# ノンパラメトリック検定

# 質的変数と質的変数の関連性を調べる－クロス表

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | 100人 | 900人 | 1000人 |
| | なし | 10人 | 990人 | 1000人 |
| 合計 | | 110人 | 1890人 | 2000人 |

行周辺度数

列周辺度数

# クロス表－行パーセント

- 各行のセルの度数を行周辺度数で割って100をかけたもの

行周辺度数

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | 10%(100人)<br>100×1/1000 | 90%(900人)<br>100×900/1000 | 1000人 |
| | なし | 1%(10人)<br>100×10/1000 | 99%(990人)<br>100×990/1000 | 1000人 |
| 合計 | | 110人 | 1890人 | 2000人 |

# クロス表－列パーセント

- 各列のセルの度数を列周辺度数で割って100をかけたもの

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | 91%(100人)<br>100×100/110 | 48%(900人)<br>100×900/1890 | 1000人 |
| | なし | 9%(10人)<br>100×10/110 | 52%(990人)<br>100×990/1890 | 1000人 |
| 合計 | | 110人 | 1890人 | 2000人 |

列周辺度数

# (復習)リスクとオッズ

- リスク
  - イベント(通常ネガティブなもの：疾患の発症、死亡など)が起こる確率(0から1の間の値をとる)
  - $\text{リスク} = \dfrac{\text{イベントが発生した人の数}}{\text{対象者の人数}}$
- オッズ
  - あるイベントが起こる確率と起こらない確率の比(0から∞の間の値)
  - $\text{オッズ} = \dfrac{\text{イベントが発生した人の数}}{\text{イベントが発生しなかった人の数}}$
- オッズとリスクの関係
  - 稀なイベントの時はリスク≒オッズ
  - リスクとオッズの関係式

$$\text{オッズ} = \frac{\text{リスク}}{1 - \text{リスク}}$$

# 関連性の指標(1)ーリスク比とオッズ比

- リスク比(相対危険度)

$$-\ \text{リスク比} = \frac{\text{要因がある群(曝露群)のリスク}}{\text{要因がない群(非曝露群)のリスク}} = \frac{\frac{A}{A+B}}{\frac{C}{C+D}}$$

$$= \frac{(C+D)\times A}{(A+B)\times C}$$

- オッズ比

$$-\ \text{オッズ比} = \frac{\text{要因がある群(曝露群)のオッズ}}{\text{要因がない群(非曝露群)のオッズ}} = \frac{\frac{A}{B}}{\frac{C}{D}}$$

$$= \frac{A\times D}{B\times C}$$

| | | イベント | |
|---|---|---|---|
| | | 発生 | 発生しない |
| 要因 | あり | A | B |
| | なし | C | D |

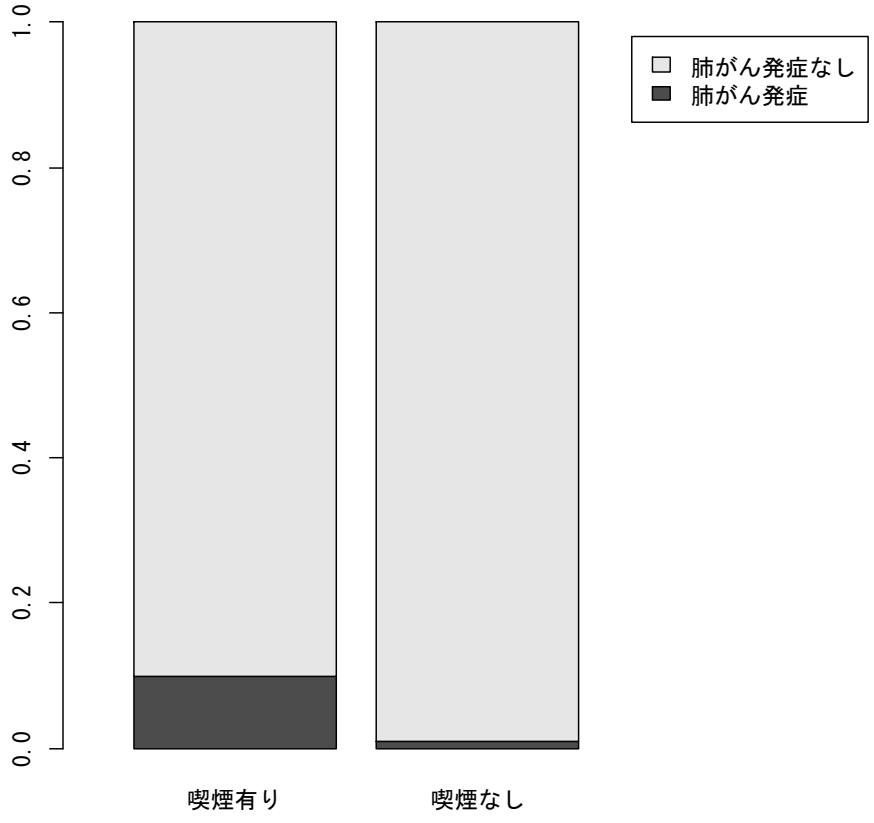
# 関連性の指標(2)ーファイ係数

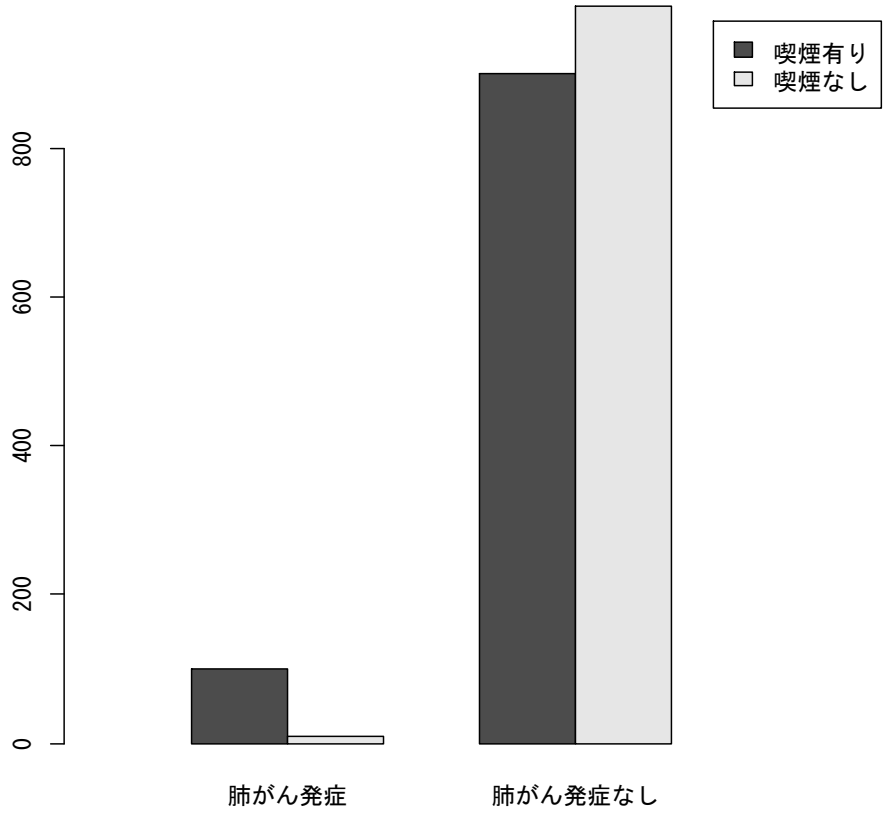
- φ(ファイ)係数
  - 相関係数の一種
  - -1から1の間の値をとる
  - 0なら関連なし

$$\varphi = \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

| | | イベント | |
|---|---|---|---|
| | | 発生 | 発生しない |
| 要因 | あり | A | B |
| | なし | C | D |

# クロス表のグラフ表現


1.0
0.8
0.6
0.4
0.2
0.0
肺がん発症なし
肺がん発症
喫煙有り
喫煙なし


800
600
400
200
0
喫煙有り
喫煙なし
肺がん発症
肺がん発症なし

# $\chi^2$検定（１）

- クロス表の検定
  - クロス表とは、２つの質的変数の関係をみる表
  - 表の中のセルには度数（人数）が入る
- クロス表で期待値から、観測値のズレがあるかどうかの検討（「ズレがない」が帰無仮説）
  - ズレがあるとその表には偏りがある
  - 偏りがあるということは２つのカテゴリカル変数の間に何らかの関係があるということになる

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | 100人 | 900人 | 1000人 |
| | なし | 10人 | 990人 | 1000人 |
| 合計 | | 110人 | 1890人 | 2000人 |

# 期待値

- 合計人数の比から計算できる各セルの期待度数
- 期待値の表は以下のようになる

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | 55人 | 945人 | 1000人 |
| | なし | 55人 | 945人 | 1000人 |
| 合計 | | 110人 | 1890人 | 2000人 |

# $\chi^2$検定（2）

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | 100人 | 900人 | 1000人 |
| | なし | 10人 | 990人 | 1000人 |
| 合計 | | 110人 | 1890人 | 2000人 |

- 帰無仮説
  - 喫煙と肺がんの発症には関係がない

→合計人数の比から計算できる期待値と観察値にズレがない

- 「ズレがあること」とは、比率計算上均等に人数が割り振られていないので偏りが生じているということ

# $\chi^2$検定（3）

- $\chi^2$検定統計量

$$\frac{(\text{観測度数}-\text{期待度数})^2}{\text{期待度数}}$$ の各セルの合計

  - これが自由度(列数－1)×(行数－1)のカイ二乗分布に従う
  - 一般に2×2表と呼ばれる2値と2値の表の検定統計量は以下のように表現できる

$$\frac{N\times|ad-bc|^2}{p\times q\times m\times n}$$

| | | 肺がん | | 合計 |
|---|---|---|---|---|
| | | 発生 | 発生しない | |
| 喫煙 | あり | a人 | b人 | m人 |
| | なし | c人 | d人 | n人 |
| 合計 | | p人 | q人 | N人 |

# 2×2表の検定の時の注意

- 各セルの期待度数が小さくなる（10以下、５以下など諸説あり）場合は以下の式で調整して計算する

$$\frac{N\times(|ad-bc|-N/2)^2}{p\times q\times m\times n}$$

- このような調整法を「イエーツの補正」という

| 経験年数 | 外科病棟 | 内科病棟 | 計 |
|---|---|---|---|
| 10年以上 | a人 | b人 | m人 |
| 10年未満 | c人 | d人 | n人 |
| 計 | p人 | q人 | N人 |

# フィッシャーの直接確率法

- クロス表の検定で各セルの値が小さいとき（4以下）に行う検定法
- 帰無仮説（二つの変数に関係がない）のもとで得られたデータより極端な場合が偶然生じる確率を直接求める
- 場合の数（コンビネーション）を用いて計算する

# 分析例

- ある病院勤務の看護師のうち９８名の経験年数（１０年以上vs１０年未満）と病棟（外科vs内科）で人数に偏りがあるのかないのかを検討する
- クロス表は以下

経験年数 と 病棟 のクロス表

度数

| | | 病棟 | | 合計 |
|---|---|---|---|---|
| | | 内科 | 外科 | |
| 経験年数 | 10年未満 | 39 | 24 | 63 |
| | 10年以上 | 30 | 5 | 35 |
| 合計 | | 69 | 29 | 98 |

# 分析結果

- 結果表で見るのは
  - χ2検定結果、連続修正（イエーツ補正）、Fisherの直接法（一般的には両側）
  - 尤度比はカイ二乗検定と同等（測定値と期待値の比の対数の二乗に測定値をかけたものの各セルの和）だが、クロス表の検定では参照しないことがほとんど
  - 線形と線形による連関は、クロスする変数が連続変量であった場合の検討方法で、一般的には参照しない

**カイ 2 乗検定**

| | 値 | 自由度 | 漸近有意確率 (両側) | 正確有意確率 (両側) | 正確有意確率 (片側) |
|---|---|---|---|---|---|
| Pearson のカイ 2 乗 | 6.122[a] | 1 | .013 | | |
| 連続修正[b] | 5.033 | 1 | .025 | | |
| 尤度比 | 6.605 | 1 | .010 | | |
| Fisher の直接法 | | | | .020 | .011 |
| 線型と線型による連関 | 6.059 | 1 | .014 | | |
| 有効なケースの数 | 98 | | | | |

a. 0 セル (.0%) は期待度数が 5 未満です。最小期待度数は 10.36 です。
b. 2x2 表に対してのみ計算

# 結果の読み取り

- 期待度数（期待値）が5未満のセルはなかったのでFisherは参考にしなくてもよい
- 一番期待度数が小さいセルの期待度数は１０年以上の外科で10.36
  - イエーツ補正を報告する（論文に記述する）かしないかは微妙なところ＝どちらでも良い
  - 補正した結果を報告する場合はイエーツ補正を実施した旨も記載すること
- いずれにせよ5%水準で有意に観測値は期待値よりも偏りがある
  - 経験年数は病棟によって均等でない
  - 外科では経験年数が１０年未満の人が多い

# マン・ホイットニーのU検定

- 2つのグループで順序変数（順序尺度）の差を検定する方法
  - 対応のない t 検定（連続変数のみ）に相当
  - 順序変数だけでなく連続変数でも使えて、検出力が高い検定といわれている
  - 連続変数でも対象者数が少なく、正規分布していない（平均と分散で表すのが妥当でない）データであればこの方法で検討する
  - 順序変数として扱う場合は中央値と範囲（最小値と最大値）を記述統計量として示す
- ウィルコクソンの順位和検定と同等の検定
  - 本によってはウィルコクソンの順位和検定として登場することもあるが、実質的に等価
  - SPSSではマンホイットニーのU検定として登場する

# 例

- 内科と外科の新人看護師それぞれ5名に現在の勤務先について、就活時の志望順位は何位だったかを聞いた
  - 内科：3位、1位、5位、6位、7位
  - 外科：4位、2位、6位、8位、7位
- 志望順位は同じレベルなのか異なるのか
  - 志望順位には差がない、が帰無仮説
  - 以下計算手順を見る（やや複雑）

# 手順１　並び替えとタイ・データ

- グループ（内科、外科）を混みにして、データを一列に並びかえる
- 同順位（タイ・データ）では平均値を与える
  - 6位と7位の平均の6.5、8位と9位の平均の8.5

| 科 | 内科 | 外科 | 内科 | 外科 | 内科 | 内科 | 外科 | 内科 | 外科 | 外科 |
|---|---|---|---|---|---|---|---|---|---|---|
| data | 1 | 2 | 3 | 4 | 5 | 6 | 6 | 7 | 7 | 8 |
| 順位 | 1 | 2 | 3 | 4 | 5 | 6.5 | 6.5 | 8.5 | 8.5 | 10 |

# 手順2　順位和と平均ランク

- 各グループ別に順位の和（順位和）を出す
  - 内科の順位和は、1+3+5+6.5+8.5=24
  - 外科の順位和は、2+4+6.5+8.5+10=31
- 順位和をデータの個数で割る（平均ランク）
  - 内科の平均ランク：　24/5=4.8
  - 外科の平均ランク：　31/5=6.2

# 2つのMann-WhitneyのUの計算

- この後の計算で用いるU値を算出する
- 2つのグループをA、Bとして、ケース数をn1,n2とすると、グループAのU（U1)の計算式は

$$\text{Mann-Whitney の } U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - \text{グループ } A \text{ の順位和}$$

- もう一つのU2は、

$$\text{Mann-Whitney の } U_2 = n_1 n_2 - (\text{Mann-Whitney の } U_1)$$

- U1＞（n1×n2）÷2のときはU2を、それ以外のときはU1をMann-WhitneyのUとして扱う

# 今回の例でのMann-WhittneyのU

- U1=5x5 + 5x(5+1)÷2 – 24 = 16
- U2=5x5 – 16 = 9

  16 > 5x5/2=12.5なので、U2がMann-WhittneyのUとしてこの後の計算式に用いることができるようになる

# 検定統計量を求める

- グループAのケース数がn1、グループBがn2
- 全対象者数はN（= n1 + n2）
- 順位iの同順位データの数をtとする

$$Z = \frac{(U - \frac{n_1 n_2}{2})}{\sqrt{\frac{n_1 n_2}{N(N-1)}(\frac{N^3 - N}{12} - \sum_i T_i)}} \qquad \text{ここで、} T_i = \frac{t^3 - t}{12}$$

- Zは検定統計量で、「標準正規分布」に従う

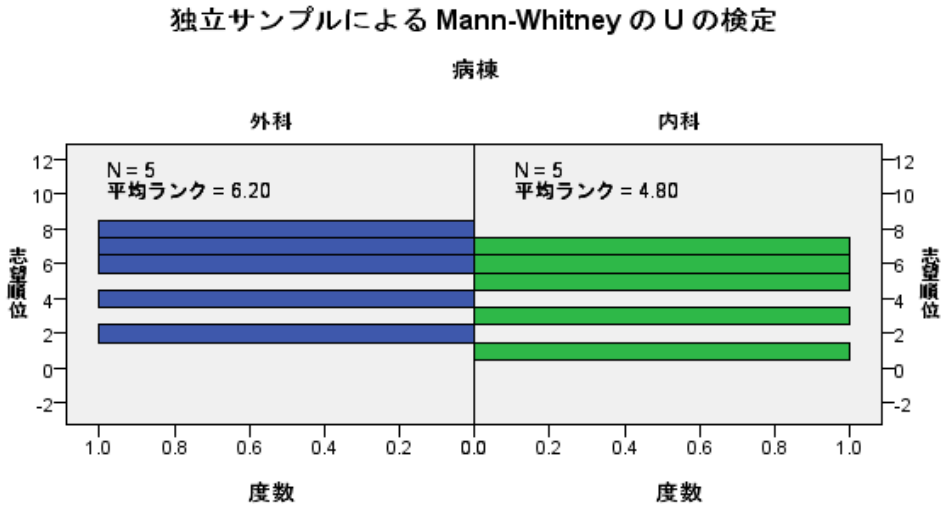
# 今回の例の検定統計量

- Zの分子は、

$$9 - 5x5/2 = -3.5$$

- Zの分母は
  - $\sqrt{\ } 5x5/10x(10-1) \times (10^3-10)/12 - (T1+T2)$
  - 同順位は２つ（6.5と8.5）ある
  - 同順位数は6.5位では2つ、8.5位でも2つなので、
  - T1=（$2^3$-2)/12=0.5、T2=（$2^3$-2)/12=0.5
  - 4.758
- Z＝-3.5÷4.758 = -0.73559

# SPSSの計算結果




| | |
|---|---|
| 合計 N | 10 |
| Mann-Whitney U | 16.000 |
| Wilcoxon W | 31.000 |
| 検定の統計 | 16.000 |
| 標準誤差 | 4.758 |
| 標準化された検定の統計 | .736 |
| 漸近有意確率 (両側検定) | .462 |
| 正確な有意確率 (両側検定) | .548 |

- 一般には漸近有意確率を参照する
  - 正確有意確率は合計30ケース以下で、同順位データが存在しないときのみ参照する
  - WilcoxsonのWはウィルコクソン検定の計算の際に使用する値・・・結果的にMann-WhitneyのU検定と同じなので特に参照せず
  - 標準誤差は先の「分母」の値
  - 標準化された「検定の統計」がZの値
- 有意にはならなかった

# ウィルコクソンの符号付順位検定

- 対応のある順序変数の検定
  →連続量の検定の対応のあるt検定に相当
- 前後で順序は変わらない、が帰無仮説
- 例
  - 看護学生5名への就職説明会の前後でA病院への志望順位がどう変わったか

| 学生 | A | B | C | D | E |
|---|---|---|---|---|---|
| 前 | 1 | 6 | 7 | 9 | 9 |
| 後 | 5 | 2 | 3 | 8 | 4 |

# 計算方法１

- 2つめ（後）の変数から１つめ（前）の変数の値を引く
- その時の差の符号（＋か－か）を示す
  - 値が同じ（差が0）の場合は分析から外す
- 差の絶対値を示す
- 差の絶対値の大きさに順位をつける(同順位は平均順位)

| 学生 | A | B | C | D | E |
|---|---|---|---|---|---|
| 前 | 1 | 6 | 7 | 9 | 9 |
| 後 | 5 | 2 | 3 | 8 | 4 |
| 差 | 4 | -4 | -4 | -1 | -5 |
| 符号 | + | - | - | - | - |
| 差の絶対値 | 4 | 4 | 4 | 1 | 5 |
| 順位 | 3 | 3 | 3 | 1 | 5 |

# 計算方法２

- 符号がプラスになった順位の合計Spと、符号がマイナスになった順位の合計Snを求める
  - Sp ＝ 3
  - Sn ＝ 1＋3＋3＋5 ＝ 12
- SpとSnのそれぞれをケース数で割って平均ランク（Xp、Xn）を求める
  - Xp ＝ 3 / 1 ＝ 3
  - Xn ＝ 12 / 4 ＝3

# 計算方法３

$$Z = \frac{min(S_p, S_n) - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \sum_j^d \frac{(t_j^3 - t_j)}{48}}}$$

- min(Sp,Sn)は、Sp, Snのうち小さい方の値
- nはケース数
- tjはj番目に同順位になった場合のデータの個数
  - 同順位が発生した個数をdとして、ｊは１からｄまで存在しうる
- このZは標準正規分布する

# 今回の例の場合

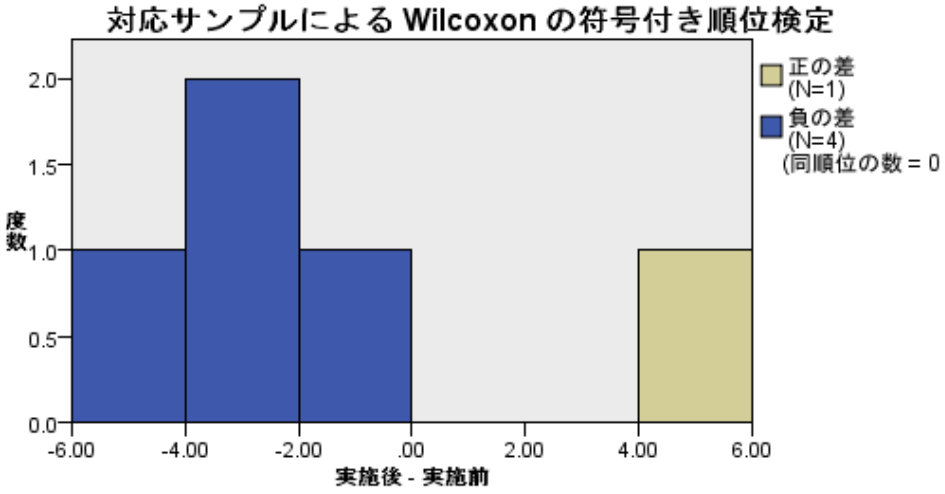
- min(Sp,Sn)は、min（3,3）で同じなので 3
- nは 5 人なので5
- 差の絶対値の同順位は1回だけあった（d=1）
  - $t_1$は3
  - 1度だけ3位が3個あった
- 計算すると

  分母は3 – (5x6)/4

  分子は√ 5x(5+1)(2x5+1)/24 – ($3^3$-3)/48

  計算すると　-1.236245

# SPSS計算結果




| | |
|---|---|
| 合計 N | 5 |
| 検定の統計 | 3.000 |
| 標準誤差 | 3.640 |
| 標準化された検定の統計 | -1.236 |
| 漸近有意確率 (両側検定) | .216 |

- 検定統計はmin(Sp,Sn)のこと
- 標準誤差は、先ほどの分子の計算結果
- 標準化された検定の統計が、検定統計量（標準正規分布を用いて検定する）
- 有意確率は0.216で有意ではなかった

# クラスカル・ウォリス検定

- 3つ以上の群で順位に差があるかどうか
  - 順位に差がない、が帰無仮説
- 例:ある病院の病棟別新人12名の看護基礎技術力スコア
  - Nが小さいのでANOVAでなくノンパラメトリック検定で検討
  - 今度は外来が加わって3カテゴリになった
  - 病棟別で技術力スコアに差があるのかどうかを検討する

| | | | | |
|---|---|---|---|---|
| 内科 | 50 | 80 | 58 | 81 |
| 外科 | 67 | 69 | 72 | 88 |
| 外来 | 54 | 62 | 75 | 77 |

# 並び替える

| | 内科 | 外来 | 内科 | 外来 | 外科 | 外科 | 外科 | 外来 | 外来 | 内科 | 内科 | 外科 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 得点 | 50 | 54 | 58 | 62 | 67 | 69 | 72 | 75 | 77 | 80 | 81 | 88 |
| 順位 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

- 小さい順に並び替えて順位を振り、グループ別に順位和と平均ランクを出す
  - 内科の順位和=25　平均ランク=6.25
  - 外科の順位和=30　平均ランク=7.50
  - 外来の順位和=23　平均ランク=5.75

| | | | | | 順位和 |
|---|---|---|---|---|---|
| 内科 | 1 | 10 | 3 | 11 | 25 |
| 外科 | 5 | 6 | 7 | 12 | 30 |
| 外来 | 2 | 4 | 8 | 9 | 23 |

# 検定統計量Hを求める

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

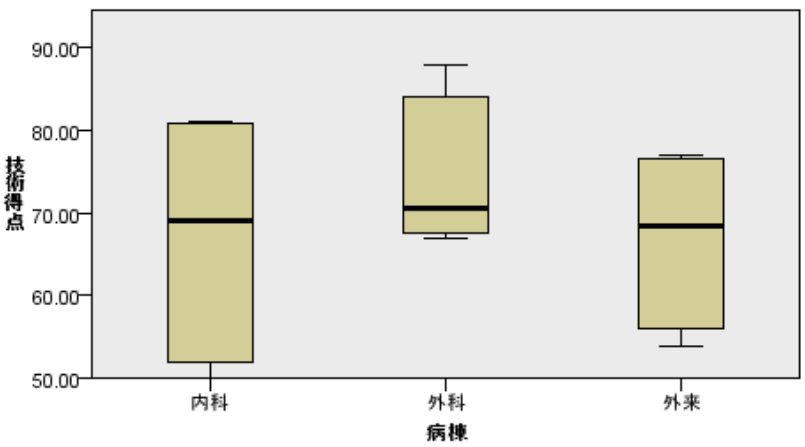
- Riはグループiの順位和、niはグループiのケース数（Nは全ケース数）
- このHはカイ二乗分布するのでカイ二乗検定
  - 自由度は、全グループ数－1
- 今回の例の場合

  12/(12x13) x (252/4 + 302/4 + 232/4) – 3x13=0.500

  自由度は3－1＝2

# SPSSの分析結果

独立サンプルによる Kruskal-Wallis の検定




| | |
|---|---|
| 合計 N | 12 |
| 検定の統計 | .500 |
| 自由度 | 2 |
| 漸近有意確率 (両側検定) | .779 |

1. 検定統計がタイに調整されています。
2. 検定全体でサンプル間の重要な差異を示さないため、複数の比較を実行できません。

- 検定統計量はHの値
- 自由度2のカイ二乗検定ではp＝0.779で有意にならなかった

# フリードマン検定

- 3つ以上の対応のあるグループで順序変数（順序尺度）に違いがあるかを検定する方法
  - 反復測定分散分析に相当
    - 3回以上測定した時系列データや反復測定データの分析に使える

| | 測定時点 | | | |
|---|---|---|---|---|
| 患者 | 1時間後 | 2時間後 | 3時間後 | 4時間後 |
| A | 9 | 17 | 12 | 16 |
| B | 1 | 21 | 16 | 11 |
| C | 7 | 19 | 6 | 9 |

# ノンパラメトリック検定と多重比較

- 3群以上の比較でどの群とどの群の間に差があるかを調べたい
  - ペアごとに検定を繰り返すと検定の多重性の問題が出るのはノンパラメトリック検定でも同じ
- 検定の多重性の問題
  - １度の検定で5%の有意水準にした→95%の確率で有意差がでない
  - 3回t検定を行うと、0.95×0.95×0.95=86%の確率で有意差がでない
  - つまり、1-0.86=0.14で、14%の確率で有意になる

  →検定を繰り返すと有意になりやすくなるため、検定を多く行うと結果が甘くなる
- ノンパラメトリック検定の多重比較の調整で使える方法
  - ボンフェローニ(Bonferroni)法
    - 有意水準を検定の回数で割って厳しくする方法．手計算でも簡単
  - スティール・ドゥワス(Steel-Dwass )法
    - テューキー(Tukey)法のノンパラメトリック版

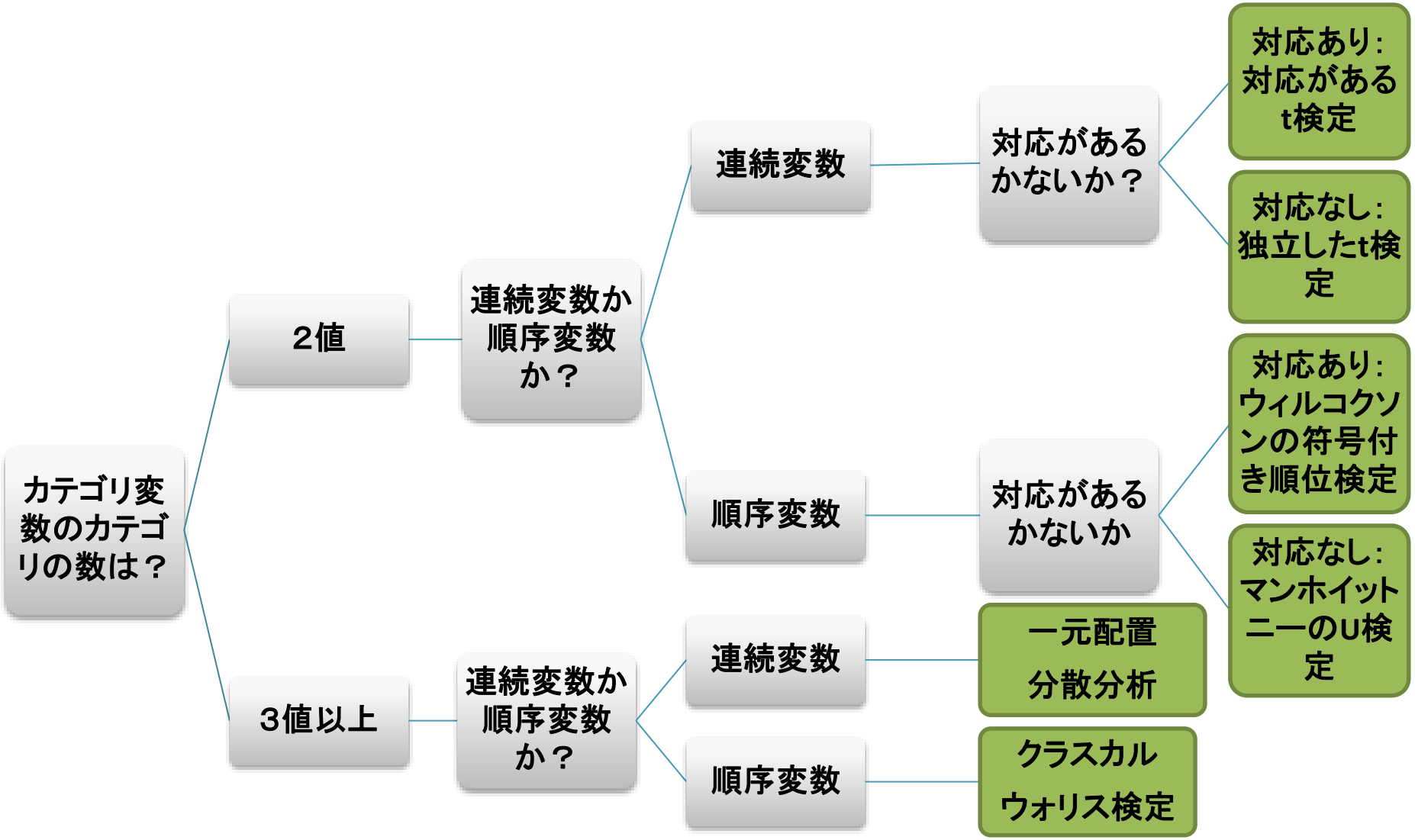
# 順位相関係数

- 順序変数同士の相関係数
- スピアマン(Spearman)とケンドール(Kendall)の2種類
  - スピアマンの順位相関係数(ρ)
    - ピアソンの相関係数(連続量同士の相関係数)の計算に使うデータを順位にしたもの
  - ケンドールの順位相関係数(ケンドールのτ)
    - 2つの順序変数の順位の組み合わせのうち，大小関係が一致するペアの割合

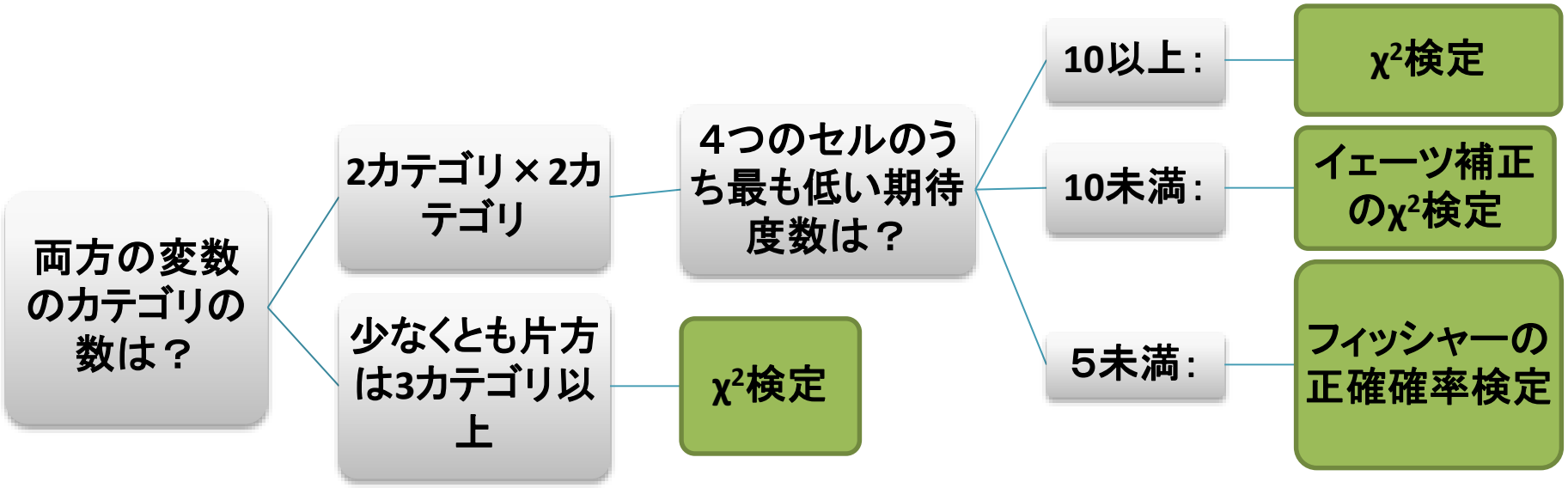
# まとめ

| | パラメトリック | ノンパラメトリック | |
|---|---|---|---|
| | | 名義尺度 | 順序尺度以上 |
| 比率 | | カイ二乗検定，フィッシャーの正確確率検定 | カイ二乗検定，フィッシャーの正確確率検定 |
| 独立した2群の代表値 | 対応のないt検定 | | マンホイットニーのU検定 |
| 対応がある2群の代表値 | 対応のあるt検定 | | ウィルコクソンの符号付き順位検定 |
| 独立した3群(以上)の代表値 | 一元配置分散分析 | | クラスカルワリス検定 |
| 対応がある3群(以上)の代表値 | 反復測定分散分析 | | フリードマン検定 |
| 相関係数 | ピアソンの積率相関係数 | | 順位相関係数 |

# 2変量検定のまとめ１
# カテゴリ変数×連続変数／順序変数



# 2変量検定のまとめ２
# カテゴリ変数×カテゴリ変数



# 参考文献


- 土田昭司　社会調査のためのデータ分析入門　有斐閣　2000
- 高木廣文、林邦彦　エビデンスのための看護研究の読み方・進め方　中山書店　2006
- 中村好一編　論文を正しく読み書くためのやさしい統計学　診断と治療社　2006