統計学II

早稲田大学政治経済学術院
西郷 浩

本日の目標

- χ^2検定
  - 適合度の検定
  - 独立性の検定
- p 値

適合度の検定(1)

- 実験
  - サイコロを6,000回投げる。
  - 結果

表1: サイコロ投げの結果

出目	1	2	3	4	5	6	合計
度数	1,041	1,010	976	1,000	985	988	6,000

  - このサイコロは公正（すべての目が等しい確率で出現する）か？
  - 不十分な解法
    - どれか一つの目に注目して、成功確率の検定を適用する。
      - 他の目の情報が活用できない。
    - それぞれの目に、成功確率の検定を適用する
      - 有意水準の制御が困難。

適合度の検定(2)

- χ^2分布にもとづく適合度の検定
  - 仮説の設定
    - H_0: $p_k = \frac{1}{6}\ (k = 1, 2, \ldots, 6)$
      - ただし、$p_k = P$(出目がkになる)
    - H_1: $p_k \neq \frac{1}{6}\ (for\ some\ k)$
  - H_0 の下での期待度数を計算する。
    - $E_k = np_{k(0)}\ (k = 1, 2, \ldots, 6)$
      - 例：$E_1 = 6000 \times \frac{1}{6} = 1000$

適合度の検定(3)

表2：観察度数と（H_0のもとでの）期待度数

出目	1	2	3	4	5	6	合計
O_k	1,041	1,010	976	1,000	985	988	6,000
E_k	1,000	1,000	1,000	1,000	1,000	1,000	6,000

– 検定統計量

- $W = \sum_{k=1}^{6} \frac{(O_k - E_k)^2}{E_k}$
  – W の性質
    » H_0 が正しいとき、W は小さくなりやすい。
    » H_0 が正しくないとき、W は大きくなりやすい。
  – H_0 が真とき、W は自由度 $5 = 6 - 1$ の χ^2 分布にしたがう。

適合度の検定(4)

– 検定手続き（有意水準0.05）
  - もし、$W_{obs} > \chi^2_{0.95}(5) = 11.07$ なら、H_0 を棄却する。
  - もし、そうでなければ、H_0 を棄却しない。
    – ただし、$\chi^2_{0.95}(5)$ は、自由度5のχ^2 分布の下側0.95（上側0.05）点

– 検定の結果
  - $W_{obs} = \frac{(1041-1000)^2}{1000} + \cdots + \frac{(988-1000)^2}{1000} = 2.73 < 11.07$
  - H_0 は棄却されない。

独立性の検定(1)

- 習慣的な朝食の摂取状況

表3：男女別朝食の摂取状況

性別	朝食の摂取状況				合計
	6-7/w	4-5/w	2-3/w	0-1/w	
男性	2,641	182	75	332	3,230
女性	3,323	223	53	218	3,817
合計	5,964	405	128	550	7,047

資料：厚生労働省「平成23年国民健康・栄養調査」

– 朝食の摂取状況に男女差があるか？

独立性の検定(2)

– 行和に対する相対度数

表4：行和に対する相対度数（%）

性別	朝食の摂取状況				合計
	6-7/w	4-5/w	2-3/w	0-1/w	
男性	81.8	5.6	2.3	10.3	100.0
女性	87.1	5.8	1.4	5.7	100.0
合計	84.6	5.7	1.8	7.8	100.0

資料：表4

- 女性の方が朝食をきちんと取る人の割合が大きい。

– はっきりとした差があるかどうかは客観的に判断すべき。

- 独立性に関するχ^2検定

独立性の検定(3)

- 独立性に関するχ^2検定
  - 仮説の設定
    - H_0: 性別(X)と朝食の摂取状況(Y)とが独立である。
    - H_1: 性別(X)と朝食の摂取状況(Y)とが独立でない。
  - 仮説の再表現
    - $p_j = P\left(X = x_j\right)$　　$x_1 =$ 男性, $x_2 =$ 女性
    - $q_k = P(Y = y_k)$　　$y_1 = 6 - 7/w$, …, $y_4 = 0 - 1/w$

独立性の検定(4)

- H_0: X と Y が独立である。
  $\Leftrightarrow P\left(X = x_j, Y = y_k\right) = P\left(X = x_j\right)P(Y = y_k) = p_j q_k$
  for all (j, k).
- H_1: X と Y が独立でない。
  $\Leftrightarrow P\left(X = x_j, Y = y_k\right) \neq P\left(X = x_j\right)P(Y = y_k) = p_j q_k$
  for some (j, k).

– H_0 は、p_j や q_k の値を指定していない。
  - データから推定できる。

独立性の検定(5)

- H_0（朝食の摂取状況に男女差がない）が正しいときのq_kの推定
  - $\hat{q}_1 = P(Y = 6-7/w$ |男女差がない$) = \frac{2641+3323}{3230+3817} = \frac{5964}{7047} = 0.846$
    - 同様にして、$\hat{q}_2 = 0.057$, $\hat{q}_3 = 0.018$, $\hat{q}_4 = 0.078$
- H_0（朝食の摂取状況に男女差がない）が正しいときのp_jの推定
  - $\hat{p}_1 = P(X =$ 男性|朝食の摂取状況に差がない$) = \frac{3230}{7047} = 0.458$
    - 同様にして、$\hat{p}_2 = 0.542$
- H_0 が正しいときの期待度数
  - $\hat{E}_{jk} = n\hat{p}_j\hat{q}_k$
    - 例：$\hat{E}_{11} = 7047 \times \frac{3230}{7047} \times \frac{5964}{7047} = 2734$

独立性の検定(6)

– 観察度数と期待度数

表5：観察度数 O_{jk} と（H_0のもとで推定された）期待度数 E_{jk}

性別	朝食の摂取状況				合計
	6-7/w	4-5/w	2-3/w	0-1/w	
男性	2,641	182	75	332	3,230
	2,734	186	59	252	
女性	3,323	223	53	218	3,817
	3,230	219	69	298	
合計	5,964	405	128	550	7,047

独立性の検定(7)

– 検定統計量

- $W = \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{\left(O_{jk} - E_{jk}\right)^2}{E_{jk}}$
  - W の性質
    - » H_0 が正しいとき、W は小さくなりやすい。
    - » H_0 が正しくないとき、W は大きくなりやすい。
  - H_0 が正しいとき、W は自由度 $(J-1) \times (K-1) = (2-1) \times (4-1) = 3$ の χ^2 分布にしたがう。

独立性の検定(8)

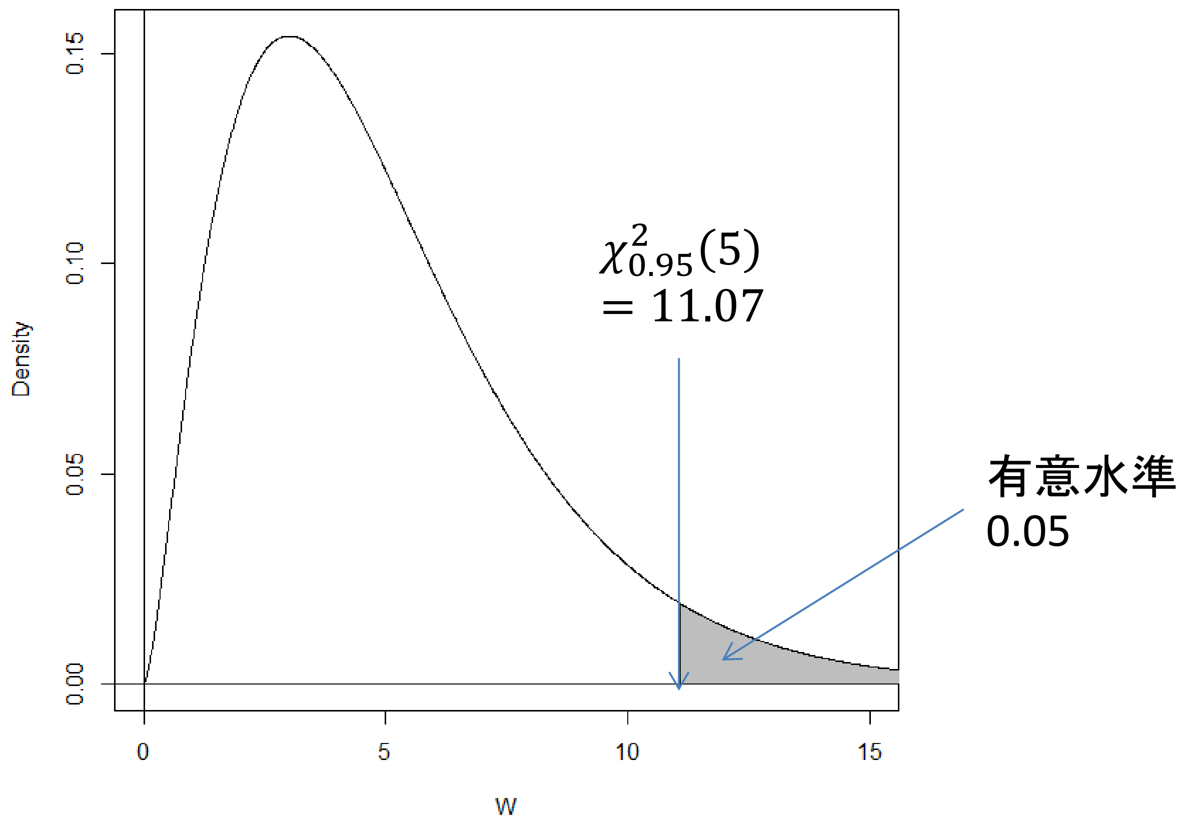
- 検定手続き（有意水準0.05）
  - もし、$W_{obs} > \chi^2_{0.95}(3) = 7.8147$ なら、H_0 を棄却する。
  - もし、そうでなければ、H_0 を棄却しない。
    - ただし、$\chi^2_{0.95}(3)$ は、自由度3のχ^2 分布の下側0.95（上側0.05）点
- 検定の結果
  - $W_{obs} = \frac{(2641-2734)^2}{2734} + \cdots + \frac{(218-298)^2}{298} = 61.1 > 7.8147$
  - H_0 は棄却される。

p値(1)

- p値
  - 右側検定でのp値の計算方法（定義）
    - $p - value = P(W > W_{obs} | H_0)$
      - H_0 が正しいときの、検定統計量の標本分布において、観察された検定統計量よりも右側の裾の面積
      - H_0 の信憑性の尺度と解釈されることもある。
        - » p値が小さいほど、H_0 が疑わしい。
  - 有意水準0.05の検定の実行方法（2つ）
    - 検定統計量 ＞ 有意水準0.05の臨界値 ⇒H_0を棄却する。
    - p値 ＜ 0.05 ⇒ H_0 を棄却する。
      - 両方ともH_0の採否については同じ結論を導く。

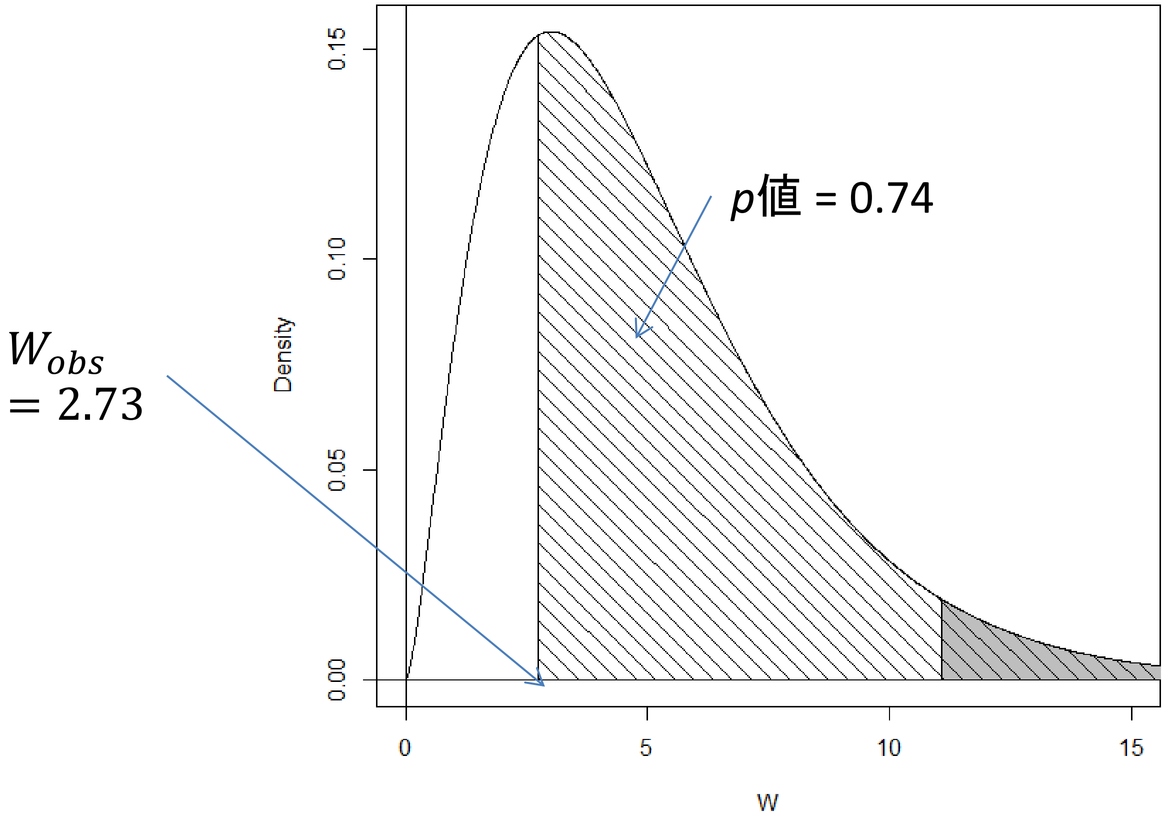
p値(2)

図1 : サイコロの実験における棄却域と有意水準



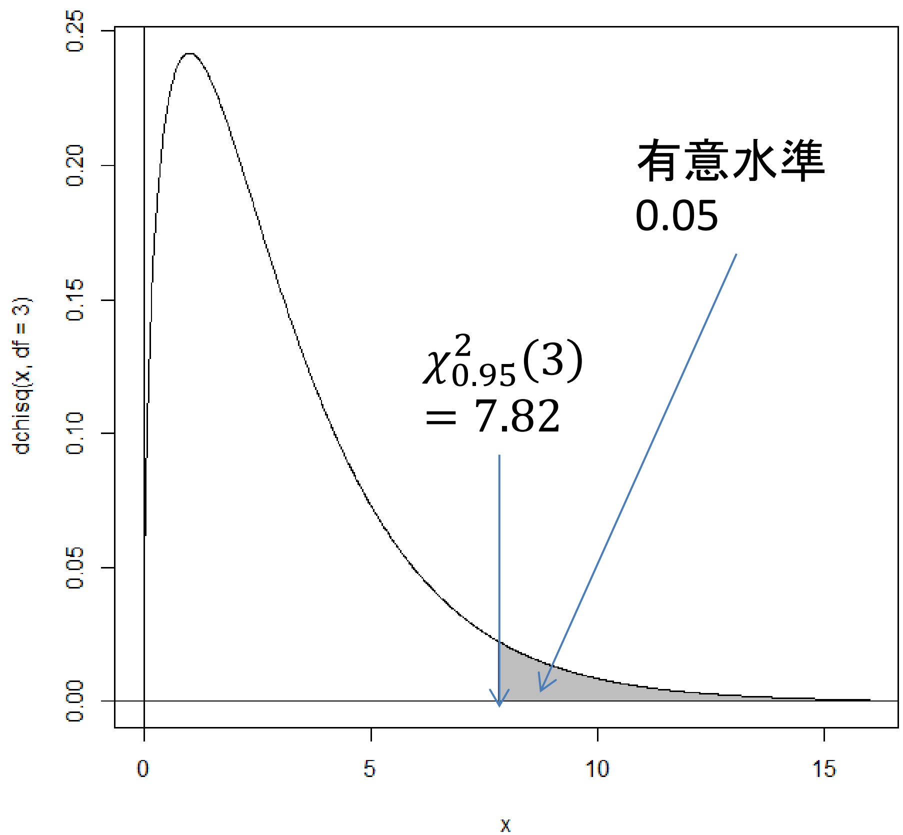
p値(3)

図2：サイコロの実験におけるp値と有意水準



p値(4)

図3：朝食摂取状況データにおける棄却域と有意水準



$\chi^2_{0.95}(3)$
$= 7.82$

有意水準
0.05

W_{obs}
$= 61.1$

p値 =
3.46×10^{-13}

p値(5)

- p値
  - 有意水準0.05の検定の実行方法（2つ）
    - p値 $< 0.05 \Rightarrow$ H_0 が棄却される。
    - p値 $\geq 0.05 \Rightarrow$ H_0 が棄却されない。
  - 注:
    - 両側検定:分布の両側を使ってp値を計算する。
      - 片側検定の2倍になる。