

数理統計：レポート 3 略解

担当：熊谷 隆

1. 帰無仮説「H: 赤, 青, 黄, 白の色の花を咲かせる率が 9:3:3:1」を検定する。 $k=4, n=160$ で

$$Z_n = \frac{(83-90)^2}{90} + \frac{(29-30)^2}{30} + \frac{(34-30)^2}{30} + \frac{(14-10)^2}{10} = 2.7111 \dots$$

であり、一方 χ^2_3 分布に従う μ について $\mu([x(0.05), \infty)) = 0.05$ となる $x(0.05)$ は 7.81 なので、 $Z_n < x(0.05)$ となり、仮説は採択される。つまり、メンデルの遺伝法則と一致していると言える。

2. 帰無仮説「H: 4 社の車の燃費に差がない」を検定する。 $k=4, n=64, S' = 17 \cdot 0.75^2 + 18 \cdot 0.85^2 + 15 \cdot 0.95^2 + 14 \cdot 1.00^2 = 50.105, \bar{X} = \frac{1}{64}(17 \cdot 13.6 + 18 \cdot 14.0 + 15 \cdot 13.4 + 14 \cdot 13.2) = 869/64 = 13.5781 \dots$ (以下 13.58 として計算), $S'' = 17 \cdot (13.6 - \bar{X})^2 + 18 \cdot (14.0 - \bar{X})^2 + 15 \cdot (13.4 - \bar{X})^2 + 14 \cdot (13.2 - \bar{X})^2 = 5.6896$ なので、

$$F := \frac{S''(n-k)}{S'(k-1)} = 341.376/150.315 = 2.27107 \dots$$

となる。一方、 $F_{60}^3([x(0.05), \infty)) = 0.05$ となる $x(0.05)$ は 2.76 なので、 $F < x(0.05)$ となり、仮説は採択される。つまり、4 社の車の燃費に差があるとは言えない。

3. 従業員数を x 人、売り上げを y 百万円とすると、 $\bar{x} = 47/8, \bar{y} = 54/8$ であり、 x^2, y^2, xy は以下の通り。よって $S_{xx} := \sum_i x_i^2 - 8\bar{x}^2 = 321 - 276.125 = 44.875, S_{xy} := \sum_i x_i y_i - 8\bar{x}\bar{y} = 354 - 317.25 = 36.75, S_{yy} := \sum_i y_i^2 - 8\bar{y}^2 = 396 - 364.5 = 31.5$ である。

	1 号店	2 号店	3 号店	4 号店	5 号店	6 号店	7 号店	8 号店	計
x^2	36	36	16	49	4	16	64	100	321
y^2	49	49	25	64	9	36	64	100	396
xy	42	42	20	56	6	24	64	100	354

従って $\hat{b} := S_{xy}/S_{xx} = 0.8189 \dots, \hat{a} := \bar{y} - \hat{b}\bar{x} = 1.9387 \dots, S_{yy} - S_{xy}^2/S_{xx} = 1.40389 \dots$ となるので、標本回帰直線は $y = 1.94 + 0.82x$ 、残差平方和は 1.40 である (少数第 3 位を四捨五入した)。

4. 帰無仮説「H: 札幌の視聴率 43.1 % ($p = 0.431$)」を、対立仮説「 $p > 0.431$ 」に対して片側検定する。 μ を $N(0, 1)$ に従う確率測度とすると、中心極限定理 (定理 1.3.3) より

$$\lim_{n \rightarrow \infty} P\left(\sqrt{\frac{n}{p(1-p)}}(X - p) > a\right) = \mu((a, \infty))$$

$\mu((a, \infty)) = 0.05$ を満たす a は 1.6449 であり、 $\mu((b, \infty)) = 0.01$ を満たす b は 2.3263 である。今 $X = 0.506, n = 200$ だから

$$\sqrt{\frac{n}{p(1-p)}}(X - p) = (28.5575 \dots) \cdot 0.075 = 2.1418 \dots$$

となり、 $a < 2.1418 < b$ なので、危険率 5 % ならば H は棄却され (つまり視聴率は関西・関東地区より高いといえる)、危険率 1 % ならば H は採択される (つまり視聴率が関西・関東地区より高いとは言えない)。