推測統計 [数理統計学]III

石綿 元

第 8 回講義

目次

1　仮説の検定 I

1.1　仮説の検定

1.1.1　検定とは

標本から母数 θ が確率 $(1-\alpha)$ で存在すると考えられる区間を推定するのが区間推定であった。これを応用して、母数 θ がとある値 θ_0 であることが適当でないか、そうとは言えないかを標本から判断するのが検定である。証明するという概念のない統計における数学で言うところの証明に相当する非常に重要な概念である。

検定と区間推定は表裏の関係にあり、統計量は共通であり、用いる確率分布も同じものを導入する。

1.2　検定の手順

■手順１　２つの仮説をたてる。

$$\begin{cases} H_0 & :\text{帰無仮説}\ \ \theta\text{は}\theta_0\text{とみなせる。} \\ H_1 & :\text{対立仮説}\ \ \theta\text{は}\theta_0\text{とみなせない。} \end{cases}$$

ここで、帰無仮説は

$$H_0 \quad \theta = \theta_0$$

とかける。一方、対立仮説は**両側検定**なら

$$H_1 \quad \theta \neq \theta_0$$

片側検定なら

$$H_1 \quad \theta > \theta_0 \ \ \text{もしくは}\ \ H_1 \quad \theta < \theta_0$$

H_1 にどれを使うかは、検定の目的によって使い分けなければならない。

■手順２　とりあえず帰無仮説 H_0 が正しいとして仮説に応じた統計量を導入する。

ここで導入する統計量は、利用する確率分布に応じた変数に変換したものである。これを特に検定統計量という。

■手順３　手順２で導入した統計量について棄却域を求める。

利用する確率分布において、ほとんど起こらない確率 α となる統計量の範囲を求める。

このとき α を**有意水準**もしくは**危険率**と呼び、その確率をとる統計量の範囲を棄却域という。有意水準は検定の目的によって変えることは可能であるが、**特に指定のない場合には通常５％を用いて行う。**

両側検定の場合、確率分布の両側に確率 $\frac{\alpha}{2}$ ずつ範囲を棄却域として設定する。この場合、同一の確率分布を用いた区間推定での信頼区間 $1-\alpha$ を除いた部分と一致する。

片側検定の場合、確率分布の上側か下側のどちらかのみに確率 α となる範囲を棄却域として設定する。

棄却域は、手順１において両側検定をするのか片側検定をするのかによって異なることに注意。

■手順４　ここまでの手順１、手順２、手順３により導入した検定統計量が棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

注意：帰無仮説は、棄却されて初めて意味を持つものであり、帰無仮説が採択されてしまうと帰無仮説が正しい訳ではなく、帰無仮説が正しくないとはいえないとしか言えない。

1.3　有意水準と両側検定・片側検定

有意水準は通常５％を用いるのであるが、両側検定では、両側に５％の棄却域を設定する。片側検定では、該当側に５％の棄却域を設定する。つまり、両側検定の場合、片側 2.5 ％ずつとなっており、片側検定の場合よりも小さい。

有意水準は、値が大きいほど棄却域が大きくなり、より棄却されやすくなる。つまり、帰無仮説が棄却されやすく、より厳しい判定となる。つまり、同じ５％の棄却域といっても片側検定と両側検定では、片側において判定の厳しさが違う。どちらを選択するかは状況により考察した上で判断するが、立場により変わってくる場合もある。

例えば、「南東京町では、ゴミステーションへの放置ゴミが多く、住民からの苦情がきた。そこで役場ではゴミ収集の回数を増やすかどうか検討するために世帯ごとの１年間のゴミ廃棄量を調査し、公開されている従来の平均と標準偏差を用いて正規分布に基づく検定を行った。このとき役場の担当課では、帰無仮説を

$$H_0：\text{ゴミ廃棄量の平均は変化していない}$$

として有意水準５％で両側検定を行い、その結果、帰無仮説を棄却できなかったと住民説明会で報告した。

しかし、納得できない住民有志たちが役場からデータの提出を受け、同じデータで、同じ有意水準を用いて片側検定を行ったところ、帰無仮説が棄却された。」

といった状況では、増えたかどうかを問題にしたい住民側では、上側の片側検定について厳しく判定したい。一方、役場は、仕事を増やしたくなければ、できるだけ棄却されたくないので両側検定で検定したい。立場によって、有意水準の選択には違いが出てしまう場合がある。

この場合、増えたかどうかの検定であれば、片側検定を選択すべきとも言えるが、一方で、両側検定で棄却できない程度であるとも言える。

そもそも、有意水準は、統計を使う者が設定するものである。通常５％を用いるものであるが、小さくもできれば、大きくもできる。どんどん小さくすると検定自体に意味がなくなっていく。

1.4 平均の検定 I（分散が既知の場合）z-検定

1.4.1 分散が既知の平均の検定

いま、母集団は正規分布 $\mathrm{N}(\mu, \sigma^2)$ に従うと仮定する。このとき標本平均 $\bar{X}$ は、$\mathrm{N}(\mu, \frac{\sigma^2}{n})$ に従う。また、母分散 σ^2 は既知とする。この条件の下で、有意水準 α として標本から母平均について検定する。

■手順 1　仮説をたてる。

$$\begin{cases} H_0 & : \text{帰無仮説} \quad \mu \text{は} \mu_0 \text{とみなせる。} \\ H_1 & : \text{対立仮説} \quad \mu \text{は} \mu_0 \text{とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順 2　統計量を導入する。いま、分散が既知なので、分散が既知の区間推定の時と同じ標準化変換した標準化変数を導入する。

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\frac{\sigma^2}{n}}}$$

この Z は、$\mathrm{N}(0, 1)$ に従う。

■手順 3　有意水準 α として導入した統計量について棄却域を求める。

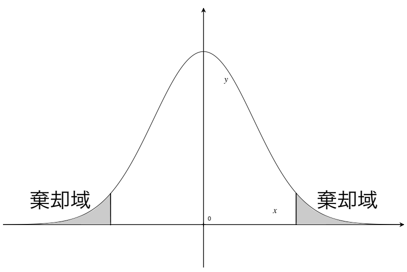


図 1　両側検定

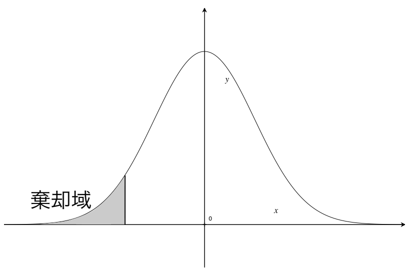


図 2　下側片側検定

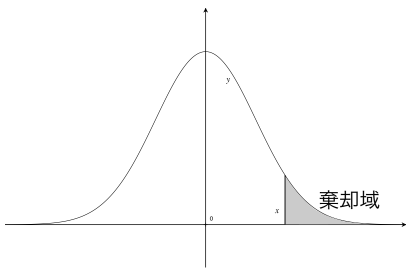


図 3　上側片側検定

■手順 4　手順 2 および手順 3 に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

1.4.2 大標本の平均の検定

大標本の平均の区間推定と同様に、ここでは母集団を $\mathrm{N}(\mu, \sigma^2)$ に従い、かつ、母分散 σ^2 を既知としたが、母集団が正規分布に従わない場合でも、n が十分大きいならば、中心極限定理により標本の平均が示す分布は正規分布と見なせ、さらに、母集団の分布にかかわらず、不偏量である不偏標本分散 s^2 は、母分散 σ^2 の推定値として扱ってよく、分散が既知の平均の検定としてよい。

例題１：平均の区間推定 I（大標本）

統計学の講義を受講している学生から無作為に 50 人選び、1 日に何時間インターネットを利用しているかアンケートを行った。50 人の平均は 3.8 時間、不偏分散 1.44 であった。このとき受講学生のインターネット利用時間の平均 μ に対して 95 %信頼区間と 98 %信頼区間をそれぞれ求めよ。

■95 %信頼区間　平均の区間推定なので 95 %信頼区間では数表から $Z_{\frac{1-\alpha}{2}} = 1.96$ であり、また、$\bar{X} = 3.8$ かつ $n = 50$ かつ $s^2 = 1.44$ なので、これらを代入すると

$$\left[\bar{X} - Z_{\frac{1-\alpha}{2}}\sqrt{\frac{s^2}{n}},\ \ \bar{X} + Z_{\frac{1-\alpha}{2}}\sqrt{\frac{s^2}{n}}\right] = \left[\bar{X} - 1.96\sqrt{\frac{s^2}{n}},\ \ \bar{X} + 1.96\sqrt{\frac{s^2}{n}}\right]$$
$$= \left[3.8 - 1.96\sqrt{\frac{1.44}{50}},\ \ 3.8 + 1.96\sqrt{\frac{1.44}{50}}\right] = [3.47,\ \ 4.13]$$

■98 %信頼区間　信頼係数が違うだけなので、98 %信頼区間では数表から $Z_{\frac{1-\alpha}{2}} = 2.33$ であり、また、$\bar{X} = 3.8$ かつ $n = 50$ かつ $s^2 = 1.44$ なので、これらを代入すると

$$\left[\bar{X} - 2.33\sqrt{\frac{s^2}{n}},\ \ \bar{X} + 2.33\sqrt{\frac{s^2}{n}}\right] = \left[3.8 - 2.33\sqrt{\frac{1.44}{50}},\ \ 3.8 + 2.33\sqrt{\frac{1.44}{50}}\right] = [3.41,\ \ 4.20]$$

例題２：平均の検定 I（大標本）

統計学の講義を受講している学生から無作為に 50 人選び、1 日に何時間インターネットを利用しているかアンケートを行った。このとき、受講学生のインターネット利用時間の平均 μ が 3.4 時間であるか有意水準 5 %と有意水準 1 %で検定せよ。なお、50 人の平均は 3.8 時間、不偏分散は 1.44 であった。

■手順１：仮説をたてる

$$\begin{cases} \text{帰無仮説}:H_0 & \mu = 3.4 \\ \text{対立仮説}:H_1 & \mu \neq 3.4 \end{cases}$$

■手順２：統計量を導入する　標本平均 $\bar{X} = 3.8$ であり、導入する統計量は、標準化変換を用いるので、

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\frac{s^2}{n}}} = \frac{3.8 - 3.4}{\sqrt{\frac{1.44}{50}}} = \frac{0.4}{0.17} = 2.35$$

有意水準 5 %で両側検定する場合

■手順３：棄却域を求める　両側検定なので、数表から $Z_{0.025} = 1.96$ で ± 1.96 の外側が棄却域。

■手順４：判定する　2.35 は棄却域に含まれるので、「平均は **3.4** 時間ではない。」といえる。

有意水準 1 %で両側検定する場合

■手順３：棄却域を求める　両側検定なので、数表から $Z_{0.005} = 2.575$ で ± 2.575 の外側が棄却域。

■手順４：判定する　2.35 は棄却域に含まれないので、「平均は **3.4** 時間ではないとはいえない。」となる。

2　仮説の検定 II

2.1　t-検定

いま、母集団は正規分布 $\mathrm{N}(\mu, \sigma^2)$ に従うと仮定する。有意水準 α として標本から母平均について検定する。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & : \text{帰無仮説}\quad \mu \text{は} \mu_0 \text{とみなせる。} \\ H_1 & : \text{対立仮説}\quad \mu \text{は} \mu_0 \text{とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。小標本の区間推定の時と同じスチューデント化変換した変数を導入する。

$$T = \frac{\bar{X} - \mu_0}{\sqrt{\frac{s^2}{n}}}$$

この T は、自由度 $n-1$ の t-分布 t_{n-1} に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

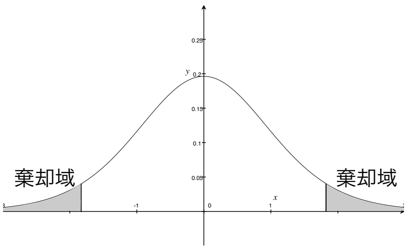


図 4　両側検定

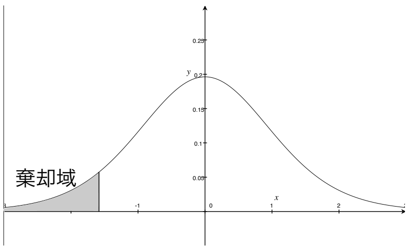


図 5　下側片側検定

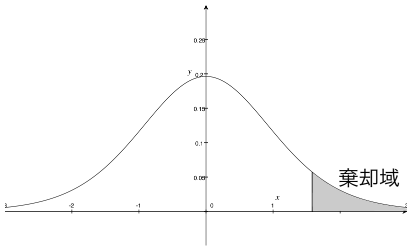


図 6　上側片側検定

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

2.2　t-検定と z-検定

「第４回講義　$t_\infty \to \mathbf{N}(0,1)$ についての理論的証明」で確認したように t-分布は n が大きくなっていくと標準正規分布に近づいていく。

t-検定は、母集団の平均についての検定である。これは、先の正規分布を用いた分散が既知の平均の検定（z-検定）と同様であるが、統計量に母数を含まない。つまり、分散が未知の場合に用いることが可能である。また、z-検定では、標本数が多い場合には、統計量の計算において σ^2 に代えて s^2 を用いても良いということであったが、z-検定において標本数が多くないのに σ^2 に代えて s^2 を用いることは適当でない。そこで、標本数が比較的少ない場合は、統計量にはじめから σ^2 が含まれない t-検定を用いる。

なお、標本数が多くても t-検定が使えるならば、t-検定を用いて良い。なぜならば、t-分布は n が大きくなっていくと標準正規分布に近づいていくのだから、n が十分に大きければ結局同じである。

2.3　分散の検定

いま、母集団は正規分布 $\mathrm{N}(\mu, \sigma^2)$ に従うと仮定する。このとき標本平均 $\bar{X}$ は、$\mathrm{N}(\mu, \frac{\sigma^2}{n})$ に従う。この条件の下で、有意水準 α として標本から母分散について検定する。

■手順 1　仮説をたてる。

$$\begin{cases} H_0 & ：帰無仮説 \quad \sigma^2 は \sigma_0^2 とみなせる。\\ H_1 & ：対立仮説 \quad \sigma^2 は \sigma_0^2 とみなせない。（両側検定もしくは片側検定） \end{cases}$$

■手順 2　統計量を導入する。分散の区間推定の時と同じ統計量を導入する。

$$X = s^2 \frac{n-1}{\sigma_0^2}$$

この X は、自由度 $n-1$ の χ^2-分布に従う。

■手順 3　有意水準 α として導入した統計量について棄却域を求める。

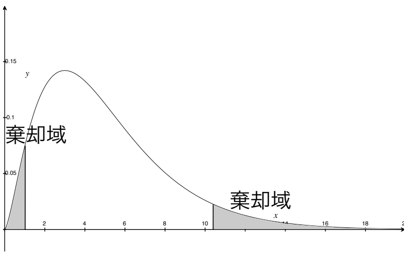


図 7　両側検定

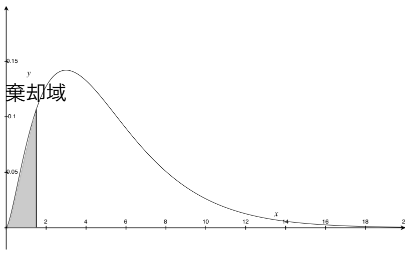


図 8　下側片側検定

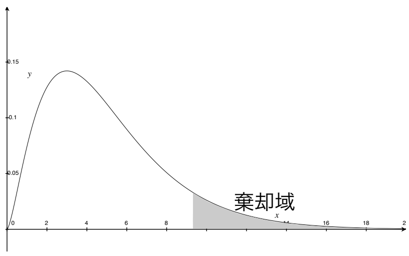


図 9　上側片側検定

■手順 4　手順 2 および手順 3 に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

2.4　p-値

ここまで、正規分布、t-分布、χ^2-分布を用いた仮説の検定を解説してきた。各確率分布の選択は、検定対象である母数に応じて使い分ける訳であるが、どの確率分布を用いる場合でも、検定手順は同じであった。

コンピュータを用いて統計解析を行う場合、統計ソフトでは多くの場合、p-値が示されることが多い。この p-値とは、仮説の検定の「手順 2」で計算した統計量（ここではこの値を x と表すとする）が、今利用している確率分布の横軸上の値として、たとえば、上側であれば、その点 x よりも大きい区間 $[x, \infty]$ の面積を表す値である。つまり、x よりも大きい区間の確率のことを p-値と呼ぶ。

したがって、p-値を用いるならば、片側検定ならば $p < \alpha$ であれば棄却域に含まれている。また、両側検定ならば $p < \frac{\alpha}{2}$ で棄却域に含まれることになる。

3　仮説の検定 III

3.1　分散比（等分散）の検定

母集団を二つ考える。母集団 I が正規分布 $\mathrm{N}(\mu_v, \sigma_v^2)$ に従い、母集団 II が正規分布 $\mathrm{N}(\mu_w, \sigma_w^2)$ に従うと仮定する。このとき、母分散 σ_v^2 と σ_w^2 が等しいかどうかの検定を行う。

次節で扱う「二つの集団の平均の差に関する検定（分散が未知の場合）」を行う際には、事前に「等分散」を検定しておく。

母集団 I からの m 個の標本 $V_1, V_2, \cdots V_m$ および母集団 II からの n 個の標本 $W_1, W_2, \cdots W_n$ を用いる。

ここで、母集団 I の標本分散は、

$$s_v^2 = \frac{1}{m-1}\sum_{i=1}^{m}(V_i - \bar{V})^2$$

母集団 II の標本分散は、

$$s_w^2 = \frac{1}{n-1}\sum_{i=1}^{n}(W_i - \bar{W})^2$$

である。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & : \text{帰無仮説}\quad \sigma_v^2 \text{は} \sigma_w^2 \text{とみなせる。} \\ H_1 & : \text{対立仮説}\quad \sigma_v^2 \text{は} \sigma_w^2 \text{とみなせない。（両側検定）} \end{cases}$$

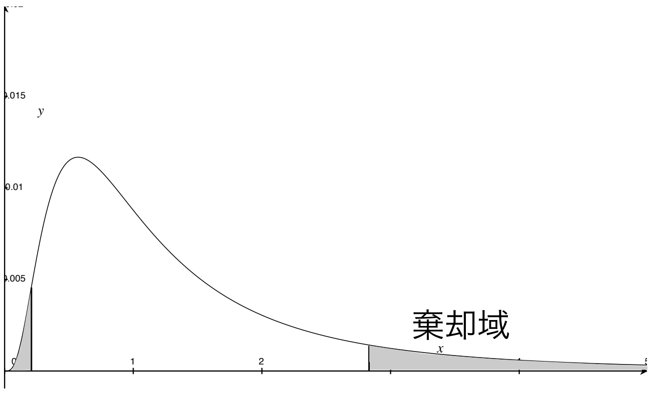
■手順２　統計量を導入する。

まず、s_v^2 と s_w^2 の大小関係を比較する。そのうえで、s_v^2 と s_w^2 のうち大きい方を分子をとして統計量を導入する。いま、$s_v^2 > s_w^2$ とすると、

$$F = \frac{s_v^2}{s_w^2}$$

この統計量は、自由度 $(m-1, n-1)$ の F-分布に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。



いま、統計量を導入する際、F-統計量の分子を大きくなるように取ったので、この統計量はいつも 1 より大きい。したがって、両側検定を行うけれども、便宜上、上側のみを用いて読み取るのでよい。

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

3.2　二つの集団の平均の差に関する検定（分散が既知の場合）

正規分布に従う 2 つの独立した分布を持つ標本の集団を考える。

- 集団 I ：$\bar{X} \sim \mathrm{N}(\mu_1, \frac{\sigma_1^2}{m})$
- 集団 II：$\bar{Y} \sim \mathrm{N}(\mu_2, \frac{\sigma_2^2}{n})$

に従っていることを仮定しておく。

ここでの検定の目的は、この 2 つの分布の平均が等しいかどうかを検定することにある。

標本集団 I の確率変数 $\bar{X}$ と標本集団 II の確率変数 $\bar{Y}$ についての差を新しい確率変数 D とする。すなわち、

$$D = \bar{X} - \bar{Y}$$

そのとき、新しい確率変数 D も正規分布に従い、

$$D \sim \mathrm{N}(\mu_3, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n})$$

と表される。ここで、$\mu_3 = \mu_1 - \mu_2$。

3.2.1　二つの集団の分散が等しく既知のとき

いま、$\sigma_1^2 = \sigma_2^2$ であって、これを σ^2 としておく。また、σ^2 が既知のときである。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & \text{：帰無仮説}\quad \mu_3\text{は 0 とみなせる。} \\ H_1 & \text{：対立仮説}\quad \mu_3\text{は 0 とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。いま、分散が既知なので、分散が既知の平均の検定と同じ標準化変換するが、導入する標準化変数は、

$$Z = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{\sigma^2}{m} + \frac{\sigma^2}{n}}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma^2}{m} + \frac{\sigma^2}{n}}}$$

この Z は、$\mathrm{N}(0, 1)$ に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

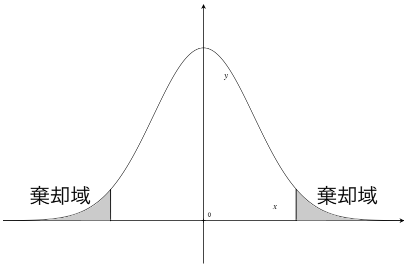


図 10　両側検定

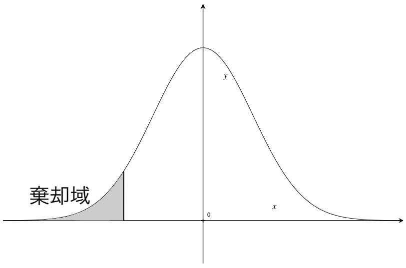


図 11　下側片側検定

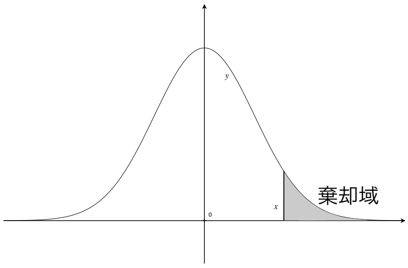


図 12　上側片側検定

■手順４　手順 2 および手順 3 に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

3.2.2　二つの集団の分散が異なっていてともに既知のとき

今度は、$\sigma_1^2 \neq \sigma_2^2$ であって、また、σ_1^2 と σ_2^2 がともに既知のときである。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & : \text{帰無仮説} \quad \mu_3\text{は } 0 \text{ とみなせる。} \\ H_1 & : \text{対立仮説} \quad \mu_3\text{は } 0 \text{ とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。いま、分散が既知なので、分散が既知の平均の検定と同じ標準化変換するが、導入する標準化変数は、

$$Z = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

この Z は、$\mathrm{N}(0, 1)$ に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

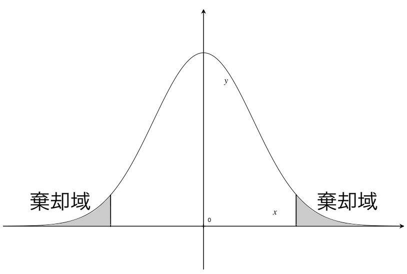


図 13　両側検定

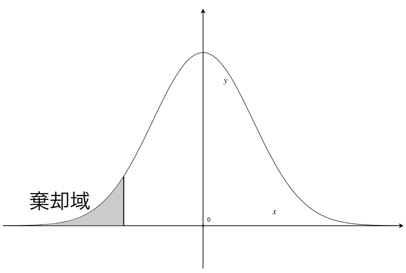


図 14　下側片側検定

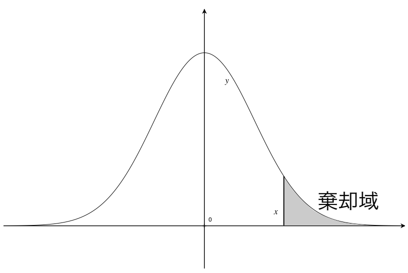


図 15　上側片側検定

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

3.3　二つの集団の平均の差に関する検定（分散が未知の場合）

正規分布に従う２つの独立した分布の標本の集団を考える。

ここでの検定の目的は、分散が既知の時と同じくこの２つの分布の平均が等しいかどうかを検定することにある。ただし、いま、標本数が少なく、母分散 σ_1^2 および σ_2^2 は未知である。このとき、t-検定を応用できる。

標本集団 I の確率変数 $\bar{X}$ と標本集団 II の確率変数 $\bar{Y}$ についての差を新しい確率変数 D とする。すなわち、

$$D = \bar{X} - \bar{Y}$$

ここで、集団 I の平均を μ_1、集団 II の平均を μ_2 として、$\mu_3 = \mu_1 - \mu_2$。

■分散が未知の場合の二つの集団の平均の差の検定で必要な事前の準備　「分散が未知の場合の二つの集団の平均の差に関する検定」の場合も、「分散が等しい場合」と「分散が等しくない場合」で手順が異なる。この場合、それぞれの母分散が共に未知なのだから、分散が等しいかどうかは、標本から判断することになる。従って、「**分散比（等分散）の検定**」を事前に行い、分散が等しいとみなせるのかどうかの判断を事前にしておく。

3.3.1　二つの集団の分散が等しく未知のとき

まず、「**分散比（等分散）の検定**」の手順の通り検定を行なった結果、「**棄却できなかった**」場合、「分散が等しく無いとまでは言えない」と判断されるが、これにより、「分散は等しいものと仮定」して扱っていく。この時、「**分散比（等分散）の検定**」と同じく、「母集団 I からの m 個の標本 $V_1, V_2, \cdots V_m$ および母集団 II からの n 個の標本 $W_1, W_2, \cdots W_n$」を用いて表現すると、標本分散 s^2 は、

$$s^2 = \frac{\sum_{i=1}^{m}(V_i - \bar{V})^2 + \sum_{i=1}^{n}(W_i - \bar{W})^2}{m+n-2}$$

として求めておく。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & \text{：帰無仮説}\quad \mu_3 \text{は } 0 \text{ とみなせる。} \\ H_1 & \text{：対立仮説}\quad \mu_3 \text{は } 0 \text{ とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。いま、分散が未知なので、分散が未知の平均の検定と同じ t-変換するが、導入する統計量は、

$$T = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{s^2}{m} + \frac{s^2}{n}}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s^2}{m} + \frac{s^2}{n}}}$$

この T は、自由度 $m+n-2$ の t-分布に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

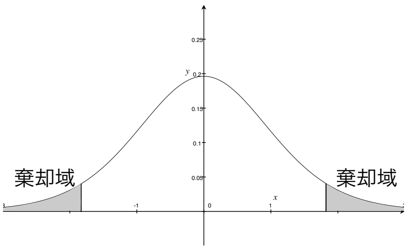


図 16　両側検定

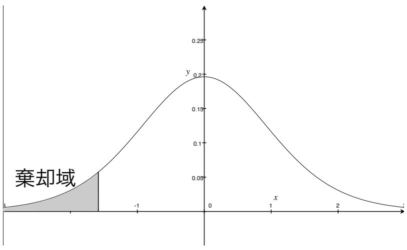


図 17　下側片側検定

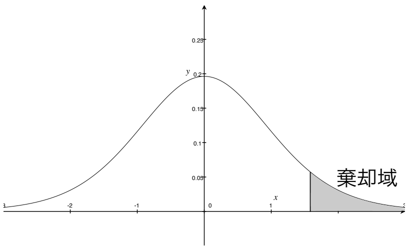


図 18　上側片側検定

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

3.3.2　二つの集団の分散が異なっていてともに未知のとき（Welch）

ここでは、「**分散比（等分散）の検定**」の手順の通り検定を行なった結果、「**棄却された**」場合を扱う。この場合「分散が等しく無い」と判断されるため、標本分散について $s_1^2 \neq s_2^2$ と区別して表現していく。

この検定は、「二つの集団の平均の差の検定」のうち「分散が未知で異なる」という状況で用いられることからよく扱われる。そのため、特に「**Welch の検定**」と呼ばれる。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & :\text{帰無仮説}\quad \mu_3\text{は }0\text{ とみなせる。} \\ H_1 & :\text{対立仮説}\quad \mu_3\text{は }0\text{ とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。いま、分散が未知なので、分散が未知の平均の検定と同じ変換するが、導入する検定統計量は、

$$T = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

この T は、自由度が、

$$\frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{\left(\frac{s_1^2}{m}\right)^2}{m-1} + \frac{\left(\frac{s_2^2}{n}\right)^2}{n-1}}$$

の t-分布に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

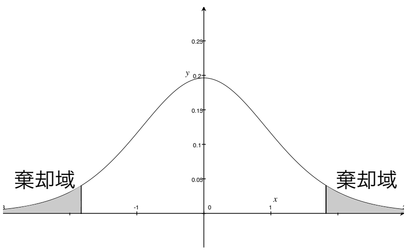


図 19　両側検定

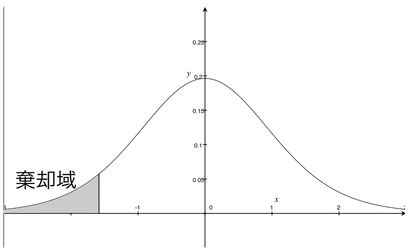


図 20　下側片側検定

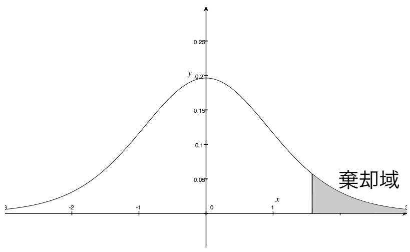


図 21　上側片側検定

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

3.4　比較することの重要性

ところで、大変危険な物質が発見されたという。

「心筋梗塞で病院に運ばれた日本人の実に 95 パーセント以上が継続的に摂取した物質で、強盗や殺人の凶悪犯人の 70 パーセント以上が 24 時間以内に摂取した物質が存在するという。」

「この物質の摂取を禁じると精神的ストレスがみられ、暴動が起こることもあるという。」

あなたはこの物質を食べますか？

実は、このような場合に本当に危険な物質かどうかは、数値だけに着目していても判断できない。

この物質を摂取したグループ（群）と、この物質の摂取を制限したグループ（群）に分け、それぞれを比較することで判断しなければ、本当に危険な物質かどうかはわからない。くだんの例の物質はごはんである。

また、創薬などの分野に置いて、新薬に効果があるかどうかを調べたり、特定保健用食品の効用を調べたりする場合にも比較することで判断する。

このとき、被験者を無作為に選び、選ばれた被験者を無作為に２群に分ける。この分け方を**無作為割り付け**と呼ぶ。このように２群に分けた後、一方の群には効用をみるための対象となる薬などを投与し、もう一方には対照薬やプラセボ（偽薬）などを与える。一定の期間を経て、この２群に差が生じているかどうかを平均の差の検定などの有意差検定を用いて判定する。さらに厳密に行われる場合には、クロスオーバー (crossover) 試験（クロスオーバー比較試験）と呼ばれる試験を行う場合もある。これは、「交差試験」とも呼ばれ、試験対象者を無作為割り付けによりＡ群とＢ群に分け、第１クールではＡ群に試験薬、Ｂ群に対照薬を投与する。しばらく期間を開けて、試験対象者の状況が戻ってから、次の第２クールではＡ群に対照薬、Ｂ群に試験薬を投与する。このように順次入れ替えていく試験を行い、結果を比較することで結論を求める。

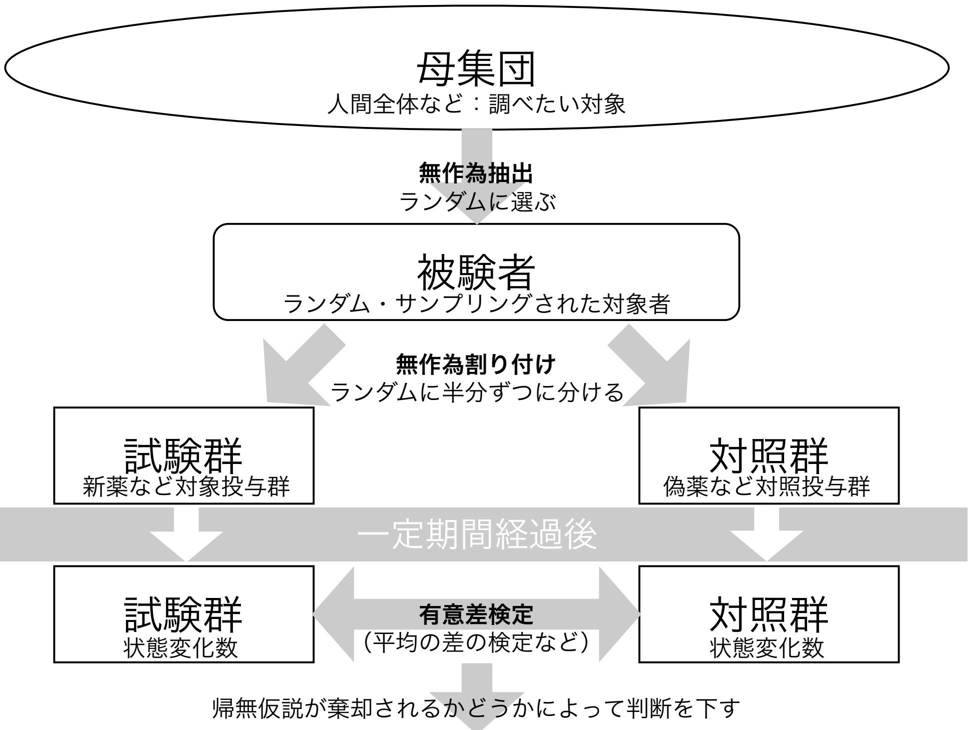


図 22　無作為抽出・無作為割り付け・有意差検定

3.5　検定の二種類の過誤

仮説の検定では、真実は H_0 が正しいにもかかわらず、検定の結果 H_0 を棄却されることは有意水準 α の確率で起こる。この誤りを第 1 種の過誤という。反対に真実が H_0 が正しくないにもかかわらず、検定の結果 H_0 を採択されてしまう場合も起こる。これを第 2 種の過誤といい、この誤りが起こる確率を β とする。

	「H_0 が本当である」が真実	「H_0 が本当でない」が真実
検定の結果 H_0 を棄却	第 1 種の過誤 (この確率が α)	OK(確率 $1-\beta$：検出力)
検定の結果 H_0 を採択	OK(確率 $1-\alpha$)	第 2 種の過誤 (この確率を β)

検定の結果正しい判定の H_0 が真実で検定の結果 H_0 を採択する判定は $1-\alpha$ の確率であり、H_0 が真実でなく検定の結果 H_0 を棄却する判定は確率 $1-\beta$ で表される。ところで帰無仮説 H_0 は、棄却されて初めて意味を持つのであるから、真実が H_0 でないときに、H_0 を棄却してくれることが意味のある結果である。このため、真実が H_0 でないときに、H_0 を棄却する確率 $1-\beta$ を**検出力**と定義する。

上記表に示した通り、検定の結果 H_0 を棄却される結果が得られたなら、たとえ、第 1 種の過誤が生じていても、その確率 α は、自分で導入した小さな確率である有意水準である。したがって、検定の結果、帰無仮説 H_0 が棄却されたときは積極的に「H_0 は棄却される」と言い切れ、「有意である」という表現ができる。一方、帰無仮説 H_0 が採択されたときは、もし第 2 種の過誤が生じていてもその確率は小さくない。したがって、検定の結果、帰無仮説 H_0 が採択されたなら、積極的な表現はできず「H_0 を棄却できなかった」程度の消極的な表現しかできない。積極的に「H_0 が採択される」や「H_0 である」とは言えないのである。

■検出力　有意水準 α と β を同時に小さくすることはできない。いま、もし有意水準 α を大きくしていくと、$1-\beta$ も大きくなっていく。しかし、有意水準 α は勝手に変えることはしないので、その条件の下で、β を小さくすることを考える。そのような場合は、どのような場合かというと、次の図に示すように帰無仮説 H_0 の分布と対立仮説 H_1 の分布を十分離すことで実現される。しかし、通常は帰無仮説 H_0 の分布と対立仮説 H_1 の分布は固定されている。その場合は、それぞれの分布において分散を小さくできれば見かけ上二つの分布は離れることになる。つまり、$1-\beta$ をできるだけ大きくすることが第 2 種の過誤が起こる確率を小さくすることになり、誤りを起こしにくい厳しい検定ということになる。

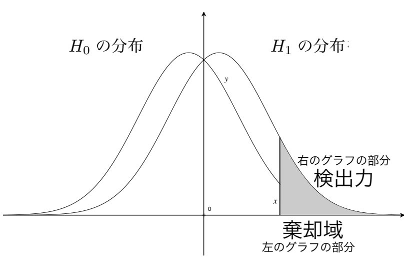


図 23　検出力例 1

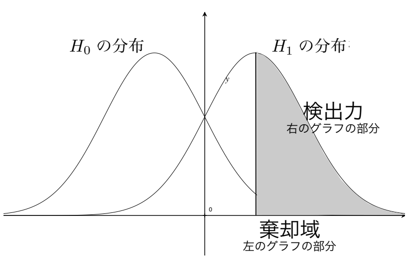


図 24　検出力例 2

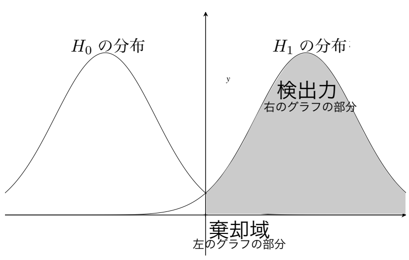


図 25　検出力例 3

検定に際して、有意水準 α を勝手に小さく変更すると第 1 種の過誤が起こる確率は小さくなるが、それに伴って検出力が低下していく。このような取り扱いは時として重大な結果を招く可能性があり、判断を間違うと非常に危険である。なお、有意水準 α を勝手に大きく変更すると検出力は上がるが、一般に何らかのコストが増大することが多い。

4　比率の検定

4.1　２項分布による正規近似のときの標準化変換 (復習)

4.1.1　事前の前提知識

正規近似

■２項分布の正規近似：　確率変数 X が $\mathrm{B}(n,p)$ に従うとき、n が十分大きい（おおむね３０以上）ならば、正規近似によって平均 np、分散 $np(1-p)$ の正規分布 $\mathrm{N}(np, np(1-p))$ に近づく。「第５回講義　**分布関数・母関数及び中心極限定理**」の講義回の通り。

中心極限定理

■その１：　母集団の確率変数 X が「平均 μ、分散 σ^2 のある分布」に従うとき、各標本は、

- $X_1 \to$「平均 μ、分散 σ^2 のある分布」に従う。
- $X_2 \to$「平均 μ、分散 σ^2 のある分布」に従う。
- $\cdots\cdots$
- $X_n \to$「平均 μ、分散 σ^2 のある分布」に従う。

と考えられ、母集団がどんな分布に従っていても中心極限定理により $\bar{X} = \frac{X_1+X_2+X_3+\cdots X_n}{n}$ は、$\mathrm{N}(\mu, \frac{\sigma^2}{n})$ に従うことになる。「第５回講義　**分布関数・母関数及び中心極限定理**」の講義回の通り。

■その２：　さらに、確率変数 X が「平均 μ、分散 σ^2 のある分布」に従うとき、$\bar{X}$ ではなく、X と同一の確率分布に従う n 個のまったく独立な確率変数 X_i の和を確率変数 S とすると、

$$S = X_1 + X_2 + X_3 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

も中心極限定理によって正規分布に近づくが、<u>**その場合の正規分布は $\mathbf{N}(n\mu, n\sigma^2)$ に近づく。**</u>

4.1.2　２項分布の正規近似を考察する

■２項分布：　母集団の確率変数 X が $\mathrm{B}(n,p)$ に従うとき、各標本とは、

- $X_1 \to$ １回目のベルヌーイ試行で、$n = 1$。
- $X_2 \to$ ２回目のベルヌーイ試行で、$n = 2$。
- $\cdots\cdots$
- $X_n \to$ n 回目のベルヌーイ試行で、$n = n$。

である。

　そこで、２項分布について改めて考えてみると、試行回数 n 回ごとに、次のように考えられる。

- $Y_1 \to$ 成功確率 p の１回目のベルヌーイ試行 X_1 そのもの。$Y_1 = X_1$
- $Y_2 \to$ 成功確率 p の２回目のベルヌーイ試行。$Y_2 = X_1 + X_2$
- $\cdots\cdots$
- $Y_n \to$ 成功確率 p の n 回目のベルヌーイ試行。$Y_n = X_1 + X_2 + \cdots + X_n$

このとき Y_i は２項分布の確率変数である。

さらに、この Y_n は「中心極限定理その２」で示した式 (??) の S そのものであるから、$\mathrm{B}(n,p)$ が近づくのは $\mathrm{N}(n\mu, n\sigma^2)$ の正規分布である。つまり、$\mathrm{B}(n,p)$ の n が中心極限定理が成立するほど大きいならば、np が近づくのは $n\mu$ に近づき、$np(1-p)$ が近づくのは $n\sigma^2$ に近づくとも言える。

■２項分布の正規近似の表現： $\mathrm{B}(n,p)$ が n が大きいとき、正規近似によって、正規分布 $\mathrm{N}(np, np(1-p))$ に近づくが、それは、別の表現を用いれば、正規分布は $\mathrm{N}(n\mu, n\sigma^2)$ に近づくと表すこともできる。要するに $\mu = p$ かつ $\sigma^2 = p(1-p)$ である。

■中心極限定理：[再掲] 母集団の確率変数 X が「平均 μ、分散 σ^2 のある分布」に従うとき、各標本は、

- $X_1 \to$「平均 μ、分散 σ^2 のある分布」に従う。
- $X_2 \to$「平均 μ、分散 σ^2 のある分布」に従う。
- $\cdots\cdots$
- $X_n \to$「平均 μ、分散 σ^2 のある分布」に従う。

と考えられ、母集団がどんな分布に従っていても中心極限定理により $\bar{X} = \frac{X_1+X_2+X_3+\cdots X_n}{n}$ は、$\mathrm{N}(\mu, \frac{\sigma^2}{n})$ に従うことになる。

■　つまり、いま、母集団 $\mathrm{B}(n,p)$ が平均 μ、分散 σ^2 であるとき、n が中心極限定理が成立するほど大きいならば、標本平均 $\bar{X} = \frac{X_1+X_2+X_3+\cdots X_n}{n}$ は、$\mathrm{N}(\mu, \frac{\sigma^2}{n})$ に従う。よって、$\mu = p$ かつ $\sigma^2 = p(1-p)$ だから、$\mathrm{N}(p, \frac{p(1-p)}{n})$ に従う。具体的に書くと、

$$f(x) = \frac{1}{\sqrt{2\pi\frac{p(1-p)}{n}}} e^{-\frac{1}{2}\frac{(x-p)^2}{\frac{p(1-p)}{n}}}$$

である。

4.1.3　２項分布による正規近似のときの標準化変換

２項分布 $\mathrm{B}(n,p)$ は試行回数 n が十分大きいなら正規分布 $\mathrm{N}\Big(p, \frac{p(1-p)}{n}\Big)$ で近似できる。その場合は、

$$f(x) = \frac{1}{\sqrt{2\pi\frac{p(1-p)}{n}}} e^{-\frac{1}{2}\frac{(x-p)^2}{\frac{p(1-p)}{n}}}$$

とできる。このとき区間 $[a,b]$ の間の確率は、

$$P(a < x \leq b) = \int_a^b \frac{1}{\sqrt{2\pi\frac{p(1-p)}{n}}} e^{-\frac{1}{2}\frac{(x-p)^2}{\frac{p(1-p)}{n}}} dx$$

このときは、

$$z = \frac{x-p}{\sqrt{\frac{p(1-p)}{n}}}$$

で変数変換を行うと、$\sqrt{\frac{p(1-p)}{n}}dz = dx$ であるから

$$P(a < x \leq b) = \int_{\frac{a-p}{\sqrt{\frac{p(1-p)}{n}}}}^{\frac{b-p}{\sqrt{\frac{p(1-p)}{n}}}} \frac{1}{\sqrt{2\pi}\sqrt{\frac{p(1-p)}{n}}} e^{-\frac{1}{2}z^2} \sqrt{\frac{p(1-p)}{n}} dz = \int_{\frac{a-p}{\sqrt{\frac{p(1-p)}{n}}}}^{\frac{b-p}{\sqrt{\frac{p(1-p)}{n}}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

とできる。これにより被積分関数は、標準正規分布にできる。なお、変数変換で用いた

$$z = \frac{x-p}{\sqrt{\frac{p(1-p)}{n}}}$$

の中にある p はすべて同一の２項分布からの成功確率 p である。

4.2　比率の検定

２項分布 $\mathrm{B}(n,p)$ の試行回数 n が十分大きいなら正規分布 $\mathrm{N}\left(p, \frac{p(1-p)}{n}\right)$ で近似できる。このとき、確率 p の推定値 $\hat{p}$ は標本を用いて

$$\hat{p} = \frac{x}{n}$$

である。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & \text{：帰無仮説}\quad p \text{ は } p_0 \text{とみなせる。} \\ H_1 & \text{：対立仮説}\quad p \text{ は } p_0 \text{とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。いま、分散が既知なので、分散が既知の平均の検定と同じ標準化変換するが、導入する標準化変数は、

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

この Z は、$\mathrm{N}(0,1)$ に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

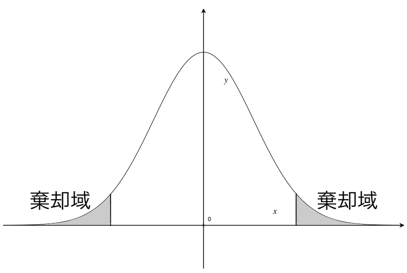


図 26　両側検定

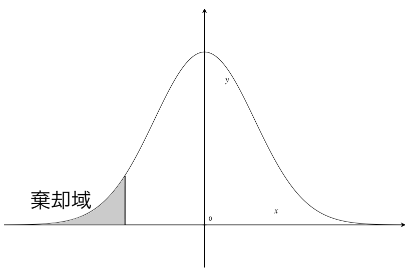


図 27　下側片側検定

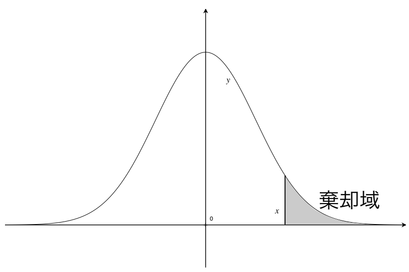


図 28　上側片側検定

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

4.3　二つの集団の比率の差に関する検定

２項分布に従う２つの標本の集団を考える。

- 集団Ｉ：$\mathrm{B}(m, p_1)$
- 集団 II：$\mathrm{B}(n, p_2)$

に従っている。

ここでの目的は、この２つの分布の比率 p_1 と p_2 が等しいかどうかを検定することにある。

いま、集団Ｉにおける母比率 p_1 についての推定値である標本比率 $\hat{p_1} = \frac{x_1}{m}$ は、$\mathrm{N}\left(\hat{p_1}, \frac{\hat{p_1}(1-\hat{p_1})}{m}\right)$ に従い、集団 II における母比率 p_2 についての推定値である標本比率 $\hat{p_2} = \frac{x_2}{n}$ は、$\mathrm{N}\left(\hat{p_2}, \frac{\hat{p_2}(1-\hat{p_2})}{n}\right)$ に従う。

よって、標本集団Ｉの確率変数 $\hat{p_1}$ と標本集団 II の確率変数 $\hat{p_2}$ についての差を新しい確率変数 D とする。すなわち、

$$D = \hat{p_1} - \hat{p_2}$$

そのとき、新しい確率変数 D も正規分布に従い、

$$D \sim \mathrm{N}\left(0, \frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n}\right)$$

と表される。

4.3.1　二つの集団の比率の差の検定

ここでの検定の目的は、この２つの分布の比率が等しいかどうかを検定することにある。

■手順１　仮説をたてる。

$$\begin{cases} H_0 & \text{：帰無仮説}\quad p_1 - p_2 \text{は } 0 \text{ とみなせる。}(p_1 = p_2) \\ H_1 & \text{：対立仮説}\quad p_1 - p_2 \text{は } 0 \text{ とみなせない。（両側検定もしくは片側検定）} \end{cases}$$

■手順２　統計量を導入する。比率の検定と同様に標準化変換するが、この場合の導入する標準化変数は、

$$Z = \frac{\hat{p_1} - \hat{p_2} - 0}{\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{m} + \frac{\hat{p_2}(1-\hat{p_2})}{n}}} = \frac{\hat{p_1} - \hat{p_2}}{\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{m} + \frac{\hat{p_2}(1-\hat{p_2})}{n}}}$$

この Z は、$\mathrm{N}(0,1)$ に従う。

■手順３　有意水準 α として導入した統計量について棄却域を求める。

■手順４　手順２および手順３に従って標本から統計量を求め、それが棄却域に含まれるか判断する。

- 棄却域に含まれるならば H_0 を棄却し、H_1 を採択する。
- 棄却域に含まれないならば H_0 を採択する。

索引