分割表, 適合度検定, 独立性検定　問題１　解答

$\boxed{1}$ 日本人の血液型の人数比は A 型が 40%, B 型が 20%, AB 型が 10%, O 型が 30%, とされている. ある地域で無作為に選ばれた 100 人の血液型について次のようなデータ表が得られた. このデータは日本人の血液型の分布と同じといえるか. 有意水準を $\alpha = 0.05$ として検定せよ.

血液型	A 型	B 型	AB 型	O 型	計
観測度数	28	27	15	30	100

[解]: 適合度の検定を行うため,

- 帰無仮説 H_0 : この地域の血液型分布は日本人の血液型の分布と一致する.
- 対立仮説 H_1 : この地域の血液型分布は日本人の血液型の分布と一致しない.

と仮説を立てる. 与えられたデータおよび帰無仮説 H_0 の仮定の上で期待度数を求めると, 以下の表となる.

血液型	A 型	B 型	AB 型	O 型	計
観測度数	28	27	15	30	100
期待度数	40	20	10	30	100

適合度の検定における検定統計量の実現値 χ^2 を計算すると,

$$\begin{aligned}\chi^2 &= \frac{(28-40)^2}{40} + \frac{(27-20)^2}{20} + \frac{(15-10)^2}{10} + \frac{(30-30)^2}{30} \\ &= 8.55.\end{aligned}$$

ここで検定統計量は自由度 3 のカイ 2 乗分布に従う. このとき棄却域 R は, カイ 2 乗分布の上側パーセント点

$$\chi_3^2(0.05) = 7.815$$

を使い,

$$R = \{x \mid 7.815 < x\}$$

と設定する. このとき $\chi^2 \in R$ より, 帰無仮説 H_0 は棄却され, 対立仮説 H_1 が採択される. したがって, 『この地域の血液型分布は日本人の血液型の分布と一致しない.』と結論できる.

$\boxed{2}$ 6 面サイコロを 600 回投げて, それぞれの目の出た回数を調べたところ, 以下のデータが得られた. このサイコロは偏っていると言えるか. 有意水準を $\alpha = 0.05$ として検定せよ.

出た目	1	2	3	4	5	6	計
回数	83	90	105	95	110	117	600

[解]: 適合度の検定を行うため,

- 帰無仮説 H_0 : サイコロの出る目は偏っていない. (出る目の確率は全て等しい)

- 対立仮説 H_1 : サイコロの出る目は偏っている.

と仮説を立てる. 与えられたデータおよび帰無仮説 H_0 の仮定の上で期待度数を求めると, 以下の表となる.

出た目	1	2	3	4	5	6	計
観測度数	83	90	105	95	110	117	600
期待度数	100	100	100	100	100	100	600

適合度の検定における検定統計量の実現値 χ^2 を計算すると,

$$\begin{aligned}\chi^2 &= \frac{(83-100)^2}{100}+\frac{(90-100)^2}{100}+\frac{(105-100)^2}{100}\\ &\quad +\frac{(95-100)^2}{100}+\frac{(110-100)^2}{100}+\frac{(117-100)^2}{100}\\ &= 8.28.\end{aligned}$$

ここで検定統計量は自由度 5 のカイ 2 乗分布に従う. このとき棄却域 R は, カイ 2 乗分布の上側パーセント点

$$\chi^2_5\,(0.05) = 11.07$$

を使い,

$$R = \{x \mid 11.07 < x\}$$

と設定する. このとき $\chi^2 \notin R$ より, 帰無仮説 H_0 は棄却されない. したがって, 『サイコロの出る目は偏っている.』とはいえない.

$\boxed{3}$ 10 面サイコロを 100 回投げて, それぞれの目の出た回数を調べたところ, 以下のデータが得られた. このサイコロは偏っていると言えるか. 有意水準を $\alpha = 0.05$ として検定せよ.

出た目	1	2	3	4	5	6	7	8	9	10	計
回数	11	11	9	13	2	12	10	16	7	9	100

[解]: 適合度の検定を行うため,

- 帰無仮説 H_0 : サイコロの出る目は偏っていない. (出る目の確率は全て等しい)
- 対立仮説 H_1 : サイコロの出る目は偏っている.

と仮説を立てる. 与えられたデータおよび帰無仮説 H_0 の仮定の上で期待度数を求めると, 以下の表となる.

出た目	1	2	3	4	5	6	7	8	9	10	計
観測度数	11	11	9	13	2	12	10	16	7	9	100
期待度数	10	10	10	10	10	10	10	10	10	10	100

適合度の検定における検定統計量の実現値 χ^2 を計算すると,

$$\begin{aligned}\chi^2 &= \frac{(11-10)^2}{10}+\frac{(11-10)^2}{10}+\frac{(9-10)^2}{10}+\frac{(13-10)^2}{10}+\frac{(2-10)^2}{10}\\ &\quad+\frac{(12-10)^2}{10}+\frac{(10-10)^2}{10}+\frac{(16-10)^2}{10}+\frac{(7-10)^2}{10}+\frac{(9-10)^2}{10}\\ &= 12.6.\end{aligned}$$

ここで検定統計量は自由度 9 のカイ 2 乗分布に従う. このとき棄却域 R は, カイ 2 乗分布の上側パーセント点

$$\chi_9^2(0.05) = 16.92$$

を使い,

$$R = \{x \mid 16.92 < x\}$$

と設定する. このとき $\chi^2 \notin R$ より, 帰無仮説 H_0 は棄却されない. したがって, 『サイコロの出る目は偏っている.』とはいえない.

$\boxed{4}$ ある市の 1 日当たりの交通事故件数は平均 2 のポアソン分布に従うと言われている. この市の交通事故件数を実際に 200 日間計測したところ, 以下のデータが得られた. このデータは平均 2 のポアソン分布に従っているか, その適合性を有意水準を $\alpha = 0.05$ として検定せよ.

事故件数	0	1	2	3	4	5 以上	計
観測度数	28	50	52	38	16	16	200

[解]: 適合度の検定を行うため,

- 帰無仮説 H_0 : 平均 2 のポアソン分布に従っている
- 対立仮説 H_1 : 平均 2 のポアソン分布に従っていない

と仮説を立てる. まず, 帰無仮説の仮定の上で, 表に対応する確率の値を計算する. 確率分布表は

$$P(X=k) = \frac{e^{-2}2^k}{k!}$$

より, 以下の表で与えられる. (小数点第 2 位まで四捨五入している)

事故件数	0	1	2	3	4	5 以上	計
確率	0.14	0.27	0.27	0.18	0.09	0.05	1.0

確率分布表と与えられたデータから期待度数を求めると, 以下の表となる.

事故件数	0	1	2	3	4	5 以上	計
観測度数	28	50	52	38	16	16	200
期待度数	28	54	54	36	18	10	200

適合度の検定における検定統計量の実現値 χ^2 を計算すると,

$$\begin{aligned}\chi^2 &= \frac{(28-28)^2}{28} + \frac{(50-54)^2}{54} + \frac{(52-54)^2}{54} + \frac{(38-36)^2}{36} + \frac{(16-18)^2}{18} + \frac{(16-10)^2}{10} \\ &= 4.304.\end{aligned}$$

ここで検定統計量は自由度 5 のカイ 2 乗分布に従う. このとき棄却域 R は, カイ 2 乗分布の上側パーセント点

$$\chi_5^2\,(0.05) = 11.07$$

を使い,

$$R = \{x \mid 11.07 < x\}$$

と設定する. このとき $\chi^2 \notin R$ より, 帰無仮説 H_0 は棄却されない. したがって, 『平均 2 のポアソン分布に従っていない』とはいえない.