母比率の片側/両側検定,p 値, 有意水準と検定力

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 I L13(2020-01-06 Mon)

最終更新: Time-stamp: "2020-01-18 Sat 09:43 JST hig"

今日の目標

- 母比率の片側/両側正規検定ができる
- 第 1 種, 第 2 種の過誤, 有意水準, 検出力の関係が説明できる

L11-Q04 L12-Q1
Quiz 解答:母平均値の区間推定 (母分散未知)

1. 重さの標本平均値は $m = 50\text{g}$. 不偏標本分散は $s^2 = \frac{1}{4-1} \cdot 14\text{g}^2$. t 分布表から 自由度 $k = n - 1 = 3$ の $t_3(0.05/2)$ を参照して, 信頼係数 0.95 の信頼区間は

$$50 - 3.182 \times \sqrt{\frac{1}{4}\frac{14}{3}} < \mu < 50 + 3.182 \times \sqrt{\frac{1}{4}\frac{14}{3}}.$$

2. 同様に, t 分布表から 自由度 $k = n - 1 = 3$ の $t_3(0.01/2)$ を参照して,

$$50 - 5.841 \times \sqrt{\frac{1}{4}\frac{14}{3}} < \mu < 50 + 5.841 \times \sqrt{\frac{1}{4}\frac{14}{3}}.$$

L12-Q2
Quiz 解答:母平均値の検定 (母分散未知)=t 検定

1. 有意水準 0.05 で, 正規分布の母平均値に対する 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 $\mu_0 = 57\mathrm{g}$ に等しい」とする.
3. サイズ $n = 5$ 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, 自由度 $5 - 1$ の t 分布に従う.
4. この標本の実現値は, $t = \frac{\bar{x} - \mu_0}{\sqrt{\frac{s^2}{n}}} = \frac{51-57}{\sqrt{\frac{1}{5}\frac{16}{5-1}}} = -3\sqrt{5} = -6.708.$
5. t 分布表より, 棄却域は $t_4(0.05/2) = 2.776 < |t|$.
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 $57\mathrm{g}$ 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 $57\mathrm{g}$ 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)

L12-Q3

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 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」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

$$T = \frac{\bar{X} - 196}{\sqrt{s^2/n}}$$

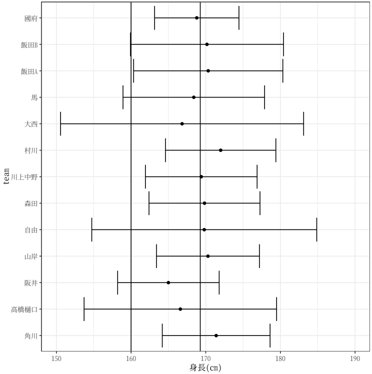
   は, 自由度 $4 - 1$ の t 分布に従う.
4. この標本の実現値は $\bar{X} = 200, s^2 = \frac{224}{4-1} = 74.7$. よって, $t = \frac{200-196}{\sqrt{\frac{1}{4}\frac{224}{3}}} = 0.92582$.
5. t 分布表より, 棄却域は $t_3(0.05/2) = 3.182 < |t|$
6. この標本の実現値は棄却域に含まれないので, 帰無仮説は棄却できない. 来店客数が変化したとは結論できない.
   (注: 結果は有意でなかった, 母平均値 μ と 196g の間には有意差がない, など).

標本抽出 → 区間推定の結果

母集団=回答者全体, 母集団サイズ=75
身長の母平均値=169.3cm, 身長の母分散=72.9cm^2, 滋賀県高校の母比率=0.31.
標本平均値 $\overline{X}$, 不偏標本分散 S^2, 標本比率 p, 標本サイズ N

team	$\overline{X}$(cm)	S^2(cm^2)	p	N
國府	168.80	20.700	0.20000	5
飯田 B	170.17	95.367	0.16667	6
飯田 A	170.33	90.667	0.33333	6
馬	168.40	58.300	0.40000	5
大西	166.83	240.567	0.50000	6
村川	172.00	49.600	0.50000	6
川上中野	169.40	36.300	0.20000	5
森田	169.83	50.167	0.33333	6
自由	169.80	147.200	0.20000	5
山岸	170.30	92.900	0.40000	10
阪井	165.00	30.000	0.20000	5
高橋樋口	166.60	107.800	0.20000	5
角川	171.40	33.800	0.20000	5

信頼係数 $1-\alpha = 0.95$ 信頼区間



ここまで来たよ

10 統計的仮説検定

11 母比率の片側/両側検定,p 値, 有意水準と検定力
- 統計的仮説検定の考え方
- 母比率の検定 (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

レポートや論文での検定の書き方

母集団を決める. 母集団の分布タイプを仮定する.

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」
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」

検定の例え話．有意水準とは?

検定とは，見逃すことはよくあるけど，発色 (検出) したら (まれな誤検出を除いて) ほぼ確かな異常検査薬のようなもの．
誤検出率$=\alpha$，ほぼ確か率 $= 1 - \alpha$.
「A 検定」はたくさんある
標本 $X \overset{\text{A 検定}}{\mapsto}$ 検定統計量 $y_A(X)$ の実現値．$y_A = T, P, \ldots$.

- 帰無仮説 (=検査の正常値) の設定
- $y_A(X)$ の実現値が境い目を越えて大きすぎたり小さすぎたりしたら (検定統計量の実現値が棄却域にはいったら) ‘発色’
- その境い目は，有意水準 α を指定して表から決める．$\alpha = 0.01, 0.05$ と小さく取り，誤検出は存在しないかのような態度をとる．

みんな性能のよい (α, β の小さい) y_A(A 検定) を本から探したり，自分で作ったりしてる．

ここまで来たよ

10 統計的仮説検定

11 母比率の片側/両側検定,p 値, 有意水準と検定力

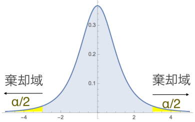
- 統計的仮説検定の考え方
- 母比率の検定 (二項検定の正規近似)
- p 値=有意確率
- 有意水準と検出力, 第 1 種の過誤, 第 2 種の過誤

母比率の検定 (二項検定の正規近似) 岩薩林 確率・統計定理 7.6(p.171)

岩薩林 確率・統計定理 7.6(p.171)

- 帰無仮説 母比率 $p = p_0$, 対立仮説 $p \neq p_0$.
- 検定統計量 $Z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$ は標準正規分布 $\mathrm{N}(0, 1^2)$ にしたがう.
  ここで $\hat{p}$ は標本比率.
- 棄却域 $|z| > t_\infty(\alpha/2) = z(\alpha/2)$.

標本サイズ n が大きく，二項分布が正規分布で近似できるとき.



L13-Q1

TA Prob and Sol:母比率の両側二項検定の正規近似

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 $p = 0.4$ 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．$p = 0.4$ でないことを結論できるか?

略解

1. 有意水準 $\alpha = 0.05$ で，母比率の両側検定を行う．
2. 帰無仮説を「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 $p = 0.4$」，対立仮説を「$p \neq 0.4$」とする．
3. サイズ $n = 68$ の標本の標本比率を $\hat{p}$ とすると，検定統計量

$$Z = \frac{\hat{p} - 0.4}{\sqrt{0.4(1 - 0.4)/68}}$$

は，標準正規分布に近似的にしたがう．

4. この標本に対して, $\hat{p} = 20/68 = 0.2941$ より, $z = -1.782$.
5. 標準正規分布表より境い目の値は $z(0.05/2) = 1.960$. (または $t_\infty(0.05/2)$) 棄却域は $|z| > 1.960$.
6. $1.960 > |-1.782|$ なので, z は棄却域に含まれない. 帰無仮説は棄却できない. 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 $p = 0.4$ でない, とは結論できない.

コメント

日本語の語尾「である」,「でない」だけで帰無仮説を定められるわけではない. 帰無仮説は「$p \neq 0.4$ でない」, 最後の文は「$p \neq 0.4$ とは結論できない」とも書ける.

日本語の語尾「である」,「でない」だけで帰無仮説を定められるわけではない. 帰無仮説は「$p \neq 0.4$ である」, 最後の文は「$p \neq 0.4$ とは結論できない」とも書ける.

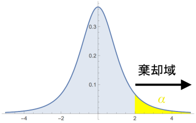
母比率の片側検定 (二項検定の正規近似) 岩薩林 確率・統計定理 7.7(p.159)

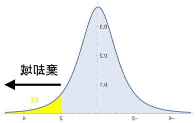
さっきのは不自然な問題設定．ふつうは $p \neq 0.4$ でなく $p < 0.4$ と言いたいでしょ．そういうときは，帰無仮説は同じ，対立仮説を変更し，両側検定のかわりに片側検定をする．

母比率の片側検定 岩薩林 確率・統計定理 7.7(p.159)

- 帰無仮説 母比率 $p = p_0$, 対立仮説 $p > p_0$ (または $p < p_0$)
- 検定統計量 $Z = \dfrac{\overline{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$ は標準正規分布 $\mathrm{N}(0, 1^2)$ にしたがう．ここで $\hat{p}$ は標本比率．
- 棄却域 $z > t_\infty(\alpha) = z(\alpha)$. (または $z < -z(\alpha)$).

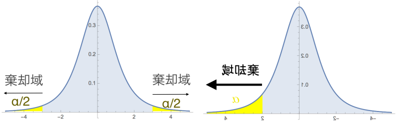
標本サイズ n が大きく，二項分布が正規分布で近似できるとき．

$z > z(\alpha)$ 

$z < -z(\alpha)$ 

片側検定と両側検定 岩薩林 確率・統計 p.150

片側検定では，どちらか片側だけに確率=面積 α の棄却域ができる．実現値がここにはいったら帰無仮説を棄却．
両側検定のときは，両側に確率=面積 $\alpha/2$ ずつの棄却域ができる．実現値がいずれかにはいったら帰無仮説を棄却．



t 検定 確率統計☆演習 I(2019)L12 にも両側検定 (やった) と片側検定がある 岩薩林 確率・統計定理 7.2p.159 .

- 片側 t 検定の対立仮説 $\mu < \mu_0$ (または $\mu > \mu_0$)
- 両側 t 検定の対立仮説 $\mu \neq \mu_0$

L13-Q2

Quiz(母比率の片側二項検定の正規近似)

「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 $p < 0.4$ である」 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. 片側二項検定の正規近似で, $p < 0.4$ を結論することを試みよう.

岩薩林 確率・統計例題 7.7(p.172), 問 8(p.172), 第 7 章演習問題 3

ここまで来たよ

10 統計的仮説検定

11 母比率の片側/両側検定,p 値, 有意水準と検定力

- 統計的仮説検定の考え方
- 母比率の検定 (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，Z や T が極端かどうか判定している

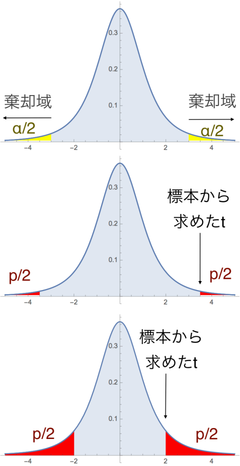
勝負			棄却する	しない	
統計量の値	片側	$-z(\alpha)$	$>$	$<$	$T=\frac{\overline{X}-\mu_0}{\sqrt{S^2/n}}$
		$+z(\alpha)$	$<$	$>$	
	両側	$+z(\alpha/2)$	$<$	$>$	$\lvert T\rvert=\left\lvert\frac{\overline{X}-\mu_0}{\sqrt{S^2/n}}\right\rvert$
	付表 1 下, 付表 2↑				付表 1 上, Excel↓
確率=面積	常時	有意水準 α	$>$	$<$	p 値

標本の p 値 (p-value)=有意確率 岩薩林 確率・統計 p.52

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

p 値 (t 分布の例)



標本の p 値 (p-value)

帰無仮説のもとで，検定統計量がこの標本よりも

$\boxed{}$ 確率.

p 値 $<$ 有意水準α のときに 帰無仮説を棄却する.

帰無仮説棄却

帰無仮説採用

座標の比較 $t^* < T$ (T は定数 t^* のどちら側?)

$\Leftrightarrow$ 右側の面積の比較 $\frac{\alpha}{2} > \frac{p}{2}$ ($\frac{p}{2}$ は定数 $\frac{\alpha}{2}$ より大?小?)

ここまで来たよ

検定の例え話. 有意水準とは? 岩薩林 確率・統計 §6.3

検定とは, 見逃すことはあるけど, 発色 (検出) したら (まれな誤検出を除いて) ほぼ確かな異常検査薬のようなもの.
見逃し率: β. 検出力:$1-\beta$.
誤検出率, 有意水準, 危険率: α.

インフルエンザ	検査陽性 (発色)	検査陰性 (発色なし)
かかってる	○ (真陽性)	はずれ (見逃し, 偽陰性)
かかってない	× (誤検出, 偽陽性)	あたり (真陰性)

	有意, 帰無仮説を棄却	陰性, 有意でない, 帰無仮説を棄却できない
正常でない, 対立仮説が成立 $\mu \neq \mu_0$	確率 $1-\alpha$	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 $1-\beta$

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 をなるべく大きくするように試みる習慣.

- 有意水準 $\alpha = \frac{\text{偽陽性}}{\text{偽陽性}+\text{真陰性}} = \frac{\text{正常}}{\text{検査薬発色}}$. 小さいほど, よい, というか発色したら間違いない検査薬.
- 検出力 $1-\beta$, $\beta = \frac{\text{偽陰性}}{\text{偽陰性}+\text{真陽性}} = \frac{\text{正常でない}}{\text{検査薬発色しない}}$. 小さいほど, よい, というか発色しなかったら間違いない検査薬. 敏感な検査薬.

$\rightsquigarrow$ 発色=正常でない, 発色せず=正常とも正常でないともわからない

有意水準と検出力

注: 二項分布，比率の検定の話をするので，母比率 p という記号を使いたいが，今日は p 値 p と紛らわしいので，r と書きます.

統計的検定

あるくじ付きお菓子は，工場で，$r_0 = 0.03$ の確率で独立に当たりを混ぜることになっている.
工場の当たりくじ混ぜ込みマシンが異常でないか調べたい.

- 対立仮説 H_1 実際の当たり確率 $r \neq r_0$
- 帰無仮説 H_0 実際の当たり確率 $r = r_0$
- 提案するマイ二項検定: 100 個からなる標本のうちの当たりくじの個数 X を検定統計量とする. 下側の境い目 $X^{**} = 0$, 上側の境い目 $X^{*} = 5$, つまり，当たりが $X = 0, 5, 6, \ldots, 100$ 個という極端な値であるときには帰無仮説を棄却する「マイ二項検定」を使ってみよう.

L13-Q3

マイ二項検定

実際の当たり確率が $r = 0.03$ であるときに，マイ二項検定で，帰無仮説を間違えて棄却してしまう確率 α を求めよう．

$$\begin{aligned}\alpha =& P(X = 0) + P(X = 5) + \cdots + P(X = 100)\\ =& 1 - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4)\\ =& 1 - {}_{100}\mathrm{C}_1 0.03^1(1 - 0.03)^{99} - \cdots - {}_{100}\mathrm{C}_4 0.03^4(1 - 0.03)^{96}\\ =& 0.230.\end{aligned}$$

このような誤りを $\boxed{}$，誤りの起こる確率 α を 検定の $\boxed{}$ という．

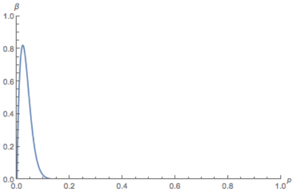
L13-Q4

マイ二項検定

実際の当たり確率が $r(\neq 0.03)$ であるときに，マイ二項検定で，帰無仮説を棄却できない確率 β を求めよう．

Solution:

$$\begin{aligned}\beta(r) =& P(X=1)+P(X=2)+P(X=3)+P(X=4)\\ =& {}_{100}\mathrm{C}_1 r^1(1-r)^{99}+\cdots+{}_{100}\mathrm{C}_4 r^4(1-r)^{96}.\end{aligned}$$



このような誤りを $\boxed{\qquad\qquad}$，誤りの起こらない確率 $1-\beta$ を検定の $\boxed{\qquad\qquad}$ という．

母比率の検定

マイ二項検定では, $(X=0)$, $(X=5,6,\cdots,100)$ を「極端扱い」した.
世の中の検定では, 先に実現すべき α が指定されており, それにあわせて, β がなるべく小さくなるように棄却域 (両側の境い目 X^{**}, X^{*}) を決める.
(それが二項検定. そしてさらに二項分布を正規分布で近似すると, 母比率の検定になる)

ネイマン-ピアソンの補題

連絡

- 2020-01-20 月 1 テスト B1. 45 分, PC 受験の予定. 持込不可. 教科書の p. xii–xv のコピーを配布. 数表も必要なら配布.
- 2019-01-20 の Trial ありません (L12 の内容は予習復習問題だけでいきなりテスト B1 に).
- 2020-01-27 月 1(定期試験期間) テスト A2,B2 実施. A1,B1 と同等の評価内容.
- 大学院進学のための学力認定試験: 2020-02-15 土. 2 年生も受験可能. 申込は樋口に. `https://wiki.math.ryukoku.ac.jp` (全学認証)
- オフィスアワー木 6(1-539) 金昼 (1-542), Math ラウンジ (1-536/538)

学籍番号での LINE 公式アカウント登録

GeoGebra 確率電卓

`https://www.geogebra.org/classic#probability`

テスト B1 計画

2020-01-20 月. 1-609 実習室. 35 ピーナッツ相当.
予習復習問題のような PC(Moodle) による回答の予定です.
採点結果は 1 日以内程度でお知らせします. 期末試験 (A2,B2) は一方または両方に参加できます. 両方とも参加しなくても合格はありえます. 成績の 70 ピーナッツ分$=\min(\max(A1, A2), \max(B1, B2))$ に従って計算するだけです.
受験準備は, 過去問題 (公開はしています) でなく, Trial, 予習復習問題を中心にすることをお勧めします. 教科書の p. xii–xv のコピーを配布するので, これにあわせて頭の中をまとめておくといいかも. 数表も必要なら配布します.
出題計画

- 正規分布 $\mathrm{N}(\mu, \sigma^2)$ にしたがう確率変数が, ある条件を満たす確率を求める (Trial L09-1,L10-1)
- 二項分布を利用して, コインに関わる確率, 母平均値, 母分散などを求める (Trial L10-2)
- 中心極限定理の内容や, n 個の独立な確率変数の和の母平均値, 母分散についての問 (L09 Quiz2, Quiz3)
- 二項分布の確率を, 中心極限定理で正規分布で近似して計算する (Trial L12-1)
- 独立同分布に従う確率変数の母期待値, 母分散の問題 (L09-Q2)
- 標本から母平均値・母比率を点推定・区間推定する (Trial L11-1, L12-2, L13-1)
- 標本から母平均値の両側 t 検定を行う (Trial L13-2)
- 標本から母比率の片側二項検定 (の正規近似) を行う (L13-Q2, Trial なし)
- 母集団と標本抽出と推定と検定の意味 (p 値, 信頼係数, 有意水準) に関する選択肢的な問 (数個 L13, Trial なし)