

予習の便宜のため、教科書の関連箇所を [かっこ内] に示した。教科書の説明や例題も読んでおくとよい。計算を追うだけでなく、下線を引いた語句の意味についても予習しておこう。
教科書: 薩摩順吉, 確率・統計, 岩波書店 (1 年次講義「確率統計学」の教科書)。

ある工場から 100 万個の部品を買った。納品書によると、部品の長さの平均値は 100 mm である。納入された部品が納品書の通りかどうか調べたい。

納入された部品の長さを知りたいければ、部品の長さを測ればよい。でも 100 万個の部品の長さを測るなんてことはやりたくない。そこで、100 万個の中から 7 個だけ取り出して長さを調べた。その結果は次の通りであった。

抜き取り検査の結果 (単位は mm) 111, 98, 99, 102, 107, 104, 97	(☆)
------------------------------------------------------	-----

このデータから、納入された部品の長さについて何が言えるか考えよう。

納入された部品と検査した部品の関係について

問題設定を、統計学の言葉で書きなおしておく。
100 万個の部品の長さ ($Y_1, Y_2, \dots, Y_{1000000}$) が 母集団 である。母集団の分布を特徴づける量 (母集団の平均値や分散など) を 母数 と呼ぶ。7 個の部品の長さ (x_1, x_2, \dots, x_7) は、母集団から取り出した 標本 である。標本を調べることによって母集団の情報を得ようというのが、統計学の目的だ (図 1)。

母数は 100 万個のデータから決まる値だ。それを 7 個のデータから決めるなんてことは、そもそも不可能だ。それでもなんとかして母数がいくらになるかをあててやろうという訳である。

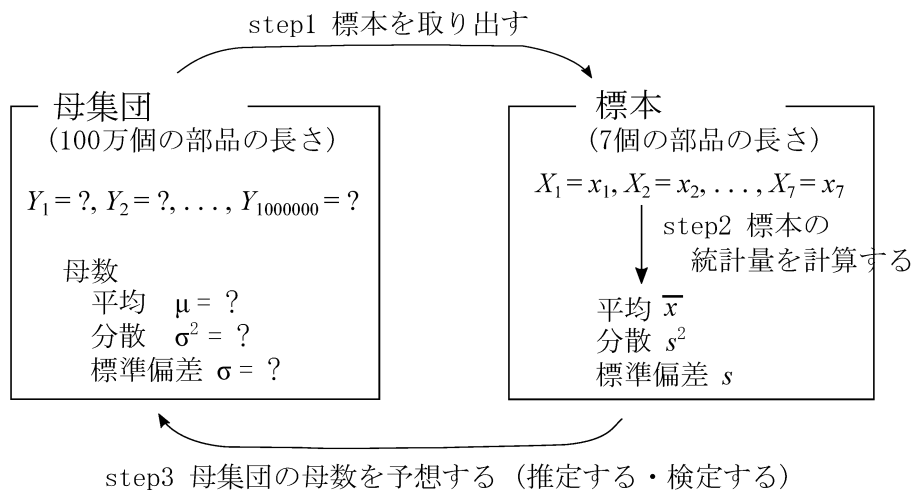


図 1: 母集団の情報 (母数) を知るための手順。

母集団についての仮定

この例題では問題を簡単にしたいので、次のことを仮定する。

- (1) 母集団のばらつき具合は正規分布に従うとする。
- (2) 例題 B2 と例題 C においては、母集団の分散 σ^2 は既知であるとする。

確率変数と実測値の区別，文字表記について

推定・検定では、標本のデータを確率変数の値であるとみなす。そのため、推定・検定の過程においては、同じものに対して2つの見方をする。ある量を、ある分布に従う確率変数と考えるときと、実測データに基づく値と考えるときとがある。確率変数と考えているときは、その値は確率的にしか決まらない(1つの値には決まらない)。実測値と考えているときは、その値は1つに決まる。

この文書では、同じ量が大きくと小文字で表されているとき、大文字は確率変数を表し、小文字は実測値を表す。例えば、図1の式「 $X_1 = x_1$ 」において、 X_1 は確率変数の名前であり、 x_1 は実測値(標本の中の1つ目のデータ 111mm)である。

例題 A 標本の統計量を計算する [§5-2 標本の整理, 例 3, 問題 5-2 の 2]

検査した7個の部品について、長さの平均，分散，標準偏差を求める。

大きさ n の標本 x_1, x_2, \dots, x_n が1個あったとき、この標本の平均 \bar{x} ，分散 s^2 ，標準偏差 s は、次式で定義される。 [§5-2 標本の整理 式 (5.1)(5.2)]

$$\text{標本平均 } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (1a)$$

$$\text{標本分散 } s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad \text{標本標準偏差 } s = \sqrt{s^2}. \quad (1b)$$

式(1)に従って、表1のように計算すると、

$$\bar{x} = \frac{718 \text{ mm}}{7} = 102.57 \text{ mm}, \quad (2a)$$

$$s^2 = \frac{157.7 \text{ mm}^2}{7} = 22.53 \text{ mm}^2, \quad s = \sqrt{22.53 \text{ mm}^2} = 4.75 \text{ mm} \quad (2b)$$

となる。

表 1: 例題 A の計算過程。

x_i/mm	111	98	99	102	107	104	97	$\sum_{i=1}^7 x_i = 718 \text{ mm}$
$(x_i - \bar{x})/\text{mm}$	8.43	-4.57	-3.57	0.57	4.43	1.43	-5.57	$\sum_{i=1}^7 (x_i - \bar{x})^2$
$(x_i - \bar{x})^2/\text{mm}^2$	71.1	20.9	12.7	0.32	19.6	2.04	31.0	$=157.7 \text{ mm}^2$

注意: 標本分散を $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ と定義する流儀もある。この演習では教科書に倣って(1b)の定義を使う。表計算ソフト・数式処理ソフトを使ったり、他の本を参考にすることは、分散の定義を確認すること。

例題 B1 点推定

[§6-1 点推定. 例題 6.1]

購入した部品の長さを推定する. 推定の方法はいろいろあるが, 本演習では「不偏推定量」を使うことにする.

用語の定義

- 不偏推定量 [§6-1]: A の期待値 $E[A]$ が母数 B に等しいとき, つまり,

$$E[A] = B \quad (3)$$

であるとき, A を母数 B の 不偏推定量 と呼ぶ.

- 期待値 [§3-3 式 (3.28)]: A が確率変数 X の関数であり, X が確率密度 $f(x)$ に従うとき, A の期待値を

$$E[A] = \int_{-\infty}^{\infty} A f(x) dx. \quad (4)$$

と定義する.

- (a) 納入された 100 万個の部品の長さについて, 平均値 μ を推定する.

標本平均を確率変数と考えたとき, 標本平均 \bar{X} の期待値は母集団の平均に等しい (これは標本平均の定義 (1a) と期待値の定義 (4) を使うと導出できる. [§5-3 式 (5.8)]).

$$E[\bar{X}] = \mu. \quad (5)$$

したがって,

$$\text{母平均 } \mu \text{ の不偏推定量} = \bar{X} \quad (6)$$

である [§6-1 式 (6.2)]. 式 (6) 右辺の標本平均として, 標本 (☆) から求めた実測値 \bar{x} を使うと, 母平均の推定値が求められる.

$$\text{母平均 } \mu \text{ の推定値} = \text{標本平均 } \bar{x} = 102.57 \text{ mm}. \quad (7)$$

標本平均の分布について (式 (5) 左辺の意味).

「 n 個の部品を取り出して長さを測った後, 元に戻す」検査を何回も繰り返すことを想像してみよう. 大きさ n の標本が何個もできる (図 2(a)). それぞれの標本から, 標本平均 $\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(k)}, \dots$ が得られる. これらの標本平均は例えば図 2(b) のような分布を持つ. $E[\bar{X}]$ は, この分布の平均値を表す.

なお, 標本平均の分布に関して次のことが成り立つ.

- 標本の大きさ n が十分大きいとき, 標本平均 \bar{X} の分布は正規分布に近づく. (母集団は正規分布でなくてもよい). [§4-3 p.86 下半分の中心極限定理]

- 母集団が正規分布 $N(\mu, \sigma^2)$ に従うとき、標本平均 \bar{X} は正規分布 $N(\mu, \sigma^2/n)$ に従う。(標本の大きさ n は小さくてもよい)。 [§5-4 正規母集団 命題 5-1]

この例題の場合、標本の大きさ ($n = 7$) は大きくないが、母集団が正規分布に従うことを仮定しているので、標本平均も正規分布に従う。

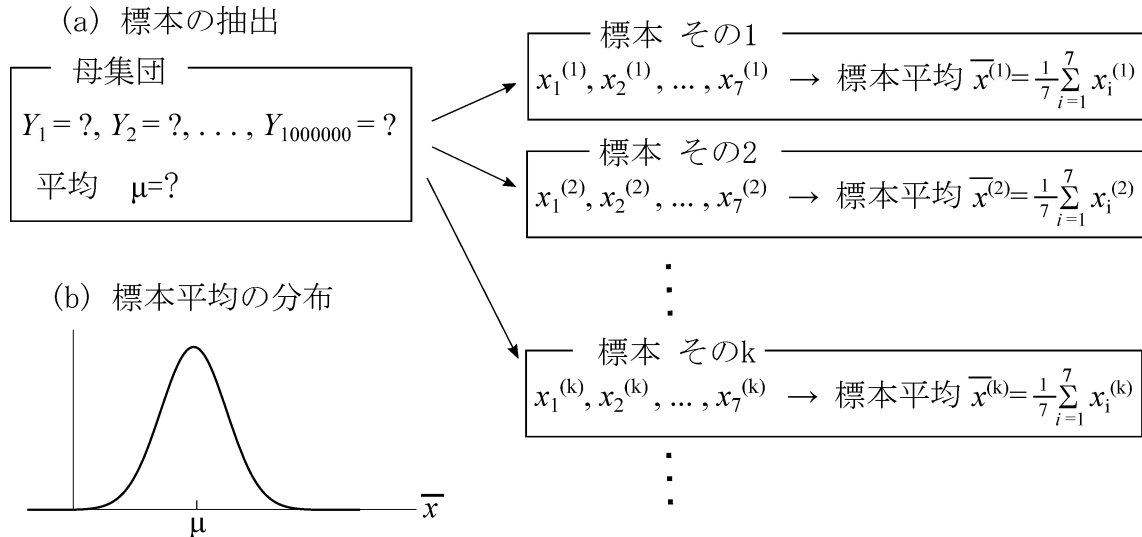


図 2: 無限個の標本と標本平均の分布 (標本の大きさ $n = 7$ の場合)。

- (b) 納入された 100 万個の部品の長さについて、分散 σ^2 を推定する。

標本分散の定義 (1b) と期待値の定義 (4) を使って、標本分散 S^2 の期待値を求めると、次式を得る [§5-3 式 (5.10)–(5.13)].

$$E[S^2] = \frac{n-1}{n} \sigma^2, \quad (8)$$

したがって、

$$E \left[\frac{n}{n-1} S^2 \right] = \sigma^2. \quad (9)$$

式 (9) から、

$$\text{母分散 } \sigma^2 \text{ の不偏推定量} = \frac{n}{n-1} S^2 \quad (10)$$

である [§6-1 式 (6.3)]. 式 (10) 右辺の分散として、標本 (☆) から求めた実測値 s^2 を使うと、母分散の推定値が求められる。

$$\text{母分散 } \sigma^2 \text{ の推定値} = \frac{n}{n-1} s^2 = \frac{7}{7-1} 22.53 \text{ mm}^2 = 26.29 \text{ mm}^2 \quad (11)$$

となる。

例題 B2 区間推定

[§6-2 区間推定. 例題 6.3]

納入された 100 万個の部品について、長さの平均値 μ は未知だが、分散だけは $\sigma^2 = 9.0\text{mm}^2$ であることがわかっているものとする。このとき、平均値 μ の信頼区間を推定する。

考え方

この例題では母集団が正規分布 $N(\mu, \sigma^2)$ に従うことを仮定している。したがって、標本平均 \bar{X} は正規分布 $N(\mu, \sigma^2/n)$ に従う [§5-4 命題 5-1]。つまり、確率密度関数

$$f(\bar{X}) = \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}} \exp\left[-\frac{(\bar{X} - \mu)^2}{2\sigma^2/n}\right] \quad (12)$$

に従う [§4-3 式(4.23)]。このとき、標本平均 \bar{X} が $X_1 < \bar{X} < X_2$ の値をとる確率は、

$$P = \int_{X_1}^{X_2} f(\bar{X}) d\bar{X} \text{ である [§3-1 式(3.7)].}$$

母平均 μ に対して、次式を満たす定数 a があつたとする (図 3(a))。

$$\int_{\mu-a}^{\mu+a} f(\bar{X}) d\bar{X} = 0.95. \quad (13)$$

このとき、 \bar{X} は 95% の確率で、

$$\mu - a < \bar{X} < \mu + a \quad (14)$$

の区間に入る。式 (14) を変形すると、

$$\bar{X} - a < \mu < \bar{X} + a \quad (15)$$

を得る。したがって、もしも、式 (13) を満たす定数 a を見つけたなら、未知数 μ が (15) の区間にあることを 95% の確率であてたことになる。ここで、式 (13) の右辺の数値を信頼水準と呼ぶ (信頼率, 信頼係数と呼ぶこともある)。不等式 (15) の区間を信頼区間と呼ぶ。上の説明では信頼水準として 95% を使ったが、90% や 99% を使うこともある。

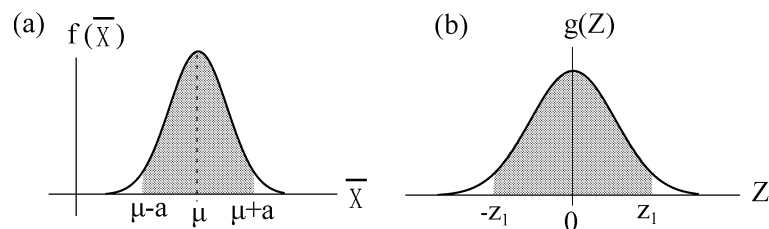


図 3: 正規分布とその積分範囲。灰色の領域の面積が 0.95 である。(a) 積分 (13), (b) 標準正規分布の積分 (18)。

計算方法

式(13)を満たす a を見つけるために、正規分布 $N(\mu, \sigma^2/n)$ を標準正規分布 $N(0, 1)$ に変換する。標準化変換を使って、 \bar{X} を

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad (16)$$

に変換すると、新しい変数 Z は $N(0, 1)$ に従う [§4-3 式(4.31)]. つまり、確率密度関数

$$g(Z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) \quad (17)$$

に従う。

次式を満たす z_1 があったなら

$$\int_{-z_1}^{z_1} g(Z) dZ = 0.95, \quad (18)$$

Z は95%の確率で以下の区間に入る (図 3(b)).

$$-z_1 < Z < z_1. \quad (19)$$

この不等式を式(16)を使って書き直せば、 μ についての信頼区間

$$\bar{X} - \frac{\sigma}{\sqrt{n}} z_1 < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_1 \quad (20)$$

を得る。

不等式(20)の両端の値を求めるためには、標本平均 \bar{X} として実測値 \bar{x} を使う。

関数(17)の積分値は数表に載っている。数表 [教科書 p.213 附表 2] によると、式(18)を満たす z_1 の値は $z_1 = 1.960$ である。この z_1 の値と例題 A で得られた \bar{x} を(20)に代入すると、信頼区間は次のようになる。

$$\begin{aligned} 102.57 \text{ mm} - \frac{\sqrt{9.0 \text{ mm}^2}}{\sqrt{7}} 1.960 < \mu < 102.57 \text{ mm} + \frac{\sqrt{9.0 \text{ mm}^2}}{\sqrt{7}} 1.960, \\ 100.35 \text{ mm} < \mu < 104.79 \text{ mm}. \end{aligned} \quad (21)$$

例題 C 検定

[§6-3 仮説と検定. §6-4 母数の検定, 例題 6.7]

納品書には「部品の長さは平均 100 mm である」と書いてあった(以下では, この値を μ_0 と記すことにする). しかし, 抜き取り検査の結果は平均が $\bar{x} = 102.57$ mm だった. 納入された 100 万個の部品の長さ μ は納品書の通りと考えてよいだろうか? あるいは納品書の記述と違うと考えるべきだろうか? この問題に答えるために, 次の仮説を検定する.

仮説 H_0 : 納入された部品の平均長さは $\mu = \mu_0$ である. ($\mu_0 = 100\text{mm}$).

対立仮説を $H_1: \mu \neq \mu_0$ とする.

仮説検定の考え方の概略を図 4 に示す. A 理論的結論と B 実測による結論を比べて, 合わないときは, 仮説 H_0 を棄却して対立仮説 H_1 が正しいと判定する. 合うときは, 仮説 H_0 を棄却しない.

注意: 仮説を「棄却しない」ことを「採択する」と表現することもある. どちらの用語を使うにしても, 仮説を「棄却しない」「採択する」の意味は, 「仮説が正しいと判定する」ということではない. 「仮説が正しくない」と判定しない」という意味である. A と B が合うときは, 仮説 H_0 が正しいとも正しくないとも言えない. 要するに役に立つような結論は何も得られない. [§6-3 p.146 帰無仮説, 特に上から 11-16 行目 (二重否定の説明)]

例題 C では母集団の分散を既知とする. 例題 B2 と同様に $\sigma^2 = 9.0\text{mm}^2$ とする. このとき, 図 4 の統計量 Θ として, 標本平均 \bar{X} を使うことができる. 検定の手順と結論は以下の通りである.

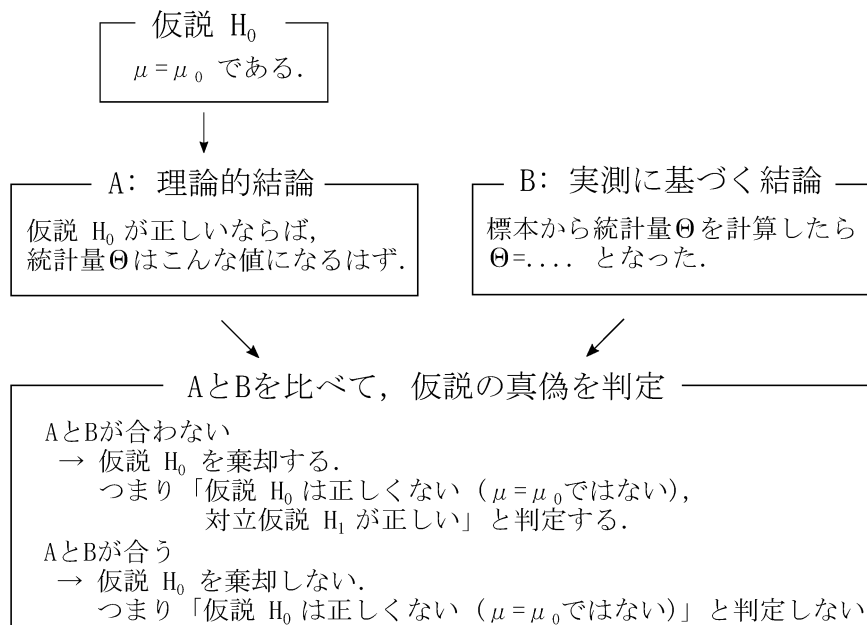


図 4: 仮説検定の考え方.

- A: 理論. 標本平均 \bar{X} を確率変数と考えると, その理論的な分布を考える.

母集団が正規分布なので, 標本平均の分布も正規分布になる. もしも仮説 H_0 が正しいなら母集団の平均は μ_0 であるから, 母集団の分布は $N(\mu_0, \sigma^2)$ である. このとき, 標本平均 \bar{X} は $N(\mu_0, \sigma^2/n)$ に従う [母集団の分布と標本平均の分布の関係については, §5-4 命題 5-1 を見よ].

\bar{X} に標準化変換

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \quad (22)$$

を施すと, Z は標準正規分布 $N(0, 1)$ に従う.

例題 B2 と同様に,

$$\int_{-z_1}^{z_1} g(Z) dZ = \gamma, \quad (\gamma = 0.95) \quad (23)$$

を満たす z_1 を考える. 数表によると $z_1 = 1.960$ なので, Z は次の区間にある.

$$\gamma (95\%) \text{ の確率で, } -1.960 < Z < 1.960, \quad (24a)$$

$$\alpha = 1 - \gamma (5\%) \text{ の確率で, } Z < -1.960, \quad 1.960 < Z. \quad (24b)$$

$\alpha = 1 - \gamma$ のことを危険率と呼ぶ (有意水準と呼ぶこともある). 区間 (24b) を (危険率 5% のときの) 棄却域 と呼ぶ (図 5).

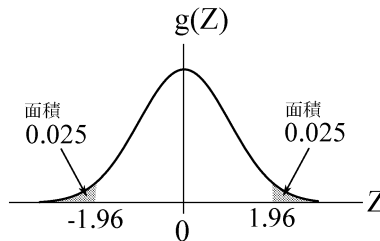


図 5: 危険率 5% の検定における棄却域. 灰色領域の面積の合計が 0.05 である.

- B: 実測. 標本平均の実測値 \bar{x} を考える.

例題 A の結果から, 標本平均の実測値は $\bar{x} = 102.57 \text{ mm}$ である. \bar{x} に対して式 (22) と同じ変換を施すと, Z の実測値 z が求まる.

$$z = \frac{102.57 \text{ mm} - 100 \text{ mm}}{\sqrt{9.0 \text{ mm}^2}/\sqrt{7}} = 2.27 \quad (25)$$

である.

- A と B を比べて、仮説の真偽を判定.

理論的結論 (24) には確率が含まれるので、そのままでは実測に基づく結論 (25) と比べることができない. そこで、仮説検定においては、次のように考える.

γ 以上の確率で起こる事柄: とても起こりやすい と考える.

α 以下の確率で起こる事柄: ほとんど起こらない と考える.

別の言い方をすると、「仮説 H_0 が正しいなら、実測値 z はまず間違いなく区間 (24a) の値をとるはずだ. 棄却域 (24b) の値をとることはほとんどあり得ない」と考える. こういう考えに基づいて、実測値 z が棄却域の値をとったときには、結論 A と結論 B が合っていないとみなす. そして、仮説 H_0 を棄却する.

この例題の場合、実測による z の値は棄却域に含まれるので、仮説 H_0 を棄却する (対立仮説 H_1 が正しいと判定する). 検定の結論は、

$$\text{仮説 } H_0 \text{ は正しくない. } \mu \text{ は } 100 \text{ mm} \text{ ではない} \quad (26)$$

となる. したがって、納品された部品の平均長さは納品書の記述と異なると考えてよい.

もしも実測による z の値が棄却域に含まれなかったらどうだろう. 例えば実測値が $z = 1.95$ だったなら、式 (24a) の範囲に含まれる. この場合、理論的結論と合うから仮説 H_0 を棄却できない. このとき、検定の結論は、

$$\text{仮説 } H_0 \text{ が正しくないとは言えない. } \mu \text{ が } 100 \text{ mm} \text{ でないとは言えない} \quad (27)$$

となる. つまり、 $\mu = 100 \text{ mm}$ かもしれないし、そうでないかもしれないということだ.

注意: 仮説 H_0 を棄却しないとき、検定の結論は、

$$\text{仮説 } H_0 \text{ は正しい. } \mu \text{ は } 100 \text{ mm} \text{ である} \quad (28)$$

とはならない. なぜかと言うと、 H_0 以外の仮説から出発しても実測値 $z = 1.95$ と合うことがあるから. 例えば $\mu = 102 \text{ mm}$ と仮定しても、理論的結論と実測による結果は合う.

例題 C 解説の補足

1. この解説では危険率 α を 5% にしたので、結論 ($\mu \neq 100 \text{ mm}$) が間違っている確率が 5% ある. 間違いになるのは、「本当は $\mu = 100 \text{ mm}$ なのに、たまたま検査した 7 個だけとても高い (あるいはとても低い) 値をとっていた」という状況だ. このような間違いをなくしてより確かな判定をしたかったら、危険率を低く設定すればよい. どの程度の危険率を使うべきか? というのは、数学的に決まることではなく、貴方の考えに応じて変わる.
2. この例題では、対立仮説を $H_1: \mu \neq \mu_0$ と設定したが、問題によっては $H_1: \mu > \mu_0$ (あるいは $\mu < \mu_0$) のように設定することもある. この場合、棄却域の取り方が変わる. 詳しくは教科書例題 6.6, 図 6-7, 図 6-8 を参照せよ.