

2018年6月 問14

一番優しい、統計学の教本
吉田

問 14 Tさんは47都道府県別のデータを用いて次の重回帰モデルを推定した。

$$\log(\text{犯罪発生率}) = \alpha + \beta_1 \times \text{失業率} + \beta_2 \times \log(\text{賃金}) + \beta_3 \times \log(\text{警察官数}) + \text{誤差項}$$

ここで、「犯罪発生率」は人口10万人当たりの刑法犯認知件数（単位は人口10万対件数）、「失業率」は完全失業率（単位は%）、「賃金」は一般労働者の一ヶ月の1人あたりの平均給与額（単位は千円）、「警察官数」は人口10万人あたりの警察官数（単位は人口10万対人数）である。

統計ソフトウェアを利用して、上記の重回帰モデルを推定したところ、次の出力結果を得た。なお、出力結果の一部を加工している。

出力結果

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.08851	1.92346	-3.685	0.000635
失業率	0.09408	0.05541	1.698	0.096773
log(賃金)	2.41815	0.31781	7.609	1.71e-09
log(警察官数)	-0.06498	0.22718	-0.286	0.776233

Residual standard error: 0.2077 on 43 degrees of freedom

Multiple R-squared: 0.6062, Adjusted R-squared: 0.5787

F-statistic: 22.06 on 3 and 43 DF, p-value: 8.353e-09

[1] 失業率が2.8, $\log(\text{賃金})$ が5.6, $\log(\text{警察官数})$ が5.3のとき, $\log(\text{犯罪発生率})$ の予測値として, 次の①～⑤のうちから最も適切なものを一つ選べ。 28

- ① 2.5 ② 3.0 ③ 4.2 ④ 5.6 ⑤ 6.4

[2] β_3 の値は大
仮説検定を行
定が棄却され
選べ。 29

- ① 0.1%

[3] 次の記述 I～III は, 出力結果に関するものである。

- I. 有意水準1%で0でない回帰係数(定数項を含む)の数は2である。
- II. 賃金が高い都道府県では, 犯罪発生率は低い傾向がある。
- III. 自由度調整済み決定係数の値は約0.58である。

記述 I～III に関して, 次の①～⑤のうちから最も適切なものを一つ選べ。

30

- ① Iのみ正しい。 ② IIIのみ正しい。 ③ IとIIのみ正しい。
④ IとIIIのみ正しい。 ⑤ IIとIIIのみ正しい。

知っておくべき知識

- 回帰分析の基礎
- 回帰分析でのT値の算出方法
- 自由度

回帰式を確認

$$\log(\text{犯罪発生率}) = \alpha + \beta_1 \times \text{失業率} + \beta_2 \times \log(\text{賃金}) + \beta_3 \times \log(\text{警察官数}) + \text{誤差}$$

出力結果

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.08851	1.92346	-3.685	0.000635
失業率	0.09408	0.05541	1.698	0.096773
log(賃金)	2.41815	0.31781	7.609	1.71e-09
log(警察官数)	-0.06498	0.22718	-0.286	0.776233

$$\log(\text{犯罪発生率}) = -7.1 + 0.09 \times \text{失業率} + 2.4 \times \log(\text{賃金}) - 0.065 \times \log(\text{警察官数}) + \text{誤差}$$

問 1 4 (1)

[1] 失業率が2.8, $\log(\text{賃金})$ が5.6, $\log(\text{警察官数})$ が5.3のとき, $\log(\text{犯罪発生率})$ の予測値として, 次の①～⑤のうちから最も適切なもの一つ選べ。 28

① 2.5

② 3.0

③ 4.2

④ 5.6

⑤ 6.4

$$\begin{aligned}\log(\text{犯罪発生率}) &= -7.1 + 0.09 \times \text{失業率} + 2.4 \times \log(\text{賃金}) - 0.065 \times \log(\text{警察官数}) + \text{誤差} \\ &= -7.1 + 0.09 \times 2.8 + 2.4 \times 5.6 - 0.065 \times 5.3 \\ &= -7.1 + 0.252 + 13.44 - 0.3445 \\ &\cong 6.4\end{aligned}$$

正解: 5

回帰分析でのT値の求め方

[2] β_3 の値は大体 -0.5 であるという主張があったとする。この説を検証するため仮説検定を行うことにした。帰無仮説 $\beta_3 = -0.5$, 対立仮説 $\beta_3 \neq -0.5$ の仮説検定が棄却される有意水準として、次の ① ~ ⑤ のうちから最も小さいもの一つ選べ。 29

- ① 0.1% ② 1% ③ 5% ④ 10% ⑤ 15%

log(警察官数)	-0.06498	0.22718	-0.286	0.776233
-----------	----------	---------	--------	----------

帰無仮説: $\beta_3=0$ の時のt-value

$$t - value = \frac{estimate - \text{帰無仮説となる数字}}{SE}$$

$$t - value(\beta_3 \text{の帰無仮説: } -0.5) = \frac{-0.06498 - (-0.5)}{0.22718} = 1.915$$

T分布表と見比べる

■T分布表を見るときに重要なこと：自由度

■自由度の概念：

□<https://best-biostatistics.com/contingency/degree-freedom.html>

出力結果

Coefficients:

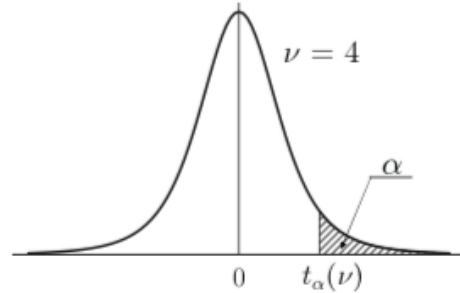
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.08851	1.92346	-3.685	0.000635
失業率	0.09408	0.05541	1.698	0.096773
log(賃金)	2.41815	0.31781	7.609	1.71e-09
log(警察官数)	-0.06498	0.22718	-0.286	0.776233

Residual standard error: 0.2077 on 43 degrees of freedom

誤差項の
自由度は43

得られた $T=1.915$ はどこに位置するか？

ν	α				
	0.10	0.05	0.025	0.01	0.005
1	3.078	6.14	12.06	31.821	63.656
2	1.886	2.20	4.03	6.965	9.925
3	1.638	2.53	3.82	4.541	5.841
4	1.533	2.32	2.76	3.747	4.604
5	1.476	2.15	2.71	3.365	4.032
6	1.440	1.95	2.47	3.143	3.707
7	1.415	1.95	2.65	2.998	3.499
8	1.397	1.60	2.06	2.896	3.355
9	1.383	1.33	2.62	2.821	3.250
10	1.372	1.12	2.28	2.764	3.169
11	1.363	1.96	2.01	2.718	3.106
12	1.356	1.82	2.79	2.681	3.055
13	1.350	1.71	2.60	2.650	3.012
14	1.345	1.61	2.45	2.624	2.977
15	1.341	1.53	2.31	2.602	2.947
16	1.337	1.46	2.20	2.583	2.921
17	1.333	1.40	2.10	2.567	2.898
18	1.330	1.34	2.01	2.552	2.878
19	1.328	1.29	2.93	2.539	2.861
20	1.325	1.25	2.86	2.528	2.845
21	1.323	1.21	2.80	2.518	2.831
22	1.321	1.17	2.74	2.508	2.819
23	1.319	1.14	2.69	2.500	2.807
24	1.318	1.11	2.64	2.492	2.797
25	1.316	1.08	2.60	2.485	2.787
26	1.315	1.06	2.56	2.479	2.779
27	1.314	1.03	2.52	2.473	2.771
28	1.313	1.01	2.48	2.467	2.763
29	1.311	1.99	2.45	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660



$$1.684 < 1.915 < 2.021$$

両側10%

両側5%

正解: 4

問 1 4 (3)

[3] 次の記述 I ~ III は、出力結果に関するものである。

- I. 有意水準 1% で 0 でない回帰係数（定数項を含む）の数は 2 である。
- II. 賃金が高い都道府県では、犯罪発生率は低い傾向がある。
- III. 自由度調整済み決定係数の値は約 0.58 である。

記述 I ~ III に関して、次の ① ~ ⑤ のうちから最も適切なものを一つ選べ。

30

- ① I のみ正しい。
- ② III のみ正しい。
- ④ I と III のみ正しい。
- ⑤ II と III のみ正

正解: 4

出力結果

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-7.08851	1.92346	-3.685	0.000635
失業率	0.09408	0.05541	1.698	0.096773
log(賃金)	2.41815	0.31781	7.609	1.71e-09
log(警察官数)	-0.06498	0.22718	-0.286	0.776233

Residual standard error: 0.2077 on 43 degrees of freedom
Multiple R-squared: 0.6062, Adjusted R-squared: 0.5787