第 10 章 χ^2 および t 分布とそれらを用いた検定

浅川 伸一

2006 年 06 月 29 日

1 χ^2 分布

1. (既出) 正規分布 $N\left(\mu,\sigma^2\right)$ に従う母集団から n 個のデータ $x_1, x_2, \ldots, x_n$ を無作為抽出すると標本平均 $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ は平均 μ 分散 $\dfrac{\sigma^2}{n}$ の正規分布 $N(\mu, \dfrac{\sigma^2}{n})$ に従う。
2. 標準正規分布 $N(0,1^2)$ に従う母集団から、無作為抽出によって 1 個のデータを得たとき、このデータの 2 乗 x^2 は、自由度 1 の χ^2 (カイ 2 乗) 分布に従う。χ^2 分布の定義。
3. 標準正規分布 $N(0,1^2)$ に従う母集団から、無作為抽出によって n 個のデータを得たとき、これらデータの 2 乗和 $\sum_{i=1}^{n} x_i^2$ は、自由度 n の χ^2 分布に従う。このことを χ^2 分布の再生性という。
4. 正規分布 $N\left(\mu,\sigma^2\right)$ に従う母集団から n 個のデータ $x_1, x_2, \ldots, x_n$ を無作為抽出すると
$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 = \left(\frac{x_1-\mu}{\sigma}\right)^2 + \left(\frac{x_2-\mu}{\sigma}\right)^2 + \cdots + \left(\frac{x_n-\mu}{\sigma}\right)^2 \tag{1}$$
は自由度 n の χ^2 分布に従う
5. 正規分布 $N\left(\mu,\sigma^2\right)$ に従う母集団から n 個のデータ $x_1, x_2, \ldots, x_n$ を無作為抽出すると
$$\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma}\right)^2 = \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{\sigma^2} \tag{2}$$
$$= \frac{n\,s^2}{\sigma^2} \tag{3}$$
は自由度 $n-1$ の χ^2 分布に従う。もう一つの χ^2 分布の定義。

6. 正規分布 $N\left(\mu,\sigma^2\right)$ に従う母集団から n 個のデータ $x_1, x_2, \ldots, x_n$ を無作為抽出すると

$$\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 = \frac{n}{\sigma^2}\left(\bar{X}-\mu\right)^2 \tag{4}$$

は自由度 1 の χ^2 分布に従う。

χ^2 分布の定義は、標準正規分布 $N(0,1^2)$ に従う確率変数の 2 乗であるが、適合度の検定としても知られている。

1.1　適合度検定

ある仮定から導き出される理論値と実測値との差を 2 乗して足し合わせると、その結果得られた数値は χ^2 に従うとされる。

1.1.1　例

メンデルは優性遺伝 (メンデルの) 法則を発見した。メンデルは、円みのあるエンドウマメと角のあるエンドウマメを交配させると次世代は円みのあるエンドウマメだけが得られ、さらにその次の世代では 75% の割合で円みのあるエンドウマメが現れることを見出した。実際のデータは円みのあるエンドウマメ 336 に対して、角のあるエンドウマメ 101 であった。メンデルの法則が正しいか否かを χ^2 による適合度検定を用いて検討してみる。帰無仮説 (H_0) は、$p = 0.75$, 対立仮説は $p \neq 0.75$ である。危険率は 0.05 とする。

メンデルの法則が正しければ総データ数 437 に対する円みのあるエンドウマメの理論値は $437 \times 0.75 = 327.75$ である。χ^2 値を計算すれば

$$\chi^2 = \left(\frac{336-327.75}{327.75}\right)^2 + \left(\frac{101-109.25}{109.25}\right)^2 \doteqdot 0.83 \tag{5}$$

を得る。これが自由度 1 の χ^2 分布に従うと仮定して検定できる。χ^2 の表を調べる (あるいは、=chidist(0.83,1) がおよそ 0.36 であり、=chiinv(0.05,1) がおよそ 3.84 である) ことにより、このようなことが起こる可能性は 36% あり、帰無仮説を棄却することはできない。すなわち危険率 0.05 で帰無仮説 $p = 0.75$ を否定することはできない。

適合度検定についての簡単な説明が、推測統計はじめの一歩 (清水誠, 2000) p.124 にある。n 個のデータがある属性によって 2 群 (データ数をそれぞれ n_1, n_2 とする) に分割され、各群の生じる可能性が確率 p (起こらない可能性が $q = (1-p)$) であると予想される場合実測値と理論値との差を理論値で

割って 2 乗したものを足し合わせると、自由度 1 の χ^2 分布に従うことを用いて検定することができる。

$$\chi^2 = \frac{(n_1 - np)^2}{np} + \frac{(n_2 - n(1-p))^2}{n(1-p)} \tag{6}$$

$$= \frac{(n_1 - np)^2}{np} + \frac{(n - n_1 - n + np)^2}{n(1-p)} \tag{7}$$

$$= \frac{(n_1 - np)^2}{np(1-p)} = \left\{ \frac{n_1 - np}{\sqrt{np(1-p)}} \right\}^2 \tag{8}$$

データ数 n が十分に大きければ中心極限定理から括弧内は $N(0, 1^2)$ の標準正規分布に近づくので、上式全体は自由度 1 の χ^2 分布に従うことになる。

1.2 分割表 (クロス表) の独立性の検定

実測値と理論値の乖離に基づく独立性の検定の例を計算方法を示す。2×2 の分割表の場合、総データ数を n とすれば、ある属性が p_1 が起こる期待値は (2 項分布より) np_1 で与えられる。同様に、もう一方の属性 p_2 が起こる場合の期待値は np_2 で与えられる。もし属性 1 が得られたとしても属性 2 が無関係であるならば np_1 の値を知ったとしても np_2 の値は独立である、という。この場合、2 の表の値の中の一つの観測値を知ると他のすべての値が定まる。従って 2 のクロス表の自由度は 1 である。

表 1: 2×2 のクロス表

	p_1	$1-p_1$	合計
p_2	np_1p_2	$n(1-p_1)p_2$	np_2
$1-p_2$	$np_1(1-p_2)$	$n(1-p_1)(1-p_2)$	$n(1-p_2)$
合計	np_1	$n(1-p_1)$	n

予防接種を受けたか、受けなかったかの違いが、ある病気にかかったか、かからなかったか、の違いに影響を及ぼすか否かを考える。次のような表を得たとして (フィッシャー,1970,p.67 より)。予防接種の効果の有無を検定するこ

表 2: 実測値

	病気にかかった者	かからなかった者	合計
予防接種を受けた者	56	6759	6815
受けなかったもの	272	11396	11668
合計	328	18156	18483

とができる。帰無仮説 (H_0) は、予防接種の効果なし、である。2 の分割表の

場合、自由度は 1 である。期待値の求め方は、周辺の値 (合計値) の積を総合計で割ることにより得られる。

表 3: 予測値

	病気にかかった者	かからなかった者	合計
予防接種を受けた者	120.94	6684.06	6815
受けなかった者	207.06	11460.94	11668
合計	328	18156	18483

$$\frac{328 \times 6815}{18483} = 120.94 \tag{9}$$

のようにして計算する (=chitest(実測値範囲, 理論値範囲))。この場合 $\chi^2 = 56.234$ なので実測値は予測値とかけ離れていたことを表す。このことから、帰無仮説は棄却されない。すなわち予防接種の効果がないとは言えない。

一般に $n \times m$ の分割表の場合では自由度 $(n-1)(m-1)$ の χ^2 分布に従う。

1.3 注意を要する適応例

χ^2 検定は適用範囲が広いが、カテゴリー分割に恣意性が残る場合は注意が必要である。例えば、ある意見が支持されるか否かを年代別にカテゴリー分けして χ^2 検定を行う場合である。カテゴリーは 10 代、20 代、30 代などと検定をする者が自由に決められる。このとき 10 代と 20 代を統合して若年層、30 代と 40 代を統合して中年層などのように分類しなおすと、前回有意であったものが有意でなくなったり、あるいはその逆が起こったりする。分類には明確な外的基準が存在することが望ましい。

2 t 分布

1. 確率変数 x と y とは互いに独立で、x は $N(0, 1^2)$ に従い、y は自由度 n の χ^2 分布に従うとき、
$$t = \frac{x}{\sqrt{\frac{y}{n}}} \tag{10}$$
は自由度 n の t 分布に従う。

2. 母集団が正規分布 $N(\mu, \sigma^2)$ であるとき、大きさ n のデータについて、標本平均を $\bar{x}$、標本分散を $s^2 = \dfrac{1}{n}\sum(x_i - \bar{x})^2$ とすれば、
$$t = \frac{\bar{x} - \mu}{\sqrt{\frac{s}{n-1}}} \tag{11}$$
は自由度 $n-1$ の t 分布に従う。

3. データが 2 群から得られたものとする。ある群は $N(\mu_1, \sigma^2)$ に従い、もう一群は $N(\mu_2, \sigma^2)$ に従うものとすれば、

$$
\begin{aligned}
t &= \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma^2}{n_1+n_2}}} && (12)\\
&= \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n_1+n_2}\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}}} && (13)
\end{aligned}
$$

は自由度 $n_1 + n_2 - 2$ の t 分布に従う。

t 分布は自由度が大きくなれば正規分布に近づく。データ数 $n = \infty$ の極限では正規分布に一致する。

2.1 小標本の平均値の有意性検定 (t 検定)

あるデータが得られたとき、母集団平均 μ についての帰無仮説 $H_0 : \mu = \mu_0$ を検定することを考える。手順としては、まず危険率を設定して対応する t の値を調べておく。次に標本平均と標本分散を計算する。標本平均と標本分散から式 (13) を用いて t の値を計算する。この値が、あらかじめ調べておいた危険率に対応する t の値よりも大きければ帰無仮説を棄却し対立仮説を採択する。そうでなければ帰無仮説を棄却しない、となる。

例 (岩原,1957,p.182 より)：A, B 2 クラスに国語のテストを行って次の結果を得た。この結果から国語の成績は A の方が優れていると判定して良い

	標本平均	標本標準偏差	人数
A	82.5	1.3	23
B	81.3	1.2	20

か。1% の有意水準で検定してみる。帰無仮説は予めどちらかとは決められないので両側検定である。表の値を式 (13) に代入すれば

$$
t = \frac{82.5 - 81.3}{\sqrt{\frac{(23)(1.3)^2+(20)(1.2)^2}{23+20-2}\left(\frac{1}{23}+\frac{1}{20}\right)}} = 3.055 \tag{14}
$$

を得る。=tinv(0.005,41) がおよそ 2.97 なので計算した t の値の方が大きい。=tdist(3.055,41,2) は 0.0039 であるから、2 つのクラスに差がないとした帰無仮説の元では、得られたデータ以上の差が生じる可能性は 0.0039 ほどであることがわかる。従って両クラスの国語の成績に差がないとは言えない。

2.2　対応のある t 検定と対応のない t 検定

表 4 のデータは 10 人の患者が、睡眠剤と考えられる 2 つの薬剤 A, B を服用したときの睡眠時間の変化を示す (Student の論文のデータ、フィッシャー,1970, p.95 より)。

表 4: 2 つの薬剤の使用による睡眠時間の延長 (時間単位)

患者	A	B	差 (B-A)
1	+0.7	+1.9	+1.2
2	-1.6	+0.8	+2.4
3	-0.2	+1.1	+1.3
4	-1.2	+0.1	+1.3
5	-0.1	-0.1	0.0
6	+3.4	+4.4	+1.0
7	+3.7	+5.5	+1.8
8	+0.8	+1.6	+0.8
9	0.0	+4.6	+4.6
10	+2.0	+3.4	+1.4
平均	+0.75	+2.33	+1.58

式 (13) は無作為抽出された 2 群のデータの標本平均の差に基づいて両群の母平均の差を検定する場合に用いられる。これを対応のない t 検定という。一方、2 群のデータの間の対に意味があるときには、対応のある t 検定と呼ばれる方法を用いる。

対応のある t 検定の場合、データ対の差が自由度 (データ対の数 -1) の t 分布に従うことを用いて検定される。

t の算出には式 (13) を用いるが自由度が異なる。自由度が $n_1 + n_2 - 2$ ではなく $n - 1$ となる。自由度はデータ対の差を検定していることになると了解すれば納得が行く。

自由度が小さくなることにより、t の値が棄却域に落ちるためには相対的に大きな値を取らねばならなくなる。すなわち対応のある t 検定は対応のない t 検定に比べて厳しい検定方法となる。

2.3　補足

t 検定を用いる場合、母集団分布の正規性、2 群間の等分散性が仮定される。統計学の教科書には、まず等分散性の検定を行って 2 群間の等分散性を検定した後、t 検定を実施すべきであると書かれている。しかし実データでは省略されることもある。

2.4 誤った使用例

1. t 検定は 2 群の母平均の差の有無を検定するものであるが、2 群に分割される場合に誤った例を見かける。例えば、ある心理検査の得点に差があるか否かを検討するために、得点の高い者と低い者とで 2 分割し、高得点群と低得点群で平均得点を比較する、などである。

   群を分割する基準そのものが、自らの得点では論理が循環してしまう。テスト得点が高いから高得点群に分類されたので、その基準そのものの差を検定しても意味がない。

2. t 検定の適用範囲は 2 群の平均値の差の検定に特化している。3 群以上の平均値の差の検定には F 検定 (未解説) を用いるか、または、多重比較 (未解説) という技法を用いる。