

データの分析 まとめ

問題プリント（全27問）

氏名 _____

得点 _____ / 27点

DRILL まとめテスト（標準）

基本（1～15）

1 中央値（データ数が奇数）

9人のバドミントン部員のサーブ練習の成功本数：
15, 22, 19, 25, 18, 20, 24, 17, 21。中央値を求めよ。

中央値

2 四分位数（データ数が奇数・11個）

次の11個のデータについて、 Q_1 ・ Q_3 ・四分位範囲を求めよ：
5, 9, 12, 15, 18, 20, 23, 26, 29, 31, 34

$Q_1 =$ $Q_3 =$ 四分位範囲

3 四分位数（データ数が偶数・8個）

次の8個のデータについて、 Q_1 ・ Q_3 ・四分位範囲を求めよ：
4, 9, 11, 14, 16, 19, 21, 25

$Q_1 =$ $Q_3 =$ 四分位範囲

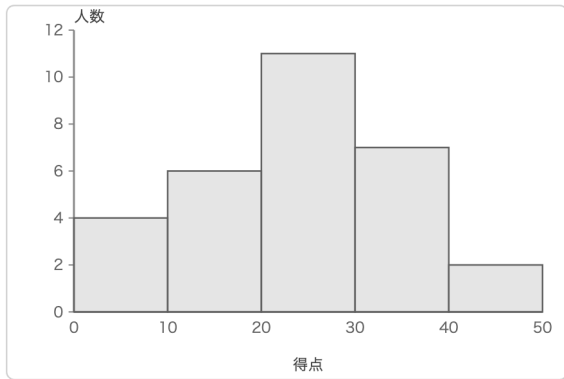
4 平均値と中央値の違い

次の7個のデータについて、平均値と中央値を求めよ：3, 4, 5, 5, 6, 7, 40

平均値 中央値

5 ヒストグラムの読み取り（最頻値）

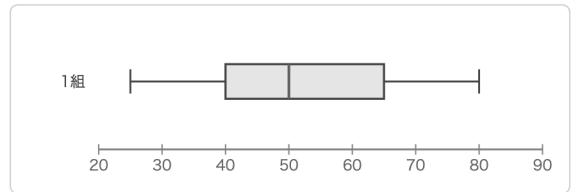
30人の小テストの得点の分布を表したヒストグラムである。最頻値が含まれる階級の階級値を求めよ。



階級値

6 箱ひげ図の読み取り

1組の小テスト結果を表した箱ひげ図である。中央値と第3四分位数を読み取れ。



中央値 第3四分位数

7 分散・標準偏差（基本）

データ7, 7, 13, 13について、平均・分散・標準偏差を求めよ。

平均 分散 標準偏差

8 公式「 $(x^2の平均)-(平均)^2$ 」を使う

データ3, 3, 3, 11, 11, 11（6個）の分散を、 $\overline{x^2} - (\bar{x})^2$ の公式で求め、標準偏差も答えよ。

分散 標準偏差

9 全員に同じ数を足す→分散はどうなる？

あるデータの平均は18、分散は9である。全員に7を加えると、新しいデータの平均・分散・標準偏差はいくらになるか。

平均 分散 標準偏差

10 全員をb倍する→分散はb²倍

あるデータの平均は8、分散は5である。全員を3倍すると、新しいデータの平均・分散・標準偏差はいくらになるか。

平均 分散 標準偏差

11 共分散の計算

4人の生徒の練習時間 x （時間）と、その後の小テストの得点 y （点）は $x = 1, 3, 5, 7$ 、 $y = 2, 5, 6, 11$ であった（生徒の順は対応している）。共分散 S_{xy} を求めよ。

$S_{xy} =$

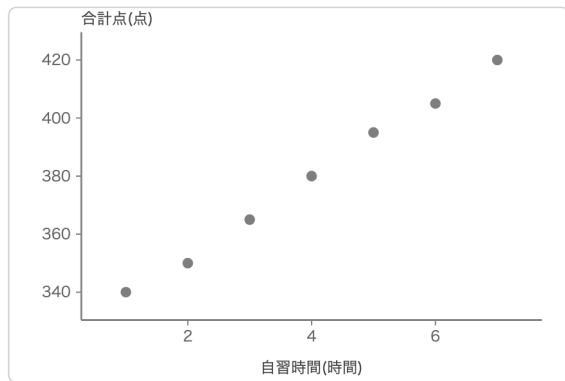
12 公式に代入するだけ

あるデータについて、 x の標準偏差 $S_x = 2.5$ 、 y の標準偏差 $S_y = 4$ 、共分散 $S_{xy} = -7$ であった。相関係数 r を求めよ。

$r =$

13 散布図の読み取り

次の散布図は、7人の1日の自習時間（時間）と定期テストの合計点（点）の関係を表したものである。相関の種類を「正」「負」「なし」のいずれかで答えよ。



相関の種類（正 / 負 / なし）

14 相関係数の範囲

次の値0.95, -1.2, 0, 1.05, -0.6のうち、相関係数の値としてありえないものを、小さい順にコンマ区切りで全て答えよ。

ありえない値

15 仮説検定：表を使った割合の計算

友達が「このコインは公正だ（表が出る確率0.5）」と言うので、20回投げてみたところ表が17回出た。公正なコインを20回投げる実験を1000回シミュレーションした結果（の一部）を下の表に示す。「20回中17回以上表が出る」割合（%）を求め、5%を基準に「このコインは公正である」という仮説を棄却できるか判断せよ。

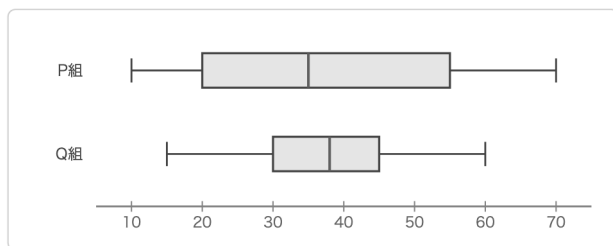
表の枚数	10	11	12	13	14	15	16	17	18	19	20
1000回中の回数	176	160	120	74	37	15	5	1	0	0	0

割合 % 判断（棄却できる / 棄却できない）

発展（16～25）

16 箱ひげ図の比較・四分位範囲

P組・Q組の小テスト結果を表した箱ひげ図である。それぞれの四分位範囲を求め、どちらの組の方が成績が「そろっている」（散らばりが小さい）と言えるか、P・Qの記号で答えよ。



P組の四分位範囲 Q組の四分位範囲
 そろっている組

17 変換の複合（ $y=bx+a$ 、 b が負の数）

あるデータの平均は12、分散は8である。全員を-2倍してから5を引いた新しいデータ（ $y = -2x - 5$ ）の平均・分散・標準偏差を求めよ。

平均 分散 標準偏差

18 2グループ合併・加重平均

生徒15人のグループAの小テストの平均点は72点、生徒25人のグループBの平均点は84点であった。AとBを合わせた全体の平均点を求めよ。

全体の平均点

19 2グループ合併・加重分散

生徒6人のグループAは平均5点・分散3、生徒4人のグループBは平均10点・分散3であった。AとBを合わせた10人全体の平均点と分散を求めよ。

全体の平均点 全体の分散

20 相関係数の計算（同じ数の集まりの並べ替え）

5人の生徒の1回目のテスト得点 x （点）と2回目のテスト得点 y （点）は $x = 10, 20, 30, 40, 50$ 、 $y = 30, 10, 50, 20, 40$ であった（生徒の順は対応している）。共分散 S_{xy} と相関係数 r を求めよ。

$S_{xy} =$ $r =$

21 相関係数の性質・総合選択

相関係数についての次の記述のうち、正しいものを全て選び、番号を小さい順にコンマ区切りで答えよ。

- 単位を変えると相関係数の値は必ず変化する
- 相関係数は常に-1以上1以下の値をとる
- 外れ値が1つ混ざっても相関係数の値は変化しない
- 相関係数の絶対値が1に近いほど、直線的な関連が強い

正しい記述の番号

22 仮説検定（別の題材・サイコロ）

友達が持ってきたサイコロについて「6の目が出やすい」と主張したので、40回投げた表は、公正なサイコロ（6の目が出る確率 $\frac{1}{6}$ ）で40回投げる実験を1000回シミュレーションした結果（の一部）を下に示す。「40回中11回以上6の目が出る」割合（%）を求め、5%を基準に「このサイコロは公正である」という仮説を棄却できるか判断せよ。

6の目の回数	6	7	8	9	10	11	12	13	14	15
1000回中の回数	167	162	134	95	59	32	16	7	3	1

割合 % 判断（棄却できる / 棄却できない）

23 四分位数+外れ値の判定（データ数が多い場合）

あるゲームアプリのユーザー13人の1週間のプレイ時間（分、小さい順）：12, 15, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 75。 Q_1 ・ Q_3 ・上側の基準値を求め、75分が外れ値と言えるか判定せよ。

$Q_1 =$ $Q_3 =$ 上側の基準値 判定

24 変換の複合・文章題（総合）

あるクラス20人の小テストの平均点は50点、標準偏差は6点であった。全員の点数を1.5倍にしたあと、10点を一律に加える特別採点を行った（先に1.5倍、その後+10）。最終的な平均点・分散・標準偏差を求めよ。

最終的な平均点 分散 標準偏差

25 章の総仕上げ：平均・四分位数・外れ値の判定

9人が参加したマラソン大会の記録（分、すでに小さい順に並んでいる）：38, 40, 42, 44, 46, 48, 50, 52, 90。(1) 平均値、(2) Q_1 ・中央値・ Q_3 ・四分位範囲、(3) 1.5×IQR基準で90分の記録が外れ値と言えるかを求めよ。

平均値 $Q_1 =$ 中央値 $Q_3 =$
 四分位範囲 上側の基準値 判定

DRILL

記述チャレンジ

答えだけでなく「なぜそうなるか」を文章と式で書く練習。模範解答と見比べて、書き方の型を身につけよう。

記1 平均値・中央値・分散・標準偏差は、どれも「データの特徴を1つの数値にまとめる」道具である。これらを「外れ値に対する強さ」という観点から比較し、どのような場面でどれを使うべきかを説明せよ。

記2 相関係数と仮説検定は、一見すると別々の技術に見えるが、どちらも「見た目の印象だけで判断せず、数値の基準で判断する」という共通の考え方にもとづいている。それぞれ具体例を1つずつ挙げながら、この共通点を説明せよ。

EXAM
共通テスト形式チャレンジ

大問1 目標時間 8分/配点 20点

花子さんのクラス（Aクラス）では、文化祭の来場者アンケートで「満足度スコア」（0点から100点）を聞いている。ランダムに選んだ5人分の記録は72, 78, 80, 84, 86点であった。先生「これを使って、データの基本的な特徴を調べてみよう。」

- (1) Aクラスの5人分のスコアについて、平均値 $\bar{x}$ = アイ点、偏差の2乗の和は ウエオ、分散は V = カキである。また、標準偏差は s = クである（根号を用いて表せ）。

アイ ウエオ カキ ク

- (2) 太郎「後日、この5人には一律5点のボーナス点を追加することになりました。」

ボーナス点を追加した後の平均点は ケコ点、分散は サシである。

ケコ サシ

- (3) 花子「隣のBクラスでも同じアンケートを5人からとったところ、75, 79, 80, 81, 85点でした。」

Bクラスの平均値は スセ点、分散は ソタである。また、Aクラス（ボーナス点追加前、分散24）とBクラスでは、どちらのスコアがより「そろっている」（散らばりが小さい）と言えるか、チに $A \cdot B$ の記号で答えよ。

スセ ソタ チ

大問2 目標時間 8分/配点 20点

太郎さんは、文化祭の模擬店の準備として、5日間の気温 x （℃）と、模擬店の来場者数 y （人）を記録した。 $x = 20, 22, 24, 26, 28, y = 34, 37, 43, 40, 46$ であった（日の順は対応している）。先生「気温と来場者数に関係がありそうか、数値で確かめてみよう。」

- (1) $\bar{x}$ = アイ℃、 $\bar{y}$ = ウエ人である。共分散は S_{xy} = オカ、相関係数は r = キである。

アイ ウエ オカ キ

- (2) 花子「相関係数が0.9ということは、かなり強い正の相関ですね。ここで、 x の単位を℃からF（華氏）に変える（ $x' = 1.8x + 32$ という変換）と、 r の値はどうなるでしょうか。」

相関係数の性質を用いて、単位を変換した後の相関係数は $r' =$ クであることを答えよ。

ク

- (3) 先生「実は、このアンケートには6日目の記録もありました。ただしこの日は大雨で模擬店を早めに閉めてしまい、気温は30℃と高かったのに、来場者数はわずか10人でした。」

花子「(1)では $r = 0.9$ というかなり強い正の相関でしたよね。この6日目のデータ（ $x = 30, y = 10$ ）を加えても、同じように強い相関になるのでしょうか。」

6日分のデータ $x = 20, 22, 24, 26, 28, 30, y = 34, 37, 43, 40, 46, 10$ について、新しい平均は $\bar{x} =$ ケコ℃、 $\bar{y} =$ サシ人、共分散は $S_{xy} =$ スである（ S_{xy} は符号もつけて入力せよ）。また、この6日分で相関係数を計算し直すと、符号は セ（正・負・なしのいずれか）になる。

ケコ サシ ス セ（正 / 負 / なし）