

医薬安全性研究会

Bulletin of Japanese Society for Biopharmaceutical Statistics.

No.1~No.8 合併号

構成

No.1 (1979年12月)	P. 1~5
No.2 (欠番)	
No.3 (1980年5月)	P. 1~4
No.4 (1980年10月)	P. 1~8
No.5 (1981年1月)	P. 1~14
No.6 (1982年4月)	P. 1~15
No.7 (1982年8月)	P. 1~29
No.8 (1983年1月)	P. 1~17

医薬安全性研究会 会報 合併号

総合目次

- No. 1
 - 安全研ニュース創刊にあたって----- 1
 - 12月8日第1回定例会開催----- 1
 - [質内コーナー] 欠測値が出た場合どうすべきか----- 3
 - 標準偏差と標準誤差の違いは?----- 4
- No. 2 (欠番)
- No. 3
 - 食品添加物の安全基準-WHO勧告(1957)を中心に----- 1
 - [質内コーナー] 共分散分析の使い方は?----- 3
 - 投与量は70キログラムか、体表面積当りか?----- 3
- No. 4
 - 動物実験における比較のための統計量----- 1-8
- No. 5
 - 動物実験における相対臓器重量の意味----- 1-14
- No. 6
 - 累積カイニ乗法の応用----- 松本一彦 /
 - [質内] 生殖試験における順位和検定について--- 佐野正樹・金子洋二 /3
 - 薬物動力学における Compartment model の定数の求め方について 香田 繁 /4
 - 事務局より 15
- No. 7
 - GLPでの有意差検定の考え方入門----- 高橋行雄 /
 - [質疑応答] 生殖試験における順位和検定について--- 金子洋二・佐野正樹 /10
 - フィッシャーの直接確率の両側仮説検定の求め方--- 田中 健 20
 - [数学研究室] 統計のParadoxについて----- 山崎昌男 23
 - ニュース 26 事務局より 29
- No. 8
 - 発癌性化学物質の定量的リスクアセスメント----- 中村晃忠 /
 - 薬物動力学における Compartment model の定数の計算法--- 増山元三郎・吉村 功 /4
 - (別解 田中 健)
 - 事務局より 17

医 薬 安 全 性 研 究 会

会 報 No. 7

1982年8月

目 次

- GLPでの有意差検定の考え方入門……高橋行雄(1)
生殖試験における順位和検定について+質疑応答
金子洋二・佐野正樹……(10)
- フィッシャーの直接確率の両側仮説検定の求め方+質疑
応答……田中 健(20)
- [数学研究室] 統計の Paradox について……山崎昌男(23)
- [ニュース](26)
- [事務局だより](29)

GLPでの有意差検定の考え方入門

高橋行雄*

1. はじめに

1976年11月に米国FDAによりGLP案が発表され、その後、公聴会におけるDiscussionを経て1978年12月にGLPが法律として制定され、1979年6月より施行された。日本でも、1978年より厚生省によって検討が開始され、1981年7月にGLP案が示され、1982年3月の薬務局通知において1982年4月より全面的に適用されることになった。

日本のGLPでも、データの解析に使用する統計学的方法については試験計画書の中に記載されていなければならないとされている。また最終報告書においても、データ解析に使用された統計手法について記載されていなければならないとされている。

これまで、医薬安全性研究会では、薬物の安全性試験にまつわる各種のデータの統計的方法について検討を続けてきた。しかしながら、どのような統計手法を用いるのが最も通しているかという結論はなかなか出せようにもない。GLPでは、試験計画書の中でどのような統計手法を用いるのか記載しなければならないとしているがどのような統計手法を用いるべきかは運営管理者の最終的な承認事項となっている。それでは、運営管理者は、どのような基準で統計手法を選定したらよいのであろうか。

ここでは、選定の目安となる各種の資料の紹介、最近、安全性試験でも用いられるようになってきた順位和検定入門、日本ロシュ研究所の安全性試験で用いられている統計手法およびその選択の理由についても紹介したい。

2. 統計手法のJIS規格

一部の統計手法にはJIS規格がある。安全性試験で最もよく用いられているのはいわゆるt検定であろう。これは、日本規格協会が出している“JISハンドブック・品質管理”の中の統計手法で取り上げられている(表1)。タイトルには“t検定”の記載はないが、規格番号Z 9049の“二つの平均値の差の検定、標準偏差未知、両側”となっているものがそうである。

この規格は、2標本のサンプルが等しく、かつ標準偏差が等しいと思われる場合について規定しており、等しくない場合については参考となっている。それは、

(1) 標準偏差が等しいと思われる場合

(2) 標準偏差が明らかに異なる場合 (サンプルの大きさは等しくても等しくなくてもよい)

の2つの場合である。さて、どの方法を選んだらよいのだろうか。

*日本ロシュ(株)

表1. JIS規格で定められている統計手法

[統計的方法]		
JIS Z 9041 (1968)	測定値の処理方法	441
JIS Z 9042 (1962)	母平均と基準値との差の検定 (標準偏差既知, 片側)	462
JIS Z 9043 (1962)	母平均と基準値との差の検定 (標準偏差既知, 両側)	464
JIS Z 9044 (1962)	母平均と基準値との差の検定 (標準偏差未知, 片側)	466
JIS Z 9045 (1962)	母平均と基準値との差の検定 (標準偏差未知, 両側)	468
JIS Z 9046 (1965)	二つの平均値の差の検定 (標準偏差既知, 片側)	470
JIS Z 9047 (1979)	二つの平均値の差の検定 (標準偏差既知, 両側)	474
JIS Z 9048 (1979)	二つの平均値の差の検定 (標準偏差未知, 片側)	477
JIS Z 9049 (1965)	二つの平均値の差の検定 (標準偏差未知, 両側)	482
JIS Z 9050 (1963)	母平均の区間推定 (標準偏差既知)	487
JIS Z 9051 (1963)	母平均の区間推定 (標準偏差未知)	489
JIS Z 9052 (1963)	二つの平均値の差の区間推定 (標準偏差既知)	491
JIS Z 9053 (1963)	二つの平均値の差の区間推定 (標準偏差未知)	494
JIS Z 9054 (1966)	未知分散と既知分散との違いの検定 (片側)	499
JIS Z 9055 (1966)	未知分散と既知分散との違いの検定 (両側)	501
JIS Z 9056 (1979)	二つの母分散の違いの検定 (片側)	503
JIS Z 9057 (1966)	二つの母分散の違いの検定 (両側)	509
JIS Z 9058 (1966)	母分散の区間推定	514
JIS Z 9059 (1966)	母分散比の区間推定	516

JIS ハンドブック
日本規格協会, 1982

3. ノンパラメトリック法の選択

ノンパラメトリック法の適応を考える上で最も参考になるのは, S. Siegel の "Nonparametric Statistics for the behavioral science" で示されている検定手法の一覧表であろう(表2)。

この表は、行方向に変数の性質

- (1) Nominal : 名義尺度
- (2) Ordinal : 順序尺度
- (3) Interval : 距離尺度

列方向に標本の性質がとられている。

- (a) One-Sample case : 1 標本の場合
- Two-Sample case : 2 標本の場合
- (b) Related Samples : 独立でない標本, 対応のある場合
- (c) Independent Samples : 独立な標本, 対応のない場合
- K-Samples case : K 標本の場合
- (d) Related Samples : 独立でない標本
- (e) Independent Samples : 独立な標本

表2 ノンパラメトリック検定法の一覧表 (S. Siegel "Nonparametric Statistics for the Behavioral Sciences" McGraw-Hill, 1956)

LEVEL OF MEASUREMENT	NONPARAMETRIC STATISTICAL TEST*				NONPARAMETRIC MEASURE OF CORRELATION (Chap. 9)
	One-sample case (Chap. 4)	Two-sample case		k-sample case	
		Related samples (Chap. 5)	Independent samples (Chap. 6)		
Nominal	Binomial test, pp. 36-42 χ^2 one-sample test, pp. 42-47	McNemar test for the significance of changes, pp. 63-67	Fisher exact probability test, pp. 96-104 χ^2 test for two independent samples, pp. 104-111	Cochran Q test, pp. 161-166	Contingency coefficient C, pp. 196-202
Ordinal	Kolmogorov-Smirnov one-sample test, pp. 47-52 One-sample run test, pp. 52-58	Sign test, pp. 68-75 Wilcoxon matched-pairs signed-ranks test,† pp. 75-83	Median test, pp. 111-116 Mann-Whitney U test, pp. 116-127 Kolmogorov-Smirnov two-sample test, pp. 127-136 Wald-Wolfowitz runs test, pp. 136-145 Moses test of extreme reactions, pp. 145-152	Friedman two-way analysis of variance, pp. 166-172	Spearman rank correlation coefficient r_s , pp. 202-213 Kendall rank correlation coefficient τ , pp. 213-223 Kendall partial rank correlation coefficient $\tau_{B,1}$, pp. 223-229 Kendall coefficient of concordance: W, pp. 229-238
Interval	Walsh test, pp. 83-87 Randomization test for matched pairs, pp. 88-92	Randomization test for two independent samples, pp. 152-156			

* Each column lists, sequentially downward, the tests applicable to the given level of measurement. For example, in the case of 1 related sample, when ordinal measurement has been achieved both the Friedman two-way analysis of variance and the Cochran Q test are applicable.

† The Wilcoxon test requires ordinal measurement and only whole pairs, as is required for the Sign test, but also of the difference between pairs. See the discussion on pp. 75-76.

(f) Nonparametric measure of correlation . . . ノンパラメトリック相関

この2つから必要な検定手法が選択できるようになっている。ただし、該当する手法は1種類とほかぎらないのが問題を残すことになる。個々のノンパラメトリック検定については、岩原信九郎の“ノンパラメトリック法”が平易に書かれている。

4. 二重盲検比較試験における統計手法

新薬の開発の最終関門、二重盲検比較試験では、どのような統計手法が一般的に用いられているのだろうか。コントローラー委員会の“薬効評価システム解説書(第1部)”によると、2標本で独立の問題については、次の3種が用いられ、

- (1) Mann-Whitney U test (U検定)
- (2) χ^2 test for independent Sample (χ^2 検定)
- (3) Fisher exact probability (直接確率法)

2標本で独立でない標本については、次の2種が用いられている。

- (4) Sign-test (サイン検定)
- (5) Wilcoxon mated pairs signed-rank test (ウィルコクソン検定)

K-Sample (多重比較)の検定手法は用いられておらず、すべて2標本の組合せの検定を行なうようになっている。次の特徴は1つの変数について2種以上の検定を組合せて同時に出している点である。このような方法についての批判もあるが、現在、臨床試験の場では一般的になっている。

5. 多重比較

安全性試験は通常 Control 群と複数の薬剤投与群が置かれている。以前に、医薬安全性研究会で次のようなメモを頂いた。フローチャートを次に示す。これは、ある安全性試験の委託試験施設が委託先より提示されたものらしい。GLPの第25条によれば、試験計画書は、試験委託先の承認も必要なので、委託先によって要求される統計手法が異なる可能性が十分にある。どのような統計手法を用いるかは常に問題となりそうである。

Bartlett 法による分散の一樣性の検定を行い、一樣の場合は一元配置型分散分析を行う。群間に有意性が認められた場合はさらに Dunnett 法 (例数一定の場合) あるいは Scheffe 法によって対照群との多重比較を行う。

分散が一樣でない場合は Kruskal-Wallis の H 検定を行う。群間に有意性が認められる場合はさらに対照群と各投与群との間で Dunnett 型 (例数一定の場合) あるいは Scheffe 型の順位検定を行う。

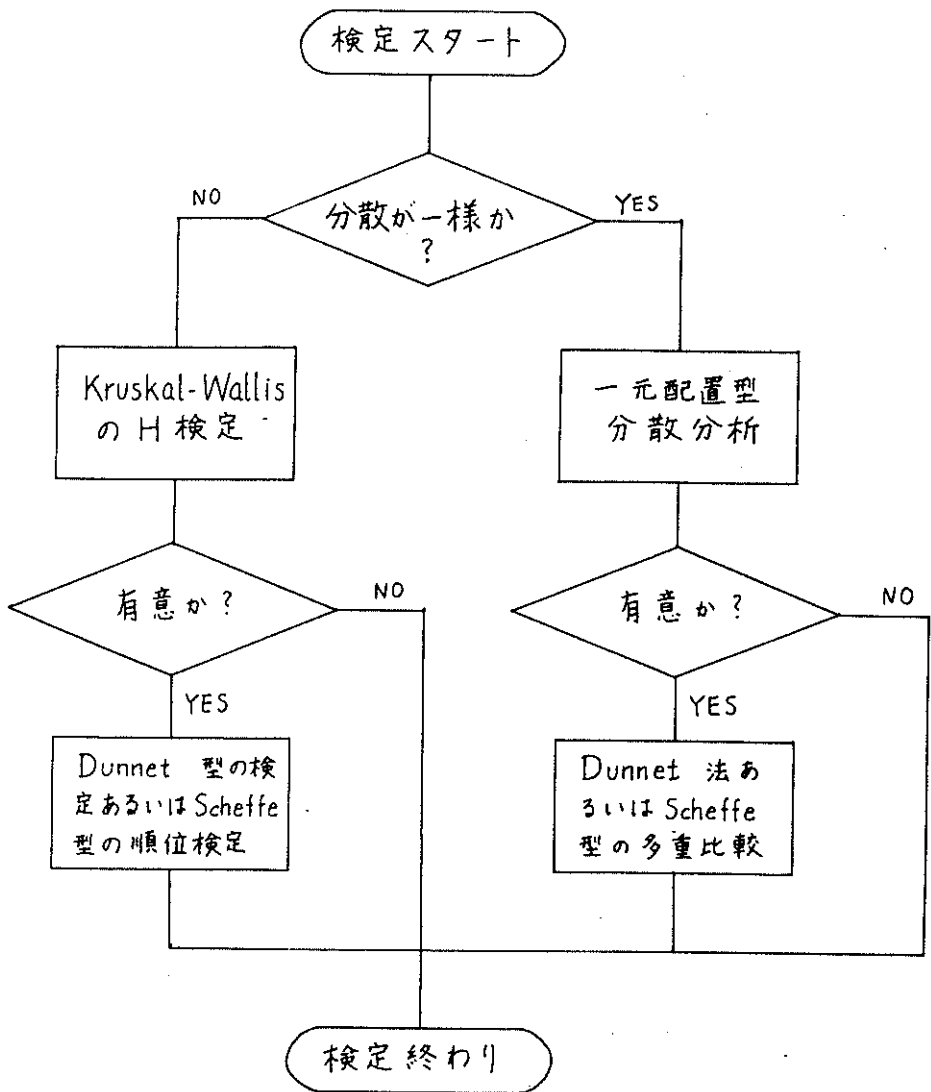


図1. 多重比較のフローチャート

6. 2 × 2 の分割表について

安全性試験でよく出てくるのは異常の有無の検定であろう。Siegel の分類表によれば、Fisher exact probability test と χ^2 test for two independent samples の 2 種類がある。このどちらを選んだらよいだろうか。あるいは両方行なうべきなのだろうか。この問題については、J.L. フライス、佐久間昭 訳『計数データの統計学』の第 2 章 4 分表の有意性検定が詳しい。

7. 順位和検定

安全性試験でも最近、順位和検定が用いられるようになってきた。多くは、Mann-Whitney の U 検定が用いられている。前出の岩原信九郎のノンパラメトリック法が

詳しい。ただしこの方法も、例えば生殖試験等のデータに用いる場合には、いろいろと問題が指摘されている。この問題については、吉村先生が詳しくふれられるので、ここではふれない。

8 日本ロシュで用いている検定方法

以上検定手法の選択上の問題を列挙してきたが、日本ロシュで、この問題についてどのように対処しているか事例を示し、御批判を頂きたい。

長期毒性試験については表3に示すように、データの種別と検定手法の一覧表が用意されており、実験No. とページおよび日付を記入して、実験計画書に組込むようになっている。繁殖試験についても表4に示すように同様なものが用意されている。

試験成績の表作成は、ほとんど電算化されている。用いている手法はすべて Fort-ran のサブルーチン形式で準備されており、X-Yプロッターのルーチンも含めて、約8000行である。

9. 統計手法の選択の考え方

次のような考えに従って手法が選ばれている。

- (1) 多重比較は一切行わない。
- (2) 差の検定はControl群と投与群間の2標本の検定に限定する。
- (3) Dose respons の検定もすべて行なう。
- (4) 順位和検定は用いない。
- (5) 検定は、直接確率を計算できるものを極力選ぶ。
- (6) t 検定は、標準偏差が異なる場合の結果を組込んでしまう。

一参考文献一

- 1) 厚生省GLP研究班：医薬品の安全性試験の実施に関する基準(案)。情報、No. 425-426(1981. 8. 24)
- 2) 日本規格協会：JISハンドブック品質管理(1982)
- 3) 滝沢 毅：毒性データ処理システムの開発(1)。第7回日本マンアスユーザーズグループ学術大会(1980)
- 4) 滝沢 毅・堀井郁夫・宇高全二：毒性データ処理システムの開発(2)。第8回日本マンアスユーザーズグループ学術大会(1981)
- 5) J. L. フライス, 佐久間 昭試：計数データの統計学。東大出版会(1975)
- 6) 薬効評価システム解説書(第1部)。Clinical Evaluation Vol. 3-No. 1(1975)
- 7) 岩原信九郎：ノンパラメトリック法。日本文化科学社
- 8) 日本工業標準調査会：二つの平均値の差の検定(JIS Z 9049)(1965)。

表3 長期毒性試験における統計手法の例

EXP. NO.: _____
 PAGE: _____
 DATE: _____

STATISTICAL METHODS :

The following statistical methods are to be employed.

- A. Randomization
 Sontage, J.M., Page, N.P., and Saffiotti, U.: Guidelines for carcinogen bioassay in small rodents, NCI, 1975.
- B. Pitman test
 Lehman: Nonparametrics, McGraw-Hill, p.43, 1974.
- C. The t-test with F-test for homogeneity of variance
 JIS: Significance test of difference between the two population means (standard deviations unknown, two-sided), Japanese Industry Standards, 29049, 1965.
- D. Fisher-Irwin test
 Fisher, R.A.: Statistical methods for research workers, (5th Ed.), Oliver & Boyd, Edinburgh, 1934.
- E. Exact rank sum test for 2 x k table
 Lehman: op. cit., p.21.
- F. Pitman test using exact generalization of the Fisher-Irwin test 2 x k table
 Cox, D.R.: The regression analysis of binary sequences, J.R. Stat. Soc., (B), 20, 215-242, 1958.
- G. Analysis of covariance
 Takizawa, T.: An unbiased comparison of organ weights when an inequality in body weight exists, Toxicology, 9, 353-360, 1978.
- H. Exact trend test
 Cox: op. cit.
- I. Dose response test (Jonckheere)
 Jonckheere, A.R.: A distribution-free k-sample test against ordered alterations, Bk, 41, 133-145, 1954.

Item	A	Treatment-control difference test					Dose response test		
		B	C	D	E	F	G	H	I
Randomization	o								
Clinical signs				o*	o**			o*	o**
Body weight		o#	o##						o
Food consumption		o#	o##						o
Hematology (general)		o#	o##						o
Hemogram (Baso.-Mono.)						o			o
Blood chemistry		o#	o##						o
Urinalysis (general)				o*	o**			o*	o**
Urinalysis (pH)						o			o
Organ weight (absolute and relative)		o#	o##						o
Adjusted organ weight							o		
Gross autopsy findings				o*	o**			o*	o**
Histopathology				o*	o**			o*	o**

* for dichotomous response, ** for graded response.

for n < 8, ## for n > 8.

All programs for statistical analysis are comprised in Tox.-DP system program documents in NRRC.

表4 生殖試験における統計手法の例

EXP. NO.: _____
 PAGE: _____
 DATE: _____

STATISTICAL METHODS :

The following statistical methods are to be employed.

- A. Randomization
 Sontage, J.M., Page, N.P., and Saffiotti, U.: Guidelines for carcinogen bioassay in small rodents, NCI, 1975.
- B. Pitman test
 Lehman: Nonparametrics, McGraw-Hill, p.43, 1974.
- C. The t-test with F-test for homogeneity of variance
 JIS: Significance test of difference between the two population means (standard deviations unknown, two-sided), Japanese Industry Standards, Z9049, 1965.
- D. Fisher-Irwin test
 Fisher, R.A.: Statistical methods for research workers, (5th Ed.), Oliver & Boyd, Edinburgh, 1934.
- E. Pitman test using exact generalization of the Fisher-Irwin test 2 x k table
 Cox, D.R.: The regression analysis of binary sequences, J.R. Stat. Soc., [B], 20, 215-242, 1958.
- F. The t-test with analysis of variance for nested design
 Healy, M.J.R.: Animal litters as experimental units, Appl. Statistics, 21, 155-159, 1972.
- G. Exact trend test
 Cox: op. cit.
- H. Dose response test (Jonckheere)
 Jonckheere, A.R.: A distribution-free k-sample test against ordered alterations, Bk. 41, 133-145, 1954.

Item	A	Treatment-control difference test					Dose response test	
		B	C	D	E	F	G	H
Randomization	o							
Body weight (dam)		o*	o**					o
Food consumption (dam)		o*	o**					o
Visceral change (dam)				o			o	
Litter data (numbers)		o*	o**					o
Litter data (rates)					o			o
Gross external change (fetus & neonate)					o			o
Microscopic visceral change (fetus)					o			o
Gross visceral change (neonate)					o			o
Skeletal change (fetus & neonate)					o			o
Ossification indices		o*	o**					o
Body weight (neonate)				o		o	o	
Survival indices Conception , viability & lactation indices					o			o
Postnatal development					o			o
Facuity test						o		o
Open field test		o*	o**					o
Learning test				o			o	
Reproductive performance test				o			o	

* for n < 8, ** for n ≥ 8.
 All programs for statistical analysis are comprised in Tox.-DP system program documents in NRRC..

[講師からの質問] 当日、講師から参加者に対し、統計手法について質問があり、挙手をもって答えてもらった。以下はその結果。

1. t検定にJIS規格があるのを知っているか？

(答) 約5人がyes, その中にはこの規格の制定委員だった人も含まれている。

2. 順位和検定, たとえばマンホイットニーのU検定を知っているか

(答) はほぼ全員yes

3. 実際に使っているか

(答) 約5人 yes

4. χ^2 検定およびイエイツの補正を知っているか

(答) 全員yes

5. いつもイエイツの補正を行っているか

(答) 約5人 yes

生殖試験における

順位和検定について

(十賢疑忘答)

金子洋二*・佐野正樹*

最近、生殖試験における検定方法として分布の型を仮定せずに済み、少数母体での異常値に検定結果が左右されにくい等の点から順位和検定が多用されている。

ところが橋敏明(1981¹⁾)は、生殖試験データに順位和検定が妥当性をもっているかをコンピュータでシミュレーションを行って調査した結果、比較的発現率の低い事象に対する通常の順位和検定では有意水準が不当に低く設定されてしまうため、ポジティブデータがネガティブなものとして評価される危険性があり、むしろ t -検定において、より妥当に有意水準が設定される傾向にあると指摘している。

そこで今回、これまでに集積された生殖試験データのうち、明らかに変化のなかった項目を除いて、 t -検定と順位和検定を同時的に行い、有意と判断される頻度を両種の検定間で比較検討した(表1~3)。

93項目についてそれぞれ検定を行ったところ、 t -検定では18項目が有意($p \leq 0.05$)、21項目が有意($p \leq 0.01$)であった。一方、順位和検定では17項目が有意($p \leq 0.05$)、19項目が有意($p \leq 0.01$)であり、その項目は t -検定でのそれとほぼ一致していた。有意な項目が全項目に占める割合は t -検定では41.9%、順位和検定で38.7%であり、ほぼ同率の値のように思われた。

これらの度数を通常の χ^2 法および累積 χ^2 法を用いて検定したところ、 t -検定と順位和検定での有意項目数には有意な差は認められなかった。これは橋氏のいう発現率の低い事象での変化のありそうな項目数が少なかったためかもしれない。

表4の死胚率についてはControl群と高用量群との間に実質科学上差異があると推測されながらも t -検定、順位和検定の両側検定ではいずれも有意とならなかった。しかし橋氏(1982²⁾)が死胚率には片側検定が適していると述べていることから、これで検定したところ、 t -検定のみで有意($p < 0.05$)となった。

したがって、比較的出現率の低い事象に対しては順位和検定より、 t -検定の方が有意となり易いが、それ以外の事象に対しては両検定のいずれを用いても同様であった。

文 献

- 1) 橋 敏明：生殖試験等の統計分析で順位和検定を用いることの危険性。先天異常, 21, 310(1981)
- 2) _____; 医薬品・食品添加物の“安全性”神話。汐文社, 61-13(1982)

表 1

项 目	Control	低用量	中用量	高用量
体重(胎仔) (g)	3.3693 ± 0.1921 No=14	3.3317 ± 0.24094 No=12 t=0.442 q=24 R=(198.5, 152.5)	3.3613 ± 0.18284 No=15 t=0.114 q=27 R=(202.5, 232.5)	2.746 ± 0.94592 No=13 t=2.328 q=12.919 R=(245, 133)
体重(新生仔) (g)	5.6250 ± 0.38646 No=20	5.6350 ± 0.44164 No=20 t=0.076 q=38 R=(409.5, 410.5)	5.2050 ± 0.36201 No=20 t=3.547 q=38 R=(523, 297)	5.1450 ± 0.23279 No=20 t=4.758 q=31.183 R=(551, 269)
体重(母体) (g)	200.6 ± 14.975 No=20	197.8 ± 20.078 No=20 t=0.499 q=38 R=(425, 395)	193.7 ± 18.085 No=20 t=1.314 q=38 R=(455.5, 344.5)	189.1 ± 20.013 No=20 t=2.057 q=38 R=(482.5, 337.5)
增体重 (g)	59.583 ± 9.2777 No=12	52.00 ± 7.3845 No=12 t=2.21 q=22 R=(184, 116)	51.833 ± 9.0034 No=12 t=2.07 q=22 R=(183, 117)	37.923 ± 12.665 No=13 t=4.33 q=23 R=(225, 100)
增体重(g)	58.929 ± 7.8883 No=14	49.083 ± 13.283 No=12 t=2.338 q=24 R=(231, 120)	55.4 ± 7.288 No=15 t=1.252 q=27 R=(234.5, 200.5)	18.077 ± 34.104 No=13 t=4.215 q=13.191 R=(276, 102)
增体重(g)	62.684 ± 13.861 No=19	53.467 ± 14.574 No=15 t=1.881 q=32 R=(395, 200)	43.533 ± 11.141 No=15 t=5.029 q=29.799 R=(406.5, 148.5)	t= q=
增体重(g)	22.435 ± 1.9866 No=20	21.340 ± 2.4286 No=20 t=1.537 q=38 R=(481.5, 338.5)	19.345 ± 2.500 No=20 t=4.326 q=38 R=(545.5, 274.5)	18.761 ± 2.7464 No=18 t=4.739 q=36 R=(523.5, 217.5)
增体重(g)	21.470 ± 1.817 No=20	21.186 ± 1.985 No=20 t=0.473 q=38 R=(432.5, 387.5)	20.135 ± 3.4715 No=20 t=1.523 q=28.685 R=(444, 376)	18.421 ± 3.9991 No=19 t=3.038 q=24.847 R=(514, 266)
增体重(g)	19.573 ± 2.818 No=20	193.06 ± 20.664 No=20 t=0.361 q=38 R=(412, 408)	183.15 ± 26.227 No=20 t=1.528 q=38 R=(458, 362)	174.9 ± 23.183 No=16 t=2.515 q=34 R=(442, 226)
增体重(g)	223.3 ± 7.3922 No=20	222.5 ± 10.476 No=20 t=0.27 q=38 R=(415.5, 404.5)	218.21 ± 16.099 No=19 t=1.25 q=24.98 R=(434, 346)	210.94 ± 15.360 No=16 t=2.956 q=20.517 R=(462, 204)
增体重(g)	124.4 ± 2.3377 No=20	128.20 ± 9.4736 No=20 t=1.346 q=38 R=(381, 439)	127.52 ± 7.0658 No=19 t=1.281 q=27 R=(371.5, 408.5)	129.71 ± 2.7448 No=17 t=1.887 q=35 R=(334, 369)
	t No=	t No=	t No=	t No=
		t= q=	t= q=	t= q=
		R=(,)	R=(,)	R=(,)
	t No=	t No=	t No=	t No=
		t= q=	t= q=	t= q=
		R=(,)	R=(,)	R=(,)

表 2

項 目	Con. in l	低用量	中用量	高用量
成長分化 腹部被毛 (田令)	7.25 ± 0.72729 No=20	2.25 ± 1.0195 No=20 t=1.428 q=38 R=(357, 463)	2.85 ± 0.96109 No=20 t=3.710 q=38 R=(229.5, 530.5)	9.5250 ± 1.175 No=20 t=5.420 q=31.693 R=(256, 564)
成長分化 脛開口 (田令)	35.500 ± 0.2717 No=20	35.190 ± 1.4765 No=20 t=0.819 q=29.856 R=(446, 374)	36.100 ± 0.2826 No=20 t=2.218 q=38 R=(374.5, 425.5)	37.000 ± 1.1921 No=20 t=4.623 q=31.693 R=(271, 549)
Open field 立上刈回数	11.950 ± 4.8933 No=20	9.350 ± 3.2235 No=20 t=1.261 q=38 R=(474, 346)	9.6316 ± 5.2727 No=19 t=1.424 q=37 R=(422.5, 330.5)	7.00 ± 4.0139 No=19 t=3.443 q=37 R=(508.5, 271.5)
Open field 立上刈回数	14.30 ± 6.514 No=20	12.895 ± 5.3426 No=19 t=0.734 q=37 R=(427, 353)	11.611 ± 4.9424 No=18 t=1.420 q=36 R=(422.5, 292.5)	9.4286 ± 3.8374 No=14 t=2.503 q=32 R=(416.5, 178.5)
Open field 立上刈回数	2.60 ± 2.28727 No=20	1.0500 ± 1.3169 No=20 t=2.193 q=26.647 R=(502, 318)	1.000 ± 1.000 No=19 t=2.345 q=23.743 R=(422, 292.5)	1.4737 ± 1.4286 No=19 t=1.561 q=22.165 R=(447.5, 322.5)
Open field 歩行量	46.625 ± 17.547 No=40	41.425 ± 20.530 No=40 t=1.217 q=78 R=(1729, 1511)	51.60 ± 12.183 No=40 t=1.245 q=78 R=(1536, 1704)	59.950 ± 22.261 No=40 t=2.973 q=78 R=(1381, 1859)
妊娠期間 (日)	21.795 ± 0.25167 No=22	21.818 ± 0.39481 No=22 t=0.230 q=35.647 R=(491, 499)	21.925 ± 0.40745 No=23 t=1.393 q=26.899 R=(462.5, 572.5)	22.318 ± 0.50111 No=22 t=4.374 q=30.960 R=(348.5, 641.5)
骨化数 (基節骨)	1.9348 ± 0.92147 No=23	2.313 ± 1.4069 No=23 t=1.057 q=44 R=(510.5, 570.5)	2.00 ± 1.3594 No=21 t=0.183 q=42 R=(522.5, 461.5)	0.75238 ± 0.66922 No=21 t=4.664 q=42 R=(684, 306)
骨化数 (胸骨板)	5.4565 ± 0.55172 No=23	5.6043 ± 0.20558 No=23 t=1.203 q=27.993 R=(527, 554)	5.4048 ± 0.34130 No=21 t=0.377 q=37.131 R=(567.5, 422.5)	4.2700 ± 0.51617 No=20 t=3.521 q=42 R=(665.5, 292.5)
	t No	t No	t No	t No
		R(,)	R(,)	R(,)
	t No	t No	t No	t No
		R(,)	R(,)	R(,)
	t No	t No	t No	t No
		R(,)	R(,)	R(,)
	t No	t No	t No	t No
		R(,)	R(,)	R(,)

表3

项目	Control	低用量	中用量	高用量
死胚率	3.333 ± 5.2678 No=12	5.067 ± 5.5636 No=12 t=0.760 q=21 R=(137, 161)	6.700 ± 7.0492 No=12 t=1.097 q=27 R=(137.5, 162.5)	10.700 ± 12.932 No=13 t=1.878 q=16.136 R=(130, 195)
死胚率	7.710 ± 9.1080 No=10	10.650 ± 15.521 No=10 t=0.516 q=18 R=(104, 106)	30.812 ± 24.955 No=8 t=2.489 q=2.496 R=(67.5, 103.5)	t= q= R=(,)
死产仔 (%)	2.000 ± 0.020 No=22	0.3087 ± 1.4305 No=22 t= - q= - R=(495, 540)	1.925 ± 2.4205 No=24 t= - q= - R=(506, 575)	16.357 ± 31.677 No=22 t= - q= - R=(418, 572)
生存率	99.00 ± 0.4721 No=20	96.250 ± 12.234 No=20 t=0.944 q=23.938 R=(421, 399)	91.995 ± 17.179 No=20 t=1.570 q=21.06 R=(450, 370)	81.665 ± 23.103 No=20 t=2.724 q=19.951 R=(482, 397)
出生率	94.655 ± 6.6522 No=20	96.335 ± 5.6589 No=20 t=0.860 q=38 R=(372.5, 442.5)	91.75 ± 7.0035 No=20 t=0.361 q=38 R=(403, 417)	90.020 ± 9.9495 No=23 t=1.731 q=38 R=(457.5, 362.5)
学习能 (T-maze) 工-回致	7.45 ± 6.4093 No=20	7.3421 ± 6.4314 No=19 t=0.190 q=37 R=(392.5, 387.5)	14.404 ± 12.524 No=18 t=2.129 q=26.805 R=(324, 417)	11.533 ± 8.9751 No=15 t=1.568 q=33 R=(318.5, 311.5)
学习能 (T-maze) 工-回致	5.1333 ± 4.7036 No=15	11.429 ± 22.125 No=14 t=0.826 q=13.478 R=(236, 199)	18.333 ± 52.146 No=15 t=0.976 q=14.227 R=(225.5, 239.5)	11.533 ± 8.9751 No=15 t=2.533 q=21.723 R=(178.5, 236.5)
骨化不全胎仔数 (肋骨) (%)	22.521 ± 21.226 No=14	22.225 ± 22.509 No=12 t=0.257 q=24 R=(187.5, 163.5)	27.179 ± 22.513 No=15 t=0.101 q=37 R=(212, 223)	15.335 ± 37.553 No=13 t=1.129 q=25 R=(202.5, 136.5)
骨格变异胎仔数 (%)	2.4421 ± 13.653 No=19	22.305 ± 25.665 No=21 t=2.191 q=31.088 R=(316, 504)	2.4619 ± 10.789 No=21 t=0.005 q=38 R=(375.5, 444.5)	t= q= R=(,)
骨格变异胎仔数 (%)	4.0176 ± 6.2654 No=17	5.9053 ± 12.311 No=19 t=0.579 q=28.806 R=(313, 353)	7.2294 ± 14.674 No=17 t=0.817 q=22.624 R=(294, 301)	9.8944 ± 7.5498 No=18 t=2.079 q=33 R=(252, 378)
内脏异常胎仔数 (%)	2.500 ± 11.180 No=20	2.500 ± 11.180 No=20 t=0.00 q=38 R=(410, 410)	5.556 ± 16.169 No=18 t=0.683 q=36 R=(379, 362)	50.00 ± 51.640 No=16 t=3.612 q=16.123 R=(294, 372)
Rota Rot 落下回数	4.0909 ± 3.1609 No=22	5.6364 ± 5.0099 No=22 t=1.223 q=35.432 R=(462, 523)	7.7619 ± 6.2527 No=21 t=2.760 q=32.516 R=(384.5, 561.5)	7.00 ± 4.6308 No=19 t=2.376 q=37 R=(384.5, 476.5)
Rota Rot 落下回数	2.8636 ± 2.7481 No=22	5.4091 ± 3.4179 No=22 t=2.722 q=47 R=(371.5, 518.5)	6.5714 ± 5.2780 No=21 t=2.269 q=28.779 R=(362, 523)	5.9244 ± 4.4122 No=15 t=2.580 q=27.279 R=(350.5, 461.5)

表 4 死 胚 率

Control

1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	7.100	7.700
2)	12.50	13.30								

Min= 0.0000 Max= 13.30 Range= 13.30 Mean= 3.393 S.D.= 5.270

No=12

低 用 量

1)	0.0000	0.0000	0.0000	0.0000	0.0000	6.300	6.700	7.100	7.100	7.100
2)	7.700	18.80								

Min= 0.0000 Max= 18.80 Range= 18.80 Mean= 5.067 S.D.= 5.564

No = 12

$$t = 0.760, \quad Q = 22, \quad R = (139, 161)$$

中 用 量

1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	12.50	12.50	13.30
2)	15.40	26.70								

Min= 0.0000 Max= 26.70 Range= 26.70 Mean= 6.700 S.D.= 9.049

No = 12

$$t = 1.097, \quad Q = 22, \quad R = (137.5, 162.5)$$

高 用 量

1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	8.300	8.300	15.40	16.70
2)	21.40	33.30	35.70							

Min= 0.0000 Max= 35.70 Range= 35.70 Mean= 10.70 S.D.= 12.93

No=13

$$t = 1.878, \quad Q = 16.136, \quad R = (130, 195)$$

分 布

Control

Class' width= 2.687

Class	Freq	%
1)-0.6650E-01- 2.620	6	66.7
2) 2.620 - 5.307	0	0.0
3) 5.307 - 7.993	2	16.7
4) 7.993 - 10.68	0	0.0
5) 10.68 - 13.37	2	16.7

低 用 量

Class' width= 3.798

Class	Freq	%
1)-0.9400E-01- 3.704	5	41.7
2) 3.704 - 7.501	5	41.7
3) 7.501 - 11.30	1	8.3
4) 11.30 - 15.10	0	0.0
5) 15.10 - 18.89	1	8.3

中 用 量

Class' width= 5.393

Class	Freq	%
1)-0.1335 - 5.260	7	58.3
2) 5.260 - 10.65	0	0.0
3) 10.65 - 16.05	4	33.3
4) 16.05 - 21.44	0	0.0
5) 21.44 - 26.83	1	8.3

高 用 量

Class' width= 7.211

Class	Freq	%
1)-0.1785 - 7.033	6	46.2
2) 7.033 - 14.24	2	15.4
3) 14.24 - 21.46	3	23.1
4) 21.46 - 28.67	0	0.0
5) 28.67 - 35.88	2	15.4

0. 增体重(g)

Control

less changed data.

11	37.00	50.00	50.00	51.00	55.00	55.00	56.00	56.00	58.00	61.00
21	41.00	59.00	59.00	69.00	70.00	71.00	73.00	79.00	101.0	

Min= 37.00 Max= 101.0 Range= 64.00 Mean= 62.48 S.D.= 13.94

$\lambda = 14$

低用量

less changed data.

11	33.00	37.00	40.00	43.00	49.00	49.00	50.00	52.00	55.00	58.00
21	36.00	52.00	44.00	51.00	52.00	59.00	59.00	62.00	65.00	68.00

Min= 33.00 Max= 62.00 Range= 59.00 Mean= 52.47 S.D.= 14.50

$N_0 = 15$

$\lambda = 1.881 \quad \lambda = 32, \quad R = (395, 200)^*$

分布

<u>Control</u>				<u>低用量</u>			
Class' width= 22.02				Class' width= 11.92			
Class	Freq	%		Class	Freq	%	
11 30.00	-	0.00	1. 0.3	11 32.71	-	0.00	0. 0.0
21 47.01	-	0.00	0. 0.0	21 44.02	-	0.00	0. 0.0
31 52.50	-	0.00	7. 2.4	31 56.54	-	0.00	3. 20.0
41 70.00	-	0.00	7. 2.4	41 80.00	-	0.00	0. 0.0
51 82.50	-	0.00	1. 0.3	51 90.00	-	0.00	1. 6.7
			1. 5.7				

28. 骨化不全胎仔数(胸骨核) %

Control

less changed data.

11	0.0000	0.0000	0.0000	0.0000	14.50	25.00	29.00	29.00	37.50	42.00
21	44.40	50.00	55.00	62.50						

Min= 0.0000 Max= 62.50 Range= 62.50 Mean= 26.52 S.D.= 21.29

$N_0 = 14$

高用量

less changed data.

11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
21	0.0000	100.0	100.0							

Min= 0.0000 Max= 100.0 Range= 100.0 Mean= 15.18 S.D.= 37.55

$N_0 = 13$

$\lambda = 1.129, \quad \lambda = 25, \quad R = (242.5, 435.5)^*$

分布

<u>Control</u>				<u>高用量</u>			
Class' width= 12.02				Class' width= 20.20			
Class	Freq	%		Class	Freq	%	
11-0.3125	-	0.00	0. 20.0	11-0.5000	-	0.00	11. 00.0
21 12.38	-	0.00	1. 7.3	21 19.70	-	0.00	0. 0.0
31 24.94	-	0.00	4. 28.0	31 39.90	-	0.00	0. 0.0
41 37.50	-	0.00	3. 21.0	41 60.10	-	0.00	0. 0.0
51 50.10	-	0.00	2. 14.3	51 80.30	-	0.00	2. 15.4

33. Rata rat (塔下回数)

Control

Line changed data	0.0000	1.0000	2.0000	3.0000	4.0000	5.0000	6.0000	7.0000	8.0000	9.0000
1) 0.0000	0.0000	0.0000	0.0000	1.0000	1.0000	2.0000	2.0000	3.0000	3.0000	4.0000
2) 2.0000	2.0000	2.0000	3.0000	3.0000	4.0000	4.0000	5.0000	5.0000	6.0000	6.0000
3) 4.0000	4.0000	5.0000	5.0000	6.0000	6.0000	7.0000	7.0000	8.0000	8.0000	9.0000

Min= 0.0000 Max= 13.00 Range= 13.00 Mean= 2.364 S.D.= 2.748

No= 22

中用量

Line changed data	0.0000	1.0000	2.0000	3.0000	4.0000	5.0000	6.0000	7.0000	8.0000	9.0000
1) 0.0000	0.0000	1.0000	2.0000	2.0000	3.0000	4.0000	4.0000	5.0000	6.0000	6.0000
2) 4.0000	4.0000	5.0000	6.0000	6.0000	7.0000	7.0000	8.0000	9.0000	10.0000	10.0000
3) 20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00

Min= 0.0000 Max= 20.00 Range= 20.00 Mean= 4.571 S.D.= 5.270

No= 21 $t = 2.869, Q = 29.79, R = (362, 584)$ **

高用量

Line changed data	0.0000	1.0000	2.0000	3.0000	4.0000	5.0000	6.0000	7.0000	8.0000	9.0000
1) 1.0000	1.0000	1.0000	2.0000	2.0000	3.0000	4.0000	4.0000	5.0000	6.0000	6.0000
2) 4.0000	4.0000	5.0000	6.0000	6.0000	7.0000	7.0000	8.0000	9.0000	10.0000	10.0000
3) 20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00

Min= 1.0000 Max= 17.00 Range= 16.00 Mean= 5.666 S.D.= 4.412

No= 18 $t = 2.580, Q = 27.279, R = (350.5, 469.5)$ **

分布

Control

Class' width= 2.026

Class	Free
1) 0.0000-2.026	13. 59.1
2) 2.026 - 4.052	7. 31.8
3) 4.052 - 6.078	2. 4.5
4) 6.078 - 8.104	0. 0.0
5) 8.104 - 10.130	1. 4.5

中用量

Class' width= 4.040

Class	Free
1) 0.0000-4.040	5. 23.0
2) 4.040 - 8.080	0. 0.0
3) 8.080 - 12.120	4. 18.6
4) 12.120 - 16.160	0. 0.0
5) 16.160 - 20.200	2. 9.5

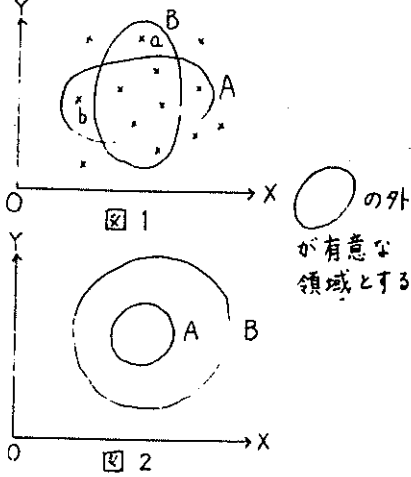
高用量

Class' width= 3.232

Class	Free
1) 0.0000-3.232	11. 01.1
2) 3.232 - 6.464	2. 11.5
3) 6.464 - 9.696	1. 5.8
4) 9.696 - 12.928	3. 16.7
5) 12.928 - 16.160	1. 5.8

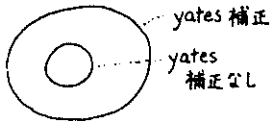
【質問1】 順位和検定とt検定ではどちらが検出力が高いか？

【答】 順位和検定とt検定との違いを理論的に正確に説明することは難しいので概念的に説明する。いま、例えば二つの特性X, Yを何人かについて計って図1の×



印のような結果を得たとする。これについて、AとBの2つの検定法で異常者を判定するものとする。このとき有意水準を同じにとれば、AとBの関係は、図1のようであって、図2のようにはならない。すなわちa点はA検定では有意であるが、B検定では有意にならないし、逆にb点は反対の結果となる。そういうように棄却感ができる。そしてA検定とB検定ではどちらが有意になりやすいかということ議論することほできないことになる。これに対して、二つの検定の棄却感が図2のようなときには、A検定の方がB検定より有意差ありと出やすいことになる。このときにはAとBのどちらが正しいかということ論ずる必要が生じる。しかし、この正しくない、間違っているということほ1+1=3という意味においてではない。それは現実にはどのような分布が成り立っているかによって異なってくるからである。相手のデータ次第でAが正しかったり、Bが正しかったりすることがあるわけである。それゆえ、現在のデータの分布が何であるかを慎重に見た上で議論を進めなくてはならない。近似計算をしなくてはならない場合、よい近似方法を用いなければならない。例えば、2×2分割表で χ^2 検定をやるときには、yatesの補正をやった方が有意差が出にくくなるが、この補正はやる方が正しい。左の図のような関係になるからである。

いかということ論ずる必要が生じる。しかし、この正しくない、間違っているということほ1+1=3という意味においてではない。それは現実にはどのような分布が成り立っているかによって異なってくるからである。相手のデータ次第でAが正しかったり、Bが正しかったりすることがあるわけである。それゆえ、現在のデータの分布が何であるかを慎重に見た上で議論を進めなくてはならない。近似計算をしなくてはならない場合、よい近似方法を用いなければならない。例えば、2×2分割表で χ^2 検定をやるときには、yatesの補正をやった方が有意差が出にくくなるが、この補正はやる方が正しい。左の図のような関係になるからである。



補正はやる方が正しい。左の図のような関係になるからである。

【質問2】 t検定をするときS-D = 0の場合、どういように検定したらよいか

Ex. ラットの妊娠期間の場合21日か22日になってしまう。S-D=0

【答】 統計的検定というものはデータ上で示された差が、誤差的なバラツキよりも明らかに大きいか否かを判断するものである。有意差ありということは誤差的な変動をこえてはっきりと差があるということの意味し、有意差なしということは、データ上の差が誤差的なバラツキのうちに入る程度に小さいということの意味する。これに対し、この有意差ありということ統計的に確かめられた差が、技術的な意味で差があるか否かということとはまた別の問題である。例えば、ぼくと彼とがゲームを2万回やって、相手の勝率が51%でぼくの勝率が49%だ、たとする。このときに、その差は統計的に有意であるといえるが、技術的に、意味のある差だとは言えないだろう。回数を大きくすると統計的変動は小さくなり、したがって誤差の変動と無視できるようになるがその差が巧みさには差があると言いうるものか否かというのは全く別の問題で、勝率51%では、5回ぐらい勝負しても、どちらが勝つかは全く予断と許さなわけ。従って両者の強さは傍目からみて、五分五分なわけである。言いかえれば、誤差の大きなデータについては、はっきりした結論は出

せない。誤差が全くないデータならば、その結果と全く技術的な差とみなしてよいというわけである。質問された例が、

G_1	20日	$S-D=0$
G_2	21日	

というようにバラツキ0ならば、それはこの測定値を100%信頼してよいことを意味し、従って $G_1 < G_2$ が言える。しかし、問題はこの1日の差が実際問題として技術的な意味のある差といえるのかである。例えば、ヒトで妊娠期間が230日と281日だといった場合、これを差があると考える人はいないであろう。これに対して、片方が200日で他方が300日だと問題になるわけである。別の場合として、

G_1	$\sigma_1=0$	20日	3匹
G_2	$\sigma_2=0$		10匹

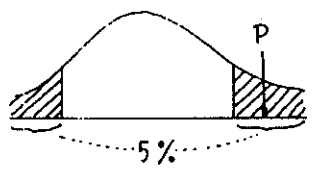
というように例を考えると、この時はサンプルサイズが違うから、 $\sigma_1=0$ といってもサンプルが小さいためににまにまバラツキの推定値が0になったと考える方が自然であろう。このときは、 $\sigma_{G_1}^2 = \sigma_{G_2}^2$ という前提をとって、全体を合わせて分散の推定をするのが自然である。つまり、

$$\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2 + \sum (y_i - \bar{y})^2}{m + n - 2} \quad (\sigma_{G_1} = \sigma_{G_2})$$

とするのである。私なら $\sigma_{G_1} = \sigma_{G_2}$ として処理する。

多重比較を使うか否かからすべての対について検定するか否か。この問題は現実問題としては、ずいぶん迷う問題である。

すべての対について検定すると、一般に検定は有意となりすぎる傾向がある。何故なら、下図のようなある分布があるものとする。この分布から一個とり出してき



たときに、それが斜線部分に入る確率は5%程度である。ところが、ここから10個とり出して、その最大値が図のPのところにくる確率は5%どころではなくてはるかに大きくなるからである。へたすると20%ぐらいになるかもしれない。極端な場合、20個とり出せば

1個ぐらい斜線部分に入るのほかに1前のごとである。すべての対について検定するということは、このようなことをやっていることに近い。今、 n 個データがあるとすれば、すべての対の数は、 $\frac{1}{2}n(n-1)$ であるから $n=5$ のときですら

$\frac{1}{2} \times 5 \times 4 = 10$, $n=10$ のときは $\frac{1}{2} \times 10 \times 9 = 45$, つまり45回検定することになる。

実は検定が独立でないからそれほど極端にはならないが、45回も検定と行なえば、帰無仮説が正しくて有意水準5%の検定をやっても2回や3回は有意となることになる。だから、あらゆる対について検定して2回か3回有意となったからといっても、それが差のあることを示すとはいいにくい。だからあらゆる種類の対について検定するときは、結果と少々割り引いて考えなくてはならない。では、どのくらい割り引きすればよいのかというのは少々難しい問題である。そのために、工夫された方法が多重比較である。つまり多重比較というのは、全部こみにして、最大と最小の差をとると、サンプルが大きいときには、サンプルが小さいときより差が大きくなるのが当然であることを考慮したやり方なのである。

増山先生の COMMENTS

今の吉村さんの話に全面的に賛成。少々思いついたことを付け足したい。それは妊娠期間の分布についてである。ラットでは妊娠期間の分布に対数正規がよくあてはまる。それから、日を単位にすると、にわたりのふ化の時間の分布は個体差が非常に小さいために差が出ない。しかし、ふ化するときに光を当てると早くふ化するという実験においては、時間単位でやると、条件が一定していると、きれいに対数正規分布となる。しかもふ化の条件に、関係なしに、タマゴの個体差は同じになる。この W_0 がお役に立てば幸いである。ヒトの場合、オランダのぼう大なデータによるとヒトの妊娠期間の分布についても対数正規分布がよく合っている。ただし、 W_0 (妊娠日数) 軸上で、下図のようなすえが見られる。これは早産である。これらのデータは医学的にみて正常と思われる場合であるが、尾をひいている。しかし、このデータが、もっともらしいのは、ヒトの場合でもラットの場合でも個体差の指標である σ がほとんど一定しているからである。



フィッシャーの直接確率の両側仮説検定の求め方

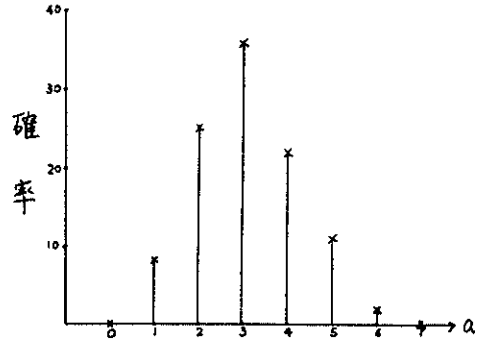
田中 健*

質問：両側仮説検定の求め方に次の2通りある。どちらを採用すべきか？

- (1) 両側水準 α の検定では、 $P \leq \frac{\alpha}{2}$ のとき有意とする。
 (増山元三郎, 佐久間昭, F.D.A. コントローラー委員会等)
- (2) H_0 の下で H_1 の方向に度数表の確率をその合計が α または、それ以下で最大になるまで加えて P を求め、 $P \leq \alpha$ のとき有意とする。
 (米虫節夫, 栗谷典量, 芥田大公望, 田中恒男,)
 (D. Mainland, P. Armsen, G.H. Freeman 等)

〈例〉 下の結果が5%で有意かどうかを見たい。

発癌性 薬剤	(あり)	(なし)	計
A	0	7	7
B	8	4	12
計	8	11	19



分割表は次のようになる。

0 7	1 6	2 5	3 4	4 3	5 2	6 1	7 0	a b
8 4	7 5	6 6	5 7	4 8	3 9	2 10	1 11	c d
56	37	18	1	20	39	58	77	ad-bc
P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	= 1.0
0.00655	0.07335	0.25673	0.36675	0.22922	0.06113	0.00611	0.00016	

(1)の場合 $P = \frac{8! \times 11! \times 7! \times 12!}{19!} \times \frac{1}{8! \times 7! \times 4!} = 0.00655$

$2 \times P = 2 \times 0.00655 = 0.0131 < 0.05$

(2)の場合 $p = \frac{8! \times 11! \times 7! \times 12!}{19!} \times \left(\frac{1}{8! \times 7! \times 4!} \times \frac{1}{6! \times 2! \times 10!} \times \frac{1}{7! \times 11!} \right)$
 $= 0.00655 + 0.00611 + 0.00016 = 0.01282 < 0.05$

* (株) 日本科学技術研修所

〈参考文献〉

- 1 増山元三郎 : 小教例のまとめ方 竹内書店(1964)
- 2 佐久間 昭 : 薬剤評価I 東京大学出版会(1977)
- 3 コントローラー委員会 : 薬剤評価システム解説書 臨床評価3巻1号(1975)
- 4 栗谷典量 : 二重盲検法の実際と盲点 医薬ジャーナル(1977)
- 5 栗谷, 米虫 : 医学のあゆみ誌くすり欄の臨床比較試験結果の要点一覧表 医薬ジャーナル 17巻3~10 (1981)
- 6 斧田大公望 : 薬効の評価法—臨床推計学の確立と— クリニカ 2巻2号(1975)
- 7 G.H.Freeman : Note on an exact treatment of contingency, goodness of fit and other problems of significance, 38, 141, Biometrika
- 8 P. Armsen : Tables for significance tests of 2×2 contingency tables, 42 494. Biometrika

【増山先生の答】

この場合の例題が悪いから両方も似た結果となっていてるのであって、論理的に言うと(2)の方が正しい。

【会員からのコメントと質問】

★コメント★ 2×2 の分割表において米虫論文を使う場合、次のような点に注意する必要があります。即ち 2×2 分割表で、

効果	有	無	①	②	③	④
A薬	a	b	0.7	0.4	70	48
B薬	c	d	84	0.7	4.8	70

と並べた場合、いずれも同じ結果になるべきであるが、実際は a, b, c, d のどこにおくかを指定しないと異なった値になる。b, c の位置に minimum かくるよう配置しなくてはならない。したがって、上記②, ③のように並べると、正しい計算結果が得られる。

★質問★ 例えは次のような結果を得たとする。

	卍	卄	十	士	一
A					
B					

このとき、 χ^2 法を適用しようとするとき、どこでクラス分けをしたらよいか明確でない。つまり卍と士のクラスで検定すれば有意となりやすいであろうし、卍&卄の組と十&士&一のクラスに適用すればまた違う結果が出るであろう。どうしたらよいか。

★答(吉村)★ 各クラスに順序がない場合は 2×5 分割表として χ^2 検定するのが本当である。順序がある場合はその順序という情報を用いて分析する方が一般には望ましい。このとき使われるのが前回の累積 χ^2 法等である。累積法の基本的考え方は順序がついているものをいくつかの部分で分けて、それぞれの分割点ごとに独立性の検定を見い出そうということである。

★質問★ 一般に χ^2 法と直接確率法ではどちらを用いるべきであろうか。

★答(吉村)★ 確率論的に言うと、直接確率が計算可能な場合であれば、直接確率法の方が正確な答を得られる。しかし、 2×2 ではなく、 3×4 とか 3×3 になると、どういう形で“すそ”の方を定義すればよいのかよく分らない場合ができてくる。とにかく独立性が成り立たないという仮説だけだから、どういう条件になったときが対立仮説に有利で帰無仮説に不利であるのかが分りにくくなる。 2×2 分割表の χ^2 検定は yeates の補正をやらないと有意に出やすい。必ず補正しなければならない。 2×3 とか 3×3 分割表においても補正が必要なのだが、そのためのテクニックが存在しない。したがって χ^2 検定より直接確率計算法の方がよいことになる。一般に χ^2 検定の方が有意差ありと出やすくなるのは近似法の関係である。

★コメント(増山)★ 佐藤長一郎さんの無相関検定というのがある。

統計のParadoxについて

— 数字に騙されないために —

山崎昌男*

読者の皆様方は統計学の手法について良く御存知で、私などの出る幕ではないと思いますが、統計数字の取扱いや、評価等について気のついた事を、二三記してみたいと思います。

1. 疾病の因果関係について

先日、高校の同窓会で、医学部の学生数人と話をしているうち、一人が、「絨毛上皮腫⁽¹⁾」の患者の既往症を調査すると、6割は破壊胞状奇胎⁽²⁾がある。これを考えると、破壊胞状奇胎は非常に恐ろしい病気と思われる。しかし、破壊胞状奇胎の患者を追跡調査すると、絨毛上皮腫に移行するものは0.1%しかないのど、それほど恐ろしくはないとも思える。実際にはどちらの考え方が正しいのだろうか。これは統計のparadoxだ。」

と言った。実際は、この数字は胞状奇胎の既往症がない者が絨毛上皮腫になる確率がごく低いことを表わしているだけで、paradoxでも何でもない。私はそれを説明した後、

「ある大学では、どの学部をとっても男性の合格率の方が女性の合格率より高いが大学全体で見ると女性の合格率の方が高い。⁽³⁾こんなことがあると思うか。」

と尋ねた。

「信じられない。」

「ところが実際にあるんだ。これから見ると、前のはparadoxのうちに入らないだろう。」

と畳みかけたが、いつの間にか「数学科 = 変人集団」の診断を下されてしまった。合格率の問題は、「女性の多くが易しい学部を受けた」と考えれば、実際に今のよ様な状況になっている例が容易に作れる。パズル好きな方は、今のヒントに基づいて実例を作ってみては如何が。

—— 考 慮 時 間 ——

今の実例は、例えば学部をA、Bの2つとして、

合格者 志願者	学部 A	学部 B	合計
男性	$\frac{18}{60}$ (30%)	$\frac{32}{40}$ (80%)	$\frac{50}{100}$ (50%)
女性	$\frac{4}{20}$ (20%)	$\frac{56}{80}$ (70%)	$\frac{60}{100}$ (60%)

とすればよい。

* 東京大学理学部数学科

この問題はまた次のように変形される。

「ある町で、寿命と飲酒の関係を調べた所、男性のうち飲酒者は6割が、非飲酒者は5割が75才以上まで生きる。飲酒は長寿に有効であるといえるか。」(調査数は十分であるとし、75才以上生きることを長寿とする)

これに対して、「この町の人々は体質Aと体質Bとに二分され、体質Aは弱い体質であり、特に酒に弱いので大半の人は酒を飲まない。一方、体質Bは強く、酒もいける口なのが多くの人酒を飲むが、飲まなければなお長生きできる」という状況を考え、前と同じ表を次のように書きかえた例が考えられる。

長寿者 全体	体質A	体質B	合計
非飲酒者	$\frac{18}{60}$ (30%)	$\frac{32}{40}$ (80%)	$\frac{50}{100}$ (50%)
飲酒者	$\frac{4}{20}$ (20%)	$\frac{56}{80}$ (70%)	$\frac{60}{100}$ (60%)

すなわち、どちらの体質にとっても飲酒が長寿に有害な例が作れる。但し、前の場合と違って、今度は仮説検定の問題だから、反例が1つあるから nonsense と言うことはできない。そこで、上の(体質を考えない)数字から、飲酒が長寿に有効であると結論する通常の手続きを考えてみよう。

まず、「飲酒は長寿に有効でない」という帰無仮説を立て、次に標本の大きさや長寿率から危険率を計算する。前の例のように、両集団の長寿率に10%の差があれば、ある程度標本が大きければ、危険率が1%以下となり、帰無仮説が否定されて「飲酒は長寿に有効」という結論が得られるのである。ここで、「危険率1%」とは、飲酒が実際には長寿と何の関係もないとき、同じやり方の仮説検定を100回行くと(平均して)99回は帰無仮説は棄却されず、1回は棄却される、という意味であった。しかし、「100回中99回は棄却されない」のは、検定の手順、特に、標本の選出が適切な場合に限られる。今の場合でいえば、飲酒者を「現在酒を飲んでいる人から抽出する」のではなく、「新生児から標本を抽出し、この標本には好みに拘わらず酒を与える」によって選び、その結果として10%の差が出て来れば帰無仮説の棄却は適当なのである。(実験の現実性は別として)

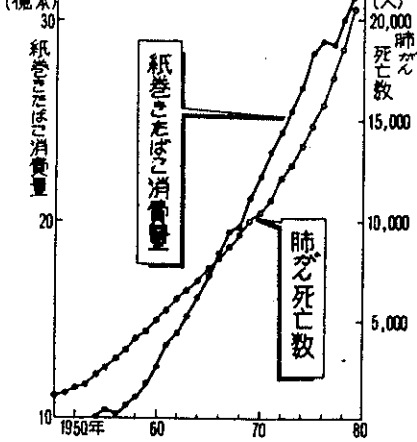
この変形された問題についても医学部の友人に質問してみたが、やはり、「数値を見ただけで、飲酒は有効であると思う。しかし、標本の抽出に問題があると言われれば、その通りだと思う。」

と言うことであつた。人々の健康を与かる立場にいる方々には、表面上の数字に惑わされずその意味を深く考え、慎重に結論を下すようになっていただきたい。

2. マスコミの統計への無理解について

次ページのグラフは先日の朝日新聞⁽⁴⁾に載せられた、「日本における紙巻きタバコ消費量と肺がん死亡数の増加ぶり」というグラフである。これを見て何に気がつくだろうか。まず、紙巻きタバコ消費量のグラフは、1950年頃にはグラフの下端にあることである。それ以前には紙巻きタバコはなかったのだろうか。そんなはずはない。そこでグラフをよく見ると、紙巻きタバコ消費量のグラフの縦軸がカットされているのがわかる。一方の肺がん死亡数の縦軸はちゃんと0から始まっている。これはおかしい。グラフを書くとき、変化を見易くするため、縦軸の一部をカット

日本における紙巻きたばこの消費量と
肺がん死亡数の増加ぶり



『喫煙の医学』（講談社）から

することはよく行なわれている。しかし、場合にもよるが、今のように伸び率を問題にした場合は、カットしない方がよいと思う。特に、今のように異なる2つのデータを同一のグラフ上に描いて比較する場合には、せめて一方をカットするならば、同じ比率で他方もカットすべきである。(例えば一方の縦軸が40~60ならば、他方は100~150というように) それではデータの正確さが損われると言われるかも知れないが、もともとグラフは大雑把な感じを掴むにのみのものであり、細かいデータと伝えるためのものではない。正確に伝えるにすれば、アナログ表示のグラフをやめデジタル式に表にすればよい。今のグラフの場合、縦軸をちゃんと直すと、紙面のものと

は逆に、肺がん死亡数の伸びの方が大きいことがわかる。(試みられよ) これは、グラフの書き方によって故意に表現を歪めることができるという事を示している。(多くの人には私のように意地悪な見方とせず、ざっと見て、たばこの消費量の伸びの方が大きいと思うだろうから) 今の場合は故意ではないと思うが、物価上昇率のグラフが今だに片対数式にならないのは何故だろうか。物価高を強調したいために、故意に不適当なグラフを描いているのだろうか。

さらに、こんな記事もあつた。⁽⁵⁾

「前略」 刑法犯少年の数は、47年~56年の10年で、全国では83%も増えたのに沖縄では7.5%と微増にとどまっている。沖縄に非行が多いとの一般的な言い方は、『統計のウソ』とも言える(後略)」

沖縄に非行が多いという主張に反論するために、刑法犯少年の数の統計を使いに行けば、刑法犯少年の数の当該年令の少年全体の数に対する比を用いるしかない。それを用いないで、伸び率だけで主張とすることは、それこそ『統計のウソ』である。

世論をリードして行くべき新聞等のマスコミは、もっと統計数字と慎重に扱うべきである。一方、マスコミにこのような記事が出る現状では、読者の方が注意して安易な表現に騙されないようにする必要がある。

〈注釈及び引用文献〉

- (1) 子宮に発生する悪性腫瘍の一種
- (2) いわゆる「ブドウ子」の重症のもの
- (3) 数学セミナー (日本評論者刊) 16-6 (1977) 24~26
- (4) 朝日新聞 昭和57年4月15日
- (5) 同上 昭和57年5月12日

なお、本文初めに出て来る医学の話に出てくる数字については、筆者の記憶違いがあるかもしれません。また、飲酒と長寿の関係に出てくる例については、筆者の創作であつて、実際の酒の薬理効果とは全く関係がないこととお断りしておきます。

ニ ュ ー ス

第2回 フッ素研究会 プログラム

- 第2回 フッ素研究会 会長 高橋 暁正 (前東京大学医学部講師)
- 日時と会場
 - 1982年8月28日(土) 15:00~17:00
新潟県小千谷市 木下十景会議室 (小千谷市城内町 ☎ 02558-2-5657)
 - 1982年8月29日(日) 9:00~17:00
新潟県小千谷市 市民会館大ホール (小千谷市土川94-3 ☎ 02558-2-9111)

■ 8月28日(土) _____ [会場=小千谷市木下十景]

15:00 ~ 15:05 開会 (高橋 暁正)

15:05 ~ 15:20 座長 藤井 彰 (日本大松戸歯学部)

• 一般演題1 ... Fの主体へ及び可影響(IV)一亜急性下経飲用時に於ける骨Ca摂取取免の血清成分値の変動(10分) 柳澤 文徳 土田 莉 (東京医歯大難治研)

• 一般演題2 ... Fによる予防効果評価のFの統計的指標 (秋上登泰)
高橋 暁正 (東京医歯大前講師)

• 一般演題3 ... Fによる加工製菓器からのF抽出 (秋上登泰)
秋庭 賢司 (東京慈恵医大歯科) 討論 5分

15:20 ~ 16:30 座長 高橋 暁正 (前東京医歯学部)

• 特別報告1 ... Fを利用したFの統計的諸問題 (25分) 吉村 功 (石塚工学部) 討論 (45分, 資料: 新潟F予防資料の論文)

16:30 ~ 17:00 座長 柳澤 文徳 (東京医歯大難治研)

• 特別報告2 ... 中国におけるFと口腔歯—スライド供覧(30分) 柳澤 文徳 (東京医歯大難治研)

17:00 ~ 21:00 全国フッ素関係者共同研修交流会 会場 市民会館大ホール

■ 8月29日(日) _____ [会場=小千谷市民会館]

9:00 開会 (高橋 暁正)

9:00 ~ 10:25 座長 柳澤 文徳 (東京医歯大難治研)

三ニホジラム — フッ素と行政

• 特別報告3 ... 小・中学校におけるF応用の状況とF(25分)
三橋 敦子 (日教職員組合 養護教員部長)

• 発言1 千葉市F洗口10校の中止とF(15分)
西巻 義通 (千葉市議会議員)

• 発言2 強硬手段で開始された新潟県におけるF洗口とF(15分)
水野 美津 (県生活改善普及係)

• 発言3 新潟県におけるF洗口の現状(10分)
小黒 洋子 (新潟県教職員組合 養護教員部長)

討論 20分

- 10 = 25 ~ 10 = 50 座長 高橋 昭正 (前東京大医学部)
- ・特別報告 4 ... 地成菌科予防の実践と効果 —— その後 (25分)
中條 幸一 (宮城町立菌科診療所長)
- 10 = 50 ~ 11 = 30 座長 中條 幸一 (宮城県立菌科診療所)
- ・一般演題 5 ... 新潟県内のフッ化物応用における歯数の変化 (15分)
第1報 小学校においてフッ素応用した児童の高校における観察について
第2報 小学校においてフッ素応用した児童の平均歯数の変化
 - ・一般演題 6 ... フッ素洗口の学童の成長と歯質影響について (15分)
成田 寛一 (新潟県守門村菌科診療所)

討論 (10分)

- 11 = 30 ~ 12 = 15 座長 滝澤 行雄 (秋田大学医学部)
- ・特別講演 ... フッ素の毒性研究 — 本邦医学界機関と中心とを文獻的考察 (4分)
柳澤 文徳 (東京歯科大難治疾患研究所・教授)
- 12 = 15 ~ 13 = 00 座長 高橋 昭正 (前東京大医学部)
- ・招待講演 ... 広州市に於ける水道水フッ素化の経緯 (90分)
郭媛珠 (中山医学院口腔内科教授)

郭媛珠教授略歴 (Prof. Kwok Yuen-Chu, D.D.S)

1921年生まれ、四川省成都華西協成大学医学院牙科卒業、米國マコーク州工科大学で歯学博士の称号DDS
取得。1945年から中山医学院口腔内科に勤務、現在付教授。

- 15 = 00 ~ 15 = 25 座長 鈴木 紀雄 (滋賀大学教育学部)
- ・特別報告 5 ... フッ素化合物とF3環境汚染 (25分)
滝澤 行雄 (秋田大学医学部公衆衛生学・教授)
- 15 = 25 ~ 15 = 50 座長 滝澤 行雄 (秋田大医学部)
- ・特別報告 6 ... フッ化酢酸系農薬 (ニッソール) の危険性 (25分)
中岡 元 (大阪大理学部・講師)
- 15 = 50 ~ 16 = 30 座長 藤井 彰 (日村松歯学部)
- ・特別報告 7 ... フッ素含有副腎皮質ホルモン軟膏の問題点 (25分)
田代 実 (大阪大医学部皮膚科・講師)

討論 15分

16 = 30 閉会

17 = 30 ~ 19 = 30 郭媛珠先生観迎・懇親会

★ 社会薬学研究会の発足 理化学研究所主任研究員 辰野 高司

昭和20年代の後半から、薬学を学ぶ者たちの間で“日本の薬学を歴史的ヒト
う位置づけ、どう発展させるべきか”について模索する動きが出はじめ、曲
折の後「若い薬学者の会」が発足した。(昭和36年)。この会は日本薬学会
年次開催に合わせ、研究条件、将来展望、薬と科学者の責任、教育問題等
を主題とする討論会を企画し注目を集めてきた。

この間、公営・業害の発生に遭遇したか、既存の学会や薬学系大学の研究

体制がほとんど無効であり、その痛感でせられるのに及んで、薬学内部の自己運動として打って出さず、医学はもちろん、社会科学領域の協力を得て幅広い取り組みをしなければ、何かとつ問題解決はあほうかない。と反省するに至る。

本年5月15日、「社会薬学研究会」の設立総会シンポジウム(東京目黒)と以上の経過を背景に行われたもので社会薬の課題と方法と主題に、辰野高司(理化研究所、薬学の立場から)、佐藤修男(東京大学医学の立場から)岡田玲一郎(社会医療研究所、医療社会学の立場から)、下山瑛二(東京都立大学、法学の立場から)、儀我社一郎(大阪市立大学、経済学の立場から)の4氏による発言と討論が行われた。

いま頃に行った研究会発足は、一種奇異の感すら与えるかもしれない。しかし、社会薬学の研究・教育の本格的実施は欧米にあり、でも近年のことであり、やはり技術革新による生物化学の進展と社会変貌との不整合という、まさに現代的情勢のもとで本研究会が誕生したものと考えられる。

薬の安全性確保は研究の主要課題である。しかし薬害と説くあまり、薬物治療上の虚無主義に陥り傾向も生れている。これをどう克服するかについては検討せねばならないし、その広範な岐にわたる活動が期待されている。

総会において、本会が場的とする薬とは、主として医薬品であるか、広く生体作用物質をも含めることが確認され、事業としては研究会・総会の開催、機関紙・誌の発行が決められた。

また、上田亨(北海道大学)、高野啓天(立命館大学)ら15名が幹事として選ばれ、その中から、会長に辰野高司、事務局長に川瀬清(東京薬科大学)会計に片平利彦(東京医科大学)、監事に吉井千代田(日本薬史学会)が任命された。事務局は東京都小王子市湊元内1432-1東京薬科大学科学史研究室内に設けられた。

(「自然」(め)巻7号、1982) P.O.B.x, p.16-17 より転載)

【事務局より】

※予定より20日ほど遅れてしまいました。会報第7号をお届けします。7月17日の第11回定例会の前に何とか発行したかったのですが、あと一歩のところまで手回取り、早くも第8号と、12回定例会の準備に忙殺されそうです。

※定例会の場で多くの資料が提供され、参加者には配布されていますが、ここに、他の会員にも紹介致します。コピーサービスを受け付けておりますので、御希望の方は御一取下さい。(但し、当日、配布されていないものもあります)

- 栗谷典量：統臨床比較試験の問題点(11). 医薬ジャーナル, 12(2) pp. 37-43 (1976) (χ^2 検定法, フィッシャーの正確確率計算法など)
- 高橋暁正：ニッソール乳剤の毒性の本質とその撒布中の中毒死の可能性に関する解析. 薬のひろば, No. 56 (1981)
- 橋 広計：2×2 分割表に関する2, 3 の話題 — 臨床試験への応用を中心として. Clin. Eval., 10(1) pp. 3-17 (1982)
- M. MASUYAMA: Human Biochemical Individual Variabilities and Their Quasi-Constancy. Selected Papers of M. Masuyama, The Research Institute of Life Insurance Welfare
- 増山元三郎：生化学的にみた個体差. 自然 33(4) pp. 26-44 (1978)
- MYLES HOLLANDER & DOUGLAS A. WOLFE: Nonparametric Statistical Method, WILEY (1973) の CHAPTER I
- J. D. GIBBONS & J. W. PRATT: P-values Interpretation and Methodology. The American Statistician, 29(1) pp. 20-25 (1975)

※第12回定例会のお知らせ

日時：10月9日(土) pm 2:00 - 5:00

場所：私学会館(市ヶ谷)

テーマ：順位和検定の有意水準と検出力について(統)

異常値データの取扱いをめぐって

フィッシャーの正確確率の両側仮説検定の求め方について(統)

時系列データの扱い方について

などを予定しています。

参加料：1000円(会員外2000円)(予定)

詳しくは、追って御案内申し上げます。御参加をお待ちしております。(大野)

医薬安全性研究会 会報 No.7

昭和57年8月1日 発行

編集・発行 医薬安全性研究会

(事務局) (株)サイエンティスト社

〒101 千代田区神田駿河台3-2

山崎ビル TEL 03(253)8992

振替 東京 8-7/335

印刷 (有)尚文社

©1982