確率と統計(O)
「仮説検定(第12章)」

- 担当教員: 杉山 将 (計算工学専攻)
- 居室: W8E-505
- 電子メール: sugi@cs.titech.ac.jp

- 授業のウェブサイト:

  http://sugiyama-www.cs.titech.ac.jp/~sugi/

講義計画(シラバス)

- 確率と統計の基礎
- 確率変数，確率分布
- 積率，積率母関数
- 離散型の確率分布の例
- 連続型の確率分布の例
- 確率不等式， 擬似乱数
- 多次元の確率分布
- 大数の法則，中心極限定理
- 統計的推定，仮説検定

仮説検定

- 仮説検定(hypothesis testing): 母集団について仮定された命題を，標本に基づいて検証すること

（例）コインを20回投げて表が17回出た.
この結果から表が出やすいといえるか？

- 帰無仮説(null hypothesis): もとの仮説

（例）コインは歪んでいない（表が出る確率 $p = 1/2$ ）

- 対立仮説(alternative hypothesis): 帰無仮説と対立する仮説

（例）表が出やすい $p > 1/2$

有意水準

- 有意水準(significance level) α :
  起こる確率がこの確率より小さければ，帰無仮説を棄却(reject)する．また，逆にこの確率以上であれば帰無仮説を採択(accept)する．
- 有意水準は5%か1%に設定することが多い．

仮説検定の例

- コインが歪んでいない($p = 1/2$)と仮定すれば，20回中17回以上表が出る確率は

$$({}_{20}C_{17} + {}_{20}C_{18} + {}_{20}C_{19} + {}_{20}C_{20}) \times (1/2)^{20} \approx 0.0013$$

- 有意水準を1%($\alpha = 0.01$)にとることにすれば，20回中17回以上表が出る確率は有意水準より小さいので帰無仮説は棄却される．
- 即ち，表が出やすい($p > 1/2$)と結論づける．

16回以上:0.0059, 15回以上:0.0207,
14回以上:0.0577, 13回以上:0.1316,…

棄却／採択と誤り

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">真実</td></tr>
<tr><td>帰無仮説が正しい</td><td>対立仮説が正しい</td></tr>
<tr><td rowspan="2">検定の結果</td><td>帰無仮説を採択</td><td>正解</td><td>第2種の誤り
(error of the second kind)</td></tr>
<tr><td>帰無仮説を棄却</td><td>第1種の誤り
(error of the first kind)</td><td>正解</td></tr>
</table>

- **第1種の誤り**

  歪んでいないコインを歪んでいると判断してしまう

- **第2種の誤り**

  歪んでいるコインを歪んでいないと判断してしまう

両側検定と片側検定

- 両側検定(two-sided test):

  得られた値が目標値と等しいかどうかを調べる

  （例）あるる装置の複製を作ったとき，もとの装置と同じ性能が得られるかを調べる

- 片側検定(one-sided test):

  得られた値が比較値よりも大きいかどうかを調べる

  （例）新しく開発した装置の性能が従来の装置よりもよいかどうかを調べる

帰無仮説の採択について

- 仮説検定で帰無仮説を棄却するときは，帰無仮説がほとんど起こらないことを証明している.
- しかし帰無仮説を採択するときは，積極的に帰無仮説が起こることを証明しているのではなく，帰無仮説が現実と矛盾することを証明するだけの十分な根拠がないと言っているだけである.
- このような論法は論理学では背理法(proof by contradiction)と呼ばれている.

例: 正規母集団の平均の検定

- 分散が既知の正規母集団に従う標本

$$X_i \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$$

から, 帰無仮説「母集団の平均は $\mu = 10$ である」を有意水準 α で検定する.

- 標本平均の分布を計算し, その出現確率を調べる:
  - 標本平均の分布は,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$$

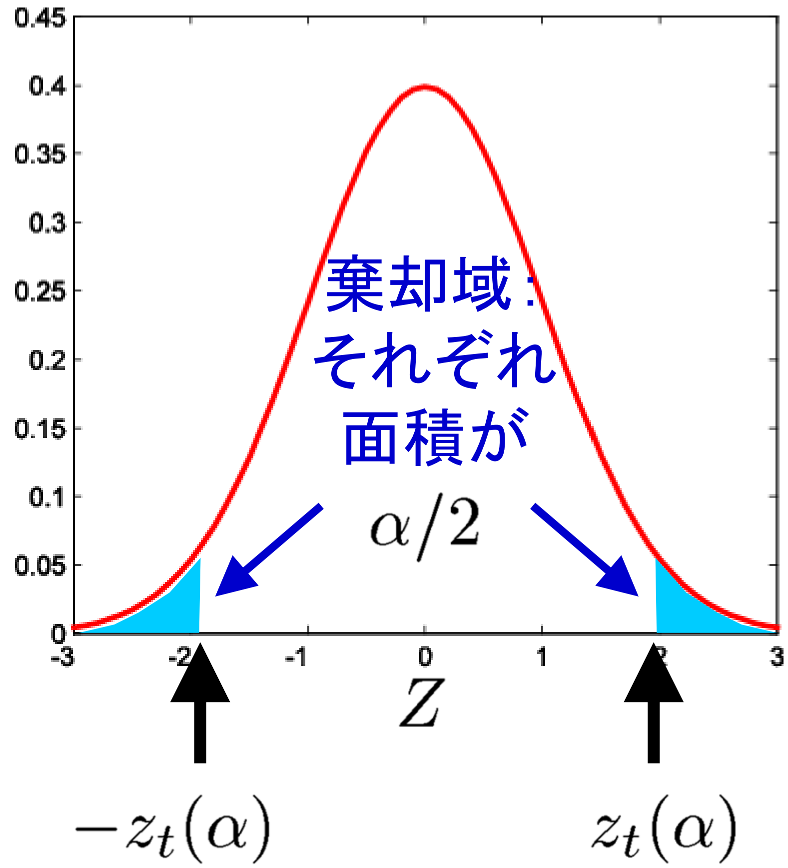
  - 適当に正規化すれば,

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

両側検定

$$Z = \frac{\overline{X} - 10}{\sigma/\sqrt{n}}$$

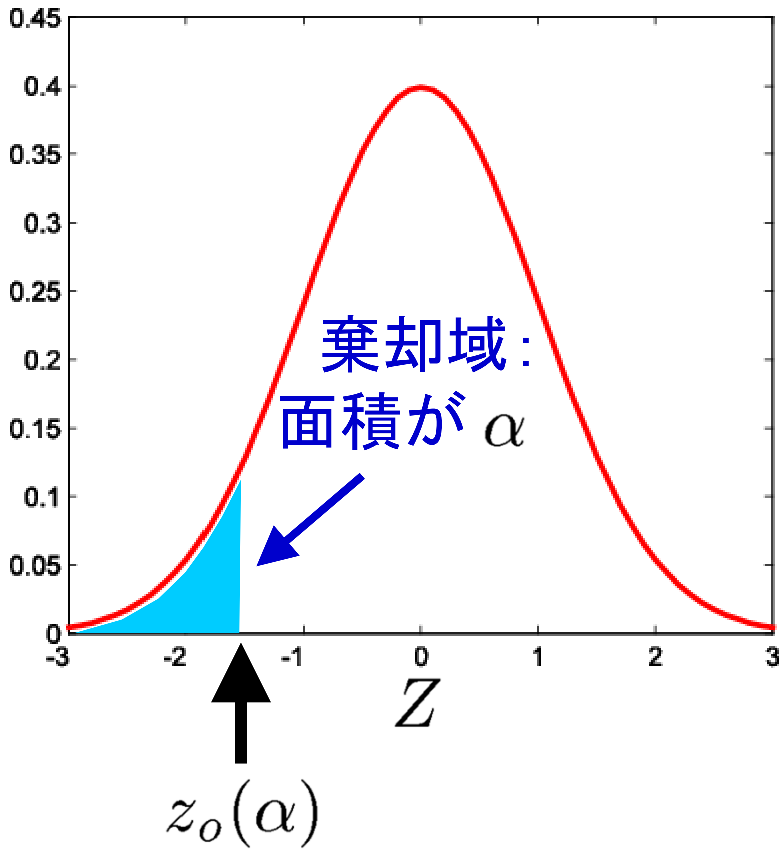
- 帰無仮説: $\mu = 10$
- 対立仮説: $\mu \neq 10$
- 棄却域(rejection region):
  有意水準 α に対して,
  確率分布の両端の面積が
  $\alpha/2$ の領域
- 現実の正規化標本平均 Z が棄却域に入ったら，帰無仮説を棄却する．即ち,
  $|Z| > z_t(\alpha)$ ならば $\mu \neq 10$ と判断する



片側検定

$$Z = \frac{\overline{X} - 10}{\sigma / \sqrt{n}}$$

- 帰無仮説: $\mu = 10$
- 対立仮説: $\mu < 10$
- 棄却域(rejection region):
  有意水準 α に対して,
  確率分布の左端の面積が
  α の領域
- 現実の正規化標本平均 Z が棄却域に入ったら，帰無仮説を棄却する．即ち，
  $Z < z_o(\alpha)$ ならば $\mu < 10$ と判断する



二標本検定

- 二標本検定(two-sample test):
  期待値がそれぞれ μ_X, μ_Y の二つの分布に従って取り出したi.i.d.標本

  $$\{X_i\}_{i=1}^{n_X}, \quad \{Y_i\}_{i=1}^{n_Y}$$

  から, 帰無仮説 $\mu_X = \mu_Y$ を検定する.

二標本検定の例

- 明日の天気を予測する問題を考える.
- いくつかのデータで, 従来法と新たに開発した手法の性能を検証してみると,
  - 従来法の平均正解率は78%
  - 新手法の平均正解率は80%

  であった.
- 新手法のほうが平均正解率が高いので, 従来法よりも優れていると単純に結論づけて良いか？
- 従来法との平均正解率の差が有意(significant)かどうかを調べたい.

正規母集団の母平均の差の検定

- 仮定:標本はそれぞれ正規分布に従う

$$X_i \overset{i.i.d.}{\sim} N(\mu_X, \sigma^2) \qquad Y_i \overset{i.i.d.}{\sim} N(\mu_Y, \sigma^2)$$

- このとき，標本平均もそれぞれ正規分布に従う.

$$\overline{X} \sim N(\mu_X, \frac{\sigma^2}{n_X}) \qquad \overline{Y} \sim N(\mu_Y, \frac{\sigma^2}{n_Y})$$

- 従って，標本平均の差も正規分布に従う.

$$\overline{X} - \overline{Y} \sim N(\mu_X - \mu_Y, \frac{\sigma^2}{n_X} + \frac{\sigma^2}{n_Y})$$

$$\overline{X} = \frac{1}{n_X}\sum_{i=1}^{n_X} X_i \qquad \overline{Y} = \frac{1}{n_Y}\sum_{i=1}^{n_Y} Y_i$$

正規母集団の母平均の差の検定（母分散が既知のとき）

- 標本平均の差を標準化すると，標準正規分布に従う.

$$\frac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{\sigma^2}{n_X}+\frac{\sigma^2}{n_Y}}} \sim N(0,1)$$

- 標準正規分布の棄却域は計算できるので，次の Z が棄却域に入るかどうかを調べればよい.

$$Z=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma^2}{n_X}+\frac{\sigma^2}{n_Y}}}$$

正規母集団の母平均の差の検定（母分散が未知のとき）

- 分散 σ^2 を標本から推定:

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^{n_X}(X_i - \overline{X})^2 + \sum_{i=1}^{n_Y}(Y_i - \overline{Y})^2}{n_X + n_Y - 2}$$

- $\widehat{Z} = \dfrac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\widehat{\sigma}^2}{n_X} + \dfrac{\widehat{\sigma}^2}{n_Y}}}$ は自由度 $\phi = n_X + n_Y - 2$ のt分布に従う.

- t分布の棄却域は計算できるので，$\widehat{Z}$ が棄却域に入るかどうかを調べればよい！
- これをt検定(t-test)と呼ぶ.

t分布

- t分布(t-distribution)：　$\phi = 1, 2, \ldots$

$$f(x) = \frac{1}{\sqrt{\phi} B(\frac{\phi}{2}, \frac{1}{2})} \left(1 + \frac{x^2}{\phi}\right)^{-\frac{\phi+1}{2}}$$

- X を標準正規分布に従う確率変数，Y を自由度 ϕ のカイ二乗分布に従う確率変数としたとき，

$$X/\sqrt{Y/\phi}$$

は自由度 ϕ のt分布に従う.

- 発見者のペンネームにちなんで，スチューデントのt分布(Student's t-distribution)と呼ぶこともある.

t分布の性質

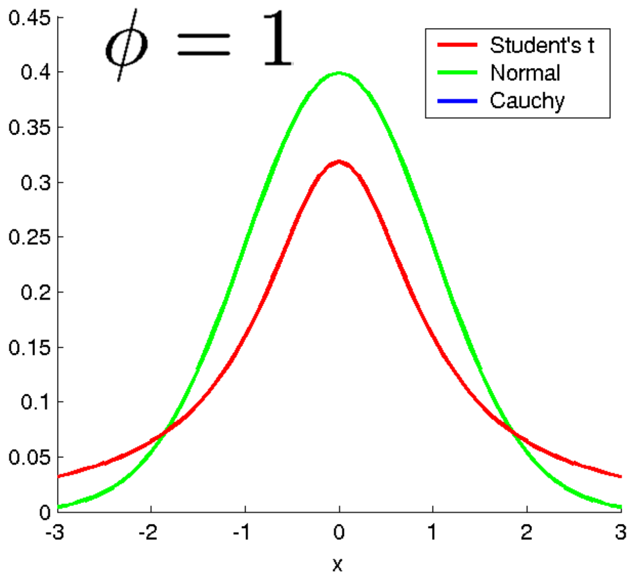
- 自由度 ϕ が1のとき, t分布はコーシー分布になる.
- 自由度 ϕ が無限大のとき, t分布は正規分布になる.
- 自由度 ϕ が2以上のとき, 期待値は

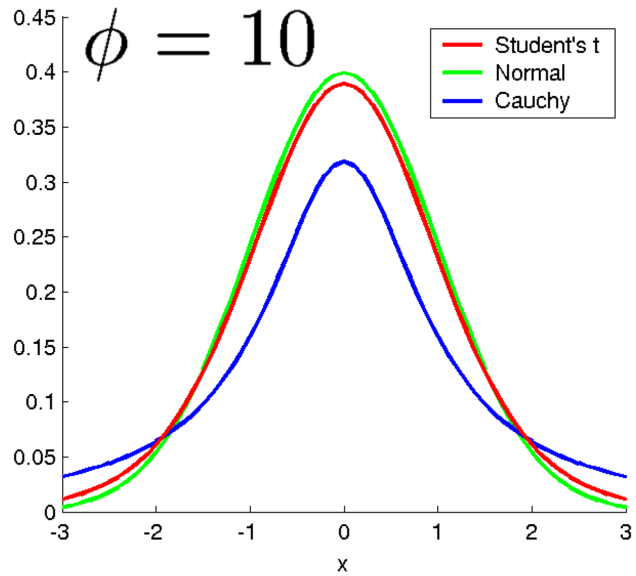
$$E(X) = 0$$

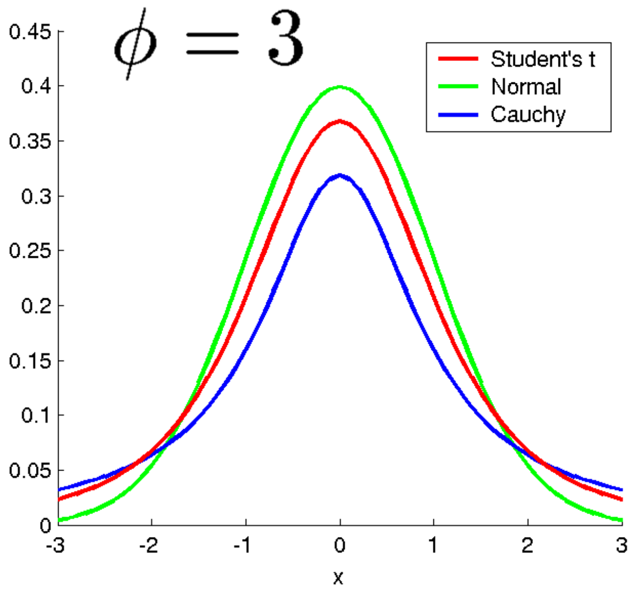
- 自由度 ϕ が3以上のとき, 分散は

$$V(X) = \frac{\phi}{\phi - 2}$$

t分布の例


$\phi = 1$
Student's t
Normal
Cauchy
x


$\phi = 10$
Student's t
Normal
Cauchy
x


$\phi = 3$
Student's t
Normal
Cauchy
x

まとめ

- 仮説検定
  - 帰無仮説と対立仮説
  - 有意水準
  - 両側検定と片側検定
- 二標本検定

試験について

- 日時:7月27日(金)3,4限
- 場所:H101
- 試験内容:
  - 専門用語の英語名を答えよ
  - 次の(a), (b), (c)からテーマを二つ選び, 自由に論ぜよ.

    (a)積率母関数, (b)中心極限定理

    (c)任意の分布に従う乱数を生成する方法

- 教科書, ノートは持ち込み不可！

今後の予定

- ７月２０日（金）
  - 試験の準備（各自）
- ７月２７日（金）
  - 試験（H101）

宿題

- 授業で扱っているトピック，授業の進め方，講義内容の難易度，講義資料，宿題などについて意見を述べ，この授業を評価せよ．
  - 分かりにくかったところがある場合は，スライドの何ページのどの部分がどのように分かりにくかったのか具体的に記述すること．
  - 悪いところはどうすれば改善するかも考えよ．
  - 良いところ，このまま続けるべきところも述べよ．

- この宿題は試験当日に集める．
- これまでの宿題の遅刻提出は試験当日まで