第 4 回：仮説検定
【教科書第 4 章第 4 節】

北村　友宏

2025 年 10 月 21 日

本日の内容

1. 仮説検定の考え方
2. 標本平均の分布
3. 母平均の検定の例題

仮説検定

- 母集団の特性に関する仮説を，標本を用いて検証することを仮説検定（hypothesis testing）という．
  - e.g.,「夏期講習実施前後でのテストの点数の変化の母平均は 0」という仮説を，市内の中学校の中から無作為に選ばれた 6 校の標本を用いて検証する．

帰無仮説

- とりあえず「真」であると想定する仮説を帰無仮説（null hypothesis）という.
  - H_0 と書くことが多い.
  - e.g., $H_0 : \mu = 0$.
  - H_0 は必ず「=」または「$\leq$ や $\geq$」を使った式.
    「$\mu > 0$」を H_0 とする検定は不可能.

- まずは H_0 が「真」であると仮定し，それを「偽」とするための証拠を探す.
  - 刑事裁判における推定無罪の原則と同様（被告人は，判決が出るまでは「無罪」として扱われ，「有罪」とするための証拠が探され裁判で提示される）.

$\Rightarrow$ 具体的には，検定統計値を計算する.

- 標本の関数を統計量（statistic）という．
  - e.g., 標本平均，標本分散など
- 検定に用いる統計量を検定統計量（test statistic）といい，その実現値を検定統計値という．
- 仮説を偽と判断して捨てることを，仮説を棄却する（reject）という．
- 仮説を偽とはいえないと判断することを，仮説を採択する（accept）という．
  - 「受容する」ともいう．
  - 仮説を偽とする証拠が不十分であるという意味で，「仮説は真」という判断ではない．

- 仮に H_0 が真であれば，計算した検定統計値が5%や 1%のわずかな確率でしか生じえない値になっている

$$\Downarrow$$

それを証拠として H_0 を偽と判断し，H_0 を棄却する.

- 仮に H_0 が真であれば，計算した検定統計値が小さすぎない確率で生じうる値になっている

$$\Downarrow$$

H_0 を偽とする証拠が不十分であり，偽とはいえないと判断し，H_0 を採択する.
  - 15%や 20%は「小さすぎない」.
  - 「H_0 は真」という判断ではない.

対立仮説

- H_0 が偽のときに代わりに採択する仮説を対立仮説（alternative hypothesis）という.
  - H_1 と書くことが多い.
  - e.g., $H_1 : \mu > 0$.
  - H_1 は「$\neq, <, >$」を使った式で設定できる.

2 種類の誤り

- H_0 が真なのに棄却することを第 1 種の誤り（type I error）という.
  - 第 1 種の過誤ともいう.
- H_0 が偽なのに採択することを第 2 種の誤り（type II error）という.
  - 第 2 種の過誤ともいう.

$$\Downarrow$$

表にすると,

	H_0 を採択	H_0 を棄却
H_0 が真	正しい判断	第 1 種の誤り
H_0 が偽	第 2 種の誤り	正しい判断

- 両方の誤りの可能性を同時になくすことは不可能である.
- 帰無仮説 H_0 の採択は「H_0 が偽とはいえない」という弱い判断なので,「H_0 が偽」と強く判断して棄却してしまう「第 1 種の誤り」のほうが「第 2 種の誤り」よりも重大である.

有意水準

- H_0 を真としたときに，検定統計値が「わずかな確率でしか生じえない値」かの判断の基準となる確率，また，許容する第 1 種の誤りの確率を有意水準（significance level）という．
  - 通常は 10%，5%，1%に設定．
  - e.g., $\boxed{\text{「有意水準 5\%で } H_0 \text{ が棄却された」}}$
    ⇒ 仮に H_0 が真であれば，そんな検定統計値が出てくる確率は 5%以下に過ぎない（H_0 を偽とする証拠）ので H_0 を棄却．
    ⇒ 言い換えると，H_0 が真のとき，「そんな検定統計値」は 5%以下の確率で出現しうる．
    ⇒ H_0 を棄却する第 1 種の誤りを犯すことが，多くとも 5%の確率でありうる．

棄却域・採択域・臨界値

- 帰無仮説を棄却する検定統計量の範囲を棄却域（rejection region）という．
- 帰無仮説を採択する検定統計量の範囲を採択域（acceptance region）という．
- 棄却域と採択域の境界の値を臨界値（critical value）という．
- 検定統計量は確率変数なので，確率分布を当てはめて，「わずかな確率でしか生じえない値」の範囲を棄却域とし，「小さすぎない確率で生じうる値」の範囲を採択域とする．

両側検定と片側検定

- 検定統計量がとりうる範囲の両側に棄却域を定める検定を両側検定（two-sided test）という.
- 両側検定問題の定式化：

$$H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu \neq 0$$

➡ H_1 の意味は,「母平均は 0 と異なる」.

- 検定統計量がとりうる範囲の片側に棄却域を定める検定を片側検定（one-sided test）という.
- 片側検定問題の定式化：

$$H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu > (<) 0$$

➡ H_1 の意味は,「母平均は 0 より大きい（小さい)」.

片側検定

- 片側検定のうち，検定統計量がとりうる範囲の，確率密度関数のグラフにおける右側に棄却域を定める検定を右片側検定（right-tailed test）という．
  - 「$H_1 : \mu > 0$」なら右片側検定となる．
- 片側検定のうち，検定統計量がとりうる範囲の，確率密度関数のグラフにおける左側に棄却域を定める検定を左片側検定（left-tailed test）という．
  - 「$H_1 : \mu < 0$」なら左片側検定となる．

仮説検定の手順

1. 帰無仮説と対立仮説を設定する.
2. 検定統計量と，それが帰無仮説のもとで従う分布を求める.
3. 有意水準を設定する.
4. 棄却域・採択域・臨界値を定める.
5. 検定統計値を計算し，それが棄却域にあるか採択域にあるかで帰無仮説の採否を判定する.
   - 検定統計値が棄却域にあれば帰無仮説を棄却し，対立仮説を採択する.
   - 検定統計値が採択域にあれば帰無仮説を採択する.

標本平均の分布

- 母集団分布を $N(\mu, \sigma^2)$ と仮定し，この母集団からの無作為標本を $(X_1, X_2, \cdots, X_n)$ とすると，任意の $i = 1, 2, \cdots, n$ について，
$$X_i \sim N(\mu, \sigma^2).$$
- $(X_1, X_2, \cdots, X_n)$ の標本平均を，
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$
とすると，
$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$
（証明は省略）

- $\bar{X}$ を標準化すると，

$$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0, 1).$$

（証明は省略）

- 実際の実証分析では，母平均 μ は未知なので，推定したり μ に対して仮説を置いて検定したりする．
- このとき，実際の実証分析では，母分散 σ^2 も未知．

- $(X_1, X_2, \cdots, X_n)$ の標本分散を，

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,$$

として，未知の母分散 σ^2 を計算可能な標本分散 S^2 で置き換えると，

$$\frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1).$$

- H_0 のもとで t 分布に従う検定統計量を t 統計量（t-statistic）という．
  - t 検定統計量ともいう．
- t 統計量を用いる検定を t 検定（t-test）という．
- t 統計量の実現値を t 値（t-ratio）という．

例題 1

母集団分布を $N(\mu, \sigma^2)$ とする．この母集団からの無作為標本を $(X_1, X_2, \cdots, X_n)$ として，

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,$$

とすると，

$$\frac{\bar{X} - \mu}{\sqrt{S^2/n}} \sim t(n-1),$$

となることを証明しなさい．ただし，$\bar{X}$ と S^2 が独立であることと，以下の性質を用いてよい．

$$\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0, 1), (n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1).$$

解法

（証明）

$$\frac{\bar{X}-\mu}{\sqrt{S^2/n}} = \frac{(\bar{X}-\mu)/\sqrt{\sigma^2/n}}{\sqrt{S^2/n}/\sqrt{\sigma^2/n}}$$

$$= \frac{(\bar{X}-\mu)/\sqrt{\sigma^2/n}}{\frac{\sqrt{S^2}}{\sqrt{n}} \cdot \frac{\sqrt{n}}{\sqrt{\sigma^2}}}$$

$$= \frac{(\bar{X}-\mu)/\sqrt{\sigma^2/n}}{\sqrt{S^2/\sigma^2}}$$

$$= \frac{(\bar{X}-\mu)/\sqrt{\sigma^2/n}}{\sqrt{(n-1)\frac{S^2}{\sigma^2}/(n-1)}}. \tag{1}$$

ここで，

$$\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1), (n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1),$$

なので，(1) は $N(0,1)$ に従う確率変数を，「$\chi^2(n-1)$ に従う確率変数を $n-1$ で割ったものの正の平方根」で割ったものである．さらに，$\bar{X}$ と S^2 は独立なので，$(\bar{X}-\mu)/\sqrt{\sigma^2/n}$ と $\sqrt{(n-1)\frac{S^2}{\sigma^2}/(n-1)}$ も独立である．よって，

$$\frac{(\bar{X}-\mu)/\sqrt{\sigma^2/n}}{\sqrt{(n-1)\frac{S^2}{\sigma^2}/(n-1)}} \sim t(n-1),$$

である．

したがって，
$$\frac{\bar{X}-\mu}{\sqrt{S^2/n}} \sim t(n-1),$$
となる（証明終）.
$$\Downarrow$$
母平均 μ の仮説検定をするには，標本平均 $\bar{X}$ を標準化したうえで母分散 σ^2 を標本分散 S^2 に置き換えたものを検定統計量（t 統計量）とし，それに自由度 $n-1$ の t 分布を当てはめればよい.

- $\bar{X}$ と S^2 が独立であることについては，以下の文献を参照.
    - Hoel, Paul G. (1984) *Introduction to Mathematical Statistics*, Fifth Edition, John Wiley & Sons, Inc., Canada.

母平均の検定の例題

例題 2

A 市は市内すべての中学校で，中学 3 年生の数学の学力向上を目的として学習塾講師による夏期講習の実施を考えている．そこで，A 市の中学校のなかから無作為に 6 校を選んで夏期講習を試験的に実施した．同じ難易度のテスト（100 点満点）を夏期講習実施前後に行ったところ，各学校でのテストの平均点の変化の，学校間での平均が 0.333 点，平均との差の 2 乗の合計を 5 で割った分散が 8.267（点 ^{2}）であった．夏期講習の効果はあったのだろうか．

1. 夏期講習実施前後で平均点が変化したかどうかを，有意水準 5%で検定しなさい．
2. 夏期講習実施後に平均点が上昇したかどうかを，有意水準 5%で検定しなさい．

考え方

- テストの平均点の変化の学校間での平均が 0.333 点なので，平均的にはテストの点数が上がっているように見える．
- テストの平均点の変化の，A 市のすべての中学校間での平均は 0 点（平均的には変化していない）で，この 6 校での結果がたまたま 0.333 点となっただけかもしれない．
- A 市内の別の 6 校で同じことを実施すると，テストの平均点の変化の学校間での平均がマイナス（平均的には下がっている）かもしれない．

⇒ 仮説検定で，テストの平均点の変化の，A 市の**すべての中学校間での平均が 0 点なのかそうでないのか**を検証．

$\boxed{\text{解法}}$

1. 夏期講習実施前後で平均点が変化したかどうか

Step 1：H_0 と H_1 を設定する

夏期講習実施前後でのテストの点数の変化の母平均を μ として，帰無仮説と対立仮説をそれぞれ,

$$H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu \neq 0$$

と設定する.

Step 2：検定統計量とその H_0 のもとでの分布を求める

標本サイズ n が 6 なので，検定統計量とその分布は,

$$T = \frac{\bar{X} - 0}{\sqrt{S^2/6}} \sim t(5).$$

Step 3. 有意水準を設定する

t 統計量を用い，有意水準 5%の両側検定を行う.

Step 4. 棄却域・採択域・臨界値を定める

t 分布表より、$\Pr(T \geq 2.571) = 0.025$, つまり $\Pr(|T| \geq 2.571) = 0.05$ なので，有意水準 5%の両側検定における，$t(5)$ に従う t 統計量の棄却域は,

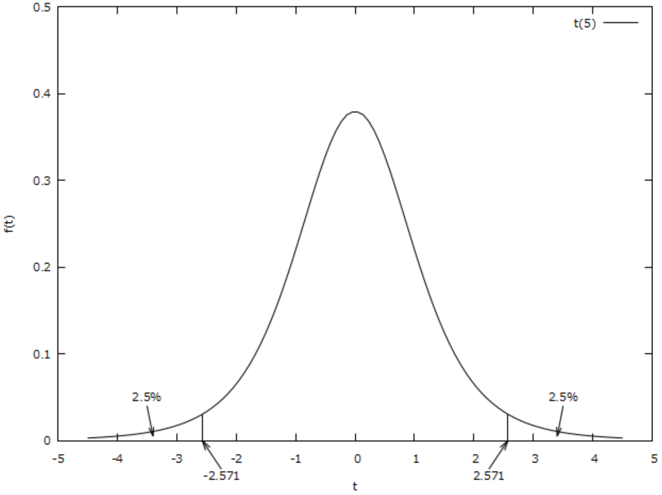
$$(-\infty, -2.571], [2.571, \infty)$$

であり，採択域は,

$$(-2.571, 2.571)$$

である．よって，臨界値は ± 2.571 となる.

自由度 5 の t 分布の確率密度関数



Step5. 検定統計値を求め，H_0 の採択・棄却を判断する

問題文より，標本平均の実現値は，$\bar{x} = 0.333$ であり，不偏性を満たす標本分散の実現値は，$s^2 = 8.267$ である．このことから，検定統計値（t 値）は，

$$t = \frac{0.333 - 0}{\sqrt{8.267/6}} \approx 0.284$$

である．$-2.571 < 0.284 < 2.571$ なので，t 値は採択域に入る．これは，仮に $\mu = 0$（H_0 が真）であれば，0.284 という t 値は小さすぎない確率で実現しうることを意味する．
したがって，$H_0 : \mu = 0$ は有意水準 5%で採択され，夏期講習実施前後でテストの平均点が変化したとはいえない．

2. 夏期講習実施後で平均点が上昇したかどうか

Step 1：H_0 と H_1 を設定する

夏期講習実施前後でのテストの点数の変化の母平均を μ として，帰無仮説と対立仮説をそれぞれ，

$$H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu > 0$$

と設定する．

Step 2：検定統計量とその H_0 のもとでの分布を求める

標本サイズ n が 6 なので，検定統計量とその分布は，

$$T = \frac{\bar{X} - 0}{\sqrt{S^2/6}} \sim t(5).$$

Step 3. 有意水準を設定する

t 統計量を用い，有意水準 5%の右片側検定を行う．

Step 4. 棄却域・採択域・臨界値を定める

t 分布表より、$\Pr(T \geq 2.015) = 0.05$ なので，有意水準 5%の右片側検定における，$t(5)$ に従う t 統計量の棄却域は,

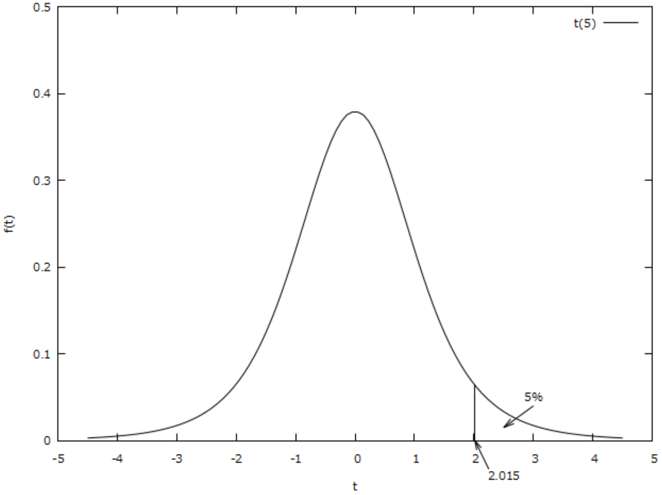
$$[2.015, \infty)$$

であり，採択域は,

$$(-\infty, 2.015)$$

である．よって，臨界値は 2.015 となる．

自由度 5 の t 分布の確率密度関数



Step5. 検定統計値を求め，H_0 の採択・棄却を判断する

問題文より，標本平均の実現値は，$\bar{x} = 0.333$ であり，不偏性を満たす標本分散の実現値は，$s^2 = 8.267$ である．このことから，検定統計値（t 値）は，

$$t = \frac{0.333 - 0}{\sqrt{8.267/6}} \approx 0.284$$

である．$0.284 < 2.015$ なので，t 値は採択域に入る．これは，仮に $\mu = 0$（H_0 が真）であれば，0.284 という t 値は小さすぎない確率で実現しうることを意味する．

したがって，$H_0 : \mu = 0$ は有意水準 5%で採択され，夏期講習実施後にテストの平均点が上昇したとはいえない．

$\boxed{\text{例題 3}}$

例題 2 の A 市について，例題 2 とは別の中学校 11 校を無作為に選んで夏期講習を試験的に実施した．同じ難易度のテスト（100 点満点）を夏期講習実施前後に行ったところ，各学校でのテストの平均点の変化の，学校間での平均が 2.909 点，平均との差の 2 乗の合計を 10 で割った分散が 12.691（点^2）であった．夏期講習の効果はあったのだろうか．

1. 夏期講習実施前後で平均点が変化したかどうかを，有意水準 5%で検定しなさい．
2. 夏期講習実施後に平均点が上昇したかどうかを，有意水準 5%で検定しなさい．

$\boxed{\text{解法}}$

1. 夏期講習実施前後で平均点が変化したかどうか

Step 1：H_0 と H_1 を設定する

夏期講習実施前後でのテストの点数の変化の母平均を μ として，帰無仮説と対立仮説をそれぞれ，

$$H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu \neq 0$$

と設定する.

Step 2：検定統計量とその H_0 のもとでの分布を求める

標本サイズ n が 11 なので，検定統計量とその分布は，

$$T = \frac{\bar{X} - 0}{\sqrt{S^2/11}} \sim t(10).$$

Step 3. 有意水準を設定する

t 統計量を用い，有意水準 5%の両側検定を行う.

Step 4. 棄却域・採択域・臨界値を定める

t 分布表より、$\Pr(T \geq 2.228) = 0.025$, つまり $\Pr(|T| \geq 2.228) = 0.05$ なので，有意水準 5%の両側検定における，$t(10)$ に従う t 統計量の棄却域は，

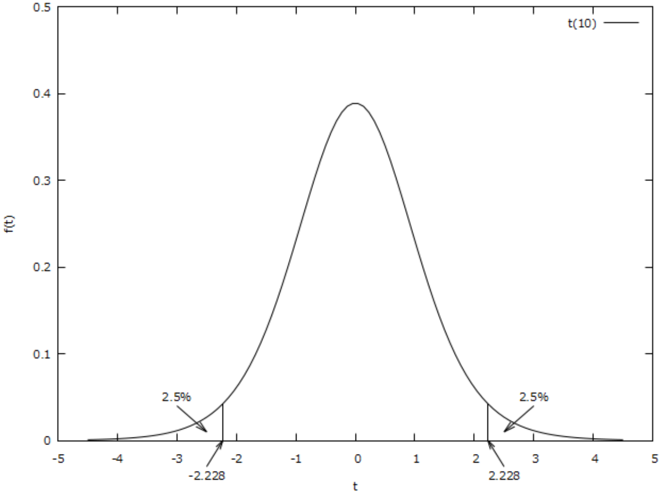
$$(-\infty, -2.228], [2.228, \infty)$$

であり，採択域は，

$$(-2.228, 2.228)$$

である．よって，臨界値は ± 2.228 となる.

自由度 10 の t 分布の確率密度関数



Step5. 検定統計値を求め，H_0 の採択・棄却を判断する

問題文より，標本平均の実現値は，$\bar{x} = 2.909$ であり，不偏性を満たす標本分散の実現値は，$s^2 = 12.691$ である．このことから，検定統計値（t 値）は，

$$t = \frac{2.909 - 0}{\sqrt{12.691/11}} \approx 2.708$$

である．$2.228 < 2.708$ なので，t 値は棄却域に入る．これは，仮に $\mu = 0$（H_0 が真）であれば，2.708 という t 値が実現する確率は 5%以下にすぎないことを意味するので，帰無仮説 H_0 は疑わしい．したがって，$H_0 : \mu = 0$ は有意水準 5%で棄却され，夏期講習実施前後でテストの平均点が変化した可能性がある．

2. 夏期講習実施後で平均点が上昇したかどうか

Step 1：H_0 と H_1 を設定する

夏期講習実施前後でのテストの点数の変化の母平均を μ として，帰無仮説と対立仮説をそれぞれ,

$$H_0 : \mu = 0 \quad \text{vs} \quad H_1 : \mu > 0$$

と設定する.

Step 2：検定統計量とその H_0 のもとでの分布を求める

標本サイズ n が 11 なので，検定統計量とその分布は,

$$T = \frac{\bar{X} - 0}{\sqrt{S^2/11}} \sim t(10).$$

Step 3. 有意水準を設定する

t 統計量を用い，有意水準 5%の右片側検定を行う.

Step 4. 棄却域・採択域・臨界値を定める

t 分布表より、$\Pr(T \geq 1.812) = 0.05$ なので，有意水準 5%の右片側検定における，$t(10)$ に従う t 統計量の棄却域は,

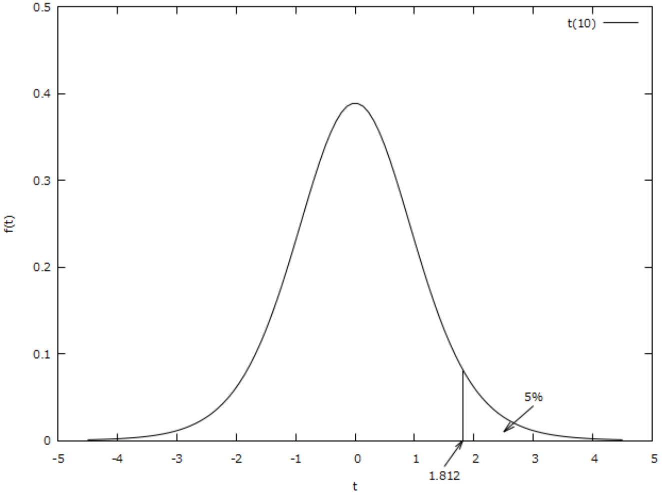
$$[1.812, \infty)$$

であり，採択域は,

$$(-\infty, 1.812)$$

である．よって，臨界値は 1.812 となる.

自由度 10 の t 分布の確率密度関数



Step5. 検定統計値を求め，H_0 の採択・棄却を判断する

問題文より，標本平均の実現値は，$\bar{x} = 2.909$ であり，不偏性を満たす標本分散の実現値は，$s^2 = 12.691$ である．このことから，検定統計値（t 値）は，

$$t = \frac{2.909 - 0}{\sqrt{12.691/11}} \approx 2.708$$

である．$1.812 < 2.708$ なので，t 値は棄却域に入る．これは，仮に $\mu = 0$（H_0 が真）であれば，2.708 という t 値が実現する確率は 5%以下にすぎないことを意味するので，帰無仮説 H_0 は疑わしい．したがって，$H_0 : \mu = 0$ は有意水準 5%で棄却され，夏期講習実施後にテストの平均点が上昇した可能性がある．

今日のキーワード

仮説検定，帰無仮説，統計量，検定統計量，検定統計値，棄却する，採択する，対立仮説，第 1 種の誤り，第 2 種の誤り，有意水準，棄却域，採択域，臨界値，両側検定，片側検定，右片側検定，左片側検定，t 統計量，t 検定，t 値

次回までの準備

- 今回の講義スライドを読み直す.
- 教科書第 5 章第 1 節〜第 2 節，第 4 節を読む.