

経済統計学 B 期末試験解説

2008 年 2 月 ; 担当 星野

- このハンドアウトは <http://stat.w3.kanazawa-u.ac.jp/econs/list.html> からダウンロード可能。
- (配点 10×4) (a) 25.9 (b) 27.7 (c) 例えば高齢の高給取りばかりの企業と給与の安い若者ばかりの企業では、前者の方が平均給与が高くなって当然である。従って年齢構成の違う企業間で給与水準を比較する場合、年齢構成を合わせて比較しないと意味がない。加重平均給与を計算する際の重み(年齢構成)が共通になっていれば正解とした。重みは B 社に合わせても C 社に合わせても、 $1/5$ ずつとして単純な算術平均を用いても良い。(d) 各年代の加重平均給与をいかに計算するかが問われている。各社 $1/3$ の重みで単純平均を使うのは良くはない。各社の従業員数を年代別に求め、それに重みを比例させる方が実態に近い。各社の従業員数は表に書いてないが、これが分かるとして計算するか、分からないので三社の規模が等しいと仮定して計算するかいずれかだろう。
 - (配点 10) 分布関数が何か理解出来ていれば容易。階段状のグラフとなる。
 - (配点 5×3) (a) $1/2$ (b) $y = 1 + \frac{1}{2}x$ (c) 0.25。教科書 p.93 を見る。 x, y の相関係数を計算出来れば、 x, y の標準偏差が等しい事から、回帰直線の式は即座に求められる。相関係数の二乗が決定係数になっている事に注意。
 - (配点 $5+10+10$) (a) 題意の誤解が生じていた。「入社三年後の」が修飾するのは「大卒男子」までのつもりであった。「入社三年後の」が「高卒女子」も修飾するように読んだように見える学生が複数居て驚いたが、「入社三年後の高卒女子の初任給」であっても初任給でしかない。もしかして初任給の意味が分からなかったのだろうか。経験年数ゼロでモデルの予測する賃金を計算しろ、という意味である。「入社三年後の大卒男子給与」は $10 + 2 \times 1 + 3 \times 1 \times 3 + 4 \times 3 + 1 \times 1 = 34$ となる。「高卒女子の初任給」は $10 + 2 \times 0 + 3 \times 0 \times 0 + 4 \times 0 + 0 \times 0 = 10$ である。従って 24 が正解。(b) モデル (2) の方が自由度修正済み決定係数が高い。故に、モデル (1) より説明力が高い。(1) は学歴を説明変数に含むが (2) は含まない。従って、学歴の給与への影響は余り無いと言えそうだ。(c) X_1 の代わりに

$$X' = \begin{cases} 2 & \text{when 大卒以上} \\ 1 & \text{when 高卒} \\ 0 & \text{when 中卒} \end{cases}$$

を使うという誤答が複数見られた。なぜこれが良くないかというと、中卒が高卒になる事の効果と高卒がそれ以上になる事の効果等しいと仮定してしまっているからである。複数のダミー変数を使うのが望ましい:

$$W_1 = \begin{cases} 1 & \text{when 大卒以上} \\ 0 & \text{when それ以外} \end{cases}, \quad W_2 = \begin{cases} 1 & \text{when 高卒} \\ 0 & \text{when それ以外} \end{cases}$$

その上でモデルが X_1 に依存する部分を変更する。 αX_1 の部分は $\alpha_1 W_1 + \alpha_2 W_2$ として、 $\beta X_1 \cdot X_2$ の部分は $\beta_1 W_1 \cdot X_2 + \beta_2 W_2 \cdot X_2$ とするのが良いだろう。

- (配点 10) 自己相関係数を時差を変えて計算し、自己相関係数が高くなる時差が周期と言える。以上。