統計学 課題その2

担当: 社会情報科学部 山本 岳洋 (<u>t.yamamoto@sis.u-hyogo.ac.jp</u>) 締切: 4月30日(金) 21:00

解答は以下の URL にある Google フォームから提出すること.

https://forms.gle/54T7WoQC7QQjThCN7

講義資料「2変数データの記述と要約」や教科書・参考書を学習し、本課題に取り組むこと、なお、本課題を解答するにあたり、電卓、Excel、プログラミング等自由に使用して良い、もちろん手計算でも良い。

(余裕のある人向け)

Excel の関数に慣れている人や関数を使ってみたい人は Excel で解答を直接求めても良いです。 Excel を用いる場合は、AVERAGE, VAR.P, STDEV.P, COVARIANCE.P, CORREL などについて調べてください。また、分散、標準偏差、共分散を用いる場合はそれぞれ VAR.P, STDEV.P,COVARIANCE.P を必ず用いてください(VAR や STDEV などは用いないでください。理由は本講義の後半で説明します)。 Excel を使用したくて、よく分からない人は「Excel 関数 AVERAGE 使い方」のようなキーワードで検索すると関数の使い方に関するページが見つかると思いますし、遠慮なく Teams やメールで質問してください。

問 1.

学生 5 名に対し数学と英語の試験を実施したところ、それぞれ以下の様な得点となった. このとき、以下の問い a.-f.に答えよ、以下、解答は、小数点第 2 位まで記載すれば十分である。手計算で求める場合は、途中計算を小数点第 3 位まで求めて計算してみよ.

	数学	英語
学生 1	8	6
学生 2	6	3
学生 3	4	4
学生 4	8	5
学生 5	10	7

- a. 数学と英語の得点の共分散:
- b. 数学と英語の得点の相関係数(講義資料 p.33 の関係式を使うと早い):

いま,新たに「学生 6」にこの数学の試験を実施したところ,9点であった.この学生 6 の 英語の得点を回帰により予測したい.これは,英語の得点を y,数学の得点を x とすると, y=ax+b となる直線の a b b の値を決めれば予測することができる.そこで,学生 $1\sim5$ の 5 名のデータに対して最小二乗法を用いて直線を求めることを考える.講義資料の p.56 の式に共分散,x の分散,x の平均,y の平均を代入すると,y=ax+b の a b b の値はそれぞれ以下の通りになる.

- c. 直線 y=ax+b の傾き a:
- d. 直線 y=ax+b の切片 b:

この直線のxに学生6の数学の得点9点を代入すると、予測される英語の得点は 以下の点となる(小数点第2位まで求めれば良い).

e.予測された学生6の英語の得点:

いま,この学生6に英語の試験を実施したところ,8点であった.いま予測した英語の得点(つまり,e.の解答)と本当の得点(つまり,8点)の差の二乗は以下の通りになる.

f. 予測した得点と本当の得点の差の二乗

問 2,

擬似相関の例を<u>自分で考えて</u>記述せよ. 擬似相関となっている 2 変数が何か,また,見えない変数が何かを明示すること. たとえば,講義資料の例に沿って記述すれば以下の様な解答例となる.

都市の「交番数」と「犯罪発生件数」には正の相関関係がみられる.しかし、これらの変数間には因果関係はなく、両者とも都市の「人口」という隠れた変数との因果関係によって「交番数」と「犯罪発生件数」が擬似相関となっていると考えられる.

問 3(自由記述):

今回の課題や、「2 変数データの要約と記述」で扱った内容について、特に説明してほしいところがあれば記述してください。また、講義への要望や感想、山本へのコメントなどもあればここに自由に記述してください。可能な限り次週の講義で補足したいと思います。コメントを匿名で講義中に表示する可能性がありますので、書きにくい内容は Teams の DM やメールなどで構いません。

問4

問1を解答するにあたり、最も使用したツールを選択してください. (成績評価とは無関係)

- 手計算・電卓
- Excel
- Google スプレッドシート
- その他 (プログラミングなど)

最後に、本課題を解く上で参考にした書籍(教科書・参考書以外)・ウェブサイト等があればここに記載してください。

以上.