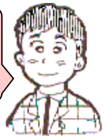
Q&A on 確率＆分布など

2021 年 5 月 15 日　加藤 厚

― 数式で諦めず筆算と電卓で辿れる所までは行ってみましょう！ ―

> God has afforded us only the twilight of probability.
>
> John Locke (1632-1704)

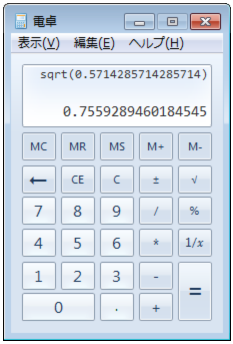


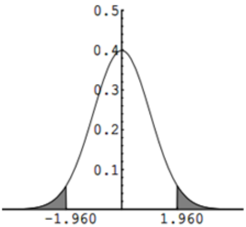
資料１　８試行の二項分布(成功確率 0.5)の各場合の数とそ

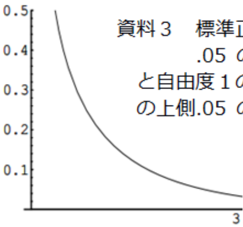
選択肢が２（成功確率 0.5）の試行でｎ回中ｒ回成功する場合
はnCr=n!/(n-r)!/r!なのでn=8の場合r=8からr=0に向けて

```
8C8=8!/0!/8!= 1 |*                                    .0039
8C7=8!/1!/7!= 8 |****                                 .0313
8C6=8!/2!/6!=28 |**************                       .1094
8C5=8!/3!/5!=56 |****************************         .2188
8C4=8!/4!/4!=70 |*********************************** .2734
8C3=8!/5!/3!=56 |****************************         .2188
8C2=8!/6!/2!=28 |**************                       .1094
8C1=8!/7!/1!= 8 |****                                 .0313
8C0=8!/8!/0!= 1 |*                                    .0039
計          256                                      1.0000
```

全ての場合の数 256（２の８乗）で割った確率は順に約.004,.031,.109,.219,.273,.219,.109,.031,.004。従ってそれ以上の偏りが偶然起こる確率は７：１や１：７で約 3.5%(.035=.031+.004)、８：０や０：８で約 0.4%です。







資料３　標準正
.05 の
と自由度１の
の上側.05 の

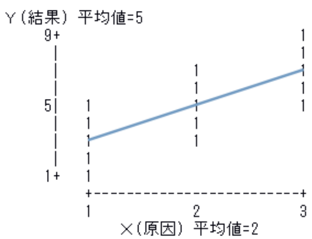
	○×	○×	○×	○×	○×	○×	○×
A	6 0	5 1	4 2	3 3	2 4	1 5	0 6
C	0 6	1 5	2 4	3 3	4 2	5 1	6 0

資料４　周辺度数固定の条件下で起こりうる全パターン（観察されたパターン：黄背景）

	○×	○×	○×	○×	○×	○×	○×
A	6 0	5 1	4 2	3 3	2 4	1 5	0 6
C	0 6	1 5	2 4	3 3	4 2	5 1	6 0
V	1.00	0.67	0.33	0.00	0.33	0.67	1.00

資料７　各パターンのクラメールのＶ(=効果量ｗ)

Y:1,2,3,4,5,3,4,5,6,7,5,6,7,8,9
X:1,1,1,1,1,2,2,2,2,2,3,3,3,3,3

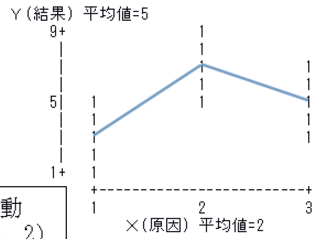


資料８　Ｘ と Ｙ の直線的散布図

資料２　指導前後の合否の集計

	指導後 ○	×	計
指導前○	6	1	7
指導前×	8	5	13
計	14	6	20

Y:1,2,3,4,5,5,6,7,8,9,3,4,5,6,7
X:1,1,1,1,1,2,2,2,2,2,3,3,3,3,3



Ｘ と Ｙ の非直線的散布図

Ｙの値	＝ 原点	＋Ｘの効果に対応する変動	＋誤差変動
3,4,5,6,7	＝(5,5,5,5,5)	＋(0, 0, 0, 0, 0)	＋(-2,-1, 0, 1, 2)
5,6,7,8,9	＝(5,5,5,5,5)	＋(2, 2, 2, 2, 2)	＋(-2,-1, 0, 1, 2)
1,2,3,4,5	＝(5,5,5,5,5)	＋(-2,-2,-2,-2,-2)	＋(-2,-1, 0, 1, 2)

目次

例えば心理学、教育学などの分野で判断や主張の数値的根拠が必要になると、記述統計に加えてカイ２乗検定、分散分析などの推測統計の処理がしばしば行われます。そして卒業研究などの場合には、手引書通りに計算したり、統計処理ソフトに計算させるのに手一杯で、確率や分布という「根拠の基盤」について思索し理解を深める時間的余裕はないのが通常でしょう。とはいえ、時間に余裕のある折々に統計学の入門書を開いても、難解な数式に目が眩んで「思索と理解の試み」を諦めてしまいがちです。

数学者でも統計学者でもない一般の統計ユーザーにとって、確率密度関数の数式などの理解はそもそも不必要です。他方「確率や分布についての基本的な理解に欠ける」のは情けなく残念なことです。そこで、基本的な用語・概念について、筆算と電卓（calc の「普通の電卓」＠表紙の右上参照）で容易に確認でき、「実感」を伴って納得できる範囲の「アリガチなＱ＆Ａ集」を作成してみました。確率と分布に加えて、相関分析と分散分析、効果量などについても少しだけ取り上げています。

不明・疑問な点、お気づきの点などありましたら atsushi.kato.1958@gmail.com までお知らせください。なお「分散分析」や「２×２分割表の分析」については、より詳しく検討している以下の資料が参考になるかもしれません：
http://mmua.html.xdomain.jp/kato/pdf/Q&A_on_ANOVA.pdf
http://mmua.html.xdomain.jp/kato/pdf/2x2_table.pdf

【用語・人名】

片側検定（one-tailed test）**と両側検定**（two-tailed test）：
01のミルクティの実験では成功＜失敗（資料１の下側）の場合の数は考慮しません。02の指導効果の例でも×→〇方向の変化にのみ注目して「実際の変化以上の場合の数／場合の数の総数」で確率を求めます（片側検定）。他方、カイ２乗分布の上側確率は「期待度数との全ての差（＝方向不問）」の確率です（両側検定）。従って、仮説に方向性がある二者の比較※なら確率は半分(10%→5%、2%→1%)とみなせます。※例：（Ｃ≠Ａではなく）Ｃ＜Ａを予想。

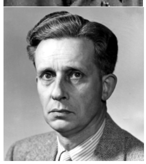
効果量（effect size）：
実質的有意性の側面から確率的有意性(実際の結果以上に極端な値が偶然に得られる確率の低さ）を補完する指標です。具体的にはカテゴリ間の連関、数量間の相関、要因の効果などの大きさを標準化した値で、07のクラメールのＶやコーエンのｗ、08～10のｒ２乗やη２乗などは0～1の範囲に収まります。

自由度（degrees of freedom）：
確率分布の形を決める変数で、カイ２乗・Ｆ・ｔなどの検定値の計算に用いる「互いに独立に正規分布するdataの個数」に対応するのだそうです。確かに周辺度数や平均値が固定されれば変動できるdataの個数は制約されますネ。

場合の数（the number of ways/events）：
表(Head)と裏(Tail)の出る確率が半々のコインを４回投げる時､出方の総数は（一投ごとに二者択一なので）2×2×2×2の16通りです。その内訳は表が4回(HHHH)と0回(TTTT)の場合が各１､1回と3回の場合が(１つが他と異なる)各４､そして2回ずつの場合が（HHTT,HTHT,HTTH,THHT,THTH,TTHH）の６です。これらの数はｎ個からｒ個を選ぶ「組合せ」と同じなので、数え上げなくても一般式のn!/(n-r)!/r!で計算できます※。
※４投中表が２回の組合せ=4!/(4-2)!/2!=4*3*2*1/(2*1)/(2*1)=6

クラメール（Harald Cramér 1893-1985）：
スウェーデンの数学者・アクチュアリー（保険数理士）・統計学者。数学者として確率論を数学的に定式化し、保険数理士の経験を踏まえ、統計学者として保険統計数学と数理統計学に貢献しました。

コーエン（Jacob Cohen 1923-1998）：
合衆国の心理学者・統計学者。確率的有意性の誤解及び過度の重視を批判し、検定力と効果量の重要性を指摘して近年のメタアナリシスの基礎を築きました。博士号は臨床心理学で取得したそうです。



マクニマー(Quinn Michael McNemar 1900-1986）：
合衆国の心理学者・統計学者。L.ターマンの下で博士号を取得し、スタンフォード・ビネー知能検査の改訂版を発表しました(1942)。McNemar検定(1947)の考案者でもあります。

01　「５％水準で有意」の意味を分かりやすく説明してください。

「統計好き」の間で有名なミルクティの実験の話があります。「味の違いで牛乳に紅茶を注いだかその逆かが分かる」※と数値的根拠に基づいて主張するためには、注ぐ順序を変えたミルクティをランダムに（≒順不同で）試す時、何回中何回言い当てればよいでしょうか？

※高温の紅茶に常温の牛乳を注ぐとタンパク質が変性して味が劣化するので「常温の牛乳に高温の紅茶を注ぐ」のが望ましいそうです。

注ぐ順序の選択肢は２つなので、当てずっぽうで（＝味の区別なしに）答えても１回目は1/2＝50％の確率で当たります。しかし、正答が偶然連続する確率は２回では1/4＝25％、３回では1/8＝12.5％という具合に毎回半分になり、４回で1/16≒6.3％、５回で1/32≒3.1％（←ここで５％未満）、６回で1/64≒1.6％、７回で1/128≒0.8％（←ここで１％未満）、８回で1/256≒0.4％になります。つまり「５％水準で有意」の意味は「偶然成功する確率が50％の試みが連続して５回成功する位まれなことが起こっている」ということなのです。

現実の世界では、成功失敗五分五分の試みを続けて３回も成功させれば「大したもの」でしょう。しかし、科学の世界ではより厳しく、五分五分の試みならその５回の連続成功、言い換えれば「偶然には32回に１回（≒３％）程度以下の確率でしか起こらないことが起こった」時にのみそれを「p<.05：５％水準で有意」※とみなすのです。

※偶然には128回に１回しか起こらない「７回連続成功」なら「１％水準」。

では、もし「８回中１回失敗・７回成功」だったらどう判断するのが正しいのでしょう？　８回全て成功する場合の数は１ですが、１回失敗する場合の数は８あります※。

※１回目を失敗、２回目を失敗（中略）７回目を失敗、８回目を失敗の８通り。

８回の試行について「全て成功」から「全て失敗」までの間の場合の数を考えると（各試行の結果は成功 or 失敗の二者択一なので）その総数は２の８乗の 256 で、各場合の数は試料１のように分布します。「１回失敗・７回成功」の場合の数だけなら８ですが、分布全体に占める「成功７回未満以外」の割合に基づけば文句の付けようがないので「全８回成功」の場合の１も加えた「７回以上成功」の９が全体に占める割合を求めると(8+1)/256≒0.03516 で約 3.5％。.035<.05 なので、この場合も「５％水準で有意」と判断してＯＫです。

この話を紹介している「実験計画法」には結果は明記されていませんが、著者のフィッシャーさんはミルクティを８杯用意し、挑戦者のご婦人はその全てを正しく言い当てたそうです。To err is human.（過つは人の常）という詩句もあります。５杯でも p<.05 になりうる実験に８杯を用意した理由は、挑戦者の「１回の失敗の可能性」への配慮だったのかもしれません。

資料１　８試行の二項分布(成功確率 0.5)の各場合の数とその全体像及び確率

選択肢が２（成功確率 0.5）の試行でｎ回中ｒ回成功する場合の数（＝組合せ）はnCr=n!/(n-r)!/r!なのでn=8 の場合 r=8 から r=0 に向けて以下の通り：

8C8=8!/0!/8!= 1	*	.0039
8C7=8!/1!/7!= 8	****	.0313
8C6=8!/2!/6!=28	**************	.1094
8C5=8!/3!/5!=56	****************************	.2188
8C4=8!/4!/4!=70	***********************************	.2734
8C3=8!/5!/3!=56	****************************	.2188
8C2=8!/6!/2!=28	**************	.1094
8C1=8!/7!/1!= 8	****	.0313
8C0=8!/8!/0!= 1	*	.0039
計　256		1.0000

全ての場合の数 256（2 の 8 乗）で割った確率は順に約.004,.031,.109,.219,.273,.219,.109,.031,.004。従ってそれ以上の偏りが偶然起こる確率は７：１や１：７で約 3.5％(.035=.031+.004)、８：０や０：８で約 0.4％です。

02　場合を数え上げる検定は実際に使われているのですか？

もちろん使われています。01 でも名前の出たフィッシャーさんの正確検定が有名ですが、まずはより単純な McNemar（マク ニマー）検定（対応のある２値のカテゴリの変化の検定）から説明しましょう。

資料２　指導前後の合否の集計

		指導後		計
指		○	×	
導	○	6	1	7
前	×	8	5	13
計		14	6	20

中学で学んだ「順列・組合せ」について、大学生 20 名を再度指導したとします。合格者（○）が事前試験では７名、事後試験では 14 名で、内訳が○→○=6、○→×=1（混乱した？）、×→○=8、×→×=5（資料２参照）だった時、指導の効果は有意でしょうか？

検討したいのは「変化」なので○→○と×→×の人数は無視して○→×と×→○の 1+8=9 名に注目します。全９名中８名以上が×→○に変化する場合の数（＝組合せ）は‥と考えると（ミルクティの実験と同様）９名中８名が×→○に変化する場合が $_9C_8=9$、９名全員が×→○に変化する場合が $_9C_9=1$ の計 10 です。９名について全員×→○（$_9C_9$）から全員○→×（$_9C_0$）までの間の場合の数の総数は２の９乗＝512 なので、８名以上が無→有に変化する確率は $(9+1)/512=0.01953125$（2%弱）、$p \fallingdotseq .02 < .05$ なので指導の効果は「５％水準で有意」です。

上の方法なら正確な確率が求められますが、変化した人数が多くて階乗の計算が大変になるとＱ＝「変化した人数の差の２乗／変化した人数の和」の値を自由度１のカイ２乗分布で検定する近似法がよく使われます。今回の例では $Q=(8-1)^2/(8+1)=49/9 \fallingdotseq 5.444$ です。

自由度１のカイ２乗分布の上側確率が.10、.05、.01 の値は順に 2.706、3.841、6.635 で、5.444 は 3.841 より大きく 6.635 より小さいので $.01 < p < .05$ となり、先ほどの「2%弱」と一致する判断が得られます※。
※この一致は「不正確」で $.005 < p < .025$ が「正確」なのですが、その理由は 05 で！

03　「カイ２乗分布」というのは何ですか？

この分布をカイ（ギリシャ文字の χ）２乗と命名したのはK. Pearsonさん。

カイ２乗分布は、01 の成功と失敗、02 の合格と不合格のようなカテゴリ※変数間の連関の近似的検定によく用いられる確率分布です。

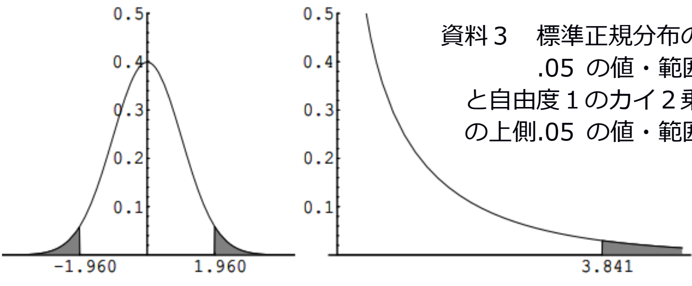
※試験得点は「数量」なので平均値が計算できます。他方「60 点以上を合格、未満は不合格」などとした合格 or 不合格は「カテゴリ」のため平均できません。

難しい定義式は使わずにそれが何かを検討するため、まず「２乗」を開いてみましょう。02 で触れたように自由度１のカイ２乗分布で上側確率が. 10、. 05、. 01 の値は順に 2. 706、3. 841、6. 635 で、これらの平方根は±1. 645、±1. 960、±2. 576 です。1. 96 という数に見覚えがありませんか？　それは標準正規分布の上側. 025＝両側. 05 の値です。同様に、1. 645 は上側. 05＝両側. 10 の、また 2. 576 は上側. 005＝両側. 01 の値です。つまり、自由度１のカイ２乗分布は平均値=0、標準偏差=1 の標準正規分布をする変数を２乗した値の分布なのです※。

※２乗するとマイナス側もプラス側になり、両側確率が上側確率になります。

自由度１のカイ２乗分布※が「標準正規分布する変数の２乗値の分布」と分かれば、Q＝「変化した人数の差の２乗／変化した人数の和」の検定値の分布がカイ２乗で近似できる理由が（ある程度）納得できることでしょう。なぜなら、まず指導が全く無効なら×→○と○→×の変化は偶然の結果なので左右均等に正規分布し、差≒0 になります。指導が無効でない（含有害）なら、その程度に応じて変化の数の差の２乗はより大きい正の値をとります。その値を「変化した人数の和」で割ることで「１件当り」に標準化した値がQだからです。

※標準正規分布変数の２乗値をｎ個合計した値の分布が自由度ｎのカイ２乗分布。



資料３　標準正規分布の両側 .05 の値・範囲（左）と自由度１のカイ２乗分布の上側.05 の値・範囲（右）

04　「フィッシャーの正確検定」はどのように行うのですか？

McNemar 検定が「対応のある２値のカテゴリの変化の検定」なのに対し、フィッシャーの正確検定は「対応のないカテゴリの件数分布の差の検定」です。「観察されたパターン及びより偏ったパターンの場合の数÷全ての場合の数」で確率を求める点は同じですが、前者が変化の件数のみに注目するのに対し、後者は全ての件数を考慮します。

ある事柄を従来の方法Ｃと新しく考案した方法Ａで６名ずつの大学生に教えたところ「納得できた＝○：納得できなかった＝×」の人数がＣでは１:５、Ａでは５:１だったとします。２つの方法の差※は有意でしょうか？※ＣかＡかと○か×かの連関(association)でもあります。

観察された「行と列の計は全て６」の条件下で可能な全パターンは資料４に示した７つです。

	○×	○×	○×	○×	○×	○×	○×
Ａ	6 0	5 1	4 2	3 3	2 4	1 5	0 6
Ｃ	0 6	1 5	2 4	3 3	4 2	5 1	6 0

資料４　周辺度数固定の条件下で起こりうる全パターン（観察されたパターン：黄背景）

Ａ・Ｃ各群内の○と×の場合の数（＝組合せ）は 6C0=6C6=1、6C1=6C5=6、6C2=6C4=15、6C3=20 で、かつＡ群とＣ群の場合の数は独立（＝無関係）なので、各パターンの場合の数は両者の積（順に 1×1=1、6×6=36、15×15=225、20×20=400、15×15=225、6×6=36、1×1=1）。また、全ての場合の数はその総和の 924。従って、各パターンの確率は順に 1/924、36/924、225/924、400/924‥※で、観察されたパターン（黄色）とより偏ったパターン（その左側）が全体に占める割合は(1＋36)/924≒0.04。.04<.05 なので「方法間の効果の差は５％水準で有意（Ｃ＜Ａ）」と言えます。

※普通は４値 a, b, c, d の(a+b)!*(c+d)!*(a+c)!*(b+d)!を n!*a!*b!*c!*d!で割って各パターンの確率を求めますが、この計算の意味はボクには理解困難です。

フィッシャーの正確検定がどんなものか、そして件数が多いとその計算がどれほど面倒になりそうか、理解・想像できたでしょうか？

05　カイ２乗による正確検定の近似について説明してください。

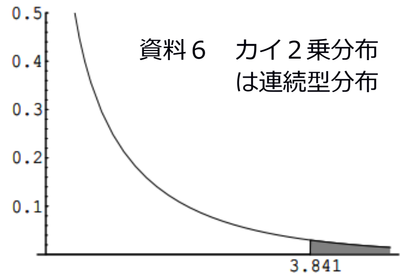
フィッシャーの正確検定もMcNemar検定と同様カイ２乗分布で近似できます。まず、方法ＣとＡの効果に差がない＝観察された差は全て偶然の結果と仮定します。その場合、方法は回答に影響しないので、回答が〇：×＝6：6という条件の下ではＣ・Ａともに3：3が期待されます（＝期待度数）。しかし、実際の度数（＝観察度数）はＣでは1：5、Ａでは5：1とかなりの差が存在するので、観察度数と期待度数の差の２乗（←方法の効果の２乗値）÷期待度数（←１件当りに標準化）の値の総和を求めて「観察された差が偶然の結果とは考えにくい程度」の指標とします。この指標＝検定値の名称が「カイ２乗」なのです※。
※そして、帰無仮説の元でのカイ２乗値の分布がカイ２乗分布。

今回の例では、カイ２乗＝(5-3)^2/3+(1-3)^2/3+(1-3)^2/3+(5-3)^2/3=4/3+4/3+4/3+4/3=16/3≒5.333。自由度１のカイ２乗分布の上側確率.05と.01の値は3.841と6.635で、5.333は3.841より大きく6.635より小さいので.01<p<.05となり、04の「約４％」と一致します。

しかし（注意深い人は気づいたことでしょうが）上の「一致」は実は不正確です。なぜなら（03で触れたように）カイ２乗分布の上側確率には（２乗した結果）標準正規分布の±両側（＝資料３の灰色領域）の確率が含まれており、片側の確率≒.04ならカイ２乗検定の結果は≒.08になるハズだからです※。（↓「２者の比較」の場合に限ります）
※予想が効果：Ｃ≠Ａなら両側、Ｃ＜ＡやＣ＞Ａなら片側の確率を使うのが適切。

今回の例は近似法で正確な確率を得るには件数が少なすぎます※。より正確に近似する確実な方法は「被験者の数を増やすこと」ですが、現在のdataに「連続性の修正」を行うのも一案です。
※適切な近似には「最小期待度数が5より大きい」とか「期待度数5未満のセルが全体の20%未満で期待度数1未満のセルが無い」とかの（専門家によって意見？の異なる）条件が必要なようです（今回の例の期待度数は全て3なので、どちらの条件も満たしていません）。

06　「連続性の修正」とは何ですか？



資料６　カイ２乗分布は連続型分布

04 の資料４に示したように、正確検定の確率は「パターンごとの確率の和」なので区間を埋めてはおらず、左右両端の外側にも分布しません。それを０～∞まで連続分布するカイ２乗分布で近似するムリを修正する手続きが「連続性の修正」（イェーツの補正）で、具体的には観察度数と期待度数の差を 0.5 ずつ０に近づけてから２乗します※。

※元々の差が 0.5 より小さい場合は０にします。

修正カイ２乗＝(2−0.5)^2/3+(−2+0.5)^2/3+(−2+0.5)^2/3+(2−0.5)^2/3=1.5^2/3+(−1.5)^2/3+(−1.5)^2/3+1.5^2/3=(2.25/3)*4=9/3=3.0。自由度１のカイ２乗分布の上側確率.10 と.05 の値は 2.706 と 3.841 で、3.0 は 2.706 より大きく 3.841 より小さいので.05<p<.10 となり、05 の「カイ２乗検定の結果は≒.08 になるハズ」と一致します※。

※効果：Ｃ＜Ａを予想しているので片側確率を使い p≒.04<.05。

今回の例では、修正カイ２乗を用いた判断は正確検定と一致しました。しかし、この修正については専門家の間でも議論があるのでこれ以上の深入りは避けます※。ともあれ「修正によりカイ２乗値は小さくなり確率は大きくなる」のは確実なので、修正カイ２乗値でも有意なら安心してカイ２乗値を示して有意と主張できます。

※「修正は２×２分割表でのみ行う」としている書籍がありますが、離散型確率分布の連続型確率分布による近似は“分割表一般に共通するムリ”なのでは？

被験者の数を増やす、連続性の修正を行う以外の方策としては（分析の意味を保てる範囲内での）セルの併合も一案です。また、以下の web 頁やＲでは２×２より大きい分割表の正確検定が可能です。

http://aoki2.si.gunma-u.ac.jp/exact/exact.html

07　「有意味」の指標は確率だけなのですか？

「マレなことが起きている＝確率的有意性」は確かに注目すべき特徴の１つです。例えば気象庁も「過去 30 年の気候に対して著しい偏りを示した天候」として異常気象を定義しています。他方、その出来事の現実的な大きさも注目すべき特徴の１つでしょう。例えば、過去 30 年間地震が起きなかった地域での地震発生は「マレ」ですが、その震度が２程度なら住民にとって「現実的な意味」は殆どありません。

統計の世界で確率的有意性を補完するのが効果量（effect size）で、カテゴリ間の連関ならクラメールのＶやコーエンのｗが使われます。ｒ行×ｃ列のクロス表のＶは次行の式で求められる値です：

Ｖ＝√（カイ２乗÷（件数×（ｒとｃの小さい方の値-1）））

例えば、04 の例の１：５と５：１ならカイ２乗=16/3、件数=12、ｒ=ｃ=2 なので、Ｖ＝√（16/3÷（12×1））=√（16/36）≒0.67 になります。ｗはＶの式の（件数×‥）を（件数）にした値で、ｒかｃが２以下ならＶと同じ値、それ以外ならＶより小さい値になります。

04 の資料４の各パターンのＶ（＝ｗ）の値は資料７に示した通りで、Ｖもｗも.10 程度なら小、.30 程度なら中、.50 程度なら大と判断します。従って、左右両端の２パターンなら「効果量は大」、その１つ内側なら「効果量は中」です（V=0 なら「効果量は 0」）。

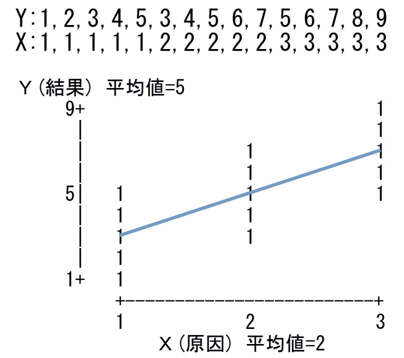
	○×	○×	○×	○×	○×	○×	○×
A	6 0	5 1	4 2	3 3	2 4	1 5	0 6
C	0 6	1 5	2 4	3 3	4 2	5 1	6 0
V	1.00	0.67	0.33	0.00	0.33	0.67	1.00

資料７　各パターンのクラメールのＶ(=効果量ｗ)

相関分析ならｒ（積率相関係数）やｒ２乗（決定係数）、分散分析ならη（イータ）（相関比）やη２乗が代表的な効果量です。小中大の目安はｒならＶやｗと同じ、ｒ２乗とη２乗なら.01、.06、.14 程度です。

08　相関分析について分かりやすく説明してください。

正確検定やカイ２乗検定は、平均値を求められないカテゴリ（例：○か×か）間の連関＝件数の分布の差をクロス表で集計した上で行う確率的有意性の検定でした。それに対し、相関分析は平均値を求められる数量（例：得点）の対の分布状況を散布図で検討した後に行う直線的関係の指標（ｒ）の計算及びその確率的有意性の検定です。



資料８　X と Y の直線的散布図

変数Ｙと変数Ｘの 15 対が順に資料８の上部の値の時、その散布図は右のようになります。Ｘの増加に対応してＹも直線的に増加しており、正の相関がありそうです。このような関係をｒ(積率相関係数)という数値で表すのに、普通はｒ=Σ(Zx*Zy)/n（∵Z は平均値を 0、標準偏差を 1 にした標準得点）というキレイな定義式を使いますが、ここではＹの散らばりを直線に乗る「当り」と乗らない「外れ」に分解してみます（補足資料８参照）。

Ｙの平均値は 5 なので散らばりは-4,−3,−2,−1,0,−2,−1,0,1,2,0,1,2,3,4。これらを（そのまま足すと 0 になってしまうので）２乗して総和すると 16+9+4+1+0+4+1+0+1+4+0+1+4+9+16=70（＝全偏差平方和）。他方、外れはＸの各水準毎に−2,−1,0,1,2 なのでその２乗の総和は 10×3 で 30（＝誤差の偏差平方和）。当りは差を取って全の 70−外れの 30=40（直接計算するなら−2^2×5+0^2×5+2^2×5=20+0+20=40）。

当り／全の値、つまりＹの全変動中のＸの値に対応する部分の割合が決定係数（ｒ２乗）で、その平方根が相関係数（ｒ）です。今回の

例ではｒ２乗＝40/70≒.571、ｒ＝√(40/70)≒.756 になります※。
※平方根の±は直線が右上がりなら＋、右下がりならーを選びます。

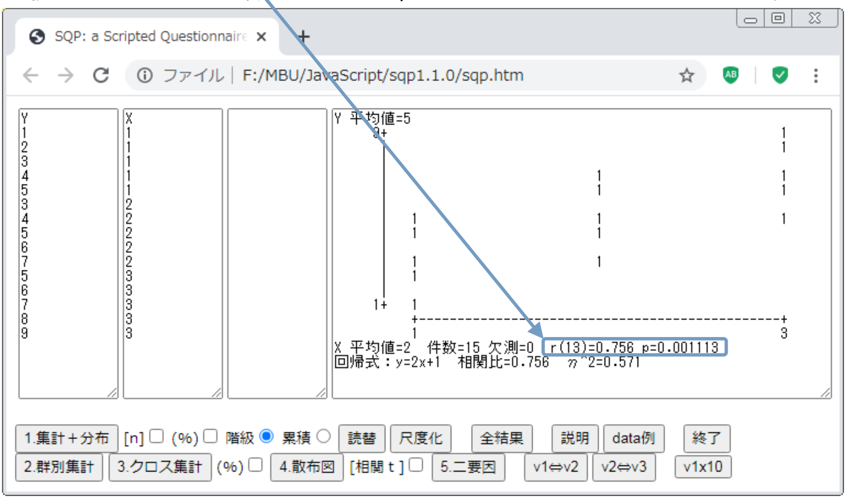
相関係数の確率の数表を見ると、自由度（＝対の数-2※）13 の行で確率.10、.05、.01 のｒの値は.441、.514、.641＃。今回のｒは.756 で.641 より大きいため p<.01 と判断できます。
※-2 の理由は「Z の計算に X と Y の平均値を使うから」と思っておきましょう。
＃両側検定の場合。仮説（例：r>0）で逆側（r<0）が無視できれば確率は半分。

補足資料８　一般的な相関係数の計算方法

1. 定義式どおりに計算（‥小数第２位に丸めて計算した場合）
Y の平均値は 5、標準偏差(SD)は 2.16 なので
Zy:-1.85,-1.39,-.93,-.46,0,-.93,-.46,0,.46,.93,0,.46,.93,1.39,1.85
X の平均値は 2、標準偏差(SD)は 0.82 なので
Zx:-1.22,-1.22,-1.22,-1.22,-1.22,0,0,0,0,0,1.22,1.22,1.22,1.22,1.22
Zx*Zy:2.26,1.7,1.13,.56,0,0,0,0,0,0,0,.56,1.13,1.7,2.26
Σ(Zx*Zy)=11.3　ｒ=11.3/15≒.753　←0.756 とホボ同値。

2. （誤差のより少ない）計算式で計算
分子が n*Σ(X*Y)－ΣX*ΣY‥15*170-75*30=2550-2250=300
分母が n*ΣX^2-(ΣX)^2 と n*ΣY^2-(ΣY)^2 の積の平方根
√(15*70-30^2)*(15*445-75^2))=√((1050-900)*(6675-5625))
=√(150*1050)=√157500≒396.863　ｒ=300/396.863≒.756

3. 統計処理ソフトで計算（SQP: Scripted Questionnaire Processor の場合）

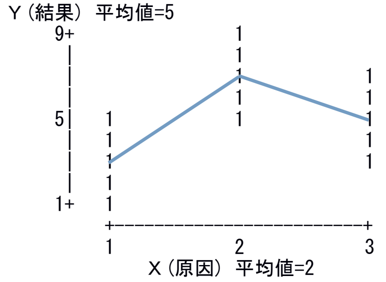


09　「関係が非直線的な場合」は何を指標にするのですか？

08 で述べたように、相関分析は「直線的関係の指標（ｒ）の計算及びその確率的有意性の検定」です。従って、散布図で検討した数量変数間の関係が直線的でない時には相関分析は適切ではありません。

資料８のＹの値をＸの２と３で入れ替えると、散布図は右のようになります。Ｙの平均値はＸが２の時７で最も高く、３では５に低下しており、関係が非直線的なためｒは.38 に留まります※。

Y:1,2,3,4,5,5,6,7,8,9,3,4,5,6,7
X:1,1,1,1,1,2,2,2,2,2,3,3,3,3,3



資料９　X と Y の非直線的散布図

※平均値も標準偏差も同じなのでＺの値は同一で順番が変わるだけです：

Zy:-1.85,-1.39,-.93,-.46,0,0,.46,.93,1.39,1.85,-.93,-.46,0,.46,.93
Zx:-1.22,-1.22,-1.22,-1.22,-1.22,0,0,0,0,0,1.22,1.22,1.22,1.22,1.22
Zx*Zy:2.26,1.7,1.13,.56,0,0,0,0,0,0,-1.13,-.56,0,.56,1.13
Σ(Zx*Zy)=5.65　　r =5.65/15≒.377

散布図で直線的でない関係が示されたら、横軸の変数Ｘを 3～7 程度に併合※して群＝カテゴリとみなし、Ｙの全偏差平方和をＸの群間の偏差平方和と群内＝誤差のそれに分解して「群間偏差平方和／全偏差平方和」の値（＝効果量）とその正の平方根（＝相関比）を求めます。
※今回の例では、X の値は最初から 1～3 なのでそのままで OK。

この分解は 08 で行った「当り」と「外れ」の分解と同じなので、全 70＝群間 40＋群内 30※で群間／全＝40/70≒.571、その正の平方根の相関比（η）≒.756（08 のｒと同値）が非直線的関係の指標です。
※Ｘの値をカテゴリとみなせば値の順番自体が無意味となるため、直線も非直線もありません。実は、資料８の X と Y は「各群の平均値が直線上に並ぶように＝要因Ｘの効果が当りと一致するように意図的に作成した値」だったのです。

10　相関比などの評価と検定はどのようにするのですか？

07 でも触れたように、相関比（η）とη２乗＝効果量の小中大の目安は.10、.30、.50 と.01、.06、.14（←２乗にしては不一貫ですが‥）程度なので、資料９の例は共に大です。そして、全偏差平方和を要因の水準に対応するそれ（群間）と誤差のそれ（群内）に分解して要因の確率的有意性を評価＝検定する方法が（有名？な）分散分析です。

資料９の例のＹの平均値は全体では 5、群 1～3 では 3,7,5 なのでＸの水準に対応する群間の変動は-2,2,0、対応しない群内の変動（＝誤差）はそれぞれ-2,-1,0,1,2 です。従って、Ｙの全ての変動は 5 を原点として「群間の-2,2,0」と「群内の-2,-1,0,1,2」に分解できます：

Ｙの値	＝	原点	＋Ｘの効果に対応する変動	＋誤差変動
1,2,3,4,5	＝	(5,5,5,5,5)	＋(−2,−2,−2,−2,−2)	＋(−2,−1,　0,　1,　2)
5,6,7,8,9	＝	(5,5,5,5,5)	＋(　2,　2,　2,　2,　2)	＋(−2,−1,　0,　1,　2)
3,4,5,6,7	＝	(5,5,5,5,5)	＋(　0,　0,　0,　0,　0)	＋(−2,−1,　0,　1,　2)

要因に対応する＝群間の偏差平方和は(−2)^2*5+2^2*5+0^2*5=40
〃しない＝誤差＝群内の〃は((−2)^2+(−1)^2+0^2+1^2+2^2)*3=30

群間の自由度は群数−1=3−1=2、群内の自由度はΣ(各群の人数−1)=4+4+4=12。「母分散の推定値＝偏差平方和／自由度」の公式に当てはめると、推定値は群間からが 40/2=20、群内からが 30/12=2.5 なので F=20/2.5=8（つまり要因の効果は誤差≒偶然の８倍）。Ｆ分布の数表の自由度 2/12 の値は上側確率.05 と.01 で 3.89 と 6.93。今回のＦは 8 で 6.93 より大きいため p<.01 と判断できます。

今回の例を SQP の[2.群別集計]で処理した結果の出力は資料 10 に示した通りです。

X

値	有効件数	Y 平均値	SD	u
3	5	5	1.414	1.581
2	5	7	1.414	1.581
1	5	3	1.414	1.581
全体	15	5	2.16	2.236

欠測=0　相関比=0.756　効果量(η^2)=0.571
F(2,12)=8 p=0.006196

資料 10　SQP の[2.群別集計]の出力

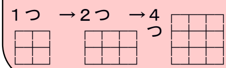
11　様々な「確率分布」にはどんな関係があるのですか？

　01～04 で紹介したミルクティの実験の判断、近似しない McNemar 検定、フィッシャーの正確検定は「二者択一のできごとの起こりかたの分布」である二項分布に基づきます。パターン毎に「場合の数」があり、全ての場合の数に占める割合がそのパターンの確率です※。
※「そのパターン及びより偏った全パターン」の確率の総和が求める確率。

　この二項分布の試行数（ｎ）を増やし、パターンが連続して間隔がなくなった状態が正規分布です。（値－平均値で）平均値を 0、（値÷標準偏差で）標準偏差を１にした（＝標準化した）正規分布を標準正規分布といい、±1.960 の外側の面積が各.025、合わせて.05 です。

　McNemar 検定やフィッシャーの正確検定で、件数が多い場合の近似に用いられるカイ２乗分布は、標準正規分布する変数の２乗値を自由度の回数足し上げた値の分布です※。２乗の結果 0～∞に分布し、平均値は自由度に一致します。
※自由度１の場合は「標準正規分布する変数の２乗値の分布」です。

大きなクロス表は、2x2 のそれが半分ズレながら重なっていると見ると納得しやすいかもしれません。



　McNemar 検定の変化やフィッシャーの正確検定の度数は、偶然の下では期待値の周囲に正規的に分布します※。Ｑは「期待値(=0)からのハズレの２乗値を件数の和で割って標準化した値」、カイ２乗は「期待値との差の２乗値を期待値で割って標準化した値の総和」ですから、件数が多ければ対応する自由度のカイ２乗分布による検定が可能です。
※偶然なら、期待値に近い値はより多く、遠い値はより少なくなります。

　10 に出てきたＦ分布は「２つの独立なカイ２乗変数の比」の分布なのだそうです。分散分析では「効果による変動から推定した母分散が誤差による変動から推定した母分散の何倍か」（＝Ｆ値）が判断の指標で、その母分散の推定は「偏差平方和／自由度」でなされるため、

資料 11 中のｎは二項分布では試行数、ｔ・カイ２乗・Ｆ分布では自由度。→

２つの推定値は共に「期待値からの正規分布する差の２乗を自由度＝平均値で割って標準化した値」＝カイ２乗値と同等です。であるなら要因の効果／偶然の変動＝Ｆは２つのカイ２乗変数の比なので、要因の自由度（＝群数−1）と偶然の自由度（＝(各群内の人数−1)の総和）で形が定まるＦ分布の上側確率で検定するのは「妥当な手続き」です。

08 の相関係数の検定はｔ分布を使って行われます（対の数がｎの時 t=(r*√￣(n−2)/√￣(1−rˆ2))）。ｔ分布は大きさｎ※の標本（‥母集団の分布は中心極限定理により不問）の平均値の標準化分布で自由度はn−1※、標準正規分布より裾が重くｎが大きいほど裾の軽い標準正規分布に近づきます。※資料 11 中のｔ分布のｎは自由度なので標本の大きさはn+1！

このような事柄について深く知る必要は、一般の統計ユーザーにはありません。とはいえ、訳が分からないまま機械的に計算する／ソフトに計算させるというのも「情けなくつまらない話」なので、筆算と電卓で辿れる範囲内でボクが「なるほどネ！」と感じた事柄のいくつかをまとめてみたのがこの小冊子です。皆さんが確率と分布などについて「若干のスッキリ感」を味わう一助となれば幸いです。