並べかえ検定とブートストラップ検定

樋口さぶろお https://hig3.net

龍谷大学大学院 理工学研究科 数理情報学専攻

理論物理学特論　L15 (2023-01-18 Wed)

最終更新: Time-stamp: ”2023-01-18 Wed 12:44 JST hig”

今日の目標

- 並べかえ検定, ノンパラメトリック検定, ブートストラップ検定の関係を説明できる
- ブートストラップ検定を利用できる

L14-Q1

Quiz 解答: ウィルコクソンの順位和検定

1. $4+3=7.$
2. $4\cdot 2+\frac{1}{2}\cdot 2\cdot(2+1)-4=7.$
3. $1+0+0+0=1.$
4. $4\cdot 2+\frac{1}{2}\cdot 4\cdot(4+1)-17=1.$
5. $\frac{1+1}{{}_6\mathrm{C}_2}=\frac{2}{15}.$

R	U_2	Y の順位	場合の数
2	9		0
3	8	$\{1,2\}$	1
4	7	$\{1,3\}$	1
5	6	$\{1,4\},\{2,3\}$	2
6	5	$\{1,5\},\{2,4\}$	2
7	4	$\{1,6\},\{2,5\},\{3,4\}$	3
8	3	$\{2,6\},\{3,5\}$	2
9	2	$\{3,6\},\{4,5\}$	2
10	2	$\{4,6\}$	1
11	2	$\{5,6\}$	1
12	0		0

6. $\frac{2}{15}>\frac{\alpha}{2}$ なので棄却できない.

ここまで来たよ

統計的仮説検定はだいたいこんな考え方 岩薩林 確率・統計 §6.3

確率変数 X は, 正規分布 $\mathrm{N}(55, 2^2)$ にしたがうという. $\sigma^2 = 2^2$ は確かだとわかってるけど, $\mu = 55$ (帰無仮説) が本当なのか疑っている. サイズ 4 の標本を抽出したところ,

$$54, 57, 57, 60$$

だった. → 標本サイズ $N = 4$, 標本平均値 $\overline{X} = \frac{1}{4}(X_1 + \cdots + X_4) = 57$.

$\overline{X} \sim \mathrm{N}(55, 2^2/4)$.
確率 α(=有意水準) でしか起きないようなまれなことが起きたら (検定統計量 $\overline{X}$ が棄却域にはいったなら)
おかしいと判定する (帰無仮説を棄却する)

検定の答案の書き方

母平均値の上側片側 t 検定

岩薩林 確率・統計 定理 7.2(p.159)

前提 母集団が正規分布 $\mathrm{N}(\mu, \text{何か})$ にしたがう.

1. 有意水準 α で母平均値の上側片側 t 検定を行う.
2. 帰無仮説を母平均値 $\mu = \mu_0$, 対立仮説を $\mu > \mu_0$ とする.
3. 帰無仮説のもとで, 検定統計量 $T = \frac{\overline{X} - \mu_0}{\sqrt{S^2/n}}$ は自由度 $n-1$ の t 分布にしたがう.
4. T の実現値 t を標本平均値 $\overline{X}$, 不偏標本分散 S^2, 標本サイズ n から計算すると $t = \cdots$.
5. 棄却域は $t > t_{n-1}(\alpha)$ である.
6. (結論) 帰無仮説を棄却する/できない, 母平均値は $\mu > \mu_0$ と結論する/できない.

ここまで来たよ

14 ブートストラップ信頼区間・ノンパラメトリック検定

15 並べかえ検定とブートストラップ検定

- 復習: 統計的仮説検定
- 対応のない等分散を仮定しない 2 標本 t 検定
- 並べ替え検定
- ブートストラップ検定
- 母平均値の相等のみを帰無仮説とするときのブートストラップ検定

Welch の t 検定

命題 (Welch の t 検定統計量)

$X_i \sim \mathrm{N}(\mu_1, \sigma_1^2), Y_j \sim \mathrm{N}(\mu_2, \sigma_2^2)$, $i = 1, \dots, m$, $j = 1, \dots, n$, それぞれ独立同分布, に対して,
$\mu_1 = \mu_2$ のとき, Welch の t

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

は自由度 ν の t 分布にしたがう.
$\overline{X}, \overline{Y}$ 標本平均値, S_1^2, S_2^2 不偏標本分散.
自由度 ν は近似的に (整数でない)

$$\nu \approx \frac{(\frac{S_1^2}{m} + \frac{S_2^2}{n})^2}{\frac{S_1^4}{m^2(m-1)} + \frac{S_2^4}{n^2(n-1)}}.$$

ここまで来たよ

並べ替え検定 permutation test

- 帰無仮説: 2 標本 $(x_1, \ldots, x_m, y_1, \ldots, y_n)$ は同一の母分布から得られたものである
- 実行可能となる前提: 標本の違いを測る検定統計量 $T(x_1, \ldots, x_m, y_1, \ldots, x_n)$ がある.

並べ替え検定の手順

1. 有意水準 α を決める.
2. 帰無仮説 (固定)「2 標本は同一の母分布から得られた」
3. 棄却域が大きい側となるような検定統計量 T を決める.
   - 小さい側や両側でも同じことができる
4. 検定統計量 T の, 与えられたサイズ m, n の標本に対する実現値を計算すると t.
5. $(x_1, \ldots, x_m, y_1, \ldots, y_n)$ の, m, n 個への並べ替えすべて $N = {}_{m+n}\mathrm{C}_m$ 個を考える. 検定統計量 t_b^* $(b = 1, \ldots, N)$ を得る.
   - 本当に $B = N$ 個ぜんぶを考える場合と, 乱数で $B < N$ 個をランダムに抽出してモンテカルロ近似する場合とがある.
6. $t_1^*, \ldots, t_B^*$ のうち, t 以上のものの個数を k とする $(0 \leq k \leq B)$.
7. p-値 $p_{\mathrm{perm}} = \frac{k}{B}$.
   - 帰無仮説のもとでは, t, t_b は同じ分布から得られたもの. t が上位 k にはいる確率は $\frac{k}{B}$.
   - 両側や下側で考えてもよい
8. $p_{\mathrm{perm}} < \alpha$ なら帰無仮説を棄却する.

例: 母平均値の差を検定統計量とする並べ替え検定

母平均値の差の検定 (Welch の t 検定) に相当するが, 母分布が正規分布であるという仮定は不要.

母平均値の差を検定統計量とする並べ替え検定

1. 有意水準 $\alpha = 0.05$
2. 帰無仮説 (固定)「2 標本は同一の母分布から得られた」
3. 分布が同じなら母平均値も同じになるはずなので, 標本平均値の差に着目する. $T = \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_1^2}{m}+\frac{S_2^2}{n}}}$ (Welch の検定統計量)
4. 与えられたサイズ n の標本に対する t を計算すると t.
5. $(x_1, \ldots, x_m, y_1, \ldots, y_n)$ の並べ替えすべて $N = {}_{m+n}\mathrm{C}_m$ 個を考える. 前 m 個が X, 後 n 個が Y とみなして検定統計量 t_b^* $(b = 1, \ldots, N)$ を得る.
   - $B = N$ 個ぜんぶを考える場合と, 乱数で $B < N$ 個をランダムに抽出してモンテカルロ近似する場合とがある.
6. $t_1^*, \ldots, t_B^*$ のうち, t 以上のものの個数を k とする $(0 \leq k \leq B)$.
   - 両側や下側で考えてもよい
7. p-値 $p_{\mathrm{perm}} = \frac{k}{B}$.
   - 帰無仮説のもとでは t, t_b は同じ分布から得られたもの. t が上位 k にはいる確率は $\frac{k}{B}$.
8. $p_{\mathrm{perm}} < \alpha$ なら帰無仮説を棄却する.

L15-Q1

Quiz(並べ替え検定)

Y の標本は $\{1.4, 3.8\}$, X の標本は $\{2.4, 4.4, 5.8, 6.8\}$ で与えられる.

1. この X, Y の標本について, Welch の t を計算しよう.
2. 並べ替えはぜんぶで何個あるか. 並べ替えの総数 $B = N$ を求めよう.
3. N 個の並べ替えのうち 1 個 (恒等変換でないもの) を具体的に書こう.
4. それに対して Welch の t を計算しよう.

ウィルコクソンの検定の並べ替え検定としての解釈

ウィルコクソンの検定は, 並べ替え検定で

- $T =$ 順位の和の差
- $N = {}_{m+n}\mathrm{C}_m$ 個すべての並べ替えを考え (モンテカルロ近似でなく), 棄却域を表にしてある

ここまで来たよ

ブートストラップ検定

- 並べ替え検定では, $\{x_1, \ldots, y_n\}$ の $m+n$ 個の元からなる集合の置換を考えている ($m+n$ 個からの $m+n$ 個の非復元抽出).
- ブートストラップ検定では, $x_i(y_j)$ を $\{x_1, \ldots, y_n\}$ から一度に 1 個, 計 $m+n$ 個取り出し $\{x_1, \ldots, y_n\}$ を作る ($m+n$ 個からの $m+n$ 個の復元抽出), それを任意回数 B 回繰りかえす.

ブートストラップ検定の手順

1. 有意水準 α を決める.
2. 帰無仮説 (固定)「2 標本は同一の母分布から得られた」
3. 棄却域が大きい側となるような検定統計量 t を決める.
   - 小さい側や両側でも同じことができる
4. 検定統計量 t の, 与えられたサイズ m, n の標本に対する実現値を計算すると t.
5. 乱数で $(x_1, \ldots, x_m, y_1, \ldots, y_n)$ のブートストラップ標本 B 個 を作る. 前 m 個が x, 後 n 個が y とみなして検定統計量 t_b^* $(b = 1, \ldots, B)$ を得る.
   - 並べ替え検定での並べ替えのモンテカルロ近似と異なり, $x_1, \ldots, x_m, y_1, \ldots, y_n$ から, 1 点ずつ $m + n$ 点を復元抽出する. 手間は小さい.
6. $t_1^*, \ldots, t_B^*$ のうち, t 以上のものの個数を k とする $(0 \leq k \leq B)$.
7. p-値 $p_{\mathrm{boot}} = \frac{k}{B}$.
   - 帰無仮説のもとでは, t, t_b は同じ分布から得られたもの. t が上位 k にはいる確率は $\frac{k}{B}$.
8. $p_{\mathrm{boot}} < \alpha$ なら帰無仮説を棄却する.

例: 平均値の差を検定統計量とするブートストラップ検定

母平均値の差の検定 (Welch の t 検定) に相当.

平均値の差を検定統計量とするブートストラップ検定の手順

1. 有意水準 $\alpha = 0.05$
2. 帰無仮説 (固定)「2 標本は同一の母分布から得られた」
3. $T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$ (Welch の t 検定統計量)
4. 与えられたサイズ m, n の標本に対する t を計算すると t.
5. 乱数で, 標本から B 個のブートストラップ標本を作り, それぞれ t を求める. t_b^* $(b = 1, \ldots, B)$ とする.
6. $t_1^*, \ldots, t_B^*$ のうち, t 以上のものの個数を k とする ($0 \leq k \leq B$).
7. p-値 $p_{\text{boot}} = \frac{k}{B}$.
   - 帰無仮説のもとでは, t, t_b は同じ分布から得られたもの. t が上位 k にはいる確率は $\frac{k}{B}$.
8. $p_{\text{boot}} < \alpha$ なら帰無仮説を棄却する.

L15-Q2

Quiz(ブートストラップ検定)

Y の標本は $\{1.4, 3.8\}$, X の標本は $\{2.4, 4.4, 5.8, 6.8\}$ で与えられる.
X と Y のしたがう分布は同じ, を帰無仮説として, ブートストラップ検定を行う.

1. この X, Y の標本について, Welch の t を計算しよう.
2. サイコロを使ってブートストラップ標本を 1 つ作ろう.
3. ブートストラップ標本に対して Welch の t を計算しよう.

ここまで来たよ

14 ブートストラップ信頼区間・ノンパラメトリック検定

15 並べかえ検定とブートストラップ検定
- 復習: 統計的仮説検定
- 対応のない等分散を仮定しない 2 標本 t 検定
- 並べ替え検定
- ブートストラップ検定
- 母平均値の相等のみを帰無仮説とするときのブートストラップ検定

帰無仮説が $\mu_X = \mu_Y$ のみであるときのブートストラップ検定

帰無仮説「X と Y の母分布が同じ」のものでは, 標本 $\{z_1, \ldots, z_{m+n}\} = \{x_1, \ldots, x_m, y_1, \ldots, y_n\}$ から X,Y 両方のブートストラップ標本を作れた.
帰無仮説が「X と Y の母平均値が同じ」と弱いとき, 標本 $\{x_1, \ldots, x_m\}$ から X の, $\{y_1, \ldots, y_n\}$ から Y の, ブートストラップ標本を作ることしかできない.

母平均値の差のブートストラップ検定の手順

対立仮説:X の母平均値 μ_1 は Y の母平均値 μ_2 より大きい を言いたい.
母分布が等しいとか, 正規分布であるという仮定は不要.

1. 有意水準 $\alpha = 0.05$
2. 帰無仮説 $\mathrm{E}[X] = \mathrm{E}[Y]$.
3. $T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$ (Welch の t 検定統計量)
4. 与えられたサイズ n の標本に対する t を計算すると t.
5. (検定力を上げるため, 標本平均値を同じ位置に調整)
   $\tilde{x}_i = x_i - \overline{x} - \overline{z}, \tilde{y}_i = x_i - \overline{y} - \overline{z}$.
6. 乱数で, 調整済 $\tilde{x}$ の標本から X の, $\tilde{y}$ の標本から Y の, B 個のブートストラップ標本を作り, それぞれ t を求める. t_b^* $(b = 1, \ldots, B)$ とする.
7. $t_1^*, \ldots, t_B^*$ のうち, t 以上のものの個数を k とする $(0 \leq k \leq B)$.
8. p-値 $p_{\mathrm{boot}} = \frac{k}{B}$.
   - 帰無仮説のもとでは, t, t_b は同じ分布から得られたもの. t が上位 k にはいる確率は $\frac{k}{B}$.
9. $p_{\mathrm{boot}} < \alpha$ なら帰無仮説を棄却する.

L15-Q3

Quiz(ブートストラップ検定)

Y の標本は $\{1.4, 3.8\}$, X の標本は $\{2.4, 4.4, 5.8, 6.8\}$ で与えられる.
X と Y の母平均値は同じ, を帰無仮説として, ブートストラップ検定を行う.

1. この X, Y の標本について, Welch の t を計算しよう.
2. 標本平均値を調整した $\overline{x}, \overline{y}$ の標本を作ろう.
3. サイコロを使ってブートストラップ標本を 1 つ作ろう.
4. ブートストラップ標本に対して Welch の t を計算しよう.