統計II (MⅠ)
第5回　適合度検定

授業担当：徳永伸一

東京医科歯科大学教養部　数学講座

[再確認] 統計IIの内容

以下の順序でやってます：

- 6.1 母分散が未知の場合の母平均の推測 (t分布)
- 6.4 対応のある2標本にもとづく母平均の差に関する推測 (t分布)
- 6.2 独立な2標本にもとづく母平均の差に関する推測
- 6.3 等分散の検定 (F分布)
- 6.5 母集団比率に関する推測 (正規近似)

(前回この辺まで)

- 6.6 適合度検定 (χ^2検定)

今日の授業の概要：

- 前回の復習（F分布・母比率に関する推測）
- 「6.6 適合度検定（χ^2検定）」のいけるところまで（結果：「独立性の検定」の手前まで）

では復習から

［復習］ 6.2 独立な2標本にもとづく母平均の差に関する推測（6）

母分散が未知だが等しい場合

次の定理を用いる：

定理6.2

分散の等しい2つの正規母集団

$$N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$$

からそれぞれ独立に無作為抽出された標本を

$$\{X_1, X_2, \cdots, X_{n_1}\}, \{Y_1, Y_2, \cdots, Y_{n_2}\}$$

とし，これらにもとづく標本平均，標本分散（不偏分散）をそれぞれX', Y' S_1^2, S_2^2 とする．

このとき，確率変数

$$T := \{(X' - Y') - (\mu_1 - \mu_2)\} / \sqrt{\{S_p^2 (1/n_1 + 1/n_2)\}}$$

は自由度n_1+n_2-2 のt分布に従う．

ただし，S_p^2はS_1^2, S_2^2の統合分散と呼ばれ，

$$S_p^2 := \{\Sigma (X_i - X')^2 + \Sigma (Y_i - Y')^2\} / (n_1+n_2-2)$$

［復習］ 6.2 独立な2標本にもとづく母平均の差に関する推測 (9)

【あらためて注意】

- 等分散の仮定は本質的.
- 分散が等しい場合について学習したが，そういうケースはたくさんあるのだろうか?
- 例題6.2で，「男女の分散が等しい」と仮定したのはそもそも妥当なのか? (標本分散の値にはかなり差があるけど許容範囲なのか?)
- 分散がほぼ同じに見えるが確信が持てない場合というのも多々あるだろう.

→だったら「2つの母集団の分散が等しい」という仮説を検定すればよい！

→「6.3 等分散の検定 (F分布)」へ.

［復習］ 6.3 等分散の検定 (F分布) (その1)

以下の定理を用いる：

<u>定理6.3</u>

分散の等しい2つの正規母集団

$$N(\mu_1, \sigma^2),\ N(\mu_2, \sigma^2)$$

からそれぞれ独立に大きさn_1, n_2の標本を無作為抽出し，その標本分散 (不偏分散) をそれぞれS_1^2, S_2^2とする.

そのとき$F:=S_1^2/S_2^2$は，

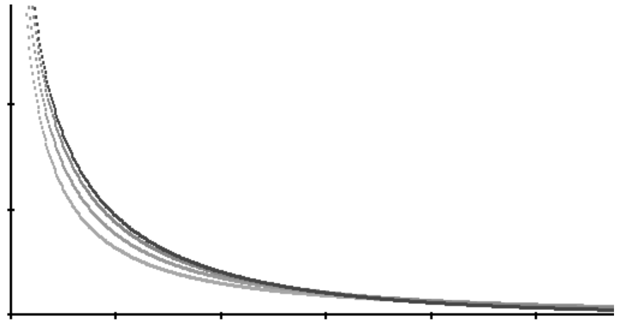
自由度対 (n_1-1, n_2-1) のF分布 (と呼ばれる分布) に従う.

[復習] 6.3 等分散の検定（F分布）（その2）

（グラフは明星大学船津好明先生のWebsiteより転載しています）

F分布の密度関数（1）

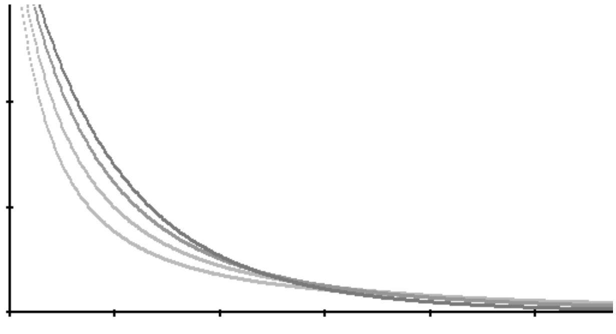
自由度対（1, 1）（1, 2）（1, 5）（1, 20）



[復習] 6.3 等分散の検定（F分布）（その3）

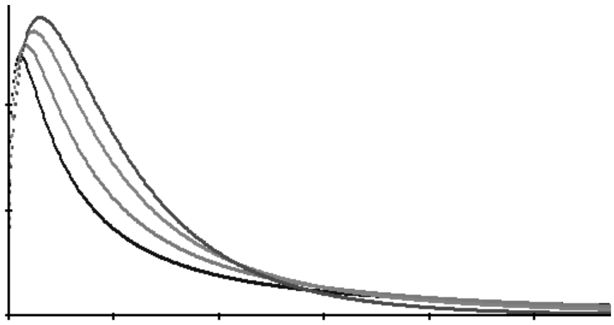
F分布の密度関数（2）

自由度対（2, 1）（2, 2）（2, 5）（2, 20）



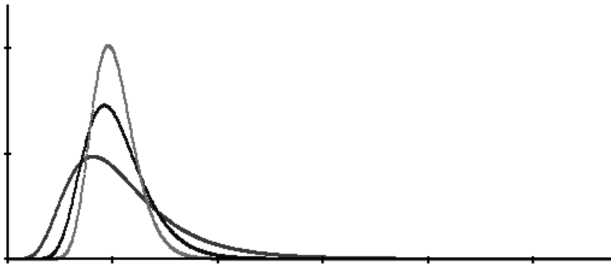
[復習] 6.3 等分散の検定（F分布）（その4）

F分布の密度関数（3）
自由度対（3, 1）（3, 2）（3, 5）（3, 20）



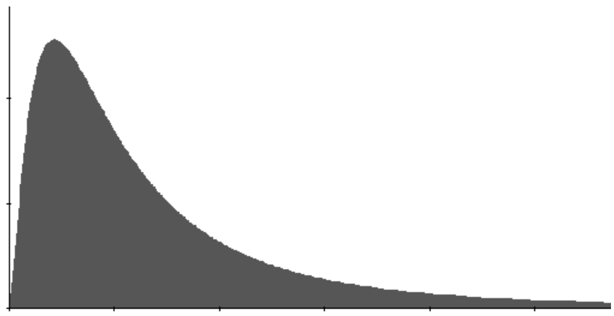
[復習] 6.3 等分散の検定（F分布）（その5）

F分布の密度関数（4）
自由度対（20,20）（50, 50）（100, 100）



[復習] 6.3 等分散の検定 (F分布) (その6)

F分布の密度関数 (5)
自由度対 (5, 5)



[復習] 6.3 等分散の検定 (F分布) (その7)

等分散の検定

- 定理6.3を利用して
  帰無仮説H_0 : $\sigma_1^2 = \sigma_2^2$, 対立仮説H_1 : $\sigma_1^2 \neq \sigma_2^2$
  として有意水準αで検定.
- F分布の密度関数は非対称だが, $F = S_1^2 / S_2^2$の分母・分子は入れ替え可能なので右側 (上側) だけわかればよい.
  - S_1^2, S_2^2 の実測値のうち大きい方を分子にして比を取る.
- 教科書巻末の付表Eを利用して
  自由度対 ($n_1 - 1, n_2 - 1$) のF分布の上側 ($\alpha/2$) 点を求める .

[復習] 6.3 等分散の検定 (F分布) (その8)

例題6.2の場合

	人数	平均値	不偏分散
男	32	15.6	1.61
女	30	13.2	3.26

F_0＝3.26 / 1.61 ＝ 2.02 とする．
有意水準α＝0.05で等分散の検定を行なう：
$F_{31}^{29}(0.05/2)$　　$F_{30}^{30}(0.05/2)$ ＝2.07＞2.02
よりH_0：$\sigma_1^2 = \sigma_2^2$は棄却できない．
よって一応両母分散は等しいと見なすことができる．

[復習] 等分散の検定に関する補足

- 通常の検定問題と違い，帰無仮説（ここでは「等分散仮説」）H_0は棄却されて欲しくない．
- 棄却されなかったからといってH_0の正しさが保証されたわけではないので，教科書でも「一応・・・とみなすことができる」という表現にとどめている．
- H_0が棄却される場合は，他の方法を用いなければならない（この授業では扱わない）．

[復習] 6.5 母集団比率に関する推測 (1)

「母集団比率に関する推測」とは

- ベルヌーイ母集団に関する推測.
  - 2項分布の応用.
  - ある程度大きな標本を扱うケースがほとんどなので，たいていは「2項分布の正規近似」を利用する.
- 「世論調査の類」．支持率調査など，身近に興味深い例が多い.
  - 「統計的な理解を深めるよいチャンス」.

[復習] 6.5 母集団比率に関する推測 (2)

2項分布の正規近似RECALL
中心極限定理により，nが十分大きいとき，
B (n,p) はN (np, np (1-p)) で近似できる.

∵ B (n,p) に従う確率変数は，B (1,p) (という同一の分布) に従う独立なn個の確率変数の和と見なせるから.

- 従って，標本比率P＝X/nの分布も，
  正規分布 N (p, p (1-p) /n) で近似できる.
- さらにPの標準化変数：

$$Z = (P-p) / \sqrt{p(1-p)/n}$$

  は近似的に標準正規分布に従う.
- 分布が決まれば，あとはこれまでと同じ考え方で進めればよい.

［復習］ 6.5 母集団比率に関する推測 (3)

［注意］母比率の推測における推定と検定の違い

母比率をp，標本比率$P=X/n$ の実現値をP_0とする．

- **推定**を行う際は，
  $Z=(P-p)/\sqrt{(p(1-p)/n)}$ の分母（Pの標準偏差）のpを近似値（推定値） P_0 で置き換えて計算．すなわち信頼度$\gamma=1-\alpha$の信頼区間は以下のようになる：

$$\left(P_0-z(\alpha/2)\sqrt{(P_0(1-P_0)/n)},\ P_0+z(\alpha/2)\sqrt{(P_0(1-P_0)/n)}\right)$$

- ただし，誤差の最大値を見積もりたいとき（例：例題6.4の（4））は$p=0.5$ を採用．
  - $p(1-p)$ は$p=0.5$のとき最大値0.25を取ることに注意（2次関数のグラフを思い出せ！）．
- **検定**の際は母比率pの値が帰無仮説で（もちろん厳密な値として）与えられるので，近似の必要がなく，上のような置き換えはまったく無意味．
  - 検定の概念をよく理解していれば間違えないところ．**ここを間違える人は進級させません**

［復習］ 6.5 母集団比率に関する推測 (4)

2項分布の正規近似に関する補足［半整数補正］

- nが大きければかなり良い近似であると思われるが，nが小さいときはどのくらい誤差が出るのだろうか？
  - たとえば5章例題5.2のケースで$B(10,0.5)$ を$N(5,2.5)$ で近似してそのまま計算すると・・・（やってみましょう）
- nが小さいときに少しでも誤差を減らす方法はないか？
- $X\sim B(n,p)$ とする．整数a,bに対し$P(a\leqq X\leqq b)$ を正規近似で求める際，$P(a-0.5\leqq X\leqq b+0.5)$ と補正してから計算した方が誤差が減る．この補正を「不連続補正」ないし「半整数補正」といい，特にnが小さいときに効果的．

  →図で確認しよう．

[復習] 6.5 母集団比率に関する推測 (6)

「半整数補正」続き

- 標本サイズn (ないしnp) が小さいとき特に有効．大標本の場合はあまり差がでない．
- 「メディカル・コメディカルの統計学」にはなぜか載っていませんが知っておくべきでしょう．

以下おまけ：

[復習] From TV News

The Governor of Tokyo said…

"I don't believe the opinion poll, because

they surveyed only 1000 people or so."

(世論調査の結果について記者から感想を尋ねられたのを受けて)

[復習] Is 1000 too small?

1000 is :

1000/127,500,000 = 1/127,000 = 0.0008% of the total population of Japan,

1000/12,140,000 = 1/1214 = 0.08% of the total population of Tokyo.

(確かに母集団と比較すれば小さいが・・・)

But!

[復習] Estimation of error (誤差の評価)

SE = $\sqrt{(p(1-p)/n)}$ depends only on p and n,

i.e., it does not depend on the size of population!

(母集団の大きさは誤差に影響しない!)

- In this case, the value of the error is around 2-3%. (計算してみましょう)
- If n=1000×4=4000, the error is reduced by half.
- In many case, the estimation by a sample with n=1000 is reliable. (←無作為抽出がうまくいっていれば)
- On the other hand, we assume random sampling in the above estimation, which can be suspected.

平成14年度試験問題より

2002年10月1日付けの朝日新聞記事によれば、最近5年間の日本棋院のプロ公式戦15000局において、黒盤勝率は51.86%であった。

(2) 日本棋院は現行ルールでは先手が有利であるとしてルールの改正方針を決定した。この判断についてどう思うか。仮説検定の考えを用いて考察せよ。

[参考] 朝日新聞の解説 (ちなみに東京版夕刊の1面):

「厳しい勝負の世界ではわずか2%の違いも無視できない」

↑「そりゃ違うだろ！」と突っ込むところ.

- 標本サイズが大きいからわずかな差でも有意となるのであって、「勝負の厳しさ」とは無関係.
  - ★適当な有意水準を設定して検定問題を解いてみましょう.
  - (帰無仮説: [先手の勝つ確率] =0.5)

6.6 適合度検定 (χ^2分布) その1

- <u>このセクションで扱う問題について:</u>
- 「観察結果が, 数量でなくて, いくつかのカテゴリーに分類され, その度数で示される場合」の検定問題 (教科書より) .
- ある意味2項分布の一般化 (多項分布) の問題.
- 各カテゴリーごとに「入るか否か」を考えると2項分布の問題となる.
  - 「ここではそうではない. 個々にではなく, 一度に総括的に検定しようというのである」(教科書より) .

6.6 適合度検定（χ^2分布）その2

例：サイコロを60回投げて，それぞれの目の出た度数を調べたところ，以下の結果を得た．

出た目	1	2	3	4	5	6	計
度数	9	13	5	11	16	6	60

この目の出方は偏っているといえるか?

→「偏りはない」という帰無仮説をたてて検定．

【注意】

- 1つの目に注目して「確率1/6で出るかどうか」を検定するなら6.5でやった「母比率の検定」．
- ここではそうではなくて，6つの目の出方を一度に扱う．

6.6 適合度検定（χ^2分布）その3

χ^2分布（と呼ばれる既知の分布）を導入．

以下の定理を用いる：

<u>定理6.4</u>

総度数nが十分大きいとき，カテゴリーの個数がmならば

X＝Σ（（観察度数－期待度数）^{2} /（期待度数））

（食い違い度）は自由度m-1のχ^2分布に従う．

【注意】

各カテゴリーの度数は総度数に縛られるので，

自由度は「（カテゴリー数）－1」．

6.6 適合度検定（χ^2分布）その4

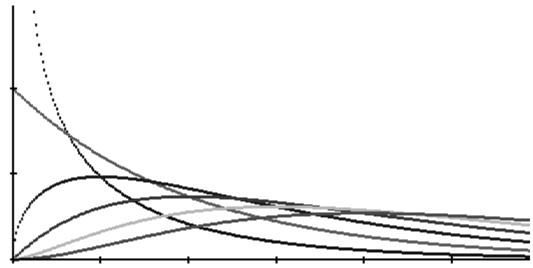
χ^2分布について：

- $Z_1, Z_2, \ldots, Z_k$ が独立かつN（0,1）に従うとき，

$$\chi^2 = Z_1^2 + Z_2^2 + \ldots + Z_k^2$$

　は自由度kの χ^2分布に従う．

- グラフ（自由度n=1～6，作成は明星大学の船津好明先生）：



n=1,2では単調減少，n≧3ではnが増大するとピークが右へ移動

6.6 適合度検定（χ^2分布）その5

冒頭のサイコロの例で検定問題を考えよう．

表6.2

出た目	1	2	3	4	5	6	計
観測度数	9	13	5	11	16	6	60
期待度数	10	10	10	10	10	10	60

- 有意水準 α ＝0.05とする．
- 帰無仮説H_0：各目の出る確率がすべて1/6

　対立仮説H_1：H_0の否定（ある目について，［出る確率］≠1/6）

- カテゴリー数m＝6より食い違い度Xは自由度6－1=5の χ^2分布に従う．
- Xの実現値X_0＝（9－10）^{2}/10＋（13－10）^{2}/10＋・・・＝8.8を自由度5の χ^2分布の上側5%点 χ^2_5（0.05）＝11.1と比較

　→ 棄却されない（偏りがあるとは言えない）

6.6 適合度検定（χ^2分布）その6

<u>一般に，適合度検定とは：</u>

- 「観測された度数が，ある特定の理論分布に適合しているか否かの検定」
  - サイコロの例・・・理論分布＝離散一様分布
- 理論分布に基づいて期待度数を算出
- 定理6.4を用いて「理論分布に適合している」という仮説を検定．
  - 食い違い度を求めて棄却域に落ちるかどうかを判定．

6.6 適合度検定（χ^2分布）その7

<u>適合度検定の例：</u>

問6.4

以下のデータ（プロシャの10軍団×20年間で，1年間に馬に蹴られて死んだ兵士の数X）はポアソン分布（注：教科書第3章に出てきたが授業では割愛した）に適合しているか？

1年間の死亡者数	0	1	2	3	4	合計
軍団の数	109	65	22	3	1	200

注：ポアソン分布の公式 $P(X=x) = e^{-\lambda}\lambda^{x}/x!$ により期待度数が計算できる．以下の手順は同様．