検定とは？

- 統計的仮説検定
(testing statistical hypothesis)

観測結果（＝標本）の統計的性質に基づいて、ある主張の正誤を合理的に判断するための手法

以後、シンプルに「検定」(test)と呼ぶ

検定とは？ (2)

- 例えば
  - あるコインを5回投げたら、5回続けて表が出た。このコインは公平だろうか？
  - ある自動車のニューモデルが発売されて一ヶ月が経った。カタログでは燃費が向上してるというが、ほんとうだろうか？
  - 音声認識のプログラムを何人かの被験者に使ってもらったところ、男性よりも女性のほうが音声認識率がよいような気がした。女声向きの新手法を発明できたのか？それともただの偶然？

検定の手順

1. 主張したい仮説と反対の仮説（帰無仮説） H_0 をたてる
2. 帰無仮説を正しいと仮定し、観測結果の確率を計算する
3. この確率が、ある値（有意水準という）よりも小さければ、帰無仮説 H_0 を棄却する

検定の手順 (2)

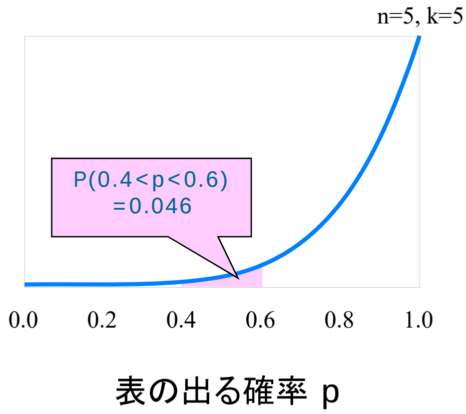
「あるコインを5回投げたら、5回続けて表が出た」

1\.

$(0.4<p<0.6)$

2\.

3\.



⇒ 「　」

ケーススタディ(1) (分散比検定-1)

「ニューモデルの燃費は向上したか？」

- 車種A 5台の燃費測定値

  18.0 18.2 15.3 17.5 17.3

- 車種B 10台の燃費測定値

  13.4 10.4 17.8 16.3 16.4

  14.9 16.4 17.7 21.5 15.3

(先週の課題と同じ値)

ケーススタディ(1) (分散比検定-2)

- 帰無仮説 H_0:

「車種Aと車種Bの間に燃費の違いはない」

この仮説のもとに、分散比 F は自由度対

$$\left[1, \sum_{i=A,B}(n^{(i)}-1)\right]$$

の F-分布に従う

ケーススタディ(1) (分散比検定-3)

- 車種A、B、全体の平均燃費

$$\bar{X}^{(A)} = \frac{1}{5}\sum_{j=1}^{5} X_j^{(A)} = 17.26$$

$$\bar{X}^{(B)} = \frac{1}{10}\sum_{j=1}^{10} X_j^{(B)} = 16.01$$

$$\bar{X} = \frac{1}{15}\sum_{i=A,B}\sum_{j=1}^{n^{(i)}} X_j^{(i)} = 16.426\cdots$$

ケーススタディ(1) (分散比検定-4)

- 級間平方和

$$s_{between}^2 = \sum_{i=A,B} n^{(i)}\left(\bar{X}^{(i)} - \bar{X}\right)^2 = 5.208\cdots$$

- 級内平方和

$$s_{within}^2 = \sum_{i=A,B}\sum_{j=1}^{n^{(i)}}\left(X_j^{(i)} - \bar{X}^{(i)}\right)^2 = 81.94\cdots$$

ケーススタディ(1) (分散比検定-5)

- 級間(不偏)分散

$$U_b^2 = s_{between}^2 / 1 = 5.208\cdots$$

- 級内(不偏)分散

$$U_w^2 = s_{within}^2 / \sum_{i=A,B} (n^{(i)} - 1) = 6.303\cdots$$

ケーススタディ(1) (分散比検定-6)

- 不偏分散比 $U_b^2 / U_w^2 = 0.826\cdots$
- 2クラスの不偏分散比は自由度対

$\left[1, \sum_{i=A,B} (n^{(i)} - 1)\right]$ の F-分布に従う

⇒ [,]

自由度対 [n1, n2] のF分布に関する境界値(有意水準5%)

- $n_1 = 1$, $n_2 = 13$ の交点の値と比較

		n1												
		1	2	3	4	5	6	7	8	9	10	11	12	13
n2	1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58

ケーススタディ(1) (分散比検定-7)

- 不偏分散比：
- 自由度対 [　,　] の F-分布の境界値（有意水準 5%）：　4.67 （>　　　）

「分散比が F-境界値より大きければ、帰無仮説を棄却できる」

⇒

「　　　」