統計的仮説検定・母平均値の両側検定

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　L14(2023-07-10 Mon)

最終更新: Time-stamp: "2023-07-10 Mon 07:47 JST hig"

今日の目標

- 最尤推定，線形モデルの推定が説明できる
- 統計的仮説検定を説明できる
- 母平均値の両側 t 検定ができる　岩薩林 確率・統計 §7.1

L12-Q3

Quiz 解答: カイ二乗分布の確率と $\chi_k^2(\alpha)$

1. $F_Z^{-1}Z(1-0.025) = z(0.025) = 1.960.$
2. 標準正規分布の確率密度関数は偶関数なので,
   $F_Z^{-1}(0.025) = z(1-0.025) = -z(0.025) = -1.960.$
3. $F_W^{-1}(1-0.05) = \chi_1^2(0.05) = 3.841$. 別解. $0.05 = P(W > w_0) = P(Z^2 > w_0) = P(Z < -\sqrt{w_0}$ or $Z > +\sqrt{w_0}) = 2 \times P(Z > \sqrt{w_0})$. よって, $\sqrt{w_0} = 1.960.$
4. $F_W^{-1}(0.05) = \chi_1^2(1-0.05) = 0.00393.$

L13-Q1

Quiz 解答: 母分散の区間推定

標本サイズは $n=9$, 自由度は $9-1$, 母分散 σ^2 の信頼係数 $1-\alpha=0.95$ の信頼区間は,

$$
\begin{aligned}
\frac{n-1}{F_W^{-1}(1-\frac{\alpha}{2})} \times S^2 < &\sigma^2 < \frac{n-1}{F_W^{-1}(\frac{\alpha}{2})} \times S^2 \\
\frac{n-1}{\chi_{n-1}^2(\frac{\alpha}{2})} \times S^2 < &\sigma^2 < \frac{n-1}{\chi_{n-1}^2(1-\frac{\alpha}{2})} \times S^2 \\
\frac{8}{17.53} \times 72 < &\sigma^2 < \frac{8}{2.180} \times 72 \\
32.85 < &\sigma^2 < 264.2
\end{aligned}
$$

L13-Q2
Quiz 解答:t 分布の確率と $t_k(\alpha)$

1. $F_{\mathrm{Z}}^{-1}(1-0.025) = \mathsf{ppf}(1-0.025) = z(0.025) = 1.960.$
2. 標準正規分布の確率密度関数は偶関数なので,
   $F_{\mathrm{Z}}^{-1}(0.025) = \mathsf{ppf}(0.025) = z(1-0.025) = -z(0.025) = -1.960.$
3. $F_{\mathrm{t}}^{-1}(1-0.025) = \mathsf{ppf}(1-0.025) = t_{40}(0.025) = 2.021.$
4. t 分布の確率密度関数は偶関数なので, $F^{-1}F_{\mathrm{Z}}^{-1}(0.025)(0.025) =$ $\mathsf{ppf}(1-0.025) = t_{40}(1-0.025) = -t_{40}(0.025) = -2.021.$

L13-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 $\overline{X} = 50\mathrm{g}$. 不偏標本分散は $S^2 = \frac{1}{4-1} \cdot 14\mathrm{g}^2$. 信頼係数 0.95 の信頼区間は

$$
\begin{aligned}
50 - F_{\mathrm{t}}^{-1}(1-0.05/2) \times \sqrt{\frac{14/3}{4}} &< \mu < 50 - F_{\mathrm{t}}^{-1}(0.05/2) \times \sqrt{\frac{14/3}{4}} \\
50 - t_3(0.05/2) \times \sqrt{\frac{14/3}{4}} &< \mu < 50 + t_3(0.05/2) \times \sqrt{\frac{14/3}{4}} \\
50 - 3.182 \times \sqrt{\frac{14/3}{4}} &< \mu < 50 + 3.182 \times \sqrt{\frac{14/3}{4}} \\
46.563 &< \mu < 53.437
\end{aligned}
$$

2. 同様に,

$$
\begin{aligned}
50 - F_{\mathrm{t}}^{-1}(1-0.01/2) \times \sqrt{\frac{14/3}{4}} &< \mu < 50 - F_{\mathrm{t}}^{-1}(0.01/2) \times \sqrt{\frac{14/3}{4}} \\
50 - t_3(0.01/2) \times \sqrt{\frac{14/3}{4}} &< \mu < 50 + t_3(0.01/2) \times \sqrt{\frac{14/3}{4}} \\
50 - 5.841 \times \sqrt{\frac{14/3}{4}} &< \mu < 50 + 5.841 \times \sqrt{\frac{14/3}{4}} \\
43.691 &< \mu < 56.309
\end{aligned}
$$

ここまで来たよ

13 母分散・母平均値の区間推定
- 最尤推定
- 線形モデルとしての回帰分析

14 統計的仮説検定・母平均値の両側検定
- 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両側 t 検定

線形モデル (統計モデルのある一族)

あるドーナツ製造機の作るドーナツの重さ Y は次のモデルに従う.

$$Y = \beta_0 + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \sigma^2)$$

Y, ϵ: 連続型確率定数, $\sigma > 0, \beta_0$: 定数=パラメタ=母数

ϵ: 誤差, ノイズ. 小文字だけど確率変数.
(隠された) パラメタ母数 μ, σ^2 を, n 個のドーナツの重さのデータから推定したい.

$$\begin{aligned} \mathrm{E}[Y] =& \beta_0 + \mathrm{E}[\epsilon] = \beta_0, \\ \mathrm{V}[Y] =& \qquad \mathrm{V}[\epsilon] = \sigma^2. \end{aligned}$$

正規分布と限定した以外は, ここしばらくやってた, 母平均値, 母分散の推定の言い換えに過ぎない.
だけど, 多数のパラメタを含む一般的なモデルにも使える考え方をする.

尤度 likelihood

$\epsilon = Y - \beta_0 \sim \mathrm{N}(0, \sigma^2)$ より，ドーナツの重さ y を得る確率密度は，

$$f(y|\beta_0, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(y-\beta_0)^2}{2\sigma^2}}$$

サイズ n の標本が $y_1, \cdots, y_n$ である確率密度は，独立同分布なので積で，

$$\begin{aligned} f(y_1, y_2, \cdots, y_n|\beta_0, \sigma) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(y_i-\beta_0)^2}{2\sigma^2}} \\ &= \left(2\pi\sigma^2\right)^{-n/2} \mathrm{e}^{-\sum_{i=1}^{n} \frac{(y_i-\beta_0)^2}{2\sigma^2}} \end{aligned}$$

n 次元正規分布　　多変量解析及び演習

この f を，y_i の確率密度関数と思わず，が測定済データ $y_1, \ldots, y_n$ が定数，β_0, σ が変数と思ったとき，尤度 (ゆうど) 関数 $L(\beta_0, \sigma)$ という．

$$L(\beta_0, \sigma) = f(y_1, y_2, \cdots, y_n|\beta_0, \sigma)$$

最尤推定 岩薩林 確率・統計 なし

最尤推定

β_0, σ の推定値として $L(\beta_0, \sigma)$ が最大になる値を選ぶ

2 変数関数の最大値 ⇝ 偏微分 微積分 II

$$0 = \frac{\partial L}{\partial \beta_0}(\beta_0, \sigma) = \frac{\partial L}{\partial \sigma}(\beta_0, \sigma)$$

ここでは最初の等式だけ解く．合成微分．

$$0 = (\text{定数})^{-n/2} \times \sum_{i=1}^{n} \frac{1}{\sigma^2}(y_i - \beta_0) \times \mathrm{e}^{\text{同じ}}$$

$$0 = \sum_{i=1}^{n}(y_i - \beta_0)$$

$$\beta_0 = \frac{1}{n}\sum_{i=1}^{n} y_i$$

ここでは最初の式だけ．合成微分．β_0(母平均値) の最尤推定 (MLE=maximum likelihood estimation) 値 $\hat{\beta}_0$ は標本平均値じゃん

ここまで来たよ

(確率変数でない) 変数 x に依存する確率変数 Y

このドーナツ製造機の作るドーナツの重さ Y は, 温度 x によるらしい.
次の線形回帰モデルを仮定する.

$$Y = \beta_0 + \beta_1 \cdot x + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \sigma^2)$$

Y, ϵ: 連続型確率定数, β_0, β_1: 回帰係数, $\sigma > 0$: 定数=パラメタ=母数

Y: 目的変数 (従属変数) ここでは確率変数
x: 説明変数 (独立変数) ここでは確率変数でない

ノイズ・誤差 $\epsilon = Y - \beta_0 + \beta_1 \cdot x \sim \mathrm{N}(0, \sigma^2)$.
$\epsilon = Y - \beta_0 - \beta_1 x \sim \mathrm{N}(0, \sigma^2)$ より, ドーナツの重さ y を得る確率密度は,

$$f(y|x, \beta_0, \beta_1, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(y-\beta_0-\beta_1 \cdot x)^2}{2\sigma^2}}$$

(正確に) 指定した温度 x_i $(i=1,\ldots,n)$ で製造したときの重さが y_i $(i=1,\ldots,n)$ である確率密度は, 独立分布なので積.

$$
\begin{aligned}
f(y_1,y_2,\cdots,y_n|x_1,\ldots,x_n,\beta_0,\beta_1,\sigma) &= \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(y_i-\beta_0-\beta_1\cdot x)^2}{2\sigma^2}} \\
&= \left(2\pi\sigma^2\right)^{-n/2}\mathrm{e}^{-\sum_{i=1}^{n}\frac{(y_i-\beta_0-\beta_1\cdot x_i)^2}{2\sigma^2}} \\
&= L(\beta_0,\beta_1,\sigma)
\end{aligned}
$$

最尤推定

測定済のデータ x_i, y_i $i = 1, \cdots, n$ を定数と思ったときの, 3 変数関数 $L(\beta_0, \beta_1, \sigma) = f(y_1, y_2, \cdots, y_n | x_1, \ldots, x_n, \beta_0, \beta_1, \sigma)$ の最大値は? $\rightsquigarrow$ 偏微分 微積分 II

$$0 = \frac{\partial L}{\partial \beta_0}(\beta_0, \beta_1, \sigma) = \frac{\partial L}{\partial \beta_1}(\beta_0, \beta_1, \sigma) = \frac{\partial L}{\partial \sigma}(\beta_0, \beta_1, \sigma)$$

ここでは最初の 2 つの等式だけ解く.

$$0 = \frac{\partial L}{\partial \beta_0} = (\text{定数}) \sum_i \frac{1}{\sigma^2}(y_i - \beta_0 - \beta_1 x_i) \times \mathrm{e}^{\text{同じ}}$$
$$0 = \frac{\partial L}{\partial \beta_0} = (\text{定数}) \sum_i \frac{1}{\sigma^2} x_i(y_i - \beta_0 - \beta_1 x_i) \times \mathrm{e}^{\text{同じ}}$$

しょせん, β_0, β_1 の連立 1 次方程式

正規方程式 岩薩林 確率・統計 §9.2

$$n\beta_0 + (\sum_i x_i)\beta_1 = \sum_i y_i$$
$$(\sum_i x_i)\beta_0 + (\sum_i x_i^2)\beta_1 = \sum_i x_i y_i$$

加減法 $\rightsquigarrow$

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx}}, \qquad y = \frac{s_{xy}}{s_{xx}} x + \overline{y} - \frac{s_{xy}}{s_{xx}} \overline{x}$$

$$\hat{\beta}_0 = \overline{y} - \frac{s_{xy}}{s_{xx}} \overline{x} \qquad y - \overline{y} = \frac{s_{xy}}{s_{xx}} (x - \overline{x})$$

ここで,

$$\overline{x} = \frac{1}{n} \sum_i x_i$$ 平均値っぽい形

$$\overline{y} = \frac{1}{n} \sum_i y_i$$ 平均値っぽい形

$$s_{xy} = \frac{1}{n} \sum_i x_i y_i - \overline{x} \cdot \overline{y} \overset{\boxed{\text{岩薩林 確率・統計 定理 1.5}}}{=} \frac{1}{n} \sum_i (x_i - \overline{x})(y_i - \overline{y})$$ 共分散っぽ

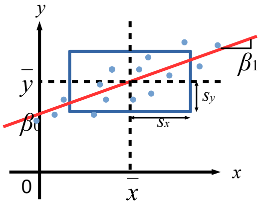
$$s_{xx} = \frac{1}{n} \sum_i x_i^2 - \overline{x}^2 \overset{\boxed{\text{岩薩林 確率・統計 定理 1.2}}}{=} \frac{1}{n} \sum_i (x_i - \overline{x})^2$$ 分散っぽい形

$Y = \hat{\beta}_0 + \hat{\beta}_1 \cdot x + \epsilon$ ($\hat{\beta}_0, \hat{\beta}_1$: 回帰係数) に代入.

回帰直線 岩薩林 確率・統計 §9.1

推定結果 $\hat{\beta}_0, \hat{\beta}_1$ を係数とする xy 平面の直線 データ分析

$$y - \overline{y} = \frac{s_{xy}}{s_{xx}}(x - \overline{x})$$



回帰係数, 予測値の信頼区間 確率統計 II,III

ここまで来たよ

13 母分散・母平均値の区間推定
- 最尤推定
- 線形モデルとしての回帰分析

14 統計的仮説検定・母平均値の両側検定
- 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両側 t 検定

推定 (estimation) と検定 (test)

点推定　μ は値 xx と推定 岩薩林 確率・統計 §6.1

区間推定　μ は値 yy と値 zz の間と推定 (信頼係数 $1-\alpha$ で) 岩薩林 確率・統計 §6.2

仮説検定　μ は値 xx と差がある と, (時々) 断言 (有意水準 α で) =見逃し多いけど発色したら正しい ($1-\alpha$ で) 血痕試験紙

岩薩林 確率・統計 §6.3

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, 違う感じ. 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⋯ 自分の方法は優れていると言いたい.

検定はだいたいこんな考え方 岩薩林 確率・統計 §6.3

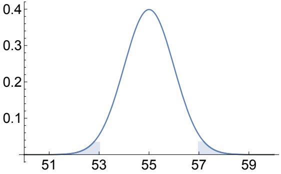
確率変数 X は, 正規分布 $\mathrm{N}(55, 2^2)$ にしたがうという. $\sigma^2 = 2^2$ は確かだとわかってるけど, $\mu = 55$ (帰無仮説) が本当なのか疑っている. というより，本当でないと言いたい
サイズ 4 の標本を抽出したところ,

$$54, 57, 57, 60$$

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)



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」

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

帰無仮説を棄却する reject
検定統計量の実現値が境目を越えて大きすぎたり小さすぎたりしたら (棄却域にはいったら) 帰無仮説 (=背理法の仮説) が偽, 対立仮説が真と結論する (試験紙が発色した)
帰無仮説を棄却しない accept
実験失敗 (背理法使おうとしたけど矛盾導けなかった). 何も言えない (発色しなかった)

ここまで来たよ

13 母分散・母平均値の区間推定
- 最尤推定
- 線形モデルとしての回帰分析

14 統計的仮説検定・母平均値の両側検定
- 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両側 t 検定

L14-Q1

Quiz(母平均値の両側検定 (母分散未知)=両側 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 Xg は, 正規分布にしたがうことがわかっている. 母平均値は 55g だと言われていたが疑っている. きょう 5 個製造したところ, 下のようだった.
50g, 50g, 51g, 46g, 48g.
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 Xg の母平均値が 55g と異なるかどうか, 有意水準 $\alpha = 0.05$ で統計的仮説検定を行おう.

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」

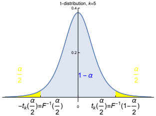
正規分布にしたがう母集団の母平均値の両側 t 検定

母平均値の両側 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 で母平均値の両側 t 検定を行う.
2. 帰無仮説を母平均値 $\mu = \mu_0$, 対立仮説を $\mu \neq \mu_0$ 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 $|t| > t_{n-1}(\frac{\alpha}{2})$ である.
6. (結論) 帰無仮説を棄却する/できない, 母平均値は $\mu \neq \mu_0$ と結論する/できない.

棄却域は $(-\infty, -u) \cup (u, +\infty)$. $P(T < -u \text{または} u < T) = \alpha$ であるように,
$u = t_{n-1}(\frac{\alpha}{2})$ と決めた.



t 分布では $F^{-1}(1 - \frac{\alpha}{2}) = -F^{-1}(\frac{\alpha}{2}) = t_k(\frac{\alpha}{2})$.

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 $\mu_0 = 55\mathrm{g}$ 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 $n = 5$ 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
$$T = \frac{\bar{X} - 55}{\sqrt{S^2/5}}$$
は, 帰無仮説のもとで, 自由度 $5 - 1$ の t 分布に従う.
4. この標本に対する検定統計量の実現値は,
$t = \frac{\overline{X} - \mu_0}{\sqrt{\frac{S^2}{n}}} = \frac{49 - 55}{\sqrt{\frac{1}{5}\frac{16}{5-1}}} = -3\sqrt{5} = -6.708.$
5. 棄却域は $|t| > t_4(0.05/2)$. rvt=stats.t(df=4) rvz.ppf(1-0.05/2) または t 分布表より $|t| > 2.776$.

6. $|-6.708| > 2.776$ であり, 実現値 t 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)

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.

岩薩林 確率・統計 例題 7.2(p.159), 問題 2(p.160)

L14-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? 有意水準 0.05 で考える．

2023-07-17 月祝は授業なし，練習問題締切あり．
ふつうじゃない Trial L14 非参照テストではなく, 2023-07-24 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.

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.
2023-07-31 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
成績評価考慮対象になるかもしれない欠席の届は 2023-07-24 月 20:00 までに提出．Trial L14, L15 の採点結果の公表は，2023-07-24 より後，2023-07-31 直前になってしまうかも．