統計的仮説検定

樋口さぶろお

龍谷大学理工学部数理情報学科

確率統計☆演習 I L12(2017-12-20 Wed)

最終更新: Time-stamp: "2017-12-25 Mon 17:31 JST hig"

今日の目標

- 統計的仮説検定の考え方が説明できる 西川確率統計 §7.1,§7.2,§7.3
- 母平均値の t 検定ができる 西川確率統計 §7.4.2
- 母比率の検定 (二項検定の正規近似) ができる 西川確率統計 §7.5.1

http://hig3.net

L11-Q1

Quiz 解答:母平均値の区間推定 (母分散既知)

1. 重さの標本平均値は $m = 50\text{g}$. よって, 信頼係数 0.95 信頼区間は
$$50 - 1.96 \times \sqrt{\frac{9}{4}} < \mu < 50 + 1.96 \times \sqrt{\frac{9}{4}}.$$
すなわち, $47.06 < \mu < 52.94$.
2. 同様に,
$$50 - 2.58 \times \sqrt{\frac{9}{4}} < \mu < 50 + 2.58 \times \sqrt{\frac{9}{4}}.$$
すなわち, $46.13 < \mu < 53.87$.

L11-Q2

L11-Q3

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 $m = 50\mathrm{g}$. 不偏標本分散は $s^2 = \frac{1}{4-1} \cdot 14\mathrm{g}^2$. 自由度 $k = n - 1 = 3$ の t 分布表を参照して, 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,

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

L11-Q4
L11-Q5
Quiz 解答:母平均値の区間推定 (母分散未知, 大標本)

1. 大標本なので, t 分布の自由度 ∞ の場合, すなわち標準正規分布で考えてよい. 信頼係数 0.95 の信頼区間は

$$51 - 1.96 \times \sqrt{\frac{4}{400}} < \mu < 51 + 1.96 \times \sqrt{\frac{4}{400}}.$$

2. 同様に,

$$51 - 2.58 \times \sqrt{\frac{4}{400}} < \mu < 51 + 2.58 \times \sqrt{\frac{4}{400}}.$$

L11-Q6

Quiz 解答:母比率の区間推定

A 候補に投票したを $X = 1$, しなかったを $X = 0$ とする.

1. 標本比率は $\hat{p} = \frac{35}{50} = 0.7$. 母比率 p を 0.7 と推定する.
2. X の母分散は $0.7 \times (1 - 0.7) = 0.21$ と推定する.
   母比率 p の信頼係数 $1 - \alpha = 0.95$ の信頼区間は,

$$0.7 - 1.96 \times \sqrt{\frac{1}{50} \cdot 0.21} < p < 0.7 + 1.96 \times \sqrt{\frac{1}{50} \cdot 0.21}$$
$$0.7 - 0.13 < p < 0.7 + 0.13$$
$$0.57 < p < 0.83$$

信頼係数 0.95 では当選ってことですね (放送用語「当選確実」で, 後であやまらなきゃいけない確率は 0.05 以下).

3. 母比率 p の信頼係数 0.99 の信頼区間は,

$$\begin{aligned} 0.7 - 2.58 \times \sqrt{0.0042} < &p < 0.7 + 2.58 \times \sqrt{0.0042} \\ 0.7 - 0.17 < &p < 0.7 + 0.17 \\ 0.53 < &p < 0.87 \end{aligned}$$

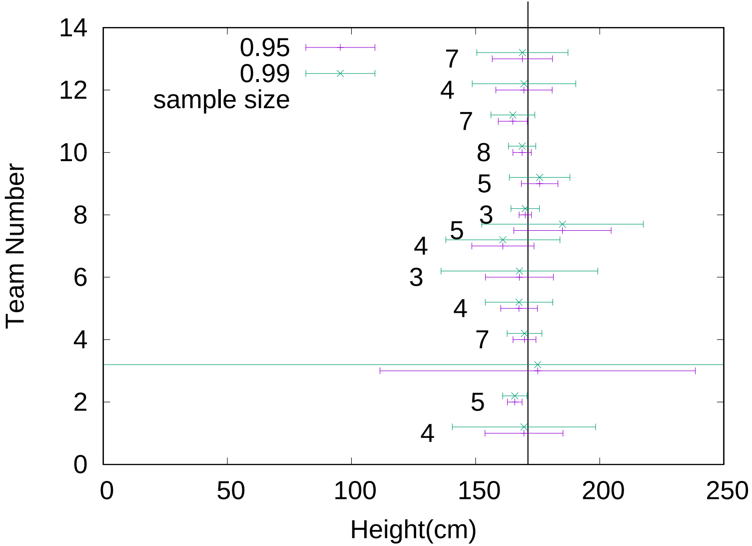
信頼係数 0.99 のほうが慎重な判断基準ですが, それでも当選ってことですね.

L11-Q7

抽出された標本のまとめ

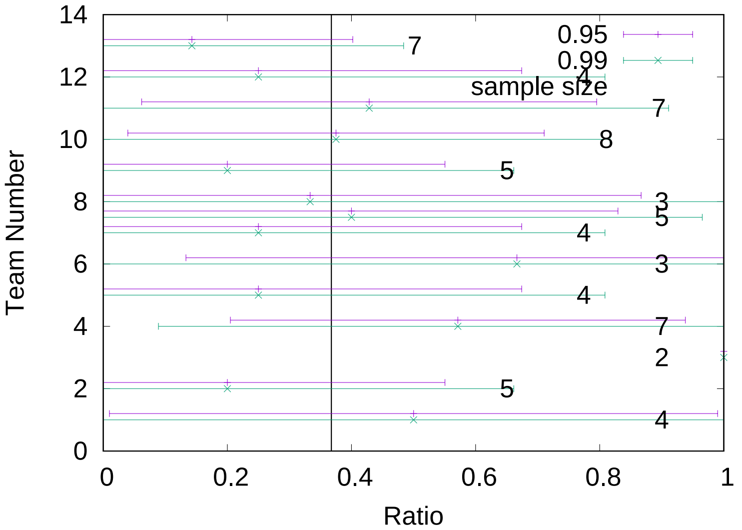
チーム	標本サイズ	滋賀県 $\sum_i Y_i$	標本平均値 $\overline{X}$(cm)	不偏標本分散 s^2 (cm^2)
1	4	2	169.5	97.7
2	5	1	165.8	5.7
3	2	2	175.0	50
4	7	4	169.7	24.9
5	4	1	167.5	21.7
6	3	2	167.7	30.3
7	4	1	161.0	62
7.5	5	2	185.0	250
8	3	1	170.0	1.0
9	5	1	175.8	35.2
10	8	3	168.8	19.6
11	7	3	165.0	39.7
12	4	1	169.5	51.0
13	7	1	168.9	171.8

母平均値の区間推定



注: 標本抽出は，「自分を含む」わけではない．母集団を類別するわけではない．

母比率の区間推定



注: 比率が 0.5 のとき, 信頼区間の幅は最大で $2 \times 1.96 \times 0.5/\sqrt{n}$.

ここまで来たよ

11 母平均値・母比率の区間推定

12 統計的仮説検定

- 統計的仮説検定の考え方
- 正規分布にしたがう母集団の母平均値の t 検定
- 母比率の検定 (二項検定の正規近似)

推定と検定 西川確率統計 §7.1

点推定 μ は値 xx と推定する

区間推定 μ は値 yy と値 zz の間と推定する (信頼係数 $1-\alpha$ で)

仮説検定 μ は値 xx と $\boxed{}$ する (‘確率的’ に=有意水準 α で) or あるかわからないと言う

あるドーナツ製造器は, 重さ X(確率変数) の母平均値が 55g であるように調整済みだという. しかし, 5 個買ってみたら, みんな軽めな感じ. これ, 本当に母平均値 55 g なの?(っていうか 55 g でないと言いたい).

ある学習法を使ってるある生徒の, 毎日のテストでの 1 か月の平均点は 63 点. 自分が別の学習法で教えた 5 日間の平均点は $\cdots$. 自分の方法は優れていると言いたい.

(例) 母比率の二項検定 西川確率統計 §7.1

瀬田生の滋賀県高校卒率は $\frac{\text{滋賀県の人口}}{\text{日本の人口}} = 0.01$ に等しい (帰無仮説), と信じてる A さんがいるので, それを論破したい (それより大きい=対立仮説). ある標本は, 10 人中 $X = 2$ 人だった. A さんの説が正しいなら, (これ以上に) まれなことが起きる確率は,

$$
\begin{aligned}
P(X \geq 2) =& P(X = 2) + P(X = 3) + \ldots + P(X = 10) \\
=& 1 - P(X = 0) - P(X = 1) \\
=& 1 - \frac{10!}{10!0!} \cdot 0.01^0 \cdot 0.99^{10} - \frac{10!}{9!1!} \cdot 0.01^1 \cdot 0.99^9 \\
=& 1 - 0.904 - 0.091 = 0.0042662.
\end{aligned}
$$

つまり, 確率 0.0042662 でしか起きない珍しい事象. あらかじめ決めておいた基準 (有意水準) $\alpha = 0.05$ より小さいので, $\boxed{\text{矛盾}}$.

この基準だと, $X = 1$ なら $\boxed{\text{矛盾}}$ にならない. 極端な値 $X = 2, 3, 4, \ldots$ で $\boxed{\text{矛盾}}$.

よって帰無仮説 $p = 0.01$ を棄却して, 対立仮説「滋賀県高校卒率 > 0.01」が $\boxed{\text{証明}}$ できた. 母比率の二項検定 西川確率統計 §7.5.1

なぜ統計的仮説検定?

心理学, 教育学, 社会科学などでは標本サイズが大きくできないことが多い. 標本サイズが小さくても Yes/No のいちおうの結論を出す, 科学業界で合意された方法が
検定 (test)=統計的仮説検定 (statistical hypothesis test)

真の母平均値は 55g と異なる, を $\boxed{\text{証明}}$ したい
しか〜し, $\neq$ の証明はやりにくい 54g である, ことが証明できれば十分だけど, 有限サイズの標本からはとうてい無理.

こういうときの常套手段は $\boxed{}$. 否定の命題「55g である」を仮定して $\boxed{\text{矛盾}}$ を導く.
注意
以下, 枠付きの $\boxed{\text{証明}}$, $\boxed{\text{矛盾}}$ は, この回の授業のローカル用語. 証明みたいなもの, 矛盾みたいなもの. 一定の確率で間違いがある.

帰無仮説と対立仮説

- H_0:帰無仮説 (null hypothesis) = 背理法の仮定 = 「真の母平均値 μ は 55g に等しい」
- H_1:対立仮説 (alternative hypothesis) = 示したい命題 = 「真の母平均値 μ は 55g でない」

上のは両側検定.
対立仮説が H_1: $\mu > 55$ という形の 片側検定 もある (最初の滋賀県高校率の例).

有意水準 significance level α

誤り (第 1 種の過誤) の確率をどれだけ許すか.
$\boxed{\text{証明}}$は確率 α で間違いを含む.
$\boxed{\text{矛盾}}$とは起きない事象 (確率 α の例外を除いて) が起きたこと.

標本で $\boxed{\text{矛盾}}$ が起きたかどうかの判定

まれな (確率 α 以下の) 事象が起きた
⇔ 検定統計量 Y を標本に対して計算したら, (確率 α 以下でしか起きない) 極端に大きな/小さな値をとった
⇔ 検定統計量 Y を標本に対して計算したら, (有意水準 α の) 棄却域に含まれる値になった

$\boxed{\text{矛盾}}$ が導かれるとき,

- H_0 を棄却 (reject) する
- H_1 を採択 (accept) する
- 標本が有意である (significant)

H_1 が $\boxed{\text{証明}}$ されたということ.

$\boxed{\text{矛盾}}$ が導けなかったとき,

- H_0 を棄却できない
- H_0 を採択する
- 標本が有意でない (not significant)

H_0 が $\boxed{\text{証明}}$ できたわけではない

ここまで来たよ

11 母平均値・母比率の区間推定

12 統計的仮説検定
- 統計的仮説検定の考え方
- 正規分布にしたがう母集団の母平均値の t 検定
- 母比率の検定 (二項検定の正規近似)

正規分布にしたがう母集団の母平均値に関する t 検定 I

L12-Q1

Quiz(母平均値の検定 (母分散未知)=t 検定)

あるドーナツ製造マシンが次々に製造するクロワッサンドーナツの重さ X_ig は, 正規分布にしたがうことがわかっている. 母平均値は 57g だと思っていたが, きょう 5 個製造したところ, 下のようだった.
$52\text{g}, 52\text{g}, 53\text{g}, 48\text{g}, 50\text{g}.$
本当にドーナツ製造マシンが次々に製造するクロワッサンドーナツの重さ X_ig の母平均値は 57g なのだろうか. 統計的仮説検定を行って判定しよう.

重さは負にならないし, 正規分布にしたがうというのはおかしな前提だが, ここは練習ってことで. 世の中には変な状況下で強引に t 検定を使う人が多くいるが, 数理の人はおかしさを認識できるように.

西川確率統計例題 7.3(p.155), 例題 7.4(p.156), 問題 7.3(p.157), 演習問題 7.1(p.162)

答案や論文での検定の書き方

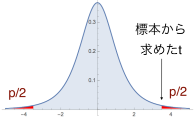
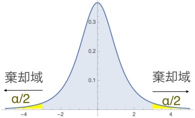
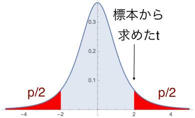
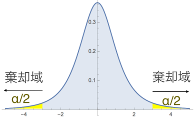
レポートもこれで.
母集団を決める. 母集団の分布タイプを仮定する.

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. 「帰無仮説のもとで検定統計量 Y は …分布にしたがう」
4. 「この標本に対して検定統計量 $y = \cdots$ である」
5. 「(y の不等式…) 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)」

検定統計量 Y この場合はこういう Y を取るとよい, というマニュアルができている. 取り方についた名前が「…検定」. たまにもっといいのを見つける人もいる.
最初のうちは, 参考書を見て, この状況ではこの検定統計量の…検定, という解法パターン的対処でいいでしょう. 不適切な検定を無理に使わないようにしよう.

不等式と棄却

p 値$=p=$ (y_1もっと極端な値を得る確率). 母比率 p とは別.

	帰無仮説を棄却	帰無仮説を棄却しない
	$\alpha > p$	$\alpha < p$
t 検定で	y^* より y_1 が極端 y が棄却域に含まれる $t_{\alpha/2}(n-1) < \|t\|$  	y^* より y_1 が極端でない $t_{\alpha/2}(n-1) > \|t\|$  

L12-Q2

Quiz(正規分布の母平均値に関する t 検定)

あるコンビニには，ドーナツ販売開始前には，9:00–10:00 に平均 196 人の客が来店していた．ドーナツ販売開始後の 4 日間，来店客数は次の通りだった．204, 208, 188, 200
来店者数は正規分布にしたがうと考える．ドーナツ販売開始後に来店客数の母平均値は変化したか?

L12-Q3

理工学部生の平均身長に関する統計的検定

日本の大学生の平均身長は 160cm であると耳にした (←教員の捏造)．理工学部生の平均身長は，これと異なるという仮説を立証したい．
理工学部生全体 (母集団) の身長が正規分布にしたがうとして，自分のチームのデータから，統計的仮説検定で立証を試みよう．

ここまで来たよ

11 母平均値・母比率の区間推定

12 統計的仮説検定
- 統計的仮説検定の考え方
- 正規分布にしたがう母集団の母平均値の t 検定
- 母比率の検定 (二項検定の正規近似)

母比率の検定 (二項検定の正規近似) 西川確率統計 §7.5.1

1. 有意水準 α で母比率の (二項) 検定の正規近似を行う
2. 帰無仮説を $p = \cdot$ とする
3. 帰無仮説のもとで検定統計量 $Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}}$ 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)
4. この標本⋯
5. 「(⋯) より帰無仮説を⋯」

L12-Q4

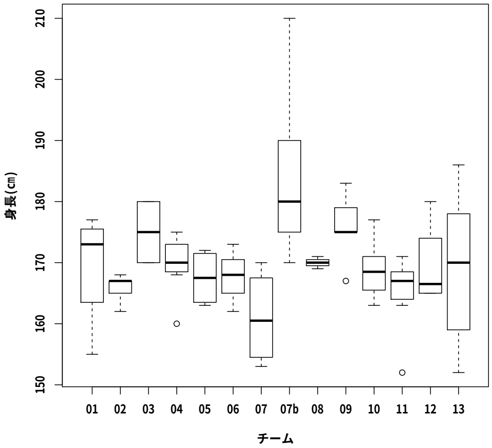
Quiz(母比率の二項検定の正規近似)

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 $p = 0.5$ 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, 25 名が滋賀県の高校を卒業していた. $p = 0.5$ でない, ことは結論できるか? (両側検定の問題にしたいので不自然な目的設定になっている, 普通は, $p > 0.5$ を示そうとして, 片側検定をするだろう).

連絡

- 予習復習問題は冬休み後の 2018-01-10 水 9:20 までです．
- t 検定のレポート．Learn Math Moodle で個人別問題を印刷して, 1–5 の全てのステップを記入．2018-01-10 水の授業, 10 水昼, 11 木昼, 15 月昼,16 火昼の Math ラウンジに提出．
- 次回は母分散の区間推定と検定とカイ二乗分布 西川確率統計 §6.4.3,§7.4.3,§8.3
- 配布資料は 1-503 向かいの引出, `http://hig3.net` で再配布．
- 加減乗除と平方根 (ルート) の使える電卓持ってきてね．関数電卓でなくてもいいです．携帯電話の機能・アプリでもかまいません．
- 樋口オフィスアワー月 3.5(1-539) 金 4(1-502), Math ラウンジ月-木昼 (1-614)

各チームの身長の分布



ここでは各標本の違いを表示したが, 箱ひげ図は, 本来は, 男子-女子, 体育会-それ以外, のような意味のある小集団 (層) の分布の違いを見るのに使う.