九州大学学術情報リポジトリ
Kyushu University Institutional Repository

使える！統計検定・機械学習 ： Ⅰ ： 2群間の有意差検定

高木, 英行
九州大学大学院芸術工学研究院

https://hdl.handle.net/2324/1467638

出版情報：システム/制御/情報 ： システム制御情報学会誌. 58 (8), pp.345-351, 2014-08
バージョン：
権利関係：

システム/制御/情報，Vol. 58, No. 8，pp. 345–351，2014

講　座

使える！ 統計検定・機械学習 — I — 2群間の有意差検定

高木　英行*

1. はじめに

新手法を提案し従来法よりも優れていることを示すためには，提案手法の実験データが従来法よりも「統計的に有意[1]によい」ことを示す必要がある．このことは誰もが知っているはずなのだが，現実には，統計的検定なしに性能平均値の大小のみから結論を導いている論文,「何はなくとも t-検定」「困ったときの分散分析頼み」的に適用しているのではないかと思われる発表が多々ある．それを理由に不採択になった論文も多いと思われる．

学部の確率統計や数理統計の授業で，仮説検定，t-検定，分散分析などを聞いたなあ，という読者は多いと思う．また，確かExcelにツールがあったから多分使えると思う，という読者も多いと思う．しかし，学部での授業は統計の数学的側面が主であり，統計検定の利用者がツールとしてすぐ使える実用的なノウハウ提供を目的にしているわけではなかろう．これが上述した現状の原因であるなら，統計検定の専門家ではないが計算知能分野での自身の研究に検定を避けて通れない筆者がユーザとしての利用ノウハウを書くだけでも読者のお役に立つかもしれないと思い，本講座解説をお引き受けした．

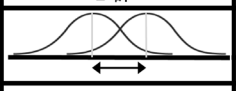
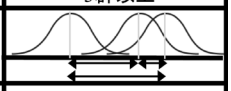
本連載では,「どのような場合に，どの検定手法を，どのように使えばよいのか」を解説する．第1図に本解説で扱う9種類の検定手法をまとめる．本連載の内容のスライドもダウンロード可能である [1].

自分の提案手法が有意に従来法よりも性能がよいことを実験的に示したいユーザは，この中からどの検定手法をどのように選ぶべきであろうか？　本連載の一番のハイライトともいえるこの解答は，第1表の3点をチェックすることである．この結果，第1図の 2^3 通りの場合分けができ，読者が選択すべき統計検定手法が確定する．

第1回目の本解説では第1図の左半分の2群の場合を扱う．二つの技術の性能値（パターン認識率，最適探索世代数，ニューラルネット学習の収束時間など）を比較し，新規提案技術が従来技術よりも有意に性能がよいかどうかを調べるような場合がその応用事例である[2]．第2回目の解説では第1図の右半分，すなわち，三つ以上の群の間に有意な差があるかどうか調べる検定手法を解説する．第3回目の解説では，主観評価実験の検定手法として，Sheffé の一対比較法を紹介する．

第1表　検定手法選択のための三つの判定条件

(1)	比較対象数が2群か3群以上か？
(2)	各群のデータが正規分布をしているか否か？
(3)	各群のデータに対応関係があるか否か？

		2群	3群以上
データ分布			
パラメトリック検定	対応関係なし	・対応関係がない場合の t–検定（3.1）	・一元配置の分散分析
	対応関係あり	・対応関係がある場合の t–検定（3.2）	・二元配置の分散分析
ノンパラメトリック検定	対応関係なし	・Mann-Whitney の U 検定（4.1）	・Kruskal-Wallis検定
	対応関係あり	・符号検定（4.2） ・Wilcoxon の符号検定（4.2）	・Friedman検定

第1図　本連載講座で扱う平均値間の差を検定する手法一覧

2. 検定手法選定のための3条件

連載第1回目の本解説では，二つの手法の間に有意な性能差があるかどうかを判定する場合，すなわち2群[3]の場合を扱うので，第1表で述べた第2，第3の判定条件で最適な統計検定手法を決定する．

第2の判定条件は，各群のデータが正規分布をしていると仮定できるかどうかである．第2図は二つの進化計算手法の探索性能を複数の初期条件で比較する例で，第 n 世代で2手法間に有意な性能差があるかどうかを示すには，第 n 世代での2手法の性能値データが各々正規分

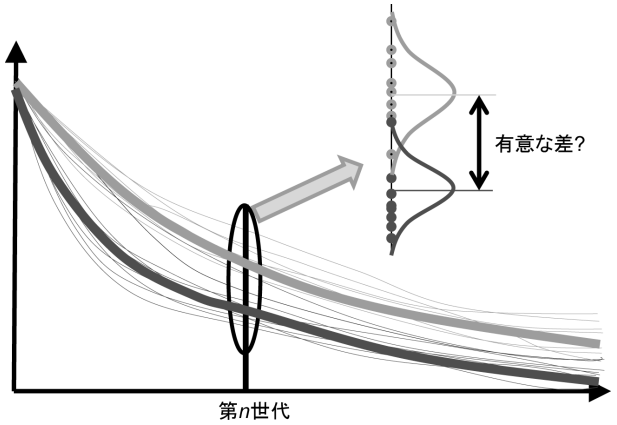
* 九州大学大学院 芸術工学研究院

Key Words: statistical tests, t-tests, U-test, sign test, Wilcoxon signed-rank test.

[1]「有意」という表現をした場合は，統計的に意味がある差であることを意味する．実験データのグラフを見て視覚的に差がありそうだというだけでこのことばを使ってはいけない．

[2]本解説では性能値の平均値の差の検定のみを扱う．品質管理のようにバラツキの大小が性能指標であるような応用事例には母分散の検定を用いる．

[3]群はグループとも標本ともいう．信号処理の標本化の感覚からすると標本＝個々のデータと勘違いしやすいので注意．標本数は群（またはグループ）の数であり，各標本に含まれるデータ個数は標本サイズという．



第 2 図　第 n 世代で性能差に有意な差があるかどうかを調べたい収束曲線の例

布をしているかどうかを検定（正規性の検定）することから始める．

正規性の検定手法には，Anderson-Darling 検定，D'Agostino-Pearson 検定，Kolmogorov-Smirnov 検定，Shapiro-Wilk 検定，Jarque-Bera 検定など，いろいろな手法がある．ネット上のフリーの Excel 用のアドインなどを探して利用するとよい[1]．

正規性があると判断できれば，**3.** の t-検定が使える．そうでなければ **4.** のノンパラメトリック検定である Mann-Whitney の U 検定か符号検定を選択する．t-検定や分散分析がパラメトリックな検定とよばれる理由は，データ分布が平均と分散のパラメータで規定される正規分布に従うことを前提に検定を行うからである．ノンパラメトリックな検定手法はデータ分布のモデルを利用できないので，データの数値を直接使う代わりにデータの順位（rank）情報を利用して検定を行うことが多い．データ分布モデルの情報が使える分，パラメトリックな検定手法の方が，有意差の検出力が高い．

第 3 の判定条件は，群間にデータの対応関係があるかどうかである．各被験者の運動前後の血圧データを例に挙げると,「運動前」の血圧データ群と「運動後」の血圧データ群は，同じ被験者同士の運動前後のデータなので対応関係がある．一方，日本人の身長データ群と米国人の身長データ群の場合は，個々のデータに対応関係がない場合である．第 3 図もデータの対応関係を視覚的に示した例である．

対応関係のない場合は，**3.1** の t-検定か **4.1** の Mann-Whitney の U 検定を選択する．対応関係がある場合は，第 2 の判定条件に基づいて，**3.2** の対応関係のある場合の t-検定（Excel の分析ツールでは,「t 検定：一対の標本による平均の検定」），あるいは，**4.2** と **4.3** の符号検定／Wilcoxon の符号検定を選択する．

データに対応関係がある場合の方が，情報量が多いため有意差検出力が高い．したがって 2 種類の手法の性能比較実験を行う場合には，実験結果のデータに対応関係をもたせるよう実験を計画すべきである．具体的には，パターン認識の場合であればテストパターンごとに 2 手法の性能を記録し，ニューラルネットや進化計算の場合であれば複数用意する学習や探索の初期条件の各々に 2 手法の性能を記録するとよい．

[1]たとえば，執筆時現在，http://www.vector.co.jp/にフリーの正規性検定ツールがある．

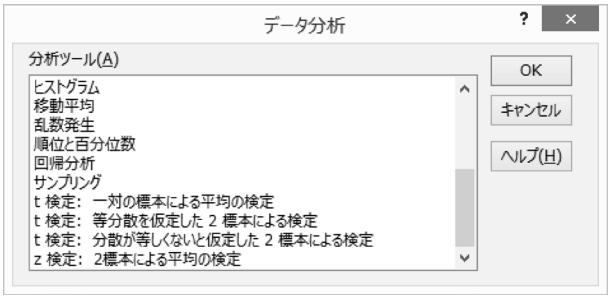
群間データに対応関係がない場合

A群	B群
4.23	2.51
3.21	3.3
3.63	3.75
4.42	3.22
4.08	3.99
3.98	3.65

群間データに対応関係がある場合

テスト番号	従来法	提案法
1	4.23	2.51
2	3.21	3.30
3	3.63	3.75
4	4.42	3.22
5	4.08	3.99
6	3.98	3.65

第 3 図　群間のデータに対応関係がない場合（左）とある場合（右）の例



第 4 図　Excel2013 で用意されているデータ分析ツール

3.　データに正規性がある場合

3.1　データに対応関係がない場合：t-検定

Excel メニューで「データ」→「データ分析」[2]を選択すると，3 種類の t-検定手法が見つかる（第 4 図）.「t 検定： 一対の標本による平均の検定」は **3.2** で用いる手法なので，データに対応関係がない場合を扱う本節では残りの 2 種類の t-検定を用いる（三つの t-検定をすべて適用し，どれでもいいから有意差を示したら「有意に提案手法の性能が従来法を上回った」としよう，などと考えないように）．

正規分布していると仮定できる 2 群のデータ分布の分散がほぼ等しい場合は「t 検定： 等分散を仮定した 2 標本による検定」を選び，そうでない場合は「t 検定： 分散が等しくないと仮定した 2 標本による検定」（Welch

[2]初めて利用する場合は，Excel の「ファイル」「オプション」「アドイン」から「分析ツール」を有効にすること．

の t-検定）を選択する[1]．等分散性の有無を判定するには，Excel の「データ分析」で用意されている F-検定を用いる．

第 2 表に，A と B の 2 グループの数値例と，これらに「t 検定：分散が等しくないと仮定した 2 標本による検定」を適用した場合の出力例を示す．この表で着目すべきは，P(T<=t) の値である．A と B の片方のデータ平均値が他方のデータ平均値より小さく（大きく）なることはない，という場合には片側検定の p 値を用い，A>B も A<B もありうる場合は両側検定の p 値を用いる．

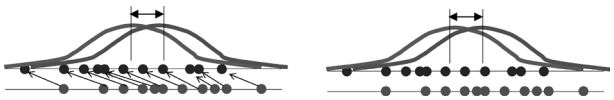
第 2 表　A 群と B 群のサンプルデータに Excel の「t-検定：分散が等しくないと仮定した 2 標本による検定」を適用した場合の出力結果

A	B
4.23	2.51
3.21	3.31
3.63	3.75
4.42	3.22
4.08	3.99
3.98	3.65
3.68	3.35
4.18	3.93
3.85	3.91
3.71	3.82
	4.01
	3.27
	2.93
	3.19

	変数 1	変数 2
平均	3.897	3.4885714
分散	0.1258233	0.2022901
観測数	10	14
仮説平均との差異	0	
自由度	22	
t	2.4841626	
P(T<=t) 片側	0.0105415	
t 境界値 片側	1.7171444	
P(T<=t) 両側	0.021083	
t 境界値 両側	2.0738731	

検定の慣習として，p 値が 0.05 以下の場合（危険率[2]5%）に有意差ありと判断する．20 回に 1 回位の間違いはあるかもしれないが，偶然とは思えないほど有意に偏っている，というわけだ．p 値が 0.01 以下の場合（危険率 1%）は，宝くじに当たることもあるわけだから偶然による偏りは否定しないがまず間違いなく意味がある偏りだ，と考えるのである．論文で有意差ありと記述する場合には，この危険率を添える．($p<0.05$) と書いてあれば危険率 5%で有意差ありと主張している．また表などでは* と ** を用いて，各々危険率 5%と 1%で有意差があるデータであることを示す場合も多い．

では，p 値が 0.05 を少し上回った場合はどのように判断すべきか？　危険率 5%というのはあくまで慣習的な目安であり 0.05 ではっきり有意差の有無[3]を断定するようなものではない．このような場合，筆者は，その p 値を示して「有意な傾向がある」と記述する場合が多々ある．



第 5 図　同じ 2 群のデータに，群間のデータに対応関係があるとして「t 検定：一対の標本による平均の検定」を適用した場合（左）と「t 検定：等分散を仮定した 2 標本による検定」を適用した場合（右）．前者は危険率 1%で有意になるが，後者では危険率 5%でも有意差は認められない．

3.2　データに対応関係がある場合：t-検定

2 群の各データに対応関係がある場合は，データの対応関係という新たな情報が加わるため，前節の t-検定に比べて有意差が検出しやすくなる．この様子を第 5 図に示す．同じデータで平均値の差が少ないように見えても，各データに図左側の矢印のような対応関係があるとすれば，少ない平均値の差であっても意味のある差である，と考えてもおかしくなかろう．

逆に本来データに対応関係がない第 5 図右側のデータに「t 検定：一対の標本による平均の検定」を適用すれば誤って有意差ありと判定してしまう．これが目をつぶって Excel の 3 種類の t-検定すべてを適用し，都合のよい結果だけをつまみ食いするような使い方をするな，という理由である．

4.　データに正規性がない場合

4.1　データに対応関係がない場合：Mann-Whitney の U 検定

データに正規性がない場合はノンパラメトリックな検定方法を選択する[4]．ノンパラメトリックな検定方法はデータ値の大小の順位関係を利用して有意に偏っているかどうかを判定する．認識率や成績の平均値が 100%や 100 点に近くデータ個数が少ない場合や，n 段階評価時の評価値が 1 点側や n 点側に偏っている場合，左右対称に拡がる正規分布にならないことが多々起きる．

データに対応関係がない場合は第 1 図から Mann-Whitney の U 検定[5]を選択する．データ個数が $n_1=4$ と $n_2=5$ からなる第 6 図の 2 群の平均値間に有意な差があるかどうかを Mann-Whitney の U 検定を用いて調べてみよう．まず片方のグループの各々のデータから見て，

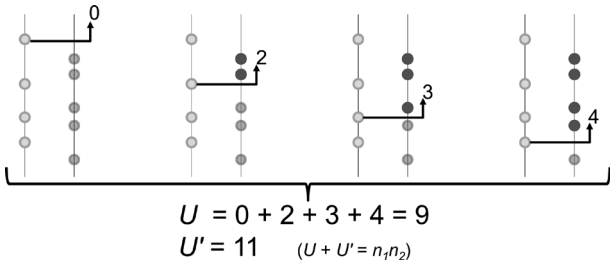
[1] Welch の t-検定は分散が等しい場合にも使えるので，等分散性の有無にかかわらず「t 検定： 分散が等しくないと仮定した 2 標本による検定」を選択してもよい，という考えもある．

[2] 危険率は有意水準ともいう．逆に，1− 危険率を信頼水準という．

[3] 帰無仮説は否定できるので「差がある」とはいえるが，帰無仮説を肯定できないので「差がない」という表現はおかしい．p 値が大きい場合は「差があるとはいえない」「差があるかどうかは判らない」などというのが正しい．

[4] ただ何でもかんでもノンパラメトリックな検定方法が適用できるわけではない．比較するグループのデータは，位置のみが異なる可能性のある二つの母集団からのサンプリングされたデータであることを前提にしているので基本的に等分散性が求められる．

[5] Wilcoxon の順位和検定，Wilcoxon-Mann-Whitney test，two sample Wilcoxon test などともよばれるので，混乱しないこと．



第 6 図 Mann-Whitney の U 検定の計算方法

その値を上回る相手側のデータ個数を数える．同じデータ値の場合は 0.5 個と数える．左グループの第 1 データを上回る右グループのデータ個数は 0 である．左グループの第 2 データを上回る右グループのデータ個数は 2 個である．以下同様に第 3，第 4 のデータを上回るグループのデータ個数は各々3 個と 4 個であるので，これらの合計値は $U=0+2+3+4=9$ である．右グループのデータをもとに同様に数えれば，$U'=11$ になる．これは $U+U'=n_1n_2$ の関係があるためで，どちらのグループを基準に数えても同じ U と U' が得られる．つぎに，U と U' の小さい方の値と U 検定表（付録第 A1 表）の数値とを比較し，検定表の数値よりも小さければ有意差ありと判定をする．

各グループのデータ個数が 20 を超える場合は，U が (1) 式の正規分布で近似できることを利用し，標準正規分布表で検定を行う．

$$N(\mu_U,\sigma_U^2)=N\left(\frac{n_1n_2}{2},\frac{n_1n_2(n_1+n_2+1)}{12}\right) \qquad (1)$$

すなわち，U が正規分布に従うのであれば，平均と標準偏差で正規化した $z=(U-\mu_U)/\sigma_U$ は標準正規分布 $N(0,1)$ に従うので，z を統計の教科書の付録等に掲載されている標準正規分布表を使って有意差検定をする．ここで，(1) 式より，$\mu_U=n_1n_2/2$，$\sigma_U=\sqrt{n_1n_2(n_1+n_2+1)/12}$ である．あるいは標準正規分布表を探す代わりに，上記 z 値を使って Excel で p 値を “=1−NORM.S.DIST(z)” [1] と直接求めた方が簡単であろう．

4.2 データに対応関係がある場合：符号検定

対応関係がある場合は，勝敗数が有意に片方に偏っているかどうかを調べる．符号検定の符号とは勝敗を + と − と考えた場合の符号のことで，たとえば A チームが B チームに対して 16 勝 4 敗であった場合，A チームの方が有意に強いといえるかどうかを検定する方法である．

二つの手法の性能値分布が正規性を示さない場合，同じ初期値を使った 2 手法の性能の勝敗を数える．その結果 n_1 対 n_2 だったとしよう．つぎに判定に使うデータ個数 N（$=n_1+n_2$）と（n_1 と n_2）の小さい方の値を使って符号検定表（付録第 A2 表）を調べる．n_1 と n_2 の小さい方の値が符号検定表の該当数値よりも小さければ有意差あり，と判断する．引き分けの場合のデータは優劣判定に使えないので無視する．パターン認識の場合，両手法とも認識に成功した，あるいは，両方とも失敗した場合は，カウントしない．

N が 90 を超える場合は，$(N-1)/2-k\sqrt{(N+1)}$ の小数点以下を切り捨てた整数を採用して比較する．ただし，危険率が 1%のときは $k=1.2879$，危険率 5%のときは $k=0.9800$ である [2,3]．

16 対 4 の場合で演習をしてみよう．$N=16+4$ の値を見ると，危険率 5%の表の値が 5 であるから，15 対 5 以上に差が開けば有意に差があるといえる．危険率 1%の表の値が 3 であるから，17 対 3 以上に差が開けば有意に差があるといえる．したがって，「16 対 4 は，危険率 5%で有意な差があるといえるが，危険率 1%では有意な差があるとはいえない」と判断する．

進化計算の探索性能やニューラルネットの学習性能では第 2 図のような探索曲線あるいは学習曲線を使って性能比較をすることが多い．2 手法の平均値を比較することはできるが，どの時点で比較するのか，平均値の差が有意か，という問題が常に付きまとい，恣意的に有意な差がある点でのみ結果を比較することすらある．客観的なデータを示すために筆者がお勧めする方法は，学習回数ごとや世代ごとに検定をして有意差の有無を示すことである．提案手法が従来手法を有意に上回る（あるいは下回る）場合，横軸の時間軸に並行して +（あるいは −）を添えることである．学習回数や進化世代によって有意な性能差があったりなかったりするであろう．それが，あるがままの姿である．

4.3 データに対応関係がある場合：Wilcoxon の符号検定

単に勝敗数を比較するだけでなく，どれだけ勝ったか負けたかの情報も利用できれば，さらに有意差の検出力を高くすることができる．第 2 図の場合でも横軸の各世代での 2 手法の差を使うことで Wilcoxon の符号検定（符号付き順位検定）が利用できる．

第 3 表の手法 A と B の性能値例を使って Wilcoxon の符号検定をしてみよう．計算手順はつぎの六つのステップである．

(step 1) A と B の差を求める．

(step 2) (step 1) の差の絶対値に対して順位を付ける[2]．

[1] Excel の「数式」メニューから「関数の挿入」を選択する．関数の挿入ウィンドウが開いたら「関数の分類」で統計を選択し，「関数名」で NORT.S.DIST を選択して「OK」をクリックする．すると標準正規分布の「関数の引数」ウィンドウが現れるので，「Z」テキストボックスに上記 z 値を，「関数形式」テキストボックスに TRUE を書いて「OK」をクリックすることで z に対する累積分布関数値が計算できる．1 からこの値を引いた値が p 値である．

[2] Excel にはデータ値から順位を求める RANK() 関数が用意されている．

(step 3) (step 2) の順位に (step 1) の符号を付ける．

(step 4) (step 3) の ± の少ない符号の順位のみを抜き出す．符号は削除．

(step 5) (step 4) の和を T とする．

(step 6) データ個数と T 値から Wilcoxon 符号検定表（**付録**第 A3 表）で検定をする．

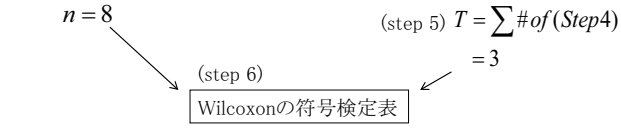
これらの計算ステップに基づいた第 3 表の場合の計算過程を第 7 図に示す．

第 3 表　Wilcoxon の符号検定のサンプルデータ

A	182	169	172	143	158	156	176	165
B	163	142	173	137	151	143	172	168

（計算例）

手法A	手法B	差分 d (step 1)	$\lvert d \rvert$の順位 (step 2)	順位に符号付与 (step 3)	少ない符号の順位抽出 (step 4)
182	163	19	7	7	
169	142	27	8	8	
172	173	−1	1	−1	1
143	137	6	4	4	
158	151	7	5	5	
156	143	13	6	6	
176	172	4	3	3	
165	168	−3	2	−2	2

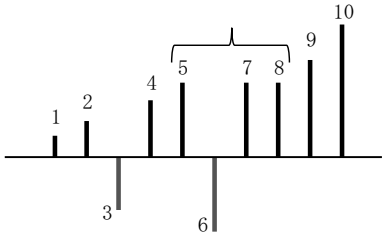


第 7 図　Wilcoxon の符号検定の計算方法

実際の計算ではつぎの 2 点がノウハウになる．

(1) 同じ数値の場合は優劣判断に使えないので，**4.2** と同様に無視し使わない．

(2) (step 2) で複数の同順位が出た場合は，平均順位を割り当てる（第 8 図）．



第 8 図　同じ順位が複数ある場合は，それらの平均値を割り振る．たとえば，(step 2) の絶対値の差が第 5 位〜第 8 位まで等しい場合，平均順位の第 6.5 位（$=(5+6+7+8)/4$）をこれら 4 データに割り振る．

データ個数が 25 を超える（$n>25$）場合は，T が (2) 式の正規分布で近似できることを利用し，標準正規分布表で検定を行うか，**4.1** で述べたように直接 Excel で p 値を計算する．

$$N(\mu_T,\sigma_T^2)=N\left(\frac{n(n+1)}{4},\frac{n(n+1)(2n+1)}{24}\right) \quad (2)$$

すなわち，T が正規分布に従うのであれば，平均と標準偏差で正規化した $z=(T-\mu_T)/\sigma_T$ は標準正規分布 $N(0,1)$ に従うので，z を統計の教科書の付録等に掲載されている標準正規分布表を使って有意差検定をするか，Excel で p 値 $=1-$NORM.S.DIST(z) を計算する．ここで，(2) 式より，$\mu_T=n(n+1)/4$，$\sigma_T=\sqrt{n(n+1)(2n+1)/24}$ である．

5.　おわりに

講座連載第 1 回目の本解説は，二つのグループのデータの平均値間に有意な差があるかどうかを検定する手法の選択方法と使い方について説明した．第 2 回目は，三つ以上のグループ間の検定手法の選択方法と使い方を，第 3 回目には，主観評価実験によく使われる検定手法について解説する予定である．

謝　辞

本解説は数理統計学がご専門の永田靖教授（早稲田大学創造理工学部）に監修を頂いた．御礼申し上げる．

(2014 年 3 月 31 日受付)

参 考 文 献

[1] http://www.design.kyushu-u.ac.jp/~takagi/TAKAGI/downloadablefileJ.html

[2] 奥津: 工場における推計学の問題とその解き方, 共立出版 (1951)

[3] 森口: 新編統計的方法: 品質管理講座, 日本規格協会 (1989)

[4] 市原: バイオサイエンスの統計学—正しく活用するための実践理論, 南江堂 (1990)

付　　録

第 A1 表　Mann-Whitney の U 検定表．n_1 と n_2 は両群の標本サイズ（[4] のデータをもとに本表を作成）

(a) 危険率 5%の両側検定，あるいは，危険率 2.5%の片側検定

n_1 \ n_2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	0	0	0	0	1	1	1	1	1	2	2	2	2
3	–	–	–	–	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	–	–	–	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	–	–	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	–	–	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	–	–	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	–	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	–	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	–	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	–	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	–	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	–	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	–	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	69	74	78	83
15	–	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	–	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	–	2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87	93	99	105
18	–	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	–	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	–	2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

(b) 危険率 1%の両側検定，あるいは，危険率 0.5%の片側検定

ackslashbox$n_1 n_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0	0
3	–	–	–	–	–	–	–	–	0	0	0	1	1	1	2	2	2	2	3	3
4	–	–	–	–	–	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	–	–	–	–	1	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	–	–	–	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	–	–	–	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	–	–	–	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	–	–	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	–	–	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	–	–	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	–	–	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	–	–	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	57	60
14	–	–	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	–	–	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	–	–	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	–	–	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	–	–	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	–	0	3	7	12	17	22	28	33	39	45	51	57	63	69	74	81	87	93	99
20	–	0	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

第 A2 表　符号検定表（両側検定時の危険率 1%および 5%，または，片側検定時の危険率 0.5%および 2.5%）

N	1%	5%	N	1%	5%	N	1%	5%	N	1%	5%	N	1%	5%
1			21	4	5	41	11	13	61	20	22	81	28	31
2			22	4	5	42	12	14	62	20	22	82	28	31
3			23	4	6	43	12	14	63	20	23	83	29	32
4			24	5	6	44	13	15	64	21	23	84	29	32
5			25	5	7	45	13	15	65	21	24	85	30	32
6		0	26	6	7	46	13	15	66	22	24	86	30	33
7		0	27	6	7	47	14	16	67	22	25	87	31	33
8	0	0	28	6	8	48	14	16	68	22	25	88	31	34
9	0	1	29	7	8	49	15	17	69	23	25	89	31	34
10	0	1	30	7	9	50	15	17	70	23	26	90	32	35
11	0	1	31	7	9	51	15	18	71	24	26			
12	1	2	32	8	9	52	16	18	72	24	27			
13	1	2	33	8	10	53	16	18	73	25	27			
14	1	2	34	9	10	54	17	19	74	25	28			
15	2	3	35	9	11	55	17	19	75	25	28			
16	2	3	36	9	11	56	17	20	76	26	28			
17	2	4	37	10	12	57	18	20	77	26	29			
18	3	4	38	10	12	58	18	21	78	27	29			
19	3	4	39	11	12	59	19	21	79	27	30			
20	3	5	40	11	13	60	19	21	80	28	30			

第 A3 表　Wilcoxon の符号検定表．(a) 両側検定時の危険率 5%，または，片側検定時の危険率 2.5%，(b) 両側検定時の危険率 1%，または，片側検定時の危険率 0.5%

n	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
(a)	0	2	3	5	8	10	13	17	21	25	29	34	40	46	52	58	65	73	81	89
(b)			0	1	3	5	7	9	12	15	19	23	27	32	37	42	48	54	61	68

著 者 略 歴

高木（たかぎ）　英行（ひでゆき）

1956 年 7 月生．1981 年九州芸術工科大学修士課程修了．1981〜1995 年松下電器産業（株），1991〜1993 年 UC Berkeley 客員研究員，1995 年九州芸術工科大学助教授，2003 年統合により九州大学助教授，現在九州大学教授．人間要素を取り込む計算知能等の研究に従事．博士（工学）．信学会篠原記念学術奨励賞 (1989)，知能情報ファジィ学会論文賞（2003），最優秀論文賞（KES'97, IIZUKA'98, ICOIN-15, ICGEC'12），功労賞（スロバキア人工知能学会 2002，IEEE SMC 学会 2003），IEEE SMC 学会 Best Associate Editor 賞（2005），2009 IEEE Most Active SMC Technical Committee 賞 (2010)，各受賞．日本ファジィ学会理事・監事（1999–2003），IEEE SMC 学会 Vice-President (2006–2009)，進化計算学会理事（2010–2012），IEEE SMC 学会日本支部長 (2014–2015).