

データサイエンスのための統計学 仮説検定

by 尚晋 (名古屋大学経済学研究科助教)

on 2025 年 7 月 01 日

仮説検定

- * 検定の考え方
- * 正規母集団に対する仮説検定
- * χ^2 を用いる検定
- * 中心極限定理による検定

仮説検定

- * 検定の考え方
- * 正規母集団に対する仮説検定
- * χ^2 を用いる検定
- * 中心極限定理による検定

» 仮説検定とは

仮説検定 統計的仮説の有意性を検定する手法である。

仮説検定の目的 母集団について仮定された命題を、標本にもとづいて
検証することである。

統計的仮説 ここで立てられた理論上の仮説を **統計的仮説**, あるいは単に**仮説**(hypothesis)という。即ち、**母集団に関する命題**を定式化したものである。

有意 標本からの観測値は理論値 (即ち母集団仮説に基づく期待値) からの**ずれが誤差の範囲外**なら、統計学では**仮説からのずれ** (簡単に、**仮説**) は**有意**(significant) であるという。

例:メンデルのエンドウ豆実験 (理論比 9:3:3:1) と観測値の比較

型	黄色・丸	黄色・しわ	緑・丸	緑・しわ	計
度数	315	101	108	32	556
理論比	9	3	3	1	16

- * 理論比とのズレが**有意かどうか**を検討する
- * **ズレが大きければ仮説を棄却**→ **ずれが有意な結果**

» 仮説検定の考え方

- * 仮説検定は”ずれ”の大きさに着目
- * ずれが標本誤差の範囲なら仮説は採択され、それ以上なら棄却される
- * 「有意である」とは「出るはずのないかなりの稀」のこと
- * 仮説検定 = 仮説を棄却するかどうかを決める判断手続き

例: コイン投げ 20 回で表が 14 回出た時、コインに歪みがないという仮説 $p = 0.5$ が妥当かどうか?

- * 表の回数 $X \sim Bi(20, 1/2)$, 仮説とのずれがどれほど稀を検証:
$$P(X \geq 14) = \sum_{i=14}^{20} {}^{20}C_i p^i q^{20-i} = 0.0577 = 5.8\%$$
- * これは仮説 $p = 0.5$ のもとでは起こりにくい現象 (5.8% しかないから)

「どの程度の確率なら“稀”と判断するか」→ 有意水準 (α)

- * 通常 $\alpha = 0.10, 0.05, 0.01$ など
- * コイン例では $P = 0.0577$:
 - * $\alpha = 0.10$ なら棄却 (稀)
 - * $\alpha = 0.01$ なら棄却されない (十分起こり得る)

» 帰無仮説と対立仮説, 帰無仮説の採択の意味

検定では、2つの仮説を設定：

- * * 帰無仮説 H_0 : 例) $p = 1/2$
- * * 対立仮説 H_1 : 例) $p \neq 1/2$
- * 検定の目的: H_0 を有意水準 α に基づき棄却するか判断
- * 有意水準 α より稀、帰無仮説を棄却した場合、対立仮説を採択したとみなす

ただし「棄却しない」からといって H_0 が真とは限らない

- * 有意性検定は「矛盾が生じたから仮説を棄却する」=背理法
- * 棄却しないからといって、仮説が正しいとは言えない!
 - * 仮説が積極的に正しい(真であることが証明された)というわけではない
 - * 帰無仮説と観測結果が矛盾しなかったというだけ
 - * → 従って、よく「帰無仮説は棄却できない」という

» 検定結果と第 1 種・第 2 種の誤り

表 12.3: 帰無仮説を棄却する,しないの決定に関しての四つの場合

	H_0 が正しい	H_0 が誤り (H_1 正しい)
H_0 を棄却しない (採択する)	①	③(Type II Error)
H_0 を棄却する	②(Type I Error)	④

誤りには以下の 2 種類がある:

- (a) 第 1 種の誤り (Type I Error): H_0 が正しいのに,それを棄却してしまう
誤り \Rightarrow ②
- (b) 第 2 種の誤り (Type II Error): H_0 が誤っているのに,それを採択する
誤り \Rightarrow ③

この構造は品質管理や裁判など現実の意思決定に適用

- * 第 1 種の誤り: 良品を不合格に (生産者のリスク)
- * 第 2 種の誤り: 不良品を合格に (消費者のリスク)

» 棄却域と両側・片側検定

両側検定

例: 標本サイズ $n = 25$, 標本平均 $\bar{X} = 13.7$, 標本標準偏差 $s = 2.3$.

- * 帰無仮説: $H_0 : \mu = 15$
- * 対立仮説: $H_1 : \mu \neq 15$ (両側検定)

t 統計量:

$$t = \frac{13.7 - 15}{2.3/\sqrt{25}} = \frac{-1.3}{0.46} = -2.83$$

有意水準 $\alpha = 0.05$ に対応する t 分布 $t(24)$ のパーセント点は $t_{0.025}(24) = 2.06$ であり, $|t| = 2.83 > 2.06$ より, 有意水準 5% で帰無仮説を棄却する.

なぜなら, ± 2.06 以上となる確率は 5% という稀なものであって, 2.83 はそれ以上に稀だからである.

したがって, 母集団に関する帰無仮説 $\mu = 15$ は成り立つと考えにくい.

» 棄却域と両側・片側検定

片側検定

- * 片側検定: $H'_1: \mu < 15$ (対立仮説が片側)

$$t = \frac{\bar{X} - 15}{s/\sqrt{n}} = -2.83$$

- * 判定: $t_{0.05}(24) = 1.71$, $t = -2.83 < -1.71$ より, 有意水準 5% で帰無仮説を棄却する。

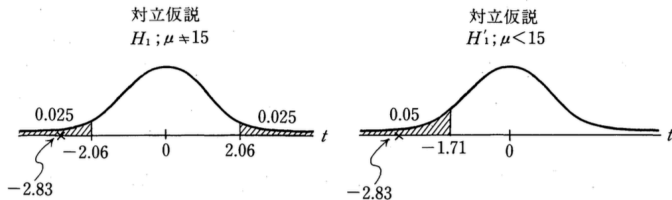


図 12.1 仮説検定のやり方

帰無仮説をまず設け, 対立仮説 H_1 を考慮して, あらかじめ選んだ有意水準(5%)に等しい確率で, 0 から遠い所に棄却域(斜線を付してある数直線の範囲)を作る. 計算された t 統計量の値(-2.83)がその棄却域の中に入るとき, その帰無仮説は棄却される.

» 棄却域と両側・片側検定

片側か両側かの選択

棄却域 帰無仮説を棄却すべき統計量の値の集合

採択域 棄却しない領域

両側検定 $|t| > t_{\alpha/2}(n-1)$

左片側検定 $t < -t_{\alpha}(n-1)$

右片側検定 $t > t_{\alpha}(n-1)$

両側検定の応用場面 母数が目標値と異なるかどうかの確認. 例: 製造機械の異常検出.

片側検定の応用場面 効果の有無など, 一方向の変化に注目する場合. 例: 特別授業による得点向上を検証する場合など.

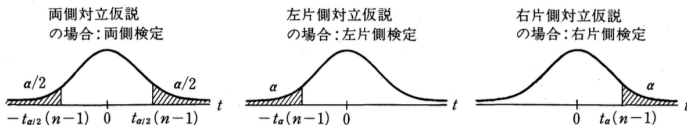


図 12.2 検定の手続例: 母平均は帰無仮説の通りか

スチューデントの t 統計量とは $t = (\bar{X} - \mu_0) / (s / \sqrt{n})$ で定義される (μ_0 はあらかじめ指定された母平均値). 関数は自由度 $n-1$ の t 分布 $t(n-1)$ の密度関数. $t=0$ は標本平均 \bar{X} と帰無仮説の μ_0 の一致を意味する. 斜線の棄却域は \bar{X} と μ_0 の十分なずれがある領域であり, 分布の端部に有意水準で決められる. 片側対立仮説および片側検定は, 右と左の場合がある.

仮説検定

- * 検定の考え方
- * 正規母集団に対する仮説検定
- * χ^2 を用いる検定
- * 中心極限定理による検定

» 母平均に関する検定 (両側)

t 検定

帰無仮説と対立仮説 $H_0: \mu = \mu_0, \quad H_1: \mu \neq \mu_0$

検定統計量 * 母分散 σ^2 が既知:

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \quad (1)$$

$Z \sim N(0, 1)$ のとき $|Z| > z_{\alpha/2}$ で H_0 を棄却

* 母分散 σ^2 が未知:

$$t = \frac{\bar{X} - \mu_0}{s / \sqrt{n}} \quad (2)$$

$t \sim t(n-1)$ のとき $|t| > t_{\alpha/2}(n-1)$ で H_0 を棄却
これをスチューデントの t 検定という。

» 母平均に関する検定（両側）

例

例：空調システムの作動状況－設定温度 25 度の検定

7 日間の温度データ：24.2, 25.3, 26.2, 25.7, 24.4, 25.1, 25.6, 有意水準 5%
でこのシステムが正しく働いているかどうかを検定せよ

» 母平均に関する検定(両側)

例

例: 空調システムの作動状況—設定温度 25 度の検定

7 日間の温度データ: 24.2, 25.3, 26.2, 25.7, 24.4, 25.1, 25.6, 有意水準 5% でこのシステムが正しく働いているかどうかを検定せよ

* 仮説: $H_0 : \mu = 25.0$, $H_1 : \mu \neq 25.0 \rightarrow$ 両側検定

* $\bar{X} = 25.21$, $s = 0.715$, $n = 7$

* 検定統計量:

$$t = \frac{25.21 - 25.0}{0.715/\sqrt{7}} = 0.777$$

* $t_{0.025}(6) = 2.447$ なので, $|t| = 0.777 < 2.447$ より、有意水準 5% では帰無仮説は棄却できない、正常に作動していると判断する。

» 母平均に関する検定 (片側)

例

帰無仮説と対立仮説

* 右片側: $H_0: \mu = \mu_0, \quad H_1: \mu > \mu_0$

* 左片側: $H_0: \mu = \mu_0, \quad H_1: \mu < \mu_0$

検定統計量と判定条件

* σ^2 既知: $Z > z_\alpha$ または $Z < -z_\alpha$

* σ^2 未知: $t > t_\alpha(n-1)$ または $t < -t_\alpha(n-1)$

例: 講義の効果検定

(前 - 後) の得点差: $-1, 3, 4, 5, 3, 0, 7, 4, 2, -2$, 有意水準 5% で講義の効果があったと言えるか

» 母平均に関する検定 (片側)

例

帰無仮説と対立仮説

$$* \text{右片側: } H_0: \mu = \mu_0, \quad H_1: \mu > \mu_0$$

$$* \text{左片側: } H_0: \mu = \mu_0, \quad H_1: \mu < \mu_0$$

検定統計量と判定条件

$$* \sigma^2 \text{ 既知: } Z > z_\alpha \text{ または } Z < -z_\alpha$$

$$* \sigma^2 \text{ 未知: } t > t_\alpha(n-1) \text{ または } t < -t_\alpha(n-1)$$

例: 講義の効果検定

(前 - 後) の得点差: $-1, 3, 4, 5, 3, 0, 7, 4, 2, -2$, 有意水準 5% で講義の効果があったと言えるか

$$* \text{仮説: } H_0: \mu = 0, \quad H_1: \mu > 0 \rightarrow \text{右片側検定}$$

$$* \bar{X} = 2.5, s = 2.5, n = 10$$

* 検定統計量:

$$t = \frac{2.5 - 0}{2.5/\sqrt{10}} = 2.82$$

* $t_{0.05}(9) = 1.833$ なので、 $2.82 > 1.833$ より、有意水準 5% において H_0 は棄却され、講義の効果があったと判断される。

» 母分散に関する検定

カイ二乗検定 (χ^2 -test)

帰無仮説と対立仮説 * 両側: $H_0 : \sigma^2 = \sigma_0^2, \quad H_1 : \sigma^2 \neq \sigma_0^2$
* 片側: $H_0 : \sigma^2 = \sigma_0^2, \quad H_1 : \sigma^2 > \sigma_0^2$, または, $\sigma^2 < \sigma_0^2$

検定統計量

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \sim \chi^2(n-1)$$

両側検定 * $\chi_{1-\alpha/2}^2(n-1) < \chi^2 < \chi_{\alpha/2}^2(n-1)$ のとき H_0 を棄却せず
* それ以外は棄却

片側検定 * 右片側: $\chi^2 > \chi_{\alpha}^2(n-1)$ のとき H_0 棄却
* 左片側: $\chi^2 < \chi_{1-\alpha}^2(n-1)$ のとき H_0 棄却

» 母分散に関する検定

例

例:能力のばらつきの検定

例年の入学テスト (25 人): 平均 50, 分散 36. 本年 (25 人): 平均 53, 分散 48, 有意水準 10% で本年度は児童の揃い方が例年と違うと見てよいか (肥田野)

» 母分散に関する検定

例

例:能力のばらつきの検定

例年の入学テスト (25 人): 平均 50, 分散 36. 本年 (25 人): 平均 53, 分散 48, 有意水準 10% で本年度は児童の揃い方が例年と違うと見てよいか (肥田野)

- * $n = 25$, 有意水準 10%, 仮説: $H_0 : \sigma^2 = 36$, $H_1 : \sigma^2 \neq 36$
- * 検定統計量: $\chi^2 = \frac{24 \cdot 48}{36} = 32$
- * $\chi_{0.05}^2(24) = 13.848$, $\chi_{0.95}^2(24) = 36.415$ であり、
 $13.848 < 32 < 36.415$ より、有意水準 10% H_0 は棄却されない。
- * 本年度の児童の質の揃い方が特に例年と変わっているとは言えない。

» 母平均の差の検定

2 標本検定

2 標本検定 2つの正規母集団の母平均 μ_1, μ_2 に差の検定のこと。実用上も非常に重要。

- * 例えば, 新しい治療法の効果を調べる場合, 患者を二つのグループに分け, 一方のみに新しい治療法を行い, その効果に差があるかを検定することがある。

処理群, 対照群 治療を施すグループを 処理群(treatment group), 比較のため治療を行わないグループを 対照群(contrast group) または 制御群(control group) という。

- * 帰無仮説: $H_0 : \mu_1 = \mu_2$
- * 対立仮説:
 - * 両側: $H_1 : \mu_1 \neq \mu_2$
 - * 片側: $H_1 : \mu_1 > \mu_2$ または $\mu_1 < \mu_2$

» 母平均の差の検定 等分散の場合 (2 標本 t 検定)仮定 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (等分散)

合併分散

$$s^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2} \quad (3)$$

検定統計量

$$t = \frac{\bar{X} - \bar{Y}}{s\sqrt{\frac{1}{m} + \frac{1}{n}}} \quad (4)$$

分布 自由度 $m+n-2$ の t 分布に従う

- * 両側検定: $|t| > t_{\alpha/2}(m+n-2)$ で H_0 を棄却
- * 片側検定: $t > t_{\alpha}(m+n-2)$ または $t < -t_{\alpha}(m+n-2)$ で H_0 を棄却

» 母平均の差の検定

等分散の場合の例

対照実験の例:投薬群と対照群の平均の差の検定

次のデータは、20 匹のラットを 10 匹ずつ 2 群に分け、一方にはふつうの食餌を与え、他方には血中の赤血球数を減らすと考えられている薬を混入した食餌を与えた場合の、血液 1mm^3 中の赤血球数である。有意水準 α は 5% とする。投薬群と対照群の平均の差の検定せよ。(単位:100 万個)。

投薬群	7.97	7.66	7.59	8.44	8.05	8.08	8.35	7.77	7.98	8.15
対照群	8.06	8.27	8.45	8.05	8.51	8.14	8.09	8.15	8.16	8.42

(出典:スネデカー,コ克蘭『統計的方法』)

» 母平均の差の検定

等分散の場合の例 (続き)

$H_0: \mu_T = \mu_C$ (薬に効果なし)

$H_1: \mu_T < \mu_C$ (薬が赤血球数を減らす, 左片側)

* $\bar{x}_T = 8.004, s_T^2 \approx 0.0761, n_T = 10$

* $\bar{x}_C = 8.230, s_C^2 \approx 0.0294, n_C = 10$

* 合併分散:

$$s_p^2 = \frac{9 \times 0.0761 + 9 \times 0.0294}{18} \approx 0.0528$$

$$s_p \approx 0.230$$

* 検定統計量:

$$t = \frac{8.004 - 8.230}{0.230 \times \sqrt{1/10 + 1/10}} \approx -2.20$$

* $t_{0.05}(18) = 1.734$ であり, $-2.20 < -1.734$ より、有意水準 5% では帰無仮説は棄却され、薬は赤血球数を有意に減少させると判断する。

» 母平均の差の検定 異分散の場合 (Welch の検定)

仮定 $\sigma_1^2 \neq \sigma_2^2$ (異分散)

検定統計量

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} \quad (5)$$

自由度

$$\nu = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}} \quad (6)$$

分布 $t(\nu^*)$ 分布に従う (ν に最も近い整数 ν^*)

- * 両側検定: $|t| > t_{\alpha/2}(\nu^*)$ で H_0 を棄却
- * 片側検定: $t > t_{\alpha}(\nu^*)$ または $t < -t_{\alpha}(\nu^*)$

» 母分散の比の検定

2つの正規母集団の分散が等しいかどうかを検定する:

- * 帰無仮説: $H_0: \sigma_1^2 = \sigma_2^2$
- * 対立仮説: $H_1: \sigma_1^2 \neq \sigma_2^2$

検定統計量:

$$F = \frac{s_1^2}{s_2^2} \quad (7)$$

F は F 分布 $F(m-1, n-1)$ に従う

検定の基準:

$F_{1-\alpha/2}(m-1, n-1) \leq F \leq F_{\alpha/2}(m-1, n-1) \Rightarrow H_0$ を棄却せず
 $F < F_{1-\alpha/2}(m-1, n-1)$ または、 $F > F_{\alpha/2}(m-1, n-1) \Rightarrow H_0$ を棄却する

例: 対照実験の例

$$F_{0.975}(9, 9) = 1/4.026 < F = \frac{0.0761}{0.0294} = 2.59 < F_{0.025}(9, 9) = 4.026$$

\Rightarrow 有意水準 5% では、母分散が等しいという仮説は棄却されない。

仮説検定

- * 検定の考え方
- * 正規母集団に対する仮説検定
- * χ^2 を用いる検定
- * 中心極限定理による検定

» 適合度の χ^2 検定

適合度の検定 (**goodness-of-fit test**):

- * 観測度数 f_i が理論度数 np_i に適合するか
- * 検定統計量:

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - np_i)^2}{np_i} \sim \chi^2(k-1) \quad (8)$$

- * 帰無仮説: $H_0 : P(A_i) = p_i$ for all i
- * 判定: $\chi^2 > \chi_\alpha(k-1)$ ならば、 H_0 を有意水準 α で棄却

表 12.6: 適合度の検定

属性 A のカテゴリー	A_1	A_2	\cdots	A_k	計
観測度数	f_1	f_2	\cdots	f_k	n
理論確率	p_1	p_2	\cdots	p_k	1
理論度数	np_1	np_2	\cdots	np_k	n

$f_i \doteq np_i$ かどうかを見るのが適合度検定である.

» 適合度の χ^2 検定

例 (メンデルの法則)

観測度数 **vs** 理論度数:

$$* f = (315, 101, 108, 32)$$

$$* np = (312.75, 104.25, 104.25, 34.75)$$

計算:

$$\begin{aligned}\chi^2 &= \frac{(315 - 312.75)^2}{312.75} + \frac{(101 - 104.25)^2}{104.25} + \\ &\quad \frac{(108 - 104.25)^2}{104.25} + \frac{(32 - 34.75)^2}{34.75} \\ &= 0.470\end{aligned}$$

臨界値: $\chi^2_{0.05}(3) = 7.815$

結論: $0.470 < 7.815$ より H_0 は棄却されない

⇒ 理論分布に適合している

» 分割表と独立性の χ^2 検定

独立性の検定 (test for independence):

- * A, B 2つの属性が独立か?
- * 観測度数: f_{ij}
- * 期待度数:

$$E_{ij} = \frac{f_{i\cdot} f_{\cdot j}}{n} \quad (9)$$

分割表の一般的な形

属性 A \ 属性 B	属性 B				
	B_1	B_2	\cdots	B_c	計
A_1	f_{11}	f_{12}	\cdots	f_{1c}	$f_{1\cdot}$
A_2	f_{21}	f_{22}	\cdots	f_{2c}	$f_{2\cdot}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
A_r	f_{r1}	f_{r2}	\cdots	f_{rc}	$f_{r\cdot}$
計	$f_{\cdot 1}$	$f_{\cdot 2}$	\cdots	$f_{\cdot c}$	n

- * 帰無仮説: $H_0 : P(A_i \cap B_j) = P(A_i)P(B_j)$ for all i, j
- * 検定統計量:

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - E_{ij})^2}{E_{ij}} \quad (10)$$

は自由度の $(r-1)(c-1)$ の χ^2 分布に従う。

- * 判定: $\chi^2 > \chi_\alpha(k-1)$ ならば、 H_0 を有意水準 α で棄却

» 独立性の検定

例

表 12.9 ある大学の工学部の期末試験の成績 (広津)

幾何 代数	優	良	可	計
優	4	2	3	9
良	8	4	6	18
可	6	3	6	15
計	18	9	15	42

表 12.9 の理論度数 (広津)

幾何 代数	優	良	可	計
優	3.86	1.93	3.21	9
良	7.71	3.86	6.43	18
可	6.43	3.21	5.36	15
計	18	9	15	42

表 12.9 から:

$$\chi^2 = \frac{(4 - 3.86)^2}{3.86} + \frac{(2 - 1.93)^2}{1.93} + \dots + \frac{(6 - 5.36)^2}{5.36} = 0.19$$

自由度: $(3 - 1)(3 - 1) = 4$

臨界値: $\chi_{0.05}^2(4) = 9.488$

結論: $0.19 < 9.488$ より H_0 の独立性の仮定は棄却されず

⇒ 代数と幾何の成績は関係あると言えない。(独立性が成立と言える)

仮説検定

- * 検定の考え方
- * 正規母集団に対する仮説検定
- * χ^2 を用いる検定
- * 中心極限定理による検定

» 中心極限定理による検定

カード色当ての例

母集団分布が正規分布でなくとも、検定統計量が近似的に正規分布に従うならば、正規分布に対する検定の手続きをそのまま使える (中心極限定理より)。

例: ある人がカード (52 枚) を黒か赤かで言い当てる実験を行い、40 枚を当てた。この結果が偶然かどうかを検定する。

- * 52 枚のカードをランダムに黒か赤かと予測するならば、これがあたる確率は $1/2$ である。あたりの成功率を $p = 1/2$ とし、帰無仮説 $H_0: p = 1/2$ が棄却されるかどうかを検討する。対立仮説: $H_1: p \neq 1/2$
- * あたり、はずれを $X_i = 1, 0$ で表すと、 $X_i \sim \text{Bi}(1, 1/2)$
- * n 枚通したあたりの枚数は、 $S_n = \sum_{i=1}^n X_n \sim \text{Bi}(n, p)$ 。中心極限定理より、

$$Z = \frac{S_n - np}{\sqrt{np(1-p)}} \quad (11)$$

は、 n が大きいとき標準正規分布 $N(0, 1)$ に近似的に従う。

- *
$$Z = \frac{S_n - np}{\sqrt{np(1-p)}} = \frac{40 - 26}{\sqrt{13}} \approx 3.883$$
- * **判定:** 両側検定を仮定し、 $z_{0.025} = 1.96$ であり、 $\Rightarrow |Z| > 1.96$ より H_0 は棄却され、あたりはずれはランダムに生じているものとは考えられない。

» 母比率の検定統計量

ここで、 $\hat{p} = \frac{S_n}{n}$ は標本における成功の比率 (相対頻度) であり、これは**母比率の検定**と考えることができる。すなわち、

$$Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \quad (12)$$

で検定が可能。これが**母比率**(母集団の比率) が p であることの検定統計量である。