母比率の片側検定・p 値・適合度のカイ二乗検定

樋口さぶろお https://hig3.net

龍谷大学 先端理工学部 数理・情報科学課程

確率統計 I L15(2024-07-29 Mon)

最終更新: Time-stamp: "2024-07-29 Mon 07:00 JST hig"

今日の目標

- 岩薩林 確率・統計 §7.3 母比率の片側検定ができる
- 岩薩林 確率・統計 p.152 p 値を使った検定ができる
- 第 1 種, 第 2 種の過誤, 有意水準と検出力を説明できる
- 岩薩林 確率・統計 §8.3 適合度のカイ二乗検定ができる

L14-Q1
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 $n = 4$ 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
$$T = \frac{\bar{X} - 196}{\sqrt{S^2/4}}$$
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 $|t| > F^{-1}(0.975) = t_3(0.05/2)$ すなわち, $|t| > 3.182$.
6. $|0.92582| < 3.182$ であり, 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).

一般の統計的仮説検定の，レポートや論文での書き方

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)「棄却域は…」
6. 「…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」

母平均値・母分散・母比率・適合度・独立性の検定

	両側	片側
母平均値	**両側 t 検定** 岩薩林 確率・統計 問題 2(p.160)	**片側 t 検定** 岩薩林 確率・統計 例題 7.2(p.159)
母分散	**両側カイ二乗検定** 岩薩林 確率・統計 例題 7.4(p.165)	**片側カイ二乗検定** 岩薩林 確率・統計 問題 5(p.166) 岩薩林 確率・統計 §7 練習問題 1(3)
母比率	**両側二項検定** 岩薩林 確率・統計 練習問題 2(p.173)	**片側二項検定** 岩薩林 確率・統計 例題 7.7(p.172) 岩薩林 確率・統計 練習問題 3(p.173)
適合度	NA	**片側カイ二乗検定** 岩薩林 確率・統計 §8.3
独立性	NA	**片側カイ二乗検定** 岩薩林 確率・統計 §8.4
母平均値の差	**両側 t 検定** 岩薩林 確率・統計 §8.1 対応なし (等分散, Welch), あり	**片側 t 検定** 岩薩林 確率・統計 §8.1 対応なし (等分散, Welch), あり
母分散の比	**両側 F 検定** 岩薩林 確率・統計 §8.2	**片側 F 検定** 岩薩林 確率・統計 §8.2

ここまで来たよ

14 統計的仮説検定・母平均値の両側 t 検定

15 母比率の片側検定・p 値・適合度のカイ二乗検定
- 母比率の検定=二項検定 (の正規近似)
- p 値=有意確率
- 仮説検定での第 1 種・第 2 種の過誤と混同行列
- 適合度のカイ二乗検定

母比率の検定 I

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

例 (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)$r = 1/5$ だと言われているが, もっと表がでちゃう気がする.
50 回投げた (サイズ $n = 50$ のサンプル) ところ, 表 $X = 15$ 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 1 回投げたときの表 (1) 裏 (0) $X_i \sim \mathrm{B}(r, 1)$. (無限母集団はベルヌイ分布)
前提の サイズ n のサンプルに対して合計 $X = \sum_{i=1}^{n} X_i$ は $X \sim \mathrm{B}(r, n)$. (二項分布).
帰無仮説 $r = \frac{1}{5}$, 対立仮説 $r > \frac{1}{5}$.

母比率の検定 II

帰無仮説 (背理法の仮定) のもとで, $X = 15$ はとても稀 (矛盾) と言いたい. 確率統計 I(2024)L14

$$P(X=15) = \frac{50!}{15!(50-15)!}\left(\frac{1}{5}\right)^{15}\left(1-\frac{1}{5}\right)^{50-15} = 0.0299 \quad \text{稀じゃん}$$

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
$1/$場合の数 $= 1/51 = 0.0196$ より大きい.
もっともらしい $50 \cdot \frac{1}{5}$ 回

$$P(X=10) = \frac{50!}{10!(50-10)!}\left(\frac{1}{5}\right)^{10}\left(1-\frac{1}{5}\right)^{50-10} = 0.1398$$

だって稀じゃん?
これ $(X = 15)$ 以上に極端なことが起きる確率を考えよう.

$$P(15 \le X \le 50) = \sum_{k=15}^{50} P(X=k)$$

母比率の検定 III

36 個も加えるの? …計算機で直接計算する手もある正確確率検定
正規近似による計算 確率統計 I(2024)L09

$\mathrm{E}[X] = 50 \cdot \frac{1}{5} = 10, \mathrm{V}[X] = 50 \cdot \frac{1}{5} \cdot \frac{4}{5} = 8.$
サンプルサイズ 50 大と思うと，中心極限定理より，近似的に $X \sim \mathrm{N}(10, 8)$. $Z = \frac{X-10}{\sqrt{8}} \sim \mathrm{N}(0, 1^2)$.

$$\begin{aligned} P(15 \le X \le 50) =& P\left(\frac{15-10}{\sqrt{8}} \le Z < +\infty\right) \\ =& \Phi(\infty) - \Phi\left(\frac{15-10}{\sqrt{8}}\right) \\ =& 1 - \mathrm{cdf}\left(\frac{15-10}{\sqrt{8}}\right) = 1 - 0.96145 = 0.0385. \end{aligned}$$

確率統計 I(2024)L09

$P(X \ge 15) = 0.0385$ (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. そのときのぎりぎりの z の値が棄却域の境い目.

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,

$$\frac{15-10}{\sqrt{8}} = \frac{\frac{15}{50} - \frac{10}{50}}{\sqrt{\frac{1}{50}\frac{1}{5}\frac{4}{5}}} = \frac{\hat{r} - r_0}{\sqrt{\frac{1}{n} \cdot r_0 \cdot (1 - r_0)}}$$

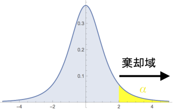
という, 母比率の区間推定の導出に似た形になる.
ただし, 平方根の中にあるのは帰無仮説の母比率 $r_0 = \frac{1}{5}$, 標本比率は $\hat{r}$.

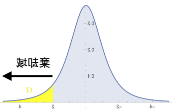
母比率の片側 (二項) 検定 (の正規近似) 岩薩林 確率・統計 定理 7.6(p.171)

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 $X \sim \mathrm{B}(r, n)$.

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 $r > r_0$ (or $r < r_0$) とする.
3. 帰無仮説のもとで，サイズ n の標本から計算する検定統計量 $Z = \frac{X - nr_0}{\sqrt{nr_0(1-r_0)}} = \frac{\hat{r} - r_0}{\sqrt{r_0(1-r_0)/n}}$ 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). $\hat{r} = \frac{X}{n}$: 標本比率.
4. 標本に対して検定統計量 Z の実現値 z は…
5. 棄却域は $z > \Phi^{-1}(1-\alpha)$ (or $z < \Phi^{-1}(\alpha)$).
6. 「(…) より帰無仮説を棄却する/できない. よって母比率 $r > r_0$ (or $r < r_0$) と結論する/できない」

上側片側 

下側片側 

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

L15-Q2

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 はこれより小さいのではないかと疑っている.
くじを 100 本ひいたところ, 5 本があたりだった.
有意水準 $\alpha = 0.01$ で母比率の検定を行おう.

ここまで来たよ

14 統計的仮説検定・母平均値の両側 t 検定

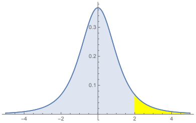
15 母比率の片側検定・p 値・適合度のカイ二乗検定
- 母比率の検定=二項検定 (の正規近似)
- p 値=有意確率
- 仮説検定での第 1 種・第 2 種の過誤と混同行列
- 適合度のカイ二乗検定

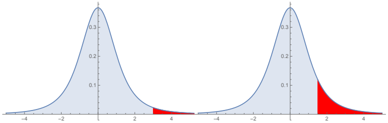
p 値=有意確率による棄却する/しない判定 岩薩林 確率・統計 p.152

検定の Step6 では, X や Z が極端かどうか判定している
$X = 15 \Leftrightarrow Z = \frac{15-10}{\sqrt{8}}$ が棄却域 ($\alpha = 0.05$) に入っているか?
$\Updownarrow$
p 値 $P(X \geq 15) = P(z \geq \frac{15-30}{\sqrt{8}})$ 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
$\alpha > p$ のとき帰無仮説を棄却





上側片側検定のとき p 値 $= 1 - F(z)$.
棄却域の考え方での棄却: $z > F^{-1}(1-\alpha)$.　　$F(z)$ は増加関数だから…
p 値の考え方での棄却: $F(z) > 1 - \alpha$ すなわち $p = 1 - F(z) < \alpha$.

母比率の片側 (二項) 検定 (の正規近似) in p 値 岩薩林 確率・統計 定理 7.6(p.171)

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 $X \sim \mathrm{B}(r, n)$.

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 $r > r_0$ (or $r < r_0$) とする.
3. 帰無仮説のもとで，サイズ n の標本から計算する検定統計量 $Z = \frac{X - nr_0}{\sqrt{nr_0(1-r_0)}} = \frac{\hat{r} - r_0}{\sqrt{r_0(1-r_0)/n}}$ 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). $\hat{r} = \frac{X}{n}$: 標本比率.
4. 標本に対して検定統計量 Z の実現値 z は⋯
5. ~~棄却域は~~ $\cancel{z > \Phi^{-1}(1-\alpha)}$ ~~(or~~ $\cancel{z < \Phi^{-1}(\alpha)}$~~).~~ 実現値 z の p 値は $p = \Phi(\infty) - \Phi(z)$. (or $p = \Phi(z) - \Phi(-\infty)$)
6. 「$p < \alpha$ より帰無仮説を棄却する (できない). よって母比率 $r > r_0$ (or $r < r_0$) と結論する/できない」

ここまで来たよ

14 統計的仮説検定・母平均値の両側 t 検定

15 母比率の片側検定・p 値・適合度のカイ二乗検定

- 母比率の検定=二項検定 (の正規近似)
- p 値=有意確率
- 仮説検定での第 1 種・第 2 種の過誤と混同行列
- 適合度のカイ二乗検定

母分布についての真実と，検定での棄却の有無

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

α, β はどちらも小さくしたいが, 両立しない.
α はほぼゼロ ($\alpha = 0.01$ or 0.05) に固定して, β をなるべく小さ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.
類似: 混同行列 確率統計 I(2024)L07

ここまで来たよ

14 統計的仮説検定・母平均値の両側 t 検定

15 母比率の片側検定・p 値・適合度のカイ二乗検定
- 母比率の検定=二項検定 (の正規近似)
- p 値=有意確率
- 仮説検定での第 1 種・第 2 種の過誤と混同行列
- 適合度のカイ二乗検定

カテゴリ変数

量的変数の確率変数 . . . 平均や分散に意味がある

- 離散型 確率関数=表 二項分布, ポアソン分布, $\cdots$, x は整数
- 連続型 確率密度関数 正規分布, カイ二乗分布, 指数分布，一様分布, $\cdots$, x は実数

質的変数の確率変数 . . . 平均や分散が考えられない
その中でも, 名義変数=カテゴリ (カル) 変数
順序や距離がなくぜんぶが対等. 例: 血液型, 性別, 携帯電話番号, チーム A 型, B 型などがカテゴリ

2 カテゴリなら, 0,1 のように番号を振って離散型と思える
3 カテゴリ以上なら, 順序や間隔によるので離散型には帰着できない.
なぜなら 自分の言葉でどうぞ あえてやるなら one-hot vector.

カテゴリ分布と多項分布 (multinomial distribution)

カテゴリに番号 $i = 1, \ldots, k$ を振る．

定義 (k カテゴリのカテゴリ分布 $\mathrm{C}(\boldsymbol{p})$)

$i = 1, \ldots, k$ に対して，

$$\text{確率関数} p(i) = P(X = i) = p_i = \begin{cases} p_1 & (i = 1) \\ p_2 & (i = 2) \\ \vdots & \\ p_k & (i = k) \end{cases}$$

$p_1 + \cdots + p_k = 1$. 略記 $\boldsymbol{p} = (p_1, \ldots, p_k)$

ベルヌイ分布 $\mathrm{B}(1, p_1)$ は $k = 2$ の $\mathrm{C}((1 - p, p))$.

定義 (k カテゴリの多項分布 $\mathrm{M}(\boldsymbol{p}, n)$)

確率変数は整数値ベクトル $(X_1, \ldots, X_k)$. パラメタ $\boldsymbol{p}, n$.

$$\text{確率関数} P(X_1 = x_1, \ldots, X_k = x_k) = f(x_1, \ldots, x_k) = \begin{cases} \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k} & (x_i \geq 0, x_1 + \cdots + x_k = n) \\ 0 & (\text{他}) \end{cases}$$

カテゴリ分布にしたがう確率変数の n 個の実現値を得て，カテゴリ $1, \ldots, k$ をそれぞれ $x_1, \ldots, x_k$ 回ずつ得る確率．
二項分布 $\mathrm{B}(n, p)$ は，$k = 2$ の，$\mathrm{M}((1 - p, p), n)$.

標本に対するピアソンの適合度基準 χ^2

カテゴリの個数 $k = 4$ の例

カテゴリ	O 型	A 型	AB 型	B 型
確率 p_i	$p_1 = 0.12$	$p_2 = 0.51$	$p_3 = 0.17$	$p_4 = 0.20$

$$\sum_{i=1}^{k} p_i = 1.$$

カテゴリ分布のサイズ $n = 12$ の標本

番号	血液型
1	B 型
2	O 型
⋮	⋮
12	A 型

$\to$ 度数分布表は多項分布 $\mathrm{M}(\boldsymbol{p}, 12)$ の実現値のはず

カテゴリ	O 型	A 型	AB 型	B 型
度数 x_i	$x_1 = 2$	$x_2 = 3$	$x_3 = 6$	$x_4 = 1$

$$\sum_{i=1}^{k} x_i = n = 12.$$

適合度を表す量

カテゴリ i の期待度数=標本サイズ × 母分布の確率=np_i

定義

ピアソンの適合度基準 χ^2 カテゴリ k 個, カテゴリ分布の確率 $p_i(i = 1, \ldots, k)$
サイズ n の標本の度数 $x_i(i = 1, \ldots, k)$
サイズ n の標本の, 母分布から考えた期待度数 $np_i(i = 1, \ldots, k)$

$$\chi^2 = \frac{(\text{度数} - \text{期待度数})^2}{\text{期待度数}} \text{の合計} = \sum_{i=1}^{k} \frac{(x_i - np_i)^2}{np_i}$$

が小さいほど, 標本は母分布から予想されるものに近い

命題

n 大のとき，ピアソンの適合度基準 χ^2 は，自由度 $k-1$ のカイ二乗分布にしたがう．

このことを利用して，帰無仮説「$\boldsymbol{p}$ のカテゴリ分布にしたがう」対立仮説「$\boldsymbol{p}$ のカテゴリ分布にしたがわない ($\Leftrightarrow p_i$ が 1 個は異なる」で検定が行える 適合度のカイ二乗検定

ふつうじゃない Trial L14 非参照テストではなく, 2024-08-05 月 20:00 までにアップロードするレポート.

- 練習問題 L14-01, L14-02 に合格すると, 個人別レポート問題の取得や提出が可能になる.
- 趣旨: 検定の長い文を (参照ありで) 書けることを確かめる必要. 非参照や短時間のテストで確認することが難しい.

ふつうじゃない Trial L15 非参照テストではなく, 2024-08-05 月 09:15 までに解答する Moodle の問題.

- 練習問題 L15-01, L15-02 に合格した後，(受験回数，受験時間に制限はあるが練習問題に似た)TrialL15-01, 作文の Trial L15-02 に回答.
- 趣旨: 対面 Trial をする回がない

定期試験期間中の任意参加授業内小テスト用途:Moodle 成績評価方法参照.
2024-08-05 月 09:15-10:15, 参照なし．7-002．事前の出欠連絡不要．理由を問わず, 欠席したときの追試や配慮はなし．Trial L01-08 の出題計画をあわせたもの．要学生証, 個人別座席指定.
すべて持込不可, ただし，岩薩林 確率・統計 pp. xii,xiii の公式集を印刷して当日配布
災害や交通機関などにより全学的に定期試験が中止・延期になるときは同様に中止・延期.
成績評価考慮対象になるかもしれない欠席の届は 2024-07-31 月 20:00 までに提出．Trial L13 までの採点結果は 2024-07-31 ごろまでに公表します．Trial L14, L15 の採点結果の公表は 2024-08-05 には間に合いません.