  
今夜、温泉につかりながら、考えてみよう!



《ワークショップ、セミナー等》

若手医師向け医学統計セミナー（聖マリ医大）

平成 21 年 3 月 5 日 開催

・プログラム

厚生労働省科学研究費補助金  
医療技術実用化総合研究事業  
(主任研究者:小林真一)

若手医師向け医学統計ワークショップ  
—OJT形式の統計学研修を目指して—

2009年3月5日(木曜)

16:30-18:00

教育棟5F セミナー室4

医学統計学分野 立浪 忍、上野隆彦

## はじめに

今般、厚生労働省科学研究費補助金医療技術実用化総合研究事業（主任研究者：小林真一）の一環として、若手医向け統計学ワークショップを開催する運びとなりました。

統計学はE BMの基本を裏打ちするものであり、医学・生物学領域で大きな発展をしつつあります。また、安価なパーソナル・コンピューターの処理能力の向上と、インターネットによる広範な情報の共有化が、複雑な統計手技の実用化とその実用性の検証に拍車をかけ、統計関連科学の発展の勢いは、今のところ、とどまる気配がありません。

このような状況下にあつて、統計科学、あるいは統計学に疎い医師や医学研究者は、国際的な舞台から脱落してしまう危険があるといつても、過言ではないでしょう。

しかし、研究の現場で応用されている統計的な手技のどれをとつても、それぞれ歴史的背景や適応の適・不適があり、瞬時にしてこれを理解し、適切な応用に供することは困難です。

今回のワークショップは、現実の諸研究分野において利用されている統計的な手技の多様性と、これを理解するための困難さとの間にあるギャップを、少しでも埋めることを目的とし、on job training 形式で問題解決能力を養うことを目指したものです。

そこで、内容は厳選し、ノンパラメトリック変量の取り扱いと、多重比較についての講義と実習によるプログラムを作りました。

実習には「4Steps エクセル統計（第2版）」、柳井久江著（星雲社）と統計解析ソフト R を用います。統計解析ソフト R はボランティアベースで開発されているフリーソフトです。今回使用する計算プログラムは群馬大学社会情報学部の青木繁伸氏により作成・公開されているものです。また付録の「ダネットの方法のための両側5%点」の表は「統計的多重比較法の基礎」、永田靖、吉田道弘著（サイエンティスト社）よりの抜粋です。

若手医師の皆様の問題解決に向けて、少しでも貢献できれば幸いです。



## I. 統計的仮説検定

仮説検定では観測されたデータをもとに仮説を確率的に検証する。

- ・得られたデータの差は偶然でも観測されるのか？
- ・データの関連は偶然か？

### 仮説検定の手順

1. 帰無仮説と対立仮説を設定する
  - ・「(本当は) 差は無い」等、帰無仮説
  - ・「差はある」、「A より B は大きい」等、対立仮説
2. 母集団からサンプルの抽出
3. 帰無仮説が正しいという仮定のもとでサンプルの得られる確率  $P$  を計算
4.  $P$  が小さい (例えば  $P < 0.05$ ) ならば  
「帰無仮説は正しく、たまたま珍しいサンプルが得られた」と考えず  
「帰無仮説は誤りであった。対立仮説が正しい」と考える。
  - ・  $P$  が大きいならば「帰無仮説は正しいかもしれない」と考える。

注意：  $P$  が大きいとき「帰無仮説が正しい」と積極的に主張できるわけではない。「帰無仮説を否定する材料は得られなかった」という立場をとる。例えば C 氏が「私はコインを投げる前に表が出るか裏が出るかわかる」と主張したとする。それならと、一回コインを投げてみる。C 氏が予想を外せば C 氏の主張は誤りであったと考える(積極的に仮説を棄却)。一方、C 氏の予想があたったなら、それは C 氏の主張を否定する材料とはならない。だからといって、C 氏の主張を積極的に肯定しようとも考えないだろう(消極的に仮説を容認)。

### 検定における2種類の判断ミス

		検定による判断	
		差あり	差なし
現実	差あり	○	× <sup>2)</sup>
	差なし	× <sup>1)</sup>	○

- 1) 「第1種の過誤」：本当は差が無いのにサンプルから差を検出してしまう(危険率)
  - 2) 「第2種の過誤」：本当は差があるのにサンプルから差を検出できない
    - ・  $1 -$  (第2種の過誤の確率) を検定の検出力(パワー)という。
    - ・ 検出力は高いほうがよい。第1種の過誤の確率を小さくすると第2種の過誤の確率は大きくなる。つまり、検出力は小さくなる。
- この検出力の意味を2群の比率の差の検定を例にして眺めておこう。

・比率の差の検定

次の分割表を考えよう。

	効果あり	効果なし	比率
標準薬	6	15	0.29
新薬 A	16	8	0.67

新薬 A と標準薬の効果を比較せよ。有意水準（危険率）は 0.05 とする。

	効果あり	効果なし	比率
標準薬	2	8	0.2
新薬 B	6	4	0.6

	効果あり	効果なし	比率
標準薬	9	22	0.29
新薬 B	15	10	0.6

新薬 B と標準薬の効果を比較せよ。

上の3つの表について R の `power.prop.test` を用いて検出力を計算すると、

A の表の検出力=0.71      `power.prop.test(n=21,p1=0.29,p2=0.67)`  
 B(左)の表の検出力=0.44      `power.prop.test(n=10,p1=0.2,p2=0.6)`  
 B(右)の表の検出力=0.60      `power.prop.test(n=25,p1=0.29,p2=0.6)`

となる。“検出力”は次のように考える。新薬 A について

$H_0$ : 「2群に差はない」、

$H_1$ : 「標準薬での比率は 0.29、新薬での比率は 0.67」

と考えたとする。検出力 0.71 の意味は「 $H_1$  が正しいときに、検定で  $H_1$  を採択できる確率が 0.71 である」ということである。一方、B(左)の表でも、

$H_0$ : 「2群に差はない」、

$H_1$ : 「標準薬での比率は 0.2、新薬での比率は 0.6」

と考えたとして、「 $H_1$  が正しいときに、検定で  $H_1$  を採択できる確率は 0.44 である」ということである。言い換えると「 $H_1$  が正しかったとしても確率  $1-0.44=0.56$  で  $H_0$  を採択する」ということになる。この場合  $H_0$  を採択（有意でない）と判定されても検出力が不十分で信頼性に欠けると考えてよいだろう。

今、標準薬群の比率は 0.3 とわかっており試験薬群の比率をパイロット試験により 0.65 と推測したとしよう。更に臨床的には試験薬群の比率が 0.6 でも意義があると考えられるが、0.6 より小さいとその意義は低くなると仮定しよう。

このとき、やはり `power.prop.test` を用いて、検定に必要な例数を計算できる。

必要例数=各群 42 例      `power.prop.test(p1=0.3,p2=0.6,power=0.8)`

となる。もし、 $p_1=0.3$ 、 $p_2=0.6$  という見込みが正しければもっと少ない例数で有意な結論

を得られるだろう（各群 20 例前後）。しかし各群 42 例のサンプルをとった場合には、たとえ有意でない結論に至ったとしても「2 群の間に有意な差は見出せなかった（power=0.8）」と記述することができる。これは「標準薬群 0.3、新薬群 0.6 と推測し、この差があれば 10 回中 8 回は有意差ありと結論づけられる大きさのサンプルをとったが、その結論は得られなかった。したがって、新薬の効果は 0.6 に満たない」という結論を意味する。

少ない例数でも有意な結論を得られれば、意味のある差が検出されたと考えてよいであろう。しかし、研究は必ずしも予定通りに“有意”となるものではない。例数設計は研究計画の段階で行うが、その目的は期待に反して有意でないという結果を得た場合にも科学的な結論を引き出すことにある。

さて新薬 A、B という系統の異なる 2 剤を併用することが提案された。検討の結果、2 種類の投与法を試してみるようになった。

	効果あり	効果なし	比率
新薬 A	100	55	0.65
新薬 B	90	65	0.58
投与 1	110	45	0.71
投与 2	120	35	0.77

この場合には、投与 1 あるいは投与 2 が新薬 A または新薬 B より有効であれば意義があると考えられる。

## II. 多重比較法

有意水準 0.05 の検定を続けて 20 回行ったらどうなるか？

→「有意水準 0.05 の検定」＝「20 回に 1 回程度判断ミスをする可能性がある検定」

→20 回の検定のうち 1 回は判断ミスをしているのでは！！

（実際は、1 回ごとに正しい判断をする確率が  $1 - 0.05 = 0.95$ 、20 回続けて正しい判断ができる確率は  $0.95^{20} = 0.36$ 、1 回以上の判断ミスをする確率が 0.64 となる。）

	有意水準 0.05 での検定回数				
	1 回	2 回	3 回	4 回	5 回
正しい判断をする確率	0.95	0.9025	0.8574	0.8145	0.7738
1 回以上判断ミスをする確率	0.05	0.0975	0.1426	0.1855	0.2262

上の表にあるように1回の検定での判断ミスの確率が0.05より小さくとも、複数回の検定を行えば判断ミスが起きる確率は0.05より大きくなる！

→それに対処する方法が多重比較法

多重比較の方法は大別すると次の2種類：

1. 有意水準を調整する方法（ボンフェローニによる方法等）
2. 固有の計算を行う方法（テューキーによる方法、ダネットによる方法等）

・ライアンによる比率の差の多重比較（ボンフェローニの改良版）

ボンフェローニによる多重比較の方法はボンフェローニの不等式に基づく方法である（と書くと難しそうだが、5回検定をする必要があり検定全体の有意水準を0.05にしたいなら $0.05 \div 5$ で各検定の有意水準を0.01で行うという方法である）。

・この方法は汎用性が高いものの一回ごとの検定の有意水準が低くなり検出力が低い。

この欠点を改善する方法の一つがライアンによる有意水準の調整法である。

例えば、5つの母集団から比率 $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ を得たとしよう（もちろん事前に

帰無仮説「すべての母集団の比率に差はない」

対立仮説「ある母集団の比率には差がある」

を設定している）。ライアンの方法では、

1. 比率を大きさ順に並べ替える。仮に $\alpha_4 \leq \alpha_1 \leq \alpha_3 \leq \alpha_5 \leq \alpha_2$ とする。

2. まず $\alpha_4$ と $\alpha_2$ に差があるか検定。

・差が無ければすべて差が無いと判断。

3. 差があれば今度は $\alpha_4$ と $\alpha_3$ 、 $\alpha_1$ と $\alpha_2$ に差があるか検定。

・ $\alpha_4$ と $\alpha_3$ に差が無ければ $\alpha_4$ から $\alpha_3$ までには差が無いと判断。

・ $\alpha_1$ と $\alpha_2$ に差が無ければ $\alpha_1$ から $\alpha_2$ までには差が無いと判断。

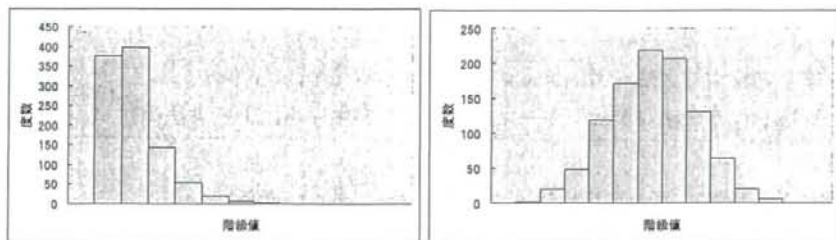
4. 差があれば今度は $\alpha_4$ と $\alpha_3$ 、 $\alpha_1$ と $\alpha_3$ 、 $\alpha_3$ と $\alpha_2$ に差があるか検定。

以下、同様。

こうして順番を入れることで一回ごとの検定の有意水準をボンフェローニ以上に保つことができる（上の2だけがボンフェローニと同じく有意水準0.005での検定。3の有意水準は0.066、4の有意水準は0.01、その次の段階では0.02となる）。

・スティールの方法、スティール・ドゥワスの方法（固有の計算を行う方法）

多重比較の方法としてよく知られているテューキーの方法やダネットの方法は母集団分布が正規分布であるときの検定手法である。データが正規分布でない場合に、それらの方法を用いるのは適切でなく、誤った判断を犯す可能性を増加させてしまう。



実際、母集団についての情報が無いまま、左のように右裾の長いデータに正規性を仮定した検定を適用したため、後日、結論が覆ったという事例もある（母集団が正規分布に従うという知見があっても、データが左の図であれば抽出の無作為性が疑われる）。正規分布に従うデータから無作為抽出を行えば右のようなデータが得られるはずである。

左のような分布をしたデータに対数変換をほどこすと、右のようなデータに変換されることはよく知られた事実である。しかし、どんなデータであっても適当な変換で正規分布のように見せるわけではない。

#### 検定手法

- ・パラメトリック — 分散の等しい正規分布に従うという仮定あり
- ・ノンパラメトリック — 正規分布に従うという仮定なし

スティールの方法とスティール・ドゥワスの方法はノンパラメトリックな検定手法

- ・全対比較がスティール・ドゥワスの方法（テューキーの方法に対応する手法）
- ・対照群と他群の比較がスティールの方法（ダネットの方法に対応する手法）

注意：ノンパラメトリックな検定手法は母集団に正規分布を仮定しないが、それぞれの群の分布が同じ形であることを仮定している。パラメトリックな検定と比較すると検出力はおちるが、外れ値（他の値とかけ離れた値）に対して頑健である。

#### スティール・ドゥワスの方法

各群の分布の代表値についてすべての対比較を行う。

帰無仮説「各群の代表値に差は無い」

対立仮説「代表値の異なる群がある」

### スティールの方法

各群の分布の代表値について対照群と他の群の対比較のみを行う。

帰無仮説「対照群と他群の代表値との間に差は無い」、

対立仮説「対照群と異なる代表値を持つ群がある」

#### スティールの方法の結果の読み方

3群の比較の場合、対照群を1群とし1群と2群、1群と3群の比較を行う。結果は

	t	rho
1:2	2.952566	0.5
1:3	1.175674	0.5

という形で提示される。付録のダネットの表の両側5%点の相関係数0.5の表を参照する。この場合3群なので  $a=3$  の一番下の数 2.212 より  $t$  が大きければ有意差ありとなる。上の例だと1と2の  $t$  は 2.95 ( $> 2.212$ ) なので1と2には有意差あり、1と3の  $t$  は 1.18 ( $< 2.212$ ) なので1と3には有意差がない。

#### 本日の実習項目

- ・ テューキーの方法 (4 Step エクセル統計による)
- ・ 比率の差の多重比較  
(`p.multi.comp`、青木繁伸 (群馬大学) が作成した R program および `pairwise.prop.test` による)
- ・ スティールの方法(Kyplot による)
- ・ スティール・ドゥワスの方法(Kyplot による)

なお、R にはノンパラメトリック多重比較法として `pairwise.wilcox.test` がある。これは ウィルコクソンの順位和検定の有意水準をホルム (その他も) の方法で調整して用いるものである。

Rの使用 (Rのインストール後)

1. 実習用模擬データをクリックして開き、子宮頸ガンのデータをコピーする。
  2. R console から `x<-read.delim("clipboard")` と入力して Enter。
  3. `x` と入力して Enter を押しデータを読み込んだことを確認。
  4. CDにある"実習用模擬データの変換"をコピーして、R console に貼り付けて Enter  
・変数 `y1,y2,y3,y4,y5` にそれぞれ normal、軽度異形成、中等度異形成、高度異形成、ガンのデータが入力された。
  5. `Steel(list(y2,y3,y4,y5))` と入力し Enter  
スティールの方法で軽度異形成からガンまでの4群を比較。  
・ダネットの方法のための両側5%点の表を参照。0.3< $\rho$ <0.5なので相関係数=0.3の表を参照する。  
 $t > d(5, \infty, 0.3, 0.05)$  ならば有意差あり (有意水準5%)。
- ・ $\rho = \rho_1$  とする。 $\rho_1 < \rho < \rho_2$  ならば、簡易的には $\rho_1$ の表を用いる。もう少しきちんと考えるなら、次の補間式を用いて計算する。

$$d(a, \infty, \rho; 0.05) = \frac{1/(1-\rho_2) - 1/(1-\rho)}{1/(1-\rho_2) - 1/(1-\rho_1)} d(a, \infty, \rho_1; 0.05) + \frac{1/(1-\rho) - 1/(1-\rho_1)}{1/(1-\rho_2) - 1/(1-\rho_1)} d(a, \infty, \rho_2; 0.05)$$

6. `Steel.Dwass(list(y1,y2,y3,y4,y5))` と入力 (簡単な方法あり) Enter  
スティール・ドゥワスの方法で5群を比較。  
スチューデント化された範囲の上側5%点の表を参照する。

$$\frac{q(5, \infty; 0.05)}{\sqrt{2}} = 2.728 \text{ が棄却の限界}$$

つまり  $t > 2.728$  ならば有意差あり (有意水準5%)。

7. 多群の比率の検定 (ライアンの方法およびテューキーの方法による)
  - 1) データの入力、各群の合計を `n1<-c(a,b,c,d)` の形で入力し Enter
  - 2) 次に効果があった人数等を `n2<-c(e,f,g,h)` の形で入力し Enter
  - 3) `p.multi.comp(n1,n2,method="ryan")` と入力し Enter
  - 4) `p.multi.comp(n1,n2,method="tukey")` と入力し Enter

Rの使用 (Rのインストール後)

1. 実習用模擬データをクリックして開き、子宮頸ガンのデータをコピーする。
2. R console から `x<-read.delim("clipboard")` と入力。
3. `x` と入力してデータを読み込んだことを確認。
4. CDにある"実習用模擬データの変換"をコピーして、R console に貼り付けて Enter  
変数 `y1,y2,y3,y4,y5` にそれぞれ normal、軽度異形成、中等度異形成、高度異形成、ガンのデータが入力された。
5. `Steel(list(y2,y3,y4,y5))` と入力し Enter  
スティールの方法で軽度異形成からガンまでの4群を比較。  
ダネットの方法のための両側5%点の表を参照。0.3< $\rho$ <0.5なので相関係数=0.3の表を参照する。  
 $t > d(5, \infty, 0.3, 0.05)$  ならば有意差あり (有意水準5%)。

・  $\rho = \text{rho}$  とする。  $\rho_1 < \rho < \rho_2$  ならば、簡易的には  $\rho_1$  の表を用いる。もう少しきちんと考えるなら、次の補間式を用いて計算する。

$$d(a, \infty, \rho; 0.05) = \frac{1/(1-\rho_2) - 1/(1-\rho)}{1/(1-\rho_2) - 1/(1-\rho_1)} d(a, \infty, \rho_1; 0.05) + \frac{1/(1-\rho) - 1/(1-\rho_1)}{1/(1-\rho_2) - 1/(1-\rho_1)} d(a, \infty, \rho_2; 0.05)$$

6. `Steel.Dwass(list(y1,y2,y3,y4,y5))` と入力 (簡単な方法あり) Enter  
スティール・ドゥワスの方法で5群を比較。  
スチューデント化された範囲の上側5%点の表を参照する。

$$\frac{q(5, \infty; 0.05)}{\sqrt{2}} = 2.728 \text{ が棄却の限界}$$

つまり  $t > 2.728$  ならば有意差あり (有意水準5%)。

7. ここからはおまけ。 `Bonfferoni(list(y1,y2,y3,y4,y5))` と入力し Enter
8. `Bonfferoni(list(y1,y2,y3,y4,y5), method="Holm")` と入力し Enter  
Holmの方法による多重比較 (有意水準は0.05), `method="Shaffer"` もある。
9. `Bonfferoni(list(y1,y2,y3,y4,y5), method="Holm", alpha=0.01)` と入力し Enter  
Holmの方法による多重比較 (有意水準は0.01) 現在、想定されるすべての状況に対応
10. `Peritz(list(y1, y2, y3, y4, y5))` と入力し Enter  
ペリの方法による多重比較 (チューキー型の比較で最も高い検出力を持つ)。



若手医師向け医学統計ワークショップ-OJT 形式の統計学研修を目指して—

---

2009年3月発行

執筆・編集

立浪 忍 (聖マリアンナ医科大学医学統計学分野)

上野 隆彦 (聖マリアンナ医科大学医学統計学分野)

---

(非売品)

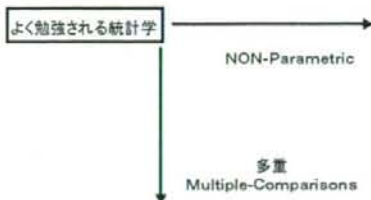
厚生労働省科学研究費補助金  
医療技術実用化総合研究事業  
(主任研究者:小林真一)

若手医師向け医学統計ワークショップ  
—OJT形式の統計学研修を目指して—

2009年3月5日(木曜日)  
16:30-18:00  
教育棟5階 セミナー室4

医学統計学分野  
立浪 忍、上野隆彦

Developmentの2大方向性



t検定は、基本的には  
計量的な変数に使う  
改善=5、やや改善=4、...

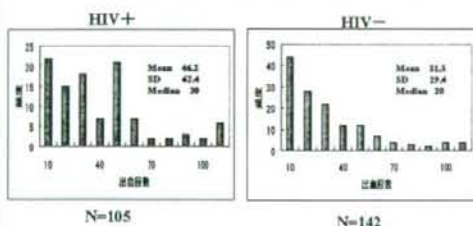
のような5, 4, 3, 2をt検定  
で比較した



正規分布してないデータ  
の比較もダメ

(正規分布:ガウス分布ともいう)

HIV感染の有無による年間出血回数の比較  
(重症型血友病A、20歳以上の症例、インシビタ症例除く)



Mann-WhitneyのU検定の結果  
P=0.017(有意差あり)



(横軸1~11の説明: 1,測定感度未満; 2,測定感度~100; 3,101~200; 4,201~300;  
5,301~400; 6,401~500; 7,501~600; 8,601~700; 9,701~800; 10,801~849; 11,≥850Copies/ml)

### 順位変数の教訓

ステージ分類	割り当てた数字
I	1
II	2
III-A	3
III-B	4

### 順位変数の平均値

ステージ分類	割り当てた数字
I	1
II	2
III-A	3
III-B	4

$$\frac{1+4}{2} = 2.5$$

問題点1  
数字2.5はステージ何か？

問題点2  
ステージIII-Aは、ステージIの3倍悪い(malignancyが高い)のか？

### 2群の比較のまとめ

	対 (対応あり)	対でない (対応なし)
パラメトリック	対応のある t 検定 (Paired-t-test)	対応のない t 検定 (Unpaired-Student's-t-test)
ノン・パラメトリック	ワイルコクソン順位和検定 (Wilcoxon-t-test)	マン・ホイットニー検定 (Mann-Whitney-U-test)

### Developmentの2大方向性



### 多群の比較はANOVA

### 有意差がでたら？



