No.17 χ^2 検定

NAKADA Masayuki

Kobe University Secondary School

January 17, 2022

仮説検定の流派

ネイマン・ピアソン流　帰無仮説（「有意差なし」）と対立仮説（「有意差あり」）による二者択一を行う（これまでに学んできた形式）.

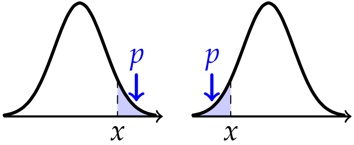
フィッシャー流　観測事実の起こりやすさをp値という確率指標で表現する．p値は証拠の確からしさであり，p値が小さいほど，強い証拠とみなされる.

p値

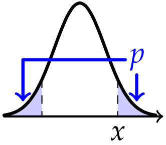
p値とは，母集団が帰無仮説（「有意差なし」）を仮定した分布に従うとするとき，観測される標本が「どの程度稀なことであるか」を示す確率である．観測された標本から得られた統計量をxとするとき，

片側p値　$p := P(X \geqq x)$（$x > 0$）または$p := P(X \leqq x)$（$x < 0$）

両側p値　$p := P(|X| \geqq |x|)$（2箇所の領域の確率の和）



片側p値



両側p値

例題（右片側検定での比較）

コインを表が出やすく加工し，20 回投げたら表が 15 回出た．加工が成功したと言えるか．

ネイマン・ピアソン流

有意水準 $\alpha = 0.05$ で右片側検定する．帰無仮説，対立仮説はそれぞれ，

$$H_0\colon p = 0,\ \ H_1\colon p > 0$$

とする．

$$P(\text{表} \geqq 15) = 0.0207 < 0.05 = \alpha$$

よって，帰無仮説 H_0 は棄却され，加工が成功したと言える．

フィッシャー流

帰無仮説，対立仮説をそれぞれ，

$$H_0\colon p = 0,\ \ H_1\colon p > 0$$

とする．帰無仮説の元では，20 回中 15 回以上表が出る確率は

$$P(\text{表} \geqq 15) = 0.0207$$

よって，片側 p 値は，$p = 0.0207$ であり，統計的に有意な差が認められた．

例題（両側検定での比較）

コインを 20 回投げたら表が 15 回出た．このコインは表裏の出方が不均等なのだろうか．

ネイマン・ピアソン流

有意水準 $\alpha = 0.05$ で両側検定する．帰無仮説，対立仮説はそれぞれ，

$$H_0 \colon p = 0,\ \ H_1 \colon p \neq 0$$

とする．

$$P(\text{表} \geqq 15) = 0.0207 < 0.025 = \alpha/2$$

よって，帰無仮説 H_0 は棄却され，コインの表裏の出方は不均等であると言える．

フィッシャー流

帰無仮説，対立仮説はそれぞれ，

$$H_0 \colon p = 0,\ \ H_1 \colon p \neq 0$$

とする．帰無仮説の元では，20 回中 15 回以上表が出る確率は

$$P(\text{表} \geqq 15) = 0.0207$$

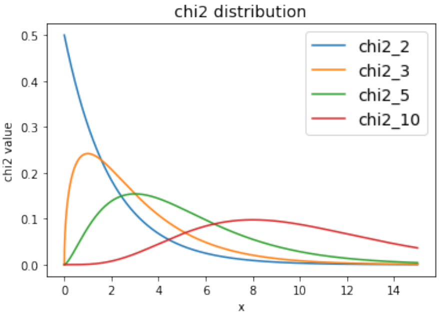
よって，両側 p 値は，$p = 2 \cdot 0.0207 = 0.0414$ であり，統計的に有意な差が認められた．

χ^2 分布

$Z_1, Z_2, \ldots, Z_n$ が標準正規分布 $N(0,1)$ に従う独立な確率変数であるとき，確率変数

$$Y = {Z_1}^2 + {Z_2}^2 + \cdots + {Z_n}^2$$

の従う分布のことを，**自由度 n の χ^2（カイ自乗）分布**といい，χ^2_n で表す（cf. No.13）.



正規分布 $N(\mu, \sigma^2)$ に従う母集団から大きさ n の標本を抽出することを考える．このとき，不偏標本分散 $\hat{S}^2 = \frac{1}{n-1} \sum_i (X_i - \overline{X})^2$ について，$\frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$ である．

※ No.13 の pp.11-13 における議論は，不偏でない標本分散 $S^2 = \frac{1}{n} \sum_i (X_i - \overline{X})^2$ についての計算であるから，その結論は $\frac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$ となることに注意.

正規分布 $N(\mu,\sigma^2)$ に従う母集団から大きさ n の標本を抽出することを考える．このとき，不偏標本分散 $\hat{S}^2 = \frac{1}{n-1}\sum_i (X_i - \overline{X})^2$ について，$\frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$ である．

例題: *母分散の検定*

ある全国規模の模擬試験の数学の素点について，母平均は 57.68，母分散は 19.49 である．この模擬試験を受験した A 高校の生徒 32 人について，標本平均は 59.24，不偏標本分散は 12.96 であった．模擬試験を受験した A 高校の生徒の数学の学力の散らばり具合は，全国の模擬試験の受験生の様子と比べて異なるとみなしてよいだろうか．

正規分布 $N(\mu,\sigma^2)$ に従う母集団から大きさ n の標本を抽出することを考える．このとき，不偏標本分散 $\hat{S}^2 = \frac{1}{n-1}\sum_i(X_i - \overline{X})^2$ について，$\frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi^2_{n-1}$ である．

例題: 母分散の検定

ある全国規模の模擬試験の数学の素点について，母平均は 57.68，母分散は 19.49 である．この模擬試験を受験した A 高校の生徒 32 人について，標本平均は 59.24，不偏標本分散は 12.96 であった．模擬試験を受験した A 高校の生徒の数学の学力の散らばり具合は，全国の模擬試験の受験生の様子と比べて異なるとみなしてよいだろうか．

模擬試験を受験した A 高校の生徒の素点の母分散 σ^2 について，

$$帰無仮説\ H_0 : \sigma^2 = 19.49,\quad 対立仮説\ H_1 : \sigma^2 \neq 19.49$$

として，有意水準 10% の両側 χ^2 検定を行う．χ^2 統計量は $\chi^2 = \frac{(32-1)\cdot 12.96}{19.49} = 20.61$ であるが，自由度 31 の χ^2 分布の下側，上側 5% 点はそれぞれ $\chi^2_{0.95}(31) = 19.28$，$\chi^2_{0.05}(31) = 44.99$ である．よって，$\chi^2_{0.95}(31) < \chi^2 < \chi^2_{0.05}(31)$ となり，有意水準 10% で帰無仮説 H_0 は棄却されない．

χ^2 検定あれこれ

χ^2 検定は，理論値と観測値の適合の度合を調べるために利用される．

一般に，事象 $A_1, A_2, \ldots, A_k$ の現れると期待される確率がそれぞれ $p_1, p_2, \ldots, p_n$ であるとする．観測によりそれぞれの事象が $n_1, n_2, \ldots, n_k$ 回現れたとすると，事象の観測度数，期待度数は右の表にまとめられる．

事象	A_1	A_2	$\cdots$	A_k
観測度数	n_1	n_2	$\cdots$	n_k
期待度数	np_1	np_2	$\cdots$	np_k

$(p_1 + \cdots + p_n = 1)$

観測度数が期待度数に適合しているかどうかを知るには，**適合度基準**

$$\boxed{\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i}}$$

を計算する．χ^2 の値が小さいほど，期待度数からのずれが小さく，よく適合しているとみなせる．適合度基準は，$n \to \infty$ で漸近的に自由度 $k-1$ の χ^2 分布 χ^2_{k-1} に従うことが知られている．

例題: 適合度の χ^2 検定

次の表は，エンドウ豆の形質の遺伝についてのメンデルの実験データである．実験データがメンデルの法則に合致しているといえるか，確かめよ．

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

帰無仮説 H_0: 観測度数は理論比に適合している
対立仮説 H_1: 観測度数は理論比に適合していない

とし，有意水準 5% の右片側検定（適合度基準 χ^2 が大きい場合に H_0 を棄却できればよい）を行う．
与えられた表から

型	黄色・丸い	黄色・しわ	緑色・丸い	緑色・しわ	計
観測度数（O）	315	101	108	32	556
確率	$9/16$	$3/16$	$3/16$	$1/16$	1
期待度数（E）	312.75	104.25	104.25	34.75	556
$O-E$	2.25	-3.25	3.75	-2.75	0

よって，適合度基準は

$$\chi^2 = \frac{(2.25)^2}{312.75} + \frac{(-3.25)^2}{104.25} + \frac{(3.75)^2}{104.25} + \frac{(-2.75)^2}{34.75} = 0.47$$

自由度 3 の χ^2 分布の上側 5% 点は $\chi^2_{0.05}(3) = 7.81$ であるから，$\chi^2 < \chi^2_{0.05}(3)$ であり，実験結果は理論値に適合しているという仮説は棄却されない．

独立性の χ^2 検定

新開発の薬剤 A の効果を検証するために，40 人を処置群（薬剤 A を投与）20 人と対照群（偽薬を投与）に分けて，効果の有無を調べた．この結果から，薬剤 A は効果があるとしてよいか．

	有効	無効	計
処置群	10	10	20
対照群	6	14	20
計	16	24	40

※対照群に偽薬を与えるのは，プラセボ効果（思い込みによる効果）の影響を排除するため．

帰無仮説 H_0: 処置群と対照群とで，薬剤 A の効果に差がない
対立仮説 H_1: 処置群と対照群とで，薬剤 A の効果に差がある

として，有意水準 5% の右片側検定を行う．
H_0 のもとでは，期待度数は括弧内のようになる．

	有効	無効	計
処置群	10(8)	10(12)	20
対照群	6(8)	14(12)	20
計	16	24	40

期待度数の計算方法:
例えば，処置群・有効の期待度数は

$$(\text{総計}) \times \frac{(\text{処置群の計})}{(\text{総計})} \cdot \frac{(\text{有効の計})}{(\text{総計})}$$

分割表における自由度:
分割表では，縦横の計の束縛で自由度が 1 ずつ減るため，$r \times c$ のサイズの分割表の自由度は $(r-1)(c-1)$ で与えられる．

よって，適合度基準は

$$\chi^2 = \frac{(10-8)^2}{8} + \frac{(10-12)^2}{12} + \frac{(6-8)^2}{8} + \frac{(14-12)^2}{12} = 1.67$$

自由度 $(2-1)(2-1) = 1$ の χ^2 分布の上側 5% 点は $\chi^2_{0.05}(1) = 3.84$ であるから，
$\chi^2 < \chi^2_{0.05}(1)$ であり，薬剤 A に効果がないとする帰無仮説は棄却されない．

よりよい検定のために

- 検定における重要な情報は，「有意水準」，「標本の大きさ」，「効果量」，「検出力」である．検定を行う際には，これらの数値が明記されるべきである．
- 上記の 4 つの情報は，相互依存関係にあり，他の 3 つの情報から算出可能である．
- 検定結果は標本に依存し，誤った結論を導く確率を 0 にすることはできない．
- 検定を行う前に，分析方法を設計する．行き当たりばったりの検定は，データの誤用を生む可能性がある．
- 帰無仮説を棄却できないとき，そこから何か積極的な結論が導き出されているわけではない．
- 検定による結果は説得力があるように見えるが，あくまでも論拠の一つである．結論は，総合的な視点から述べること．