

統計学 1 中間課題 答案（再提出）

文理学部 情報科学科

5419045 高林 秀

2022 年 12 月 14 日

概要

本稿は、後期総合教育科目である統計学 1 の中間課題として与えられた、健康診断データの分析に関するレポートである。

平均や、標準偏差といった各種指標数値の計算には Python のライブラリである Pandas を用いて計算を行った。また、グラフの描画には matplotlib を使用した。

はじめに、各数値データ（年齢、身長、体重、最大血圧、最小血圧）に関して、それぞれ平均や標準偏差、度数分布表、ヒストグラムを作成し、元データ全体がどのような分布になっているかについて考察した。その後、各数値（量的）データと質的データ（以降は「血圧判定」「心電図判定」を指す）の関連性を調べるため、相関比を算出し、結果から導かれることを考察した。その後、質的データ同士の関連性をクラメールの関連係数を用いて考察した。

最後に、算出した各種指標値の結果から、どの属性が質的データに影響を与えているのかを考察した。

目次

1	各数値データの概要	1
2	各数値データの相関比	1
3	各属性のクラメールの関連係数	1

- 1 各数値データの概要
- 2 各数値データの相関比
- 3 各属性のクラメールの関連係数

付録

計算機環境

図表一覧

ソースコード