

経済統計学講義ノート No.8

仮説検定

蛭川雅之

2024 年 11 月 7 日

目 次

1. 仮説検定とは
2. 平均に関する検定
3. 無相関検定
4. χ^2 検定

1 仮説検定とは

1.1 仮説検定の手順

- **仮説検定**とは、母集団に関して初めに仮定したことが正しいかどうかを標本に照らしあわせて客観的に判断する方法である。
 - － 「初めに仮定したこと」を**帰無仮説**といい、記号 H_0 で表す。
 - * 例：母集団の平均 μ はゼロである。 $\Rightarrow H_0 : \mu = 0$
 - － 帰無仮説の一部または全部の否定を**対立仮説**といい、記号 H_1 で表す。
 - * 例 1：母集団の平均 μ はゼロでない。 $\Rightarrow H_1 : \mu \neq 0$
 - * 例 2：母集団の平均 μ は正である。 $\Rightarrow H_1 : \mu > 0$
 - * 例 3：母集団の平均 μ は負である。 $\Rightarrow H_1 : \mu < 0$

- 仮説検定は以下の手順で実行される。
 1. 帰無仮説 H_0 および対立仮説 H_1 を立てる。
 2. 帰無仮説 H_0 が正しいと仮定して検定統計量の確率分布を求める。
 3. 有意水準（例：5%, 1%）を設定し、この有意水準に基づいて2の確率分布に対する臨界値および棄却域を定める。
 - － 標本から計算した検定統計量の値が棄却域に入る場合、帰無仮説 H_0 を棄却し対立仮説 H_1 を採択する。

■帰無仮説 H_0 を棄却できない場合の解釈

- 「 H_0 を棄却するに足りる事実（証拠）が見つからなかった」に過ぎない。
 - － 「 H_0 は正しい」とは断言できない（これも可能性の一つ）。

1.2 両側検定と片側検定

- 帰無仮説 $H_0 : \mu = 0$ と対立仮説 H_1 の組がどのように設定されているかにより、棄却域は異なる。
 1. $H_1 : \mu \neq 0$ の場合、棄却域は検定統計量の確率分布の両側にある。
 2. $H_1 : \mu < 0$ の場合、棄却域は検定統計量の確率分布の左側にある。
 3. $H_1 : \mu > 0$ の場合、棄却域は検定統計量の確率分布の右側にある。
- 1 は両側検定、2-3 は片側検定の例である。

1.3 仮説検定における 2 種類の誤り

判断 \ 真偽	H_0 が真 (H_1 が偽)	H_0 が偽 (H_1 が真)
H_0 を採択 (H_1 を棄却)	○	× (第 2 種の過誤)
H_0 を棄却 (H_1 を採択)	× (第 1 種の過誤)	○

- $\alpha = \text{Pr}(\text{第 1 種の過誤})$ および $\beta = \text{Pr}(\text{第 2 種の過誤})$ をともにゼロにすることはできない。
 - α と β はトレード・オフの関係にある。
 - 例えば、 $\alpha = 0$ とするには、 H_0 の真偽にかかわらず常に H_0 を採択すればよい。 $\Rightarrow \beta \uparrow 1$
- 現実には、 α (=有意水準) を小さい値 (例: 5%, 1%) に設定した上で、 β をできる限り小さくするように検定統計量を作成している。

2 平均に関する検定

2.1 正規母集団の平均に関する検定

問題 1 ある週刊誌の編集長はしばらく前に交代した。新しい編集長は販売部数を伸ばすことにかけて定評のある人物である。編集長が交代して以降の無作為に選ばれた 16 週について、この週刊誌の 1 週当たり販売部数を調べたところ、平均は 26.5 万部、標準偏差は 3.2 万部であった。この週刊誌の 1 週間の目標販売部数が 25 万部であるとして、この編集長は目標を上回る成果を収めたといっってよいかどうかを有意水準 5% で検定せよ。なお、この週刊誌の 1 週間当たり販売部数は正規分布に従うものとし、与えられた標準偏差は不偏分散に基づくものとする。

解法 2 正規母集団 $N(\mu, \sigma^2)$ の平均 μ に関する検定であることに注意する。具体的な手順は以下の通りである。

1. 帰無仮説 H_0 および対立仮説 H_1 を以下のように設定する（単位：万部/週）。

$$H_0 : \mu = 25, H_1 : \mu > 25$$

2. 観測値の個数を n とすると、 H_0 の下で検定統計量

$$t = \frac{\sqrt{n}(\bar{X} - \mu)}{s} = \frac{\sqrt{n}(\bar{X} - 25)}{s}$$

は自由度 $(n - 1)$ の t 分布に従う。

3. $n = 16$, ($\bar{X}$ の実現値 =) $\bar{x} = 26.5$, $s = 3.2$ により $t = 1.875$ が得られ、一方、 $t(15)$ の右側 5% 点（=有意水準 5% の臨界値）は $t_{0.05}(15) = 1.753$ である。

- $t = 1.875 > 1.753 = t_{0.05}(15)$ から H_0 は有意水準 5% で棄却され、この編集長は目標を上回る成果を収めたといえる。

2.2 母集団比率に関する検定

問題 3 A 予備校は新聞・テレビ等で「昨年度の大学受験生の 72% は同校の模擬試験を受験していた」と宣伝している。昨年度の大学受験生から無作為に選ばれた 350 名に実際に A 予備校の模擬試験を受験したことがあるかどうか尋ねたところ、231 名が受験したと回答した。A 予備校の宣伝は誇大であるかどうかを有意水準 5% で検定せよ。

解法 4 昨年度の大学受験生全体を母集団とし、A 予備校の模擬試験を受験した比率を p 、その標本比率を $\hat{p}$ と表記する。具体的な手順は以下の通りである。

1. 帰無仮説 H_0 および対立仮説 H_1 を以下のように設定する。

$$H_0 : p = 0.72, H_1 : p < 0.72$$

2. 観測値の個数 n が十分大きい場合、 H_0 の下で検定統計量

$$z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} = \frac{\hat{p} - 0.72}{\sqrt{0.72(1-0.72)/n}}$$

は近似的に $N(0, 1)$ に従う。

3. $n = 350$, $\hat{p} = 0.66$ により $z = -2.5$ が得られ、一方、 $N(0, 1)$ の^{左側} 5% 点 (=有意水準 5% の臨界値) は -1.65 である。

- $z = -2.5 < -1.65$ から H_0 は有意水準 5% で棄却され、A 予備校の宣伝は誇大であるといえる。

2.3 母集団比率の差に関する検定

問題 5 ある映画の鑑賞後に「面白かった」と答えた観客は、男性 1800 人中 1116 人、一方、女性 900 人中 504 人であった。この映画の評価に男女差があるかどうかを有意水準 5% で検定せよ。

- この問題は実は 2 つの母集団比率が等しいかどうかの検定である。
 - この映画を鑑賞した男性客および女性客全体が母集団である。
- 検定の詳細を説明するため、以下の表記を導入する。

p_1 : 男性客全体に占める「面白かった」と評価した者の比率

p_2 : 女性客全体に占める「面白かった」と評価した者の比率

n_1 : 男性の標本の大きさ

n_2 : 女性の標本の大きさ

x_1 : 男性の標本中「面白かった」と評価した人数

x_2 : 女性の標本中「面白かった」と評価した人数

- 帰無仮説 H_0 および対立仮説 H_1 を

$$H_0 : p_1 = p_2, H_1 : p_1 \neq p_2$$

と設定し、これらを以下のように書き改める。

$$H_0 : p_1 - p_2 = 0, H_1 : p_1 - p_2 \neq 0$$

- 中心極限定理により、男性・女性の標本比率について

$$\hat{p}_1 \overset{A}{\sim} N \left(p_1, \frac{p_1 (1 - p_1)}{n_1} \right), \hat{p}_2 \overset{A}{\sim} N \left(p_2, \frac{p_2 (1 - p_2)}{n_2} \right)$$

であり、 $\hat{p}_1$ と $\hat{p}_2$ は独立であることから、正規分布の再生性により

$$\hat{p}_1 - \hat{p}_2 \overset{A}{\sim} N \left(p_1 - p_2, \frac{p_1 (1 - p_1)}{n_1} + \frac{p_2 (1 - p_2)}{n_2} \right)$$

が成り立つ。

– H_0 の下では $p_1 = p_2 = p$ であるから

$$\hat{p}_1 - \hat{p}_2 \overset{A}{\sim} N \left(0, p (1 - p) \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right)$$

となり、

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{p (1 - p) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \overset{A}{\sim} N (0, 1)$$

である。

- ここで、

$$\hat{p}_1 - \hat{p}_2 = \frac{x_1}{n_1} - \frac{x_2}{n_2}$$

であり、 p の一致推定量

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

を用いると、最終的に H_0 の下で近似的に $N(0, 1)$ に従う検定統計量

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

が得られる。

解法 6 以下の手順に従う。

1. 帰無仮説 H_0 および対立仮説 H_1 を以下のように設定する。

$$H_0 : p_1 = p_2, H_1 : p_1 \neq p_2 \Rightarrow H_0 : p_1 - p_2 = 0, H_1 : p_1 - p_2 \neq 0$$

2. 観測値の個数 (n_1, n_2) が十分大きい場合、 H_0 の下で検定統計量

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

は近似的に $N(0, 1)$ に従う。

3. $(n_1, n_2, x_1, x_2) = (1800, 900, 1116, 504)$ により $z = 3.0$ が得られ、一方、 $N(0, 1)$ の両側 5% 点 (=有意水準 5% の臨界値) は 1.96 である。

- $|z| = 3.0 > 1.96$ から H_0 は有意水準 5% で棄却され、この映画

の評価に男女間で差があるといえる。

3 無相関検定

- 標本相関係数 r_{XY} を用いて母集団の相関係数 $Corr(X, Y)$ に関する検定を行うことができる。
- 帰無仮説を $H_0 : Corr(X, Y) = 0$ と設定する。
 - － 対立仮説 H_1 を以下のいずれかから選ぶ。

$$H_1 : Corr(X, Y) \neq 0 (\Rightarrow \text{両側検定})$$

$$H_1 : Corr(X, Y) < 0 (\Rightarrow \text{左側検定})$$

$$H_1 : Corr(X, Y) > 0 (\Rightarrow \text{右側検定})$$

- 観測値の個数を n とすると、 H_0 の下で検定統計量

$$t = \frac{r_{XY} \sqrt{n-2}}{\sqrt{1-r_{XY}^2}}$$

は自由度 $(n-2)$ の t 分布に従う。

4 χ^2 検定

4.1 適合度検定

- **適合度検定**とは、与えられた度数データが特定の理論的度数分布と一致しているかどうかを検定する手法である。
- 確率変数 X が属する母集団は k 個のカテゴリー $A_1, \dots, A_k$ に分類できるものとする（例：サイコロの目 $\Rightarrow k = 6$ ）。
 - 仮説として X がカテゴリー A_i に属する確率を p_i とすると、帰無仮説 H_0 および対立仮説 H_1 は以下のようなになる。

$$H_0 : \Pr(X \in A_i) = p_i, i = 1, \dots, k$$

$$H_1 : H_0 \text{は正しくない}$$

- 大きさ n の標本において、
 1. H_0 の下でのカテゴリー A_i に属する期待度数（＝理論値）は $E_i = np_i$ である。
 2. 実際に観測されたカテゴリー A_i に属する度数を O_i とする。
 - － 観測値の個数

$$n = \sum_{i=1}^k O_i$$

が十分大きい場合、 H_0 の下で検定統計量

$$Q = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

は近似的に自由度 $(k - 1)$ の χ^2 分布に従う。

■ Q の確率分布に関する直感的な説明

- 単純化のため、 $k = 2$ とし、 X をベルヌーイ確率変数とする。
 - 添え字を $0 (= \text{失敗})$, $1 (= \text{成功})$ とすると、

$$O_0 = n - O_1, p_0 = 1 - p_1$$

が成り立つ。

- 中心極限定理により、観測値の個数 n が十分大きい場合、

$$\frac{\hat{p}_1 - p_1}{\sqrt{p_1(1-p_1)/n}} = \frac{O_1/n - p_1}{\sqrt{p_1(1-p_1)/n}} = \frac{O_1 - np_1}{\sqrt{np_1(1-p_1)}} \stackrel{A}{\approx} N(0, 1)$$

であるから、

$$\frac{(O_1 - np_1)^2}{np_1(1-p_1)} \stackrel{A}{\approx} \chi^2(1)$$

となる。

- 一方、

$$(O_0 - np_0)^2 = \{(n - O_1) - n(1 - p_1)\}^2 = (O_1 - np_1)^2$$

であるから、

$$\begin{aligned} Q &= \frac{(O_0 - E_0)^2}{E_0} + \frac{(O_1 - E_1)^2}{E_1} \\ &= \frac{(O_0 - np_0)^2}{np_0} + \frac{(O_1 - np_1)^2}{np_1} \\ &= \frac{(O_1 - np_1)^2}{n(1 - p_1)} + \frac{(O_1 - np_1)^2}{np_1} \\ &= \frac{(O_1 - np_1)^2}{n} \left(\frac{1}{1 - p_1} + \frac{1}{p_1} \right) \\ &= \frac{(O_1 - np_1)^2}{np_1(1 - p_1)} \stackrel{A}{\sim} \chi^2(1) \end{aligned}$$

が確かに成り立つ。

問題 7 日本人の血液型の分布は A 型、O 型、B 型、AB 型の順にそれぞれ 40%、30%、20%、10% であると言われている。無作為に抽出された 200 人の血液型について次のようなデータが得られたとする。このデータは日本人の血液型の分布に従っているといえるかどうか有意水準 5% で検定せよ。【ヒント：Excel の関数 `chiinv` が χ^2 分布の $100q\%$ 分位点を与える。】

血液型	A 型	O 型	B 型	AB 型	計
度 数	98	54	31	17	200

■仮定する分布の未知母数をデータから推定する場合の自由度修正

問題 8 あるデパートの売り場において、1 日当たり苦情件数を無作為に選ばれた 100 日について調べたところ、次の結果を得た。なお、苦情件数が「4 以上」であった 5 日とも、苦情件数は丁度 4 件であったとする。このとき、以下の問いに答えよ。

1 日当たり苦情件数	0	1	2	3	4 以上	計
日 数	29	41	18	7	5	100

- (1) このデータから 1 日当たり苦情件数の平均および分散をそれぞれ小数第 2 位まで求めよ。
- (2) このデータにポアソン分布をあてはめた結果に対し、有意水準 5% で適合度検定を行え。なお、(1) で求めた平均をポアソン分布の平均発生回数 λ の推定値 $\hat{\lambda}$ として使用せよ。

- 問題 7 では各血液型の割合、問題 8 では 1 日当たり平均苦情件数が母数である。
 - 前者では母数を与えられている（＝推定は不要である）のに対し、後者では母数を推定する必要がある。
- 母数の推定を伴わない場合、 H_0 の下で検定統計量 Q の漸近分布は自由度 $(k - 1)$ の χ^2 分布であった。
 - データから未知母数を推定する必要がある場合、推定する母数の数を m とすると、 H_0 の下で検定統計量 Q は近似的に自由度 $(k - m - 1)$ の χ^2 分布に従う。

4.2 独立性検定

- **独立性検定**は適合度検定の 2 次元拡張版であり、2 つの要因の関連性の有無を検定する。
 - － 連続変数を取り扱う無相関検定と異なり、各次元がカテゴリーに分類できる場合を想定する。
 - * 「ある俳優の好き嫌い」に「男女」差があるか？
 - * 「企業の業績の良し悪し」に「企業の分野」間で差があるか？
 - － アンケート調査の分析によく使われる。

- 一組の確率変数 (X, Y) が属する母集団は $k \times \ell$ 個のカテゴリーに分類できるものとする。
 - (X, Y) が特定のカテゴリー $(A_i, B_j), i = 1, \dots, k, j = 1, \dots, \ell$ に属する確率を

$$p_{ij} = \Pr(X \in A_i, Y \in B_j)$$

とする。

- 対応する X および Y の周辺確率をそれぞれ

$$p_{i\cdot} = \Pr(X \in A_i), p_{\cdot j} = \Pr(Y \in B_j)$$

とすると、帰無仮説 H_0 および対立仮説 H_1 は以下のようになる。

$$H_0 : p_{ij} = p_{i\cdot} p_{\cdot j}, i = 1, \dots, k, j = 1, \dots, \ell$$

$$H_1 : H_0 \text{は正しくない}$$

- 差し当たり、 $p_{i\cdot}$ および $p_{\cdot j}$ は既知と仮定して議論を進める。
 - 大きさ n の標本において、
 1. H_0 の下でのカテゴリー (A_i, B_j) に属する期待度数（＝理論値）は $E_{ij} = np_{i\cdot}p_{\cdot j}$ である。
 2. 実際に観測されたカテゴリー (A_i, B_j) に属する度数を O_{ij} とする。
 - 多くの場合、観測データは k 行 $\times$ ℓ 列の分割表（クロス集計表）で与えられる。

－ 観測値の個数

$$n = \sum_{j=1}^{\ell} \sum_{i=1}^k O_{ij}$$

が十分大きい場合、 H_0 の下で検定統計量

$$Q = \sum_{j=1}^{\ell} \sum_{i=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

は近似的に自由度 $k\ell - 1$ の χ^2 分布に従う。

- 適合度検定の場合と異なり、 $p_{i\cdot}$ および $p_{\cdot j}$ の具体的な数値は H_0 で与えられていないがどうするか？
 - 次の手順で $p_{i\cdot}$ および $p_{\cdot j}$ を推定する。

$$\hat{p}_{i\cdot} = \frac{1}{n} \sum_{j=1}^{\ell} O_{ij}, \quad \hat{p}_{\cdot j} = \frac{1}{n} \sum_{i=1}^k O_{ij}$$

- 最終的に、カテゴリー (A_i, B_j) に属する期待度数を $\hat{E}_{ij} = n\hat{p}_{i\cdot}\hat{p}_{\cdot j}$ に修正した検定統計量

$$\hat{Q} = \sum_{j=1}^{\ell} \sum_{i=1}^k \frac{(O_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}$$

を考える。

- H_0 の下で $\hat{Q}$ は近似的に自由度 $(k-1)(\ell-1)$ の χ^2 分布に従う。

問題 9 次の表は 2×2 分割表の一例で、無作為に選ばれた東京・大阪在住の成人男性各 100 人に対し、自家用車の色として赤を選ぶかどうかを調査した結果を一覧表にまとめたものである。

都市 \ 赤を	選 ぶ	避ける	計
東京	40	60	100
大阪	55	45	100
計	95	105	200

- (1) 「都市」「自家用車の色」2つの属性に関係が認められるとってよいかどうか有意水準 5% で独立性検定を行え。
- (2) 独立性検定に関する検定統計量 $\hat{Q}$ と（赤を選ぶ）母集団比率の差に関する検定統計量 z との間に $\hat{Q} = z^2$ が成り立つことを確認せよ。

■2 つの要因の関連性の強さ

- 2 つの要因の関連性の有無を検定するのと併せて、それらの関連性の強さを知りたい場合どうするか？
 - クラメール連関係数 (Cramér's coefficient of association) は、 k 行 $\times$ ℓ 列の分割表における行要素と列要素の関連の強さを示す代表的な指標である。
- クラメール連関係数 V は

$$V = \sqrt{\frac{\hat{Q}}{n \{\min(k, \ell) - 1\}}}$$

で与えられる。

- 特に、 2×2 分割表（より正確には、いずれか一方の要因のカテゴリー数が 2）の場合、

$$V = \sqrt{\frac{\hat{Q}}{n}}$$

となる。

- $0 \leq V \leq 1$ であり、 $V > 0.25$ が 2 つの要因の関連を示す大まかな目安とされる。

問題 10 問題 9 の「都市」「自家用車の色」2 つの属性に関し、クラメール連関係数 V を計算せよ。