

RjpWiki アーカイブス

【Rの古典的検定関数一覧 (06.06.25)】

1 Ansari-Bradley 検定 (対応の無い二標本)

```
ansari.test(x, y, alternative = c("two.sided", "less", "greater"),
            exact = NULL, conf.int = FALSE, conf.level = 0.95, ...)
ansari.test(formula, data, subset, na.action, ...)
```

- 密度関数のスケールパラメータのくいちがいに対する二標本検定 (独立二標本の等分散の検定)
- ノンパラメトリック手法
- x と y がそれぞれ密度関数 $f((t-m)/s)/s$ と $f(t-m)$ を持つ分布からの独立標本で m は未知の局外母数で、スケール比 s が興味のあるパラメータ
- 帰無仮説 $s=1$ (分布は分散だけが異なる)
- 両側対立仮説 $s \neq 1$ (x の分散が異なる)
- 片側対立仮説 $s > 1$ (x の分散が大) または $s < 1$ (小さい)
- 正確計算と正規近似計算を指定可能
- オプションで、 s のノンパラメトリックな信頼区間と、推定値を計算

2 Bartlett 検定 (対応の無い多標本)

```
bartlett.test(x, g, ...)
bartlett.test(formula, data, subset, na.action, ...)
```

- 複数標本の分散の同一性の同時検定
- 正規データを仮定するパラメトリック手法
- 帰無仮説: 分散は全て同一
- χ 自乗分布による近似検定

3 ベルヌイ試行の成功確率 p の検定 (一標本 : 母比率の検定)

```
binom.test(x, n, p = 0.5, alternative = c("two.sided", "less", "greater"),
           conf.level = 0.95)
```

- パラメトリック手法
- 単純帰無仮説 $p = p_0$ (p_0 は特定の確率値)
- 両側対立仮説 $p \neq p_0$
- 片側対立仮説 $p > p_0$ または $p < p_0$
- 正確検定 (exact test)
- 信頼区間も計算

4 χ 自乗検定 (分割表の独立性の検定)

```
chisq.test(x, y = NULL, correct = TRUE,
           p = rep(1/length(x), length(x)),
           simulate.p.value = FALSE, B = 2000)
```

- (本来は) ノンパラメトリック手法
- χ 自乗分布による漸近近似検定
- 2 x 2 分割表に対する連続補正使用の有無を選択可能
- 帰無仮説 引数がベクトルの場合には標本比率が母比率に等しいかどうかの検定
- 帰無仮説 引数が 2 x 2 以上の行列なら二要因の独立性の検定
- モンテカルロ法による p 値の近似計算

5 相関係数の検定 (対標本)

```
cor.test(x, y, alternative = c("two.sided", "less", "greater"),
         method = c("pearson", "kendall", "spearman"),
         exact = NULL, conf.level = 0.95, ...)
cor.test(formula, data, subset, na.action, ...)
```

- パラメトリック手法 (ピアソン法) とノンパラメトリック手法 (ケンドールの τ 、スピアマンの ρ 法)
- ケンドールの τ 法に対しては正確計算指定可能 (既定では正規近似による近似値を計算)
- ピアソン法 (いわゆる相関係数に基づく) は正規データを仮定する t 分布を用いる正確計算
- 帰無仮説 「母相関ゼロ (無相関)」
- 対立仮説 「母相関はゼロでない (相関あり)」
- ピアソン法では、フィッシャーの Z 変換による近似信頼区間を計算
- NA 値があっても可

6 フィッシャーの正確検定 (二次元分割表において周辺和を固定することによる行と列の独立性の検定)

```
fisher.test(x, y = NULL, workspace = 200000, hybrid = FALSE,
            or = 1, alternative = "two.sided", conf.level = 0.95)
```

- パラメトリック手法
- 帰無仮説 「二属性は独立」
- 対立仮説 「二属性は従属」
- 2x2 分割表に対しては片側対立仮説 「正のオッズ比」、「負のオッズ」
- 近似計算も選択可能
- 2x2 分割表に対しては信頼区間も計算

7 Fligner-Killeen (のメディアンによる) 検定 (対応の無い多標本)

```
fligner.test(x, g, ...)  
fligner.test(formula, data, subset, na.action, ...)
```

- 多標本の分散の同一性の同時検定
- あるシミュレーション結果によると正規性仮定のずれに対して最もロバストである
- ランクにもとづくノンパラメトリック手法
- 帰無仮説 「分散は全て同一」
- 対立仮説 「分散は同一でない」

8 Friedman のランク和検定 (繰り返しの無い完備ブロックデザイン型データ)

```
friedman.test(y, groups, blocks, ...)  
friedman.test(formula, data, subset, na.action, ...)
```

- ノンパラメトリック検定
- 帰無仮説 「block 効果を除けば、各 group 毎の位置パラメータは同一である」
- 帰無仮説の下で検定統計量の分布を χ 自乗分布で近似する (漸近近似)
- NA 値不可

9 Kruskal-Wallis のランク和検定 (多標本 ; 3 標本以上)

```
friedman.test(y, groups, blocks, ...)  
friedman.test(formula, data, subset, na.action, ...)
```

- ノンパラメトリック検定
- 帰無仮説「各グループ毎の位置パラメータは同一」
- 対立仮説「各グループ毎の位置パラメータは同一でない」
- 帰無仮説が成立する場合の検定統計量の分布を χ 自乗分布で近似する (漸近近似)
- NA 値不可

10 Kolmogorov-Smirnov 検定 (一標本, 二標本)

```
ks.test(x, y, ..., alternative = c("two.sided", "less", "greater"),  
        exact = NULL)
```

- ノンパラメトリック検定 (経験分布関数に基づく)
- 一標本の場合の帰無仮説「データの分布は指定された連続分布 (理論分布) に等しい」
- 二標本の場合の帰無仮説「二組のデータの連続分布は等しくない」
- 両側・片側対立仮説を選択可能
- データにタイ (同順位) があると警告
- 対立仮説「各標本の位置パラメータは同一でない」
- 正確な p 値は二標本、両側対立仮説で、タイが無いときのみ計算可能
- それ以外は近似計算を用い、データ数が少ないときは不正確である恐れあり

11 Cochran-Mantel-Haenszel χ 自乗検定 (分割表型データ)

```
mantelhaen.test(x, y = NULL, z = NULL,  
                alternative = c("two.sided", "less", "greater"),  
                correct = TRUE, exact = FALSE, conf.level = 0.95)
```

- 3次元の分割表型データ (各次元は少なくとも2つの水準を持ち、最後の次元は層、もしくは少なくとも2水準を持つ因子)
- 帰無仮説「二組の名義尺度データが各層毎に条件付きで独立である」(3次の交互作用無しと仮定)
- $2 \times 2 \times k$ 分割表の場合、両側・片側対立仮説を選択可能
- $2 \times 2 \times k$ 分割表の場合、連続補正をするかどうか選択可能
- $2 \times 2 \times k$ 分割表の場合、Mantel-Haenszel 検定をするか、正確な条件付き検定をするかどうか選択可能 $2 \times 2 \times k$ 分割表の場合、信頼区間を計算
- NA 値不可
- 正確検定以外は帰無分布を χ 自乗分布で近似
- $2 \times 2 \times k$ 分割表の場合、帰無仮説の下での共通のオッズ比の推定値を与える

12 McNemar の χ 自乗検定 (2次元分割表型データの行と列の対称性 ; 2×2 分割表のときは、対応のある比率の差の検定)

```
mcnemar.test(x, y = NULL, correct = TRUE)
```

- 帰無仮説「行と列の対称性, i.e. 全ての i, j で $x[i, j] = x[j, i]$
- 2×2 分割表に対しては連続補正の有無を指定可能
- 帰無分布を χ 自乗分布で近似

13 Mood の二標本検定 (スケールパラメータの同一性)

```
mood.test(x, y, alternative = c("two.sided", "less", "greater"), ...)  
mood.test(formula, data, subset, na.action, ...)
```

- 帰無仮説 「 x データの密度関数 $f(x-l)$, y データの密度関数 $f((x-l)/s)$, l は共通位置パラメータ, s はスケールパラメータ、で $s = 1$ 」
- 両側対立仮説 「 $s \neq 1$ 」
- 片側対立仮説 「 $s > 1$ 」 または 「 $s < 1$ 」
- ノンパラメトリック手法
- もっと役に立つ他の検定があるとの注あり
- NA 値可

14 正規データの平均の同一性検定 ; 一元配置分散分析 (独立二標本, 多標本)

```
oneway.test(formula, data, subset, na.action, var.equal = FALSE)
```

- パラメトリック手法
- 正規標本、分散が異なっても良い
- `var.equal = TRUE` ならば一元配置分散分析による単純な F 検定、`FALSE` なら多標本に対する Welch (近似) 検定の拡張を用いる
- 帰無仮説 「平均は全て同一」
- NA 値可

15 prop.test による (多重比較補正つき) 比率の対比較検定 (ベルヌイ試行型多標本データ)

```
pairwise.prop.test(x, n, p.adjust.method=p.adjust.methods, ...)
```

- 複数のベルヌイ試行の成功比率 (母比率) の prop.test 関数を用いた対比較。多重比較に対する補正を行う (4 種類の補正法については p.adjust 参照)
- 帰無仮説「成功比率の同一性」

16 t 検定による平均の差の (多重比較補正付き) グループ間の対比較 (正規多標本)

```
pairwise.t.test(x, g, p.adjust.method=p.adjust.methods, pool.sd=TRUE, ...)
```

- t.test を用いたグループ毎の平均差の比較。多重比較に対する補正を行う (4 種類の補正法については p.adjust 参照)
- 帰無仮説「平均の同一性」

17 多重比較用検定関数を用いた p 値表作成

```
pairwise.table(compare.levels, level.names, p.adjust.method)
```

18 Wilcoxon 検定を用いた平均の差のグループ毎の対比較 (多重比較補正付き)

```
pairwise.wilcox.test(x, g, p.adjust.method=p.adjust.methods, ...)
```

- wilcoxon.test を用いたグループ毎の平均差の比較。多重比較に対する補正を行う (4 種類の補正法については p.adjust 参照)

19 ANOVA に対する検定の検出力を計算

```
power.anova.test(groups = NULL, n = NULL, between.var = NULL,  
                 within.var = NULL, sig.level = 0.05, power = NULL)
```

20 prop.test に対する検定の検出力を計算

```
power.prop.test(n=NULL, p1=NULL, p2=NULL, sig.level=0.05, power=NULL,  
               alternative=c("two.sided", "one.sided"), strict=FALSE)
```

21 t.test に対する検出力の計算

```
power.t.test(n=NULL, delta=NULL, sd=1, sig.level=0.05, power=NULL,  
            type=c("two.sample", "one.sample", "paired"),  
            alternative=c("two.sided", "one.sided"), strict=FALSE)
```

22 ベルヌイ試行の成功確率に対する検定 (一標本、多標本)

```
prop.test(x, n, p = NULL, alternative = c("two.sided", "less", "greater")  
         conf.level = 0.95, correct = TRUE)
```

- 帰無仮説 一標本なら 「 $p = p_0$ 」 (特定の比率 p_0 に一致するか)
- 帰無仮説 多標本なら 「成功確率 (母比率) は同一」
- 両側対立仮説
- 一標本・二標本ケースでは二種類の片側対立仮説 「より大」、「より小」
- 帰無仮説の下での検定統計量の分布は χ 自乗分布で近似
- Yates の補正をするかどうか指定可能
- 信頼区間の計算

23 集団内のトレンドに対する χ 自乗検定 (ベルヌイ試行多標本)

```
prop.trend.test(x, n, score = 1:length(x))
```

- 集団内で対数オッズ比の変化があるという対立仮説に対する検定
- 漸近的に最適との注釈

24 Quade 検定 (繰り返しの無いブロック化データ)

```
quade.test(y, groups, blocks, ...)  
quade.test(formula, data, subset, na.action, ...)
```

- 正規性を仮定出来ない場合の検定
- 帰無仮説 「ブロック効果を除き位置パラメータが同一」

25 正規性に関する Shapiro-Wilk 検定 (一標本)

```
shapiro.test(x)
```

- 標本数は 3 - 5000 と注釈 (本当に 3 でいいの!?)
- 帰無仮説は 「母集団分布は正規分布」
- パラメトリック検定 (?)

26 正規データの平均、平均差の t 検定 (一標本、二標本)

```
t.test(x, y = NULL, alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = FALSE, var.equal = FALSE, conf.level = 0.95, ...)
t.test(formula, data, subset, na.action, ...)
```

- 正規分布仮定のパラメトリック検定
- 帰無仮説 「平均 = μ_0 」 (一標本)、「平均差 = 0」 (二標本)
- 両側対立仮説 「平均 $\neq \mu_0$ 」 (一標本)、「平均差 $\neq 0$ 」 (二標本)
- 片側対立仮説 「平均 $>, < \mu_0$ 」 (一標本)、「平均差 $>, < 0$ 」 (二標本)
- 二標本ケースでは分散が同一とみなせるか `var.equal=TRUE`、分散が同一とみなさない `var.equal=FALSE`、の選択可
- 一標本および分散が同一とみなせる場合は t 分布による正確検定 (無論正規性の仮定可)、分散が同一とみなさない場合は近似検定である Welch 検定を使用 (この場合も正規性は仮定)
- 信頼区間、推定値も計算
- `paired=TRUE` なら対応のあるデータとみなし、対応データの差からなる一標本ケースとして処理

注：分散が異なる正規二標本の平均の差に関する検定問題は Behrens-Fisher 問題 と呼ばれ、正確検定が無いことが知られている。Welch 検定等のいくつかの近似的検定が知られている。(私が子供時代は Welch 検定は one-of-them だったが、いまや代表?)

27 正規二標本の分散の同一性に対する F 検定

```
var.test(x, y, ratio = 1, alternative = c("two.sided", "less", "greater"),
         conf.level = 0.95, ...)
var.test(formula, data, subset, na.action, ...)
```

- 正規分布に従う二標本に対するパラメトリック検定
- 帰無分布 F 分布による正確検定
- 帰無仮説 「分散比 $s = 1$ 」 ... 等分散である
- 両側対立仮説 「 $s > 1$ 」, 「 $s < 1$ 」 ... 等分散ではない
- 分散比の信頼区間、推定値計算
- 多標本 ($k > 2$) の場合のパラメトリック検定は `bartlett.test`、ノンパラトリック検定は `ansari.test` と `mood.test` が使える

28 位置の差に関する Wilcoxon 検定 (一標本、対応のある二標本、独立二標本)

```
wilcox.test(x, y = NULL, alternative = c("two.sided", "less", "greater"),
            mu = 0, paired = FALSE, exact = NULL, correct = TRUE,
            conf.int = FALSE, conf.level = 0.95, ...)
wilcox.test(formula, data, subset, na.action, ...)
```

- ノンパラメトリック検定
- (a) x データだけが与えられると Wilcoxon signed rank test で検定
- (b) x, y データ双方を与え、paired=TRUE を指定すると、対応のある二標本として対応データの差に関する Wilcoxon signed rank test で検定
- (c) それ以外は Wilcoxon rank sum test (Mann-Whitney test と同値) を使い検定
- (a) のケースでは帰無仮説は「母集団分布が指定値 mu に関して対称」
 - (b) のケースでは帰無仮説は「対応データ差からなる一標本の母集団分布が指定値 mu に関して対称」
 - (c) のケースでは帰無仮説は「二標本の母集団分布が指定値 mu だけ位置がずれている」
 - 両側対立仮説、片側対立仮説指定可能
 - 正確な p 値を計算 exact=TRUE するか、近似値を計算 exact=FALSE するか指定可能
 - p 値を計算する際、正規近似を高めるため連続補正を行うかどうか指定可能
 - ノンパラメトリックな信頼区間と推定値を計算できる