

2018年6月 問12

一番優しい、統計学の教本

吉田

問題

問 12 次の表は、2017 年度プロ野球におけるリーグ毎の球団別ホームゲーム年間入場者数（単位は万人）である。

セントラル・リーグの球団別年間入場者数

球団 A	球団 B	球団 C	球団 D	球団 E	球団 F	平均	偏差平方和
218	303	198	296	201	186	233.7	13,549

パシフィック・リーグの球団別年間入場者数

球団 G	球団 H	球団 I	球団 J	球団 K	球団 L	平均	偏差平方和
209	177	167	145	161	253	185.3	7,763

資料：日本野球機構

各リーグ内において入場者数は独立で同一の分布に従い、かつ、セントラル・リーグとパシフィック・リーグの各球団の年間入場者数の母分散は等しいと見なし、両リーグの球団別年間入場者数の母平均に差があるかどうかを2つの方法で検定したい。

[1] 2つの母平均の差に関する t 検定を行う。 t -値として、次の ①～⑤のうちから最も適切なものを一つ選べ。 24

- ① 0.07 ② 0.33 ③ 1.05 ④ 1.82 ⑤ 2.00

[2] 同様の帰無仮説・対立仮説に対して、一元配置分散分析を行うことを考える。一元配置分散分析における F -値として、次の ①～⑤のうちから最も適切なものを一つ選べ。 25

- ① 0.14 ② 1.11 ③ 1.66 ④ 3.30 ⑤ 4.01

知っておくべき知識

- t値の式

- 分散分析表の作り方

t値の計算方法

$$t = \frac{\text{群1の平均値} - \text{群2の平均値}}{\sqrt{\text{2群のプールした分散} \left(\frac{1}{\text{群1の例数}} + \frac{1}{\text{群2の例数}} \right)}}$$

$$\text{2群のプールした分散} = \frac{(\text{群1の例数} - 1) \times \text{群1の不偏分散} + (\text{群2の例数} - 1) \times \text{群2の不偏分散}}{\text{群1の例数} + \text{群2の例数} - 2}$$

セリーグ、パリーグそれぞれの不偏分散を計算

$$s_c^2 = \frac{(218 - 233.7)^2 + (303 - 233.7)^2 + (198 - 233.7)^2 + (296 - 233.7)^2 + (201 - 233.7)^2 + (186 - 233.7)^2}{6 - 1}$$
$$= \frac{13549}{6 - 1} = 2709.8$$

$$s_p^2 = \frac{(209 - 185.3)^2 + (177 - 185.3)^2 + (167 - 185.3)^2 + (145 - 185.3)^2 + (161 - 185.3)^2 + (253 - 185.3)^2}{6 - 1}$$
$$= \frac{7763}{6 - 1} = 1552.6$$

T値を計算！

$$\text{2群のプールした分散} = \frac{(\text{群1の例数} - 1) \times \text{群1の不偏分散} + (\text{群2の例数} - 1) \times \text{群2の不偏分散}}{\text{群1の例数} + \text{群2の例数} - 2}$$

$$= \frac{(6 - 1) \times 2709.8 + (6 - 1) \times 1552.6}{6 + 6 - 2} = 2131.2$$

$$t = \frac{\text{群1の平均値} - \text{群2の平均値}}{\sqrt{\text{2群のプールした分散} \left(\frac{1}{\text{群1の例数}} + \frac{1}{\text{群2の例数}} \right)}}$$

$$= \frac{233.7 - 185.3}{\sqrt{2131.2 \left(\frac{1}{6} + \frac{1}{6} \right)}} = 1.82$$

正解：4

分散分析表とは？

要因	平方和S	自由度df	不偏分散v	F値
群	S(群)	df(群) (群の数-1)	v(群) (=S(群)/df(群))	v(群)/v(残)
残差	S(残)	df(残) (全データ-群の数-1)	v(残) (=S(残)/df(残))	
全体	S(全)	df(全) (全データ-1)		

それぞれの平方和を計算！

	セリーグ	パリーグ	全体
		218	209
		303	177
		198	167
		296	145
		201	161
		186	253
データ数	6	6	12
平均	233.7	185.3	209.5
標準偏差	52.1	39.4	50.7
分散	2709.9	1552.7	2574.6
平方和	13549.3	7763.3	28321

分散分析表とは？

要因	平方和S	自由度df	不偏分散v	F値
群	S(群)	df(群) (群の数-1)	v(群) (=S(群)/df(群))	v(群)/v(残)
残差	S(残)	df(残) (全データ-群の数-1)	v(残) (=S(残)/df(残))	
全体	S(全)	df(全) (全データ-1)		

要因	平方和S	自由度df	不偏分散v	F値
群	7009	1	7009	3.29
残差	21312	10	2131.2	
全体	28321	11		

正解：4