

経済統計：前期期末試験

村澤 康友

2011 年 8 月 3 日

注意：3 問とも解答すること。

- (20 点) 以下の用語の定義を式または言葉で書きなさい (各 20 字程度)。
 - 統計的仮説
 - 対立仮説
 - 第 2 種の誤り
 - 有意水準
- (30 点) $\text{Bin}(1, p_X)$, $\text{Bin}(1, p_Y)$ から独立に抽出した大きさ m, n の無作為標本の標本比率 (標本平均) を \hat{p}_X, \hat{p}_Y とする。次の片側検定問題を考える。

$$H_0 : p_X = p_Y \quad \text{vs.} \quad H_1 : p_X > p_Y.$$

- \hat{p}_X, \hat{p}_Y の漸近分布を求めなさい。また $\hat{p}_X - \hat{p}_Y$ の漸近分布を求めなさい。
 - 検定統計量を与えなさい (プールした標本比率 \hat{p} を使っても使わなくてもよい)。それは H_0 の下でどのような分布に近似的に従うか？
 - 検定統計量の値は 1.78 であった。p 値を求めなさい。また有意水準 5 % の検定を実行しなさい。
3. (50 点) 男女の相性は血液型で決まるとの俗説がある。その真偽を科学的に検証したい。そこで無作為に選んだ 117 組の夫婦の血液型を調べたところ、次表の結果が得られた (数値は百分率を四捨五入)。

夫 \ 妻	A	O	B	AB	計
A	0.15	0.14	0.06	0.07	0.41
O	0.10	0.07	0.10	0.03	0.30
B	0.08	0.09	0.04	0.01	0.22
AB	0.04	0.00	0.03	0.00	0.07
計	0.37	0.30	0.23	0.10	1.00

夫の血液型を X , 妻の血液型を Y とし, 夫婦の血液型の同時確率関数を $p_{X,Y}(\cdot, \cdot)$, その周辺確率関数を $p_X(\cdot), p_Y(\cdot)$ とする。

- 検定問題を定式化しなさい (言葉でなく数式で)。
- H_0 の下で期待される各欄の相対度数を表で示しなさい (小数第 4 位まで)。
- 独立性の χ^2 検定統計量は H_0 の下でどのような分布に近似的に従うか? (証明不要)
- 有意水準 5 % の検定の棄却域を定めなさい。
- χ^2 検定統計量の値は 14.2309624 であった。有意水準 5 % の検定を実行しなさい。

解答例

1. 仮説検定の基本用語

(a) 母集団分布に関する仮説 .

- 「母集団」がなければ 0 点 .

(b) 帰無仮説を棄却するとき代わりに採択する仮説 .

- 「帰無仮説以外の仮説」は不十分なので 2 点 (例えば片側検定) .

(c) H_1 が真なのに H_0 を採択する誤り .

(d) 許容する第 1 種の誤りの確率 .

2. 2 標本の母比率の差の検定

(a)

$$\hat{p}_X \stackrel{a}{\sim} N\left(p_X, \frac{p_X(1-p_X)}{m}\right),$$
$$\hat{p}_Y \stackrel{a}{\sim} N\left(p_Y, \frac{p_Y(1-p_Y)}{n}\right).$$

両者は独立なので

$$\hat{p}_X - \hat{p}_Y \stackrel{a}{\sim} N\left(p_X - p_Y, \frac{p_X(1-p_X)}{m} + \frac{p_Y(1-p_Y)}{n}\right).$$

- 各 5 点 .
- 後半は前半と整合的なら OK.
- 母数を統計量にしたら 0 点 .

(b) 検定統計量は

$$Z := \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}_X(1-\hat{p}_X)/m + \hat{p}_Y(1-\hat{p}_Y)/n}},$$

または

$$Z := \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\hat{p}(1-\hat{p})(1/m + 1/n)}}.$$

H_0 の下で $Z \stackrel{a}{\sim} N(0, 1)$.

- 統計量で 5 点 , 分布で 5 点 .
- 統計量に未知母数を残したら 0 点 .
- 厳密には誤りではないかもしれないが , t 分布は不可とする .

(c) p 値は 0.037538 . p 値が有意水準以下なので H_0 は棄却 .

3. 独立性の χ^2 検定

(a)

$$H_0 : p_{X,Y}(\cdot, \cdot) = p_X(\cdot)p_Y(\cdot) \quad \text{vs.} \quad H_1 : p_{X,Y}(\cdot, \cdot) \neq p_X(\cdot)p_Y(\cdot).$$

(b) 次表の通り .

夫 \ 妻	A	O	B	AB	計
A	0.1517	0.1230	0.0943	0.0410	0.41
O	0.1110	0.0900	0.0690	0.0300	0.30
B	0.0814	0.0660	0.0506	0.0220	0.22
AB	0.0259	0.0210	0.0161	0.0070	0.07
計	0.37	0.30	0.23	0.10	1.00

(c) $\chi^2(9)$.

- 自由度がなければ 0 点.

(d) $[16.919, \infty)$.

- 前問の解答と整合的なら OK (ただし χ^2 分布のみ).

(e) 検定統計量の値が採択域に入るので H_0 を採択.

- 前問の解答と整合的なら OK (ただし χ^2 分布の右側棄却域のみ).

答案は返却します. 採点や成績に関する質問にも応じます. オフィスアワーの時間に研究室まで来てください (夏季休業中は随時).