統計的仮説検定

目次

本スライドの内容

このスライドは，次の書籍の第 7 章「統計的仮説検定」の内容に基づく．

- 『ガイダンス 確率統計：基礎から学び本質の理解へ』，発行：サイエンス社，ISBN：978-4-7819-1526-5．

書籍に関する最新の情報は，以下の URL から入手することができます．

`https://www.saiensu.co.jp`

この URL は，サイエンス社が運営しているホームページです．

概要

統計的仮説検定とは，母集団分布に対する仮説を立て，その仮説が妥当か否かを，母集団から得られた標本データを用いて統計的に検証する方法である．このスライドでは，統計的仮説検定における論証方法の概要と，様々な仮説に応じた標準的な検定方法を紹介する．

はじめに

統計的仮説検定とは，母集団分布のパラメータに関する「疑わしい仮説」が正しいか否かを，母集団から得られた標本データに基づいて判断する統計手法である．統計的仮説検定は，単に検定とよばれることも多いため，以降では検定とよぶ．検定によって否定したい「疑わしい仮説」は帰無仮説とよばれ，帰無仮説が正しいか否かを判断するために利用する統計量は検定統計量とよばれる．検定では，(確率論的な) 背理法の考え方を用いる．すなわち，検定では，「帰無仮説が正しいと仮定した上で，標本データに基づいて計算した検定統計量の実現値が，確率的に現れにくい “極端な値” であれば，最初に仮定した帰無仮説が正しくないと考える」という背理法の論理に基づいて，統計学的観点から帰無仮説が誤りであると主張することを意図している．

検定における論証方法の概要 (1)

母数を θ，母集団分布を D_θ と表し，$X_1, X_2, \cdots, X_n$ は D_θ 母集団からの大きさ n の無作為標本とする．母数 θ が取り得る値全体からなる集合を Θ で表し，パラメータ空間とよぶ．

Θ を，部分集合 Θ_0 と $\Theta_1 = \Theta \setminus \Theta_0$ に分割し，帰無仮説は「母数 θ が Θ_0 に含まれる」($\theta \in \Theta_0$) と表されるとする．このとき，帰無仮説は $H_0 : \theta \in \Theta_0$ と略記する．一方で，「母数 θ が Θ_1 に含まれる」($\theta \in \Theta_1$) という，帰無仮説と反対の仮説は対立仮説とよばれ，$H_1 : \theta \in \Theta_1$ と略記する．

$$\text{帰無仮説 } H_0 : \theta \in \Theta_0, \quad \text{対立仮説 } H_1 : \theta \in \Theta_1.$$

検定における論証方法の概要 (2)

以降では，母数 θ が実数であり，Θ_0 が 1 点 θ_0 からなる単純仮説の場合 ($\Theta_0 = \{\theta_0\}$) で考察する．この場合，帰無仮説は $H_0 : \theta = \theta_0$ と表され，対立仮説 H_1 は次の 2 つのいずれかであることが多い．

$$H_1 : \theta \neq \theta_0, \tag{7.1}$$

$$H_1 : \theta > \theta_0 \quad (\text{または } H_1 : \theta < \theta_0). \tag{7.2}$$

なお，(7.1) の場合は両側検定とよばれ，(7.2) の場合は片側検定とよばれる．パラメータ空間 Θ と対立仮説 H_1 に応じて，Θ_1 は次のように表せる．

$$\begin{aligned}
\Theta = (0, \infty), &\quad H_1 : \theta \neq \theta_0 &\iff&\quad \Theta_1 = (0, \theta_0) \cup (\theta_0, \infty), \\
\Theta = (0, \theta_0], &\quad H_1 : \theta < \theta_0 &\iff&\quad \Theta_1 = (0, \theta_0), \\
\Theta = [\theta_0, \infty), &\quad H_1 : \theta > \theta_0 &\iff&\quad \Theta_1 = (\theta_0, \infty).
\end{aligned}$$

検定における論証方法の概要 (3)

まず，検定統計量 $T_n = T(X_1, X_2, \cdots, X_n)$ を適切に決める．次に，「H_1 のもとでの現れやすさと相対的に比較すると，H_0 のもとでは現れにくい T_n の実現値の範囲」を棄却域とよび，記号 W で表す．棄却域 W の決め方の詳細は後程解説する．「帰無仮説 H_0 のもとで検定統計量 T_n が棄却域 W に含まれる確率」は有意水準とよばれる．有意水準 α $(0 < \alpha < 1)$ は小さい値を想定し，0.1, 0.05 や 0.01 が習慣的によく使われる．以上より，検定統計量 T_n，棄却域 W および有意水準 α は，次の関係式をみたす．

$$P(T_n \in W|\theta_0) = \alpha. \qquad (7.3)$$

ただし，左辺の記号 $P(T_n \in W|\theta_0)$ は，条件付き確率ではなく，「母数 θ_0 のもとでの確率 $P(T_n \in W)$」を表す．検定統計量 T_n の決め方は，個々の検定方法に応じて解説する．

検定における論証方法の概要 (4)

最後に，標本変量 $X_1, X_2, \cdots, X_n$ の標本データ $x_1, x_2, \cdots, x_n$ を用いて検定統計量 T_n の実現値 $t_n = T(x_1, x_2, \cdots, x_n)$ を計算し，この実現値 t_n が棄却域 W に含まれるか否かを判定すると，次のいずれかの結論を得る．

1. t_n が棄却域 W に含まれるとき $(t_n \in W)$，「H_1 のもとでの現れやすさと相対的に比較すると，H_0 のもとでは現れにくい実現値 t_n が得られた」と考える．よって，このとき「H_1 と比較して相対的に H_0 は誤りであり，H_0 より H_1 を支持する」と結論付け，このことを「有意水準 $100\alpha\%$ で帰無仮説 H_0 を棄却する」という．
2. t_n が棄却域 W に含まれないとき $(t_n \notin W)$，「H_1 と比較して相対的に H_0 が誤りである」とは判断できない．このことを「有意水準 $100\alpha\%$ で帰無仮説 H_0 を受容する」という．なお，H_0 を受容することは，H_0 が正しいことを意味するものではない．

このように，H_0 を棄却するか，受容するかの判定を行うことを「有意水準 $100\alpha\%$ で検定を行う」という．なお，有意水準 α が小さいほど，棄却域 W は狭くなり，H_0 は棄却されにくくなる．

検定統計量の尤度比関数と棄却域 (1)

母数 θ のもとでの検定統計量 T_n の分布が，密度関数 $f_\theta(x)$ から定まる分布か，離散分布かに応じて，T_n の尤度関数 $L(\theta|x)$ をそれぞれ

$$L(\theta|x) = f_\theta(x), \quad L(\theta|x) = P(T_n = x|\theta)$$

と定義する．ただし，記号 $P(T_n = x|\theta)$ は，条件付き確率ではなく，「母数 θ のもとでの確率 $P(T_n = x)$」を表す．T_n の尤度関数 $L(\theta|x)$ の値の大きさは，「母数 θ のもとでの T_n の実現値 x の現れやすさ」を表す．

「帰無仮説 H_0 のもとでは T_n の実現値が現れにくい範囲」W_0 は，「$L(\theta_0|x)$ の値が小さい x の集合」である．そのため，たとえば定数 $c_0 > 0$ を用いて W_0 を

$$W_0 = \{x \mid L(\theta_0|x) \leq c_0\} \qquad (7.4)$$

と定義する決め方があり得る．

検定統計量の尤度比関数と棄却域 (2)

「対立仮説 H_1 のもとでは T_n の実現値が現れにくい範囲」W_1 は，「どのように母数 $\theta \in \Theta_1$ を選んでも，母数 θ のもとでは T_n の実現値として表れにくい x の集合」である．つまり，W_1 は「どのように $\theta \in \Theta_1$ を選んでも $L(\theta|x)$ の値が小さい x の集合」である．よって，W_1 は「$\sup_{\theta\in\Theta_1} L(\theta|x)$ の値が小さい x の集合」である．ここで，$\sup_{\theta\in\Theta_1} L(\theta|x)$ は，集合 $\{L(\theta|x) \mid \theta \in \Theta_1\}$ の上限とよばれる．なお，最大値 $\max_{\theta\in\Theta_1} L(\theta|x)$ が存在するとき，上限 $\sup_{\theta\in\Theta_1} L(\theta|x)$ と最大値 $\max_{\theta\in\Theta_1} L(\theta|x)$ は一致する．そのため，たとえば定数 $c_1 > 0$ を用いて W_1 を

$$W_1 = \left\{ x \;\middle|\; \sup_{\theta\in\Theta_1} L(\theta|x) \leq c_1 \right\} \qquad (7.5)$$

と定義する決め方があり得る．

検定統計量の尤度比関数と棄却域 (3)

棄却域 W は「H_1 のもとでの現れやすさと相対的に比較すると，H_0 のもとでは現れにくい T_n の実現値の範囲」であった．本項でのこれまでの議論から，

$$T_n \text{ の尤度比関数}: \Lambda(x) = \frac{L(\theta_0|x)}{\sup_{\theta\in\Theta} L(\theta|x)} \qquad (\Theta = \{\theta_0\} \cup \Theta_1)$$

は，「H_0 のもとでの T_n の実現値 x の現れやすさ」を相対的に表す．したがって，棄却域 W とは「T_n の尤度比関数 $\Lambda(x)$ の値が小さい x の集合」である．T_n の尤度比関数 $\Lambda(x)$ の取り得る値は 0 以上かつ 1 以下の実数である．そのため，たとえば定数 c $(0 < c < 1)$ を用いて棄却域 W を

$$W = \left\{ x \;\middle|\; \frac{L(\theta_0|x)}{\sup_{\theta\in\Theta} L(\theta|x)} \leq c \right\} \qquad (7.7)$$

と定義する決め方があり得る．

正規母集団の母平均の検定：はじめに

正規母集団 $N(\mu, \sigma^2)$ からの大きさ n の無作為標本 $X_1, X_2, \cdots, X_n$ を考える．また，既知の定数 μ_0 に対し帰無仮説 H_0 が

$$H_0 : \mu = \mu_0$$

で与えられ，対立仮説 H_1 が次のいずれかで与えられる場合を考える．

$$\text{(両側検定)}\quad H_1 : \mu \neq \mu_0, \tag{7.9}$$

$$\text{(片側検定)}\quad H_1 : \mu > \mu_0 \quad (\text{または } H_1 : \mu < \mu_0). \tag{7.10}$$

定理 6.1.1 と定理 6.1.2 より，統計量の分布に関して次が成り立つ．

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \sim N(0,1), \quad \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{U_n^2}} \sim t(n-1). \tag{7.11}$$

例 7.2.1 と例 7.2.2 では，有意水準 α $(0 < \alpha < 1)$ は固定し，検定統計量 T_n を適切に決め，対立仮説 (7.9) と (7.10) のそれぞれについて，対応する棄却域 W の決め方を解説する．

例 7.2.1

ここでは，検定統計量 T_n と，$N(0,1)$ に従う確率変数 $\widetilde{T}_n$ を

$$T_n = \frac{\sqrt{n}(\overline{X}_n - \mu_0)}{\sigma}, \quad \widetilde{T}_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \sim N(0,1) \qquad (7.12)$$

と定める．このとき，H_0 のもとで $T_n = \widetilde{T}_n$ が成り立つため，H_0 のもとで T_n は $N(0,1)$ に従う．母平均 μ のもとでの T_n の密度関数を記号 $f_\mu(x)$ で表すと，$f_\mu(x)$ は

$$f_\mu(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left(x - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma} \right)^2 \right\} \qquad (7.14)$$

と計算することができる．

例 7.2.1 (1) H_1 : $\mu \neq \mu_0$ の場合

対立仮説が H_1 : $\mu \neq \mu_0$ の場合，(7.14) より，T_n の尤度比関数は

$$\Lambda(x) = \frac{f_{\mu_0}(x)}{\sup_{\mu \in \mathbb{R}} f_\mu(x)} = e^{-\frac{x^2}{2}} \qquad (x \in \mathbb{R})$$

と計算できる．よって，$0 < c < 1$ を用いて棄却域 W を

$$W = \{x \in \mathbb{R} \mid \Lambda(x) \leq c\} = \left(-\infty, -\sqrt{-2\log c}\right] \cup \left[\sqrt{-2\log c}, \infty\right)$$

と表すと，関係式 (7.3) より，$\sqrt{-2\log c} = z_{\alpha/2}$ となる：

$$W = (-\infty, -z_{\alpha/2}] \cup [z_{\alpha/2}, \infty). \qquad (7.18)$$

例 7.2.1 (2) H_1 : $\mu > \mu_0$ の場合

対立仮説が H_1 : $\mu > \mu_0$ の場合，(7.14) より，T_n の尤度比関数は

$$\Lambda(x) = \frac{f_{\mu_0}(x)}{\sup_{\mu \geq \mu_0} f_\mu(x)} = \begin{cases} e^{-\frac{x^2}{2}} & (x > 0) \\ 1 & (x \leq 0) \end{cases}$$

と計算できる．よって，$0 < c < 1$ を用いて棄却域 W を

$$W = \{x \in \mathbb{R} \mid \Lambda(x) \leq c\} = \left[\sqrt{-2\log c}, \infty\right)$$

と表すと，関係式 (7.3) より，$\sqrt{-2\log c} = z_\alpha$ となる：

$$W = [z_\alpha, \infty). \qquad (7.19)$$

例 7.2.1 (3) H_1 : $\mu < \mu_0$ の場合

対立仮説が H_1 : $\mu < \mu_0$ の場合，(7.14) より，T_n の尤度比関数は

$$\Lambda(x) = \frac{f_{\mu_0}(x)}{\sup_{\mu \leq \mu_0} f_\mu(x)} = \begin{cases} 1 & (x > 0) \\ e^{-\frac{x^2}{2}} & (x \leq 0) \end{cases}$$

と計算できる．よって，$0 < c < 1$ を用いて棄却域 W を

$$W = \{x \in \mathbb{R} \mid \Lambda(x) \leq c\} = \left(-\infty, -\sqrt{-2\log c}\right]$$

と表すと，関係式 (7.3) より，$-\sqrt{-2\log c} = -z_\alpha$ となる：

$$W = (-\infty, -z_\alpha]. \qquad (7.20)$$

例 7.2.2

ここでは，$T_n = \sqrt{n}(\overline{X}_n - \mu_0)/\sqrt{U_n^2}$ と定める．H_0 のもとで

$$T_n = \sqrt{n}(\overline{X}_n - \mu)/\sqrt{U_n^2} \sim t(n-1)$$

が成り立つ．例 7.2.1 と同様に，対立仮説 (7.9) と (7.10) のそれぞれについて，対応する棄却域 W を次のように決める．

$$H_1 : \mu \neq \mu_0 \quad \Longleftrightarrow \quad W = \left(-\infty, -t_{\alpha/2}^{(n-1)}\right] \cup \left[t_{\alpha/2}^{(n-1)}, \infty\right), \tag{7.21}$$

$$H_1 : \mu > \mu_0 \quad \Longleftrightarrow \quad W = \left[t_{\alpha}^{(n-1)}, \infty\right), \tag{7.22}$$

$$H_1 : \mu < \mu_0 \quad \Longleftrightarrow \quad W = \left(-\infty, -t_{\alpha}^{(n-1)}\right]. \tag{7.23}$$

注意 7.2.1

棄却域 W が定まり，標本変量 $X_1, X_2, \cdots, X_n$ の標本データ $x_1, x_2, \cdots, x_n$ が得られたとする．

このとき，$\overline{X}_n$ や U_n^2 の実現値を具体的に計算できるため，例 7.2.2 の検定統計量 T_n の実現値も具体的に計算できる．

一方で，母分散 σ^2 の値が未知の場合は，例 7.2.1 の検定統計量 T_n の実現値を具体的に計算できない．そのため，σ^2 の値が未知の場合は，例 7.2.1 で紹介した方法を用いても，帰無仮説 H_0 を棄却するか，または受容するかを判定できない．

例題 7.2.1

例題 7.2.1

ある部屋に設置された空調システムでは，設定温度を 25.0 度として作動したとき，室内温度 X は正規分布 $N(\mu, \sigma^2)$ に従うとする．7 日間にわたり，設定温度を 25.0 度として作動し，室内温度 $X_1, X_2, \cdots, X_7$ を測定したところ，室内温度の標本平均 $\overline{X}_7$ の実現値は 25.21 であり，不偏標本分散の正の平方根 $\sqrt{U_7^2}$ の実現値は 0.715 であった．このとき，この空調システムは正しく作動しているといえるか，有意水準 5% で検定せよ．ただし，$X_1, X_2, \cdots, X_7$ は独立であるとする．

例題 7.2.1

[解答]　$\mu_0 = 25.0$ とおく．空調システムが正しく作動しているとは，$\mu = \mu_0$ が成り立つことを意味し，正しく作動していないとは，$\mu \neq \mu_0$ であることを意味する．よって，

$$\text{帰無仮説 } H_0 : \mu = \mu_0, \quad \text{対立仮説 } H_1 : \mu \neq \mu_0$$

について，$T_7 = \sqrt{7}(\overline{X}_7 - \mu_0)/\sqrt{U_7^2}$ を用いて，有意水準 5% で検定する．まず，(7.21) と表 C.3 より，棄却域 W は

$$W = \left(-\infty, -t_{0.025}^{(6)}\right] \cup \left[t_{0.025}^{(6)}, \infty\right) = (-\infty, -2.447] \cup [2.447, \infty)$$

である．一方で，検定統計量 T_7 の実現値 t_7 は

$$t_7 = \frac{\sqrt{7}(25.21 - 25)}{0.715} = 0.777$$

と計算できる．このとき，$t_7 \notin W$ であるため，帰無仮説 H_0 を受容する．したがって，この空調システムが正しく作動していないとはいえない．

正規母集団の母分散の検定：はじめに

正規母集団 $N(\mu, \sigma^2)$ からの大きさ n の無作為標本 $X_1, X_2, \cdots, X_n$ を考える．また，既知の定数 σ_0^2 に対し帰無仮説 H_0 が

$$H_0 : \sigma^2 = \sigma_0^2$$

で与えられ，対立仮説 H_1 が次のいずれかで与えられる場合を考える．

$$\text{(両側検定)} \quad H_1 : \sigma^2 \neq \sigma_0^2, \tag{7.24}$$

$$\text{(片側検定)} \quad H_1 : \sigma^2 > \sigma_0^2 \quad (\text{または } H_1 : \sigma^2 < \sigma_0^2). \tag{7.25}$$

定理 6.1.1 と定理 6.1.2 より，統計量の分布に関して次が成り立つ．

$$\frac{n\widehat{s}_n^2}{\sigma^2} \sim \chi^2(n) \quad (n \geq 1), \quad \frac{(n-1)U_n^2}{\sigma^2} \sim \chi^2(n-1) \quad (n \geq 2). \tag{7.26}$$

次の例 7.3.1 と例 7.3.2 では，有意水準 α $(0 < \alpha < 1)$ は固定し，検定統計量 T_n を適切に選び，対立仮説 (7.24) と (7.25) のそれぞれについて，対応する棄却域 W の決め方を解説する．

例 7.3.1

ここでは，$n \geq 2$ とし，検定統計量 T_n と確率変数 $\widetilde{T}_n$ を

$$T_n = \frac{(n-1)U_n^2}{\sigma_0^2}, \quad \widetilde{T}_n = \frac{(n-1)U_n^2}{\sigma^2} \sim \chi^2(n-1) \qquad (7.27)$$

と定める．このとき，H_0 のもとで $T_n = \widetilde{T}_n$ が成り立つため，H_0 のもとで検定統計量 T_n は $\chi^2(n-1)$ に従う．$m = n-1$ とおき，母分散 σ^2 のもとでの T_n の密度関数を記号 $f_{\sigma^2}^{(m)}(x)$ $(x > 0)$ で表すと，

$$f_{\sigma^2}^{(m)}(x) = \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} x^{\frac{m}{2}-1} g_x\left(\frac{\sigma_0^2}{\sigma^2}\right) \quad (x > 0) \qquad (7.29)$$

が得られる．ただし，$x > 0$ に対して，関数 $g_x(c)$ を

$$g_x(c) = c^{\frac{m}{2}} \exp\left\{-\frac{xc}{2}\right\} \quad (c > 0) \qquad (7.30)$$

と定義した．

例 7.3.1

また，準備のため，関数 $h_m(x)$ を

$$h_m(x) = \left(\frac{x}{m}\right)^{\frac{m}{2}} e^{\frac{m-x}{2}} \quad (x > 0)$$

と定義する．この関数 $h_m(x)$ の導関数 $\frac{d}{dx}h_m(x)$ と符号について，

$$\frac{d}{dx}h_m(x) = \frac{m-x}{2x}\left(\frac{x}{m}\right)^{\frac{m}{2}} e^{\frac{m-x}{2}} \quad (x > 0),$$
$$\frac{d}{dx}h_m(x) > 0 \quad (0 < x < m) \quad \text{および} \quad \frac{d}{dx}h_m(x) < 0 \quad (m < x)$$

が成り立つため，$h_m(x)$ は $0 < x < m$ で単調に増加し，$x > m$ で単調に減少し，$h_m(m) = 1$ が極大値である．また，$h_m(0) = 0$ が成り立つ．次に，ロピタルの定理を複数回適用すると，$\lim_{x\to\infty} h_m(x) = 0$ も得られる．

例 7.3.1 (1) $H_1 : \sigma^2 \neq \sigma_0^2$ の場合

対立仮説が $H_1 : \sigma^2 \neq \sigma_0^2$ のとき，(7.29) より，T_n の尤度比関数は

$$\Lambda(x) = \frac{f_{\sigma_0^2}^{(m)}(x)}{\sup_{\sigma^2>0} f_{\sigma^2}^{(m)}(x)} = h_m(x) \quad (x > 0) \qquad (7.38)$$

と計算できる．よって，$0 < r_1 < m < r_2 < \infty$ をみたす実数 r_1, r_2 を用いて，棄却域を $W = (0, r_1] \cup [r_2, \infty)$ と決めればよい．ここで，表 C.4 より，有意水準 α が $0 < \alpha \leq 0.2$ をみたすとき，不等式

$$c_{1-\alpha/2}^{(n-1)} < m = n - 1 < c_{\alpha/2}^{(n-1)}$$

が成り立つ．したがって，このとき関係式 (7.3) より，$r_1 = c_{1-\alpha/2}^{(n-1)}$ と $r_2 = c_{\alpha/2}^{(n-1)}$ を採用し，棄却域 W を次のように決める．

$$W = \left(0, c_{1-\alpha/2}^{(n-1)}\right] \cup \left[c_{\alpha/2}^{(n-1)}, \infty\right). \qquad (7.39)$$

例 7.3.1 (2) $H_1 : \sigma^2 > \sigma_0^2$ の場合

対立仮説が $H_1 : \sigma^2 > \sigma_0^2$ のとき，(7.29) より，T_n の尤度比関数は

$$\Lambda(x) = \frac{f_{\sigma_0^2}^{(m)}(x)}{\sup_{\sigma^2 \geq \sigma_0^2} f_{\sigma^2}^{(m)}(x)} = \begin{cases} 1 & (0 < x \leq m) \\ h_m(x) & (x > m) \end{cases} \qquad (7.40)$$

と計算できる．よって，実数 $r\ (> m)$ を用いて，$[r, \infty)$ を棄却域 W として決めればよい．表 C.4 より，有意水準 α が $0 < \alpha \leq 0.1$ をみたすとき，$m = n - 1 < c_\alpha^{(n-1)}$ が成り立つ．したがって，このとき関係式 (7.3) より，$r = c_\alpha^{(n-1)}$ を採用し，棄却域 W を次のように決める．

$$W = \left[c_\alpha^{(n-1)}, \infty \right). \qquad (7.41)$$

例 7.3.1 (3) $H_1 : \sigma^2 < \sigma_0^2$ の場合

対立仮説が $H_1 : \sigma^2 < \sigma_0^2$ のとき，(7.29) より，T_n の尤度比関数は

$$\Lambda(x) = \frac{f_{\sigma_0^2}^{(m)}(x)}{\sup_{0<\sigma^2\le\sigma_0^2} f_{\sigma^2}^{(m)}(x)} = \begin{cases} h_m(x) & (0 < x \le m) \\ 1 & (x > m) \end{cases} \qquad (7.42)$$

と計算できる．よって，$0 < r < m$ をみたす実数 r を用いて，$(0, r]$ を棄却域 W として決めればよい．表 C.4 より，有意水準 α が $0 < \alpha \le 0.1$ をみたすとき，$c_{1-\alpha}^{(n-1)} < m = n-1$ が成り立つ．したがって，このとき関係式 (7.3) より，$r = c_{1-\alpha}^{(n-1)}$ を採用し，棄却域 W を次のように決める．

$$W = \left(0, c_{1-\alpha}^{(n-1)}\right]. \qquad (7.43)$$

例題 7.3.1

例題 7.3.1

ある機械で生産される製品の重量 X は正規分布 $N(\mu, \sigma^2)$ に従うとする．あるとき無作為に 10 個の製品を抽出して重量 $X_1, X_2, \cdots, X_{10}$ を測定したところ，不偏標本分散の正の平方根 $\sqrt{U_{10}^2}$ の実現値が 1.535 g であった．このとき，母標準偏差 σ は 1.0 g より大きいと考えてよいか，有意水準 5% で検定せよ．

例題 7.3.1

[解答]　$\sigma_0 = 1.0$ とおく．帰無仮説 $H_0 : \sigma = \sigma_0$ と対立仮説 $H_1 : \sigma > \sigma_0$ について，検定統計量

$$T_{10} = (10-1)U_{10}^2/\sigma_0^2$$

を用いて有意水準 5% で検定する．まず，(7.41) と表 C.4 より，棄却域は

$$W = [c_{0.05}^{(9)}, +\infty) = [16.92, +\infty)$$

である．一方で，検定統計量 T_{10} の実現値 t_{10} は

$$t_{10} = 9 \cdot (1.535)^2/(1.0)^2 = 21.21$$

と計算できる．このとき，$t_{10} \in W$ であるため，帰無仮説 H_0 を棄却する．したがって，母標準偏差 σ は 1.0 g より大きいと考えてよい．

二項母集団の母比率の検定

二項母集団 $Be(p)$ からの大きさ n の無作為標本 $X_1, X_2, \cdots, X_n$ を考える．また，既知の定数 p_0 $(0 < p_0 < 1)$ に対し帰無仮説 H_0 が

$$H_0 : p = p_0$$

で与えられ，対立仮説 H_1 が次のいずれかで与えられる場合を考える．

$$(\text{両側検定}) \quad H_1 : p \neq p_0, \tag{7.47}$$

$$(\text{片側検定}) \quad H_1 : p > p_0 \quad (\text{または } H_1 : p < p_0). \tag{7.48}$$

以下では $\overline{p}_n := \overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ とおく．なお，$\overline{p}_n$ は標本比率ともよばれる．このとき，例 3.2.7 より，$E(\overline{p}_n) = p$ かつ $V(\overline{p}_n) = p(1-p)/n$ である．よって，$\overline{p}_n$ の標準化は次のように計算できる．

$$\frac{\overline{p}_n - E(\overline{p}_n)}{\sqrt{V(\overline{p}_n)}} = \frac{\sqrt{n}(\overline{p}_n - p)}{\sqrt{p(1-p)}} = \frac{X_1 + X_2 + \cdots + X_n - np}{\sqrt{np(1-p)}}.$$

二項母集団の母比率の検定

したがって，定理 5.2.1 (中心極限定理) より，任意の $a < b$ に対して，n が十分大きければ，近似式

$$P\left(a \leq \frac{\sqrt{n}(\overline{p}_n - p)}{\sqrt{p(1-p)}} \leq b\right) \approx \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \qquad (7.49)$$

が成り立つ．ここで，検定統計量 T_n と確率変数 $\widetilde{T}_n$ を

$$T_n = \frac{\sqrt{n}(\overline{p}_n - p_0)}{\sqrt{p_0(1-p_0)}}, \quad \widetilde{T}_n := \frac{\sqrt{n}(\overline{p}_n - p)}{\sqrt{p(1-p)}} \qquad (7.50)$$

と定める．このとき，(7.49) より，n が十分大きいとき，$\widetilde{T}_n$ の分布の形は，標準正規分布 $N(0, 1)$ の形に近づく．以下では n は十分大きいと仮定し，有意水準 α $(0 < \alpha < 1)$ を固定し，対立仮説 (7.47) と (7.48) のそれぞれについて，対応する棄却域 W の決め方を解説する．

二項母集団の母比率の検定

まず，H_0 のもとで $T_n = \widetilde{T}_n$ が成り立つため，H_0 のもとで T_n の分布の形は，$N(0,1)$ の形に近い．よって，母比率 p_0 のもとでの T_n の密度関数 $f_{p_0}(x)$ が，次式 (7.51) で与えられるとする．

$$f_{p_0}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (x \in \mathbb{R}). \qquad (7.51)$$

このとき，2 つの関数 $c(p)$ と $d_n(p)$ を

$$c(p) := \frac{\sqrt{p(1-p)}}{\sqrt{p_0(1-p_0)}}, \quad d_n(p) := \frac{\sqrt{n}(p-p_0)}{\sqrt{p_0(1-p_0)}}$$

と定めると，母比率 p のもとでの T_n の密度関数 $f_p(x)$ は

$$f_p(x) = \frac{1}{\sqrt{2\pi} \cdot c(p)} \exp\left(-\frac{(x-d_n(p))^2}{2c(p)^2}\right) \qquad (7.52)$$

で与えられる．

二項母集団の母比率の検定：H_1 ：$p \neq p_0$ の場合

対立仮説が H_1 ：$p \neq p_0$ のとき，(7.52) より，実数 x に対して n が十分大きければ，T_n の尤度比関数は次のように近似計算できる．

$$\Lambda(x) = \frac{f_{p_0}(x)}{\sup_{p \in (0,1)} f_p(x)} \approx e^{-\frac{x^2}{2}}. \qquad (7.59)$$

よって，$0 < c < 1$ を用いて棄却域 W を

$$W = \{x \in \mathbb{R} \mid \Lambda(x) \leq c\} = \left(-\infty, -\sqrt{-2\log c}\right] \cup \left[\sqrt{-2\log c}, \infty\right)$$

と表すと，関係式 (7.3) より，$\sqrt{-2\log c} = z_{\alpha/2}$ となる：

$$W = (-\infty, -z_{\alpha/2}] \cup [z_{\alpha/2}, \infty). \qquad (7.60)$$

二項母集団の母比率の検定：H_1 ： $p > p_0$ の場合

対立仮説が H_1 ： $p > p_0$ のとき，(7.52) より，実数 x に対して n が十分大きければ，T_n の尤度比関数は次のように近似計算できる．

$$\Lambda(x) = \frac{f_{p_0}(x)}{\sup_{p \in [p_0, 1)} f_p(x)} \approx \begin{cases} e^{-\frac{x^2}{2}} & (x > 0) \\ 1 & (x \leq 0). \end{cases} \qquad (7.61)$$

よって，$0 < c < 1$ を用いて棄却域 W を

$$W = \{x \in \mathbb{R} \mid \Lambda(x) \leq c\} = \left[\sqrt{-2 \log c}, \infty\right)$$

と表すと，関係式 (7.3) より，$\sqrt{-2 \log c} = z_\alpha$ となる：

$$W = [z_\alpha, \infty). \qquad (7.62)$$

二項母集団の母比率の検定：H_1 ： $p < p_0$ の場合

対立仮説が H_1 ： $p < p_0$ のとき，(7.52) より，実数 x に対して n が十分大きければ，T_n の尤度比関数は次のように近似計算できる．

$$\Lambda(x) = \frac{f_{p_0}(x)}{\sup_{p \in (0, p_0]} f_p(x)} \approx \begin{cases} 1 & (x > 0) \\ e^{-\frac{x^2}{2}} & (x \leq 0). \end{cases} \tag{7.63}$$

よって，$0 < c < 1$ を用いて棄却域 W を

$$W = \{x \in \mathbb{R} \mid \Lambda(x) \leq c\} = \left(-\infty, -\sqrt{-2\log c}\right]$$

と表すと，関係式 (7.3) より，$-\sqrt{-2\log c} = -z_\alpha$ となる：

$$W = (-\infty, -z_\alpha]. \tag{7.64}$$

例題 7.4.1

例題 7.4.1

不良率が 0.05 であるとされていた工程を，不良品が少なくなるように対策した．対策後に改善されたか否かを調べるため無作為に 500 個抽出して調べたところ，不良品が 15 個あった．このとき，この対策により工程は改善されたといえるか，有意水準 5% で検定せよ．

例題 7.4.1

[解答]　工程が改善されたとは，対策後の真の不良率 p が $p_0 = 0.05$ より小さいことを意味する．よって，帰無仮説 $H_0 : p = p_0$ と対立仮説 $H_1 : p < p_0$ について，500 個の無作為抽出における不良率 $\overline{p}_{500}$ と，(7.50) で定めた検定統計量 T_{500} を用いて，有意水準 5% で検定する．まず，(7.64) と表 C.2 より，棄却域は

$$W = (-\infty, -z_{0.05}] = (-\infty, -1.645].$$

一方で，$\overline{p}_{500}$ の実現値が $15/500$ であるため，T_{500} の実現値 t_{500} は

$$t_{500} = \frac{\sqrt{500}(\frac{15}{500} - 0.05)}{\sqrt{0.05(1-0.05)}} = -2.052$$

と計算できる．このとき，$t_{500} \in W$ であるため，帰無仮説 H_0 を棄却する．したがって，この対策により工程は改善されたといえる．