母比率の検定・p 値・統計的仮説検定の考え方

樋口さぶろお　https://hig3.net

龍谷大学理工学部数理情報学科

確率統計☆演習　L15(2021-01-18 Mon)

最終更新: Time-stamp: "2021-02-06 Sat 19:32 JST hig"

今日の目標

- 母比率の片側・両側検定ができる 岩薩林 確率・統計 §7.3
- 棄却域 岩薩林 確率・統計 p.149 ,p 値 岩薩林 確率・統計 p.152 両方の考え方で検定ができる
- 有意水準 α・検出力 $1-\beta$ と混同行列の関係を説明で

L14-Q1

Quiz 解答:母平均値の両側検定 (母分散未知)=両側 t 検定

1. 有意水準 0.05 で, 正規分布の母平均値に対する両側 t 検定を行う.
2. 帰無仮説を「ドーナツの重さの母平均値 μ が $\mu_0 = 55\text{g}$ に等しい」すなわち「$\mu = \mu_0$」とする. 対立仮説を「$\mu \neq \mu_0$」とする.
3. サイズ n の標本の標本平均値を $\bar{X}$, 不偏標本分散を s^2 とするとき, 検定統計量
$$T = \frac{\bar{X} - \mu_0}{\sqrt{s^2/n}}$$
は, 帰無仮説のもとで, 自由度 $n-1$ の t 分布に従う.
4. この標本の実現値は, $t = \frac{\bar{x}-\mu_0}{\sqrt{\frac{s^2}{n}}} = \frac{49-55}{\sqrt{\frac{1}{5}\frac{16}{5-1}}} = -3\sqrt{5} = -6.708.$
5. t 分布表より, 棄却域は $|t| > t_4(0.05/2)$ すなわち $|t| > 2.776.$

6. t の実現値は棄却域に含まれるので，帰無仮説を棄却する．ドーナツの重さの母平均値は 55g と異なる，と結論する．
(注: このことを，「有意」significant という言葉で表現する人もいる．結果は有意である，母平均値 μ は 55g と有意に異なる，母平均値 μ と 55 の間には有意差がある，有意な標本である，など)

L14-Q2
Quiz 解答:正規分布の母平均値に関する t 検定

1. 有意水準 0.05 で，正規分布の母平均値に対する両側 t 検定を行う．
2. 帰無仮説を「ドーナツ販売開始後の，来店客数の母平均値 μ は 196 に等しい」，すなわち $\mu = 196$ とする．対立仮説を $\mu \neq 196$ とする．

3. サイズ n の標本の標本平均値を $\bar{X}$, 不偏標本分散を s^2 とすると, 検定統計量
$$T = \frac{\bar{X} - 196}{\sqrt{s^2/n}}$$
は, 帰無仮説のもとで, 自由度 $n-1$ の t 分布に従う.
4. この標本の実現値は $\bar{X} = 200, s^2 = \frac{224}{4-1} = 74.7$. よって, $t = \frac{200-196}{\sqrt{\frac{1}{4}\frac{224}{3}}} = 0.92582.$
5. t 分布表より, 棄却域は $|t| > t_3(0.05/2)$ すなわち, $|t| > 3.182$.
6. この標本の実現値は棄却域に含まれないので, 帰無仮説は棄却できない. 来店客数の母平均値が変化したとは結論できない.
(注: 結果は有意でなかった, 母平均値 μ と 196g の間には有意差がない, など).

L14-Q3

Quiz 解答:母平均値の片側検定 (母分散未知)=片側 t 検定

1. 有意水準 $\alpha = 0.05$ で, 正規分布の母平均値に対する片側 t 検定を行う.
2. 帰無仮説を「ドーナツの重さの母平均値 μ が $\mu_0 = 55$g と等しい」すなわち「$\mu = \mu_0$」とする. 対立仮説を「$\mu < \mu_0$」とする.
3. サイズ n の標本の標本平均値を $\bar{X}$, 不偏標本分散を s^2 とするとき, 検定統計量
$$T = \frac{\bar{X} - \mu_0}{\sqrt{s^2/n}}$$
は, 帰無仮説のもとで, 自由度 $n-1$ の t 分布に従う.
4. この標本の実現値は, $t = \frac{\bar{x}-\mu_0}{\sqrt{\frac{S^2}{n}}} = \frac{53-55}{\sqrt{\frac{1}{5}\frac{13}{5-1}}} = 4\sqrt{5/13} = -2.4807.$
5. t 分布表より, 棄却域は $t < t_4(0.95)$ すなわち $t < -t_4(0.05)$ すなわち $t < -2.132$.
6. t の実現値は棄却域に含まれるので, 帰無仮説を棄却する. ドーナツの重さの母平均値は 55g より小さい, と結論する.

ここまで来たよ

14 統計的仮説検定—母平均値の両側・片側検定

15 母比率の検定・p 値・統計的仮説検定の考え方
- 母比率の (二項) 検定 (の正規近似)
- p 値=有意確率
- 有意水準と検出力, 第 1 種の過誤, 第 2 種の過誤

母比率の検定 I

p と書いてた母比率, 二項分布の $\mathrm{B}(p,n)$ の p を, 今日は r と書きます

Example (母比率の検定のアイデア)

このマジック用コインは表の出る確率 (母比率)$r = 3/5$ だと言われているが, もっと表が出る気がする.
50 回投げた (サイズ $n = 50$ のサンプル) ところ, 表 $T = 35$ 回表が出た.
こんな「稀な」ことが起きるってことは, そうに違いない!

$\rightsquigarrow$ これ, 統計的仮説検定として定式化できる考え方
前提 $T \sim \mathrm{B}(r,n)$.
帰無仮説 $r = \frac{3}{5}$, 対立仮説 $r > \frac{3}{5}$.
帰無仮説 (背理法の仮定) のもとで, $T = 35$ はとても稀 (矛盾) と言いたい. 確率統計☆演習 I(2020)L08

$$P(T=35) = \frac{50!}{35!(50-35)!}\left(\frac{3}{5}\right)^{35}\left(1-\frac{3}{5}\right)^{50-15} = 0.04155 \quad \text{稀じゃん}$$

母比率の検定 II

ちょっと待て. 特定の回数が出ることじたい稀. 上の確率は
$1/51 = 0.0196$ より大きい.
もっともらしい

$$P(T = 30) = \frac{50!}{35!(50-35)!}\left(\frac{3}{5}\right)^{35}\left(1 - \frac{3}{5}\right)^{50-15} = 0.11456$$

だって稀じゃん?
これ (=35) 以上に極端なことが起きる確率

$$P(35 \leq T \leq 50) = \sum_{k=35}^{50} P(T = k) = 16 \text{ 個も加えるの?}$$

母比率の検定 III

正規近似による計算 確率統計☆演習 I(2020)L11

$\mathrm{E}[T] = 50 \cdot \frac{3}{5} = 30, \mathrm{V}[T] = 50 \cdot \frac{3}{5} \cdot \frac{2}{5} = 12.$
中心極限定理より, ‘近似的に’ $T \sim \mathrm{N}(30, 12)$. $Z = \frac{T-30}{\sqrt{12}} \sim \mathrm{N}(0, 1^2)$.

$$\begin{aligned} P(35 \le T \le 50) =& P(\tfrac{35-30}{\sqrt{12}} \le Z \le +\infty) \\ =& I(+\infty) - I(\tfrac{35-30}{\sqrt{12}}) \\ =& \tfrac{1}{2} - I(1.44) = 1 - 0.4251 = 0.0749. \end{aligned}$$

確率統計☆演習 I(2020)L10

$P(T \ge 35) = 0.0749$ (これがあとで出てくる p 値) が稀と思うかは主観. けんかにならないように, 計算し始める前に基準を決めておく. それが, 有意水準 $\alpha = 0.05$ or 0.01.

母比率の検定 IV

上で出てきた, Z は, 分子分母を $n = 50$ で割ると,

$$\frac{35-30}{\sqrt{12}} = \frac{\frac{35}{50} - \frac{30}{50}}{\sqrt{\frac{1}{50}\frac{3}{5}\frac{2}{5}}} = \frac{\hat{r} - r_0}{\sqrt{\frac{1}{n} \cdot r_0 \cdot (1 - r_0)}}$$

という, 母比率の区間推定の導出に似た形になる.
ただし, 平方根の中にあるのは帰無仮説の母比率 r_0, 標本比率は $\hat{r}$.

一般の統計的仮説検定の，レポートや論文での書き方

確率統計☆演習 I(2020)L14

母集団を決める．母集団の分布タイプを仮定する．使う検定を決める/決まる (将来は自作できる).

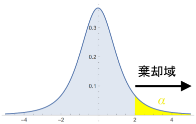
1. 「有意水準 $\alpha = \cdots$ で」「…検定を行う」(2,3 を名前で予告する)
2. 「帰無仮説を…とする」「対立仮説を…とする」
3. 「帰無仮説のもとでナントカ検定統計量 Y は …分布にしたがう」
4. 「この標本に対してナントカ検定統計量の実現値は $y = \cdots$ である」
5. (棄却域の境い目の値を計算しておく)
6. 「'y 不等号 (境い目)' より帰無仮説を棄却する/棄却できない」「よって母ナントカは…である/とはいえない」

母比率の片側 (二項) 検定 (の正規近似)

岩薩林 確率・統計 定理 7.6(p.171)

前提 $T \sim \mathrm{B}(r, n)$.

1. 有意水準 α で正規近似による母比率の片側 (二項) 検定を行う.
2. 帰無仮説を母比率 $r = r_0$, 対立仮説を母比率 $r > r_0$ とする.
3. 帰無仮説のもとで検定統計量 $Z = \frac{T - nr_0}{\sqrt{nr_0(1-r_0)}} = \frac{\hat{r} - r_0}{\sqrt{r_0(1-r_0)/n}}$ は近似的に標準正規分布 $\mathrm{N}(0, 1^2)$ にしたがう (n 大のとき).
   $\hat{r} = \frac{T}{n}$: 標本比率, r_0: (帰無仮説の) 母比率.
4. この標本に対して検定統計量の実現値は…
5. 棄却域は $z > z(\alpha)$.
6. 「(…) より帰無仮説を棄却する/できない. よって母比率 $r > r_0$ と結論する/できない.



帰無仮説として適切であるかどうかは, 日本語の語尾「である」,「でない」だけで決まるわけではに.
帰無仮説は「$r \neq \frac{3}{5}$ でない」, 最後の文は「$r \leq \frac{3}{5}$ でないと結論する」なども書ける.
帰無仮説は $=$ 的なもの.

L15-Q1

Quiz(母比率の片側検定)

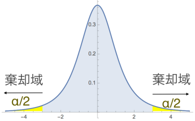
あるスピードくじ (「あたり」と「はずれ」だけがある) は, あたりの母比率 r は $\frac{1}{10}$ に等しいと言われている. しかし, 実際の r はこれより大きいのではないかと疑っている.
くじを 100 本ひいたところ, 15 本があたりだった.
有意水準 $\alpha = 0.05$ で母比率の検定を行おう.

岩薩林 確率・統計 例題 7.7, 問題 8(p.172), 第 7 章練習問題 2,3

母比率の両側 (二項) 検定 (の正規近似)

岩薩林 確率・統計 定理 7.6(p.171)

1. 有意水準 α で正規近似による母比率の両側 (二項) 検定を行う.
2. 帰無仮説を母比率 $r = r_0$, 対立仮説を母比率 $r \neq r_0$ とする.
3. 帰無仮説のもとで検定統計量 $Z = \frac{T - nr_0}{\sqrt{nr_0(1-r_0)}} = \frac{\hat{r} - r_0}{\sqrt{r_0(1-r_0)/n}}$ は近似的に標準正規分布 $\mathrm{N}(0, 1^2)$ にしたがう (n 大のとき). $\hat{r} = \frac{T}{n}$: 標本比率, r_0: (帰無仮説の) 母比率.
4. この標本に対して検定統計量の実現値は…
5. 棄却域は $|z| > z(\frac{\alpha}{2})$.
6. 「(…) より帰無仮説を棄却する/できない. よって母比率 $r \neq r_0$ と結論する/できない.



L15-Q2

Quiz(母比率の両側 (二項) 検定)

瀬田学舎生のうち, 滋賀県の高校を卒業した人の母比率は $r = 0.4$ でない, ことを示すため, サイズ 68 の標本を抽出したところ, 20 名が滋賀県の高校を卒業していた. $r = 0.4$ でないことを結論できるか?

ここまで来たよ

14 統計的仮説検定—母平均値の両側・片側検定

15 母比率の検定・p 値・統計的仮説検定の考え方
- 母比率の (二項) 検定 (の正規近似)
- p 値=有意確率
- 有意水準と検出力, 第 1 種の過誤, 第 2 種の過誤

p 値=有意確率による棄却する/しない判定 岩薩林 確率・統計 p.152

検定の Step6 では, T や Z が極端かどうか判定している

$T = 35 \Leftrightarrow Z = \frac{35-30}{\sqrt{12}}$ が棄却域 ($\alpha = 0.05$) に入っているか?

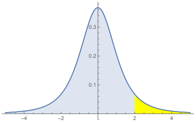
$\Updownarrow$

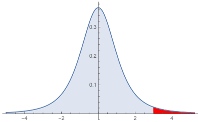
p 値 $P(T \geq 35) = P(z \geq \frac{35-30}{\sqrt{12}})$ が有意水準 α より小さいか?

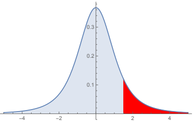
標本の p 値 (p-value)=有意確率 岩薩林 確率・統計 p.152

帰無仮説のもとで, 検定統計量がこの標本よりも極端な値をとる確率=端側の面積

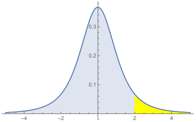
$\alpha > p$ のとき帰無仮説を棄却

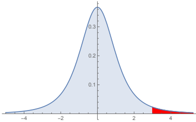


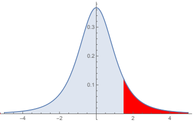




勝負			棄却する /	しない	
統計量の値	片側	$+z(\alpha)$	$<$	$>$	$Z = \frac{\hat{r}-r_0}{\sqrt{r_0(1-r_0)/n}}$
		$-z(\alpha)$	$>$	$<$	
	両側	$+z(\frac{\alpha}{2})$	$<$	$>$	$\lvert Z\rvert = \boxed{\frac{\hat{r}-r_0}{\sqrt{r_0(1-r_0)/n}}}$
確率＝面積		α	$>$	$<$	p 値







統計的仮説検定での p 値による棄却判定

1. 有意水準 α で正規近似による母比率の片側 (二項) 検定を行う.
2. 帰無仮説を母比率 $r = r_0$, 対立仮説を母比率 $r > r_0$,
3. 帰無仮説のもとで検定統計量 $Z = \frac{T - nr_0}{\sqrt{nr_0(1-r_0)}} = \frac{\hat{r} - r_0}{\sqrt{r_0(1-r_0)/n}}$ は近似的に標準正規分布 $\mathrm{N}(0, 1^2)$ にしたがう (n 大のとき). $\hat{r}$ 標本比率, r_0 (帰無仮説の) 母比率
4. この標本に対して検定統計量の実現値は⋯
5. ~~棄却域は $z > z(\alpha)$.~~ p 値は $P(Z > \frac{T - nr_0}{\sqrt{nr_0(1-r_0)}})$.
6. 「~~(⋯)~~ p 値 $<$ / $\geq$ α より帰無仮説を棄却する/できない. よって母比率 $r > r_0$ と結論する/できない.

p 値で再度やってみよう.
L15-Q1

Quiz(母比率の片側検定)

あるスピードくじ(「あたり」と「はずれ」だけがある)は, あたりの母比率 r は $\frac{1}{10}$ に等しいと言われている. しかし, 実際の r はこれより大きいのではないかと疑っている.
くじを 100 本ひいたところ, 15 本があたりだった.
有意水準 $\alpha = 0.05$ で母比率の検定を行おう.

岩薩林 確率・統計 例題 7.7, 問題 8(p.172), 第 7 章練習問題 2,3

ここまで来たよ

14. 統計的仮説検定—母平均値の両側・片側検定
15. 母比率の検定・p 値・統計的仮説検定の考え方
    - 母比率の (二項) 検定 (の正規近似)
    - p 値=有意確率
    - 有意水準と検出力, 第 1 種の過誤, 第 2 種の過誤

検定とインフルエンザ検査 岩薩林 確率・統計 §6.3

	検査陽性 (発色)	検査陰性 (発色なし)
病気かかってる	○ (TP 真陽性)	× (FN 見逃し, 偽陰性)
病気かかってない	× (FP 誤検出, 偽陽性)	○ (TN 真陰性)

真/偽=True/False, 陽性/陰性=Postive/Negative
これに人数を記入したものが混同行列,confusion matrix

検査薬の性能 1 個の数値だけで測れない.
どれも大きい方がいいが, 一方を大きくすると他が小さくなる.
精度=適合率=precision= $\frac{\text{TP}}{\text{TP+FP}}$
感度=真陽性率=recall=sensitivity= $\frac{\text{TP}}{\text{TP+FN}} = 1-\beta$
特異度=真陰性率=specificity= $\frac{\text{TN}}{\text{TN+FP}} = 1-\alpha$

	帰無仮説を棄却, 有意	帰無仮説を棄却できない, 有意でない
正常でない, 対立仮説が成立 $\mu \neq \mu_0$	確率 $1-\beta$	第 2 種の過誤 確率 β
正常, 帰無仮説が成立 $\mu = \mu_0$	第 1 種の過誤 確率 α	確率 $1-\alpha$

$1-\alpha$, $1-\beta$ はどちらも大きくしたいが, 両立しない.
α はほぼゼロ ($\alpha = 0.01$ or 0.05) に固定して, $1-\beta$ をなるべく大きくするように試みる習慣.
見逃し率: β. 検出力:$1-\beta$.
誤検出率, 有意水準, 危険率: α.