

データのばらつきを表す値

樋口さぶろお

龍谷大学工学部数理情報学科

使える統計! L03(2013-10-09 Wed)

今日の目標

- ① 分散・標準偏差の意味が説明できる
- ② データから分散・標準偏差が計算できる
- ③ $X + b$, $a \times X$, の平均値・分散・標準偏差が, X の平均値・分散・標準偏差から求められる.



<http://hig3.net>

ここまで来たよ

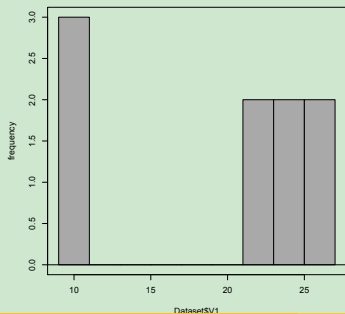
- 1 復習:データの位置を代表する値
 - 代表値
- 2 データのばらつきを表す値
 - データのばらつきを表す値
 - 平均偏差と分散
 - 標準偏差
 - 分散の意味
 - 平均値・分散・標準偏差の変換

Q1

Quiz(平均値中央値最頻値)

次のヒストグラムから求めよう.

- 1 中央値
- 2 最頻値
- 3 平均値

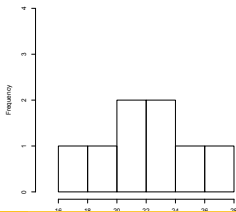
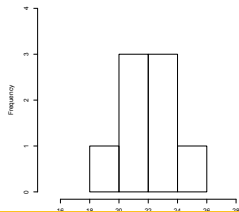
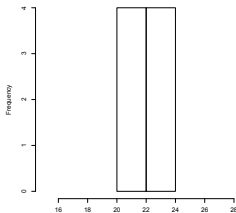
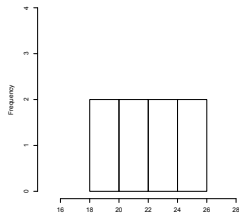


ここまで来たよ

- 1 復習:データの位置を代表する値
 - 代表値
- 2 データのばらつきを表す値
 - データのばらつきを表す値
 - 平均偏差と分散
 - 標準偏差
 - 分散の意味
 - 平均値・分散・標準偏差の変換

データの位置がすべてじゃない!

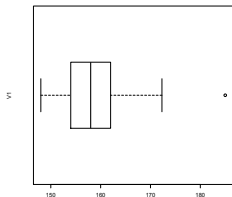
代表値データの位置を1個に要約した値: 中央値, 最頻値, 平均値.
平均値が同じである分布



データのばらつきを表す値

● 範囲 (range) =

● 四分位範囲 (IQR=inter quartile range)=



Q2

Quiz(範囲)

次のデータの、範囲, 四分位範囲を求めよう。

14 14 15 16 18 18 18 25

ここまで来たよ

- 1 復習:データの位置を代表する値
 - 代表値
- 2 データのばらつきを表す値
 - データのばらつきを表す値
 - 平均偏差と分散
 - 標準偏差
 - 分散の意味
 - 平均値・分散・標準偏差の変換

平均偏差と分散

(1 個のデータの) **偏差** (deviation) = データの値 - 平均値

- **平均偏差** (mean deviation): 偏差の絶対値 (負なら正に直したもの) の平均
- **分散** (variance): (偏差)² の平均

(例) データ: 9 10 10 12 14

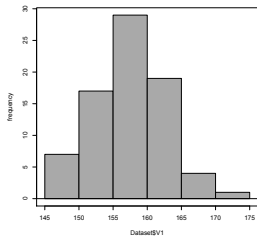
| | データ | 偏差 | 偏差の絶対値 | (偏差) ² |
|-------|-----|----|--------|-------------------|
| | 14 | | | |
| | 12 | | | |
| | 10 | | | |
| | 10 | | | |
| | 9 | | | |
| 合計 | | | | |
| 合計/個数 | | | | |

ここまで来たよ

- 1 復習:データの位置を代表する値
 - 代表値
- 2 データのばらつきを表す値
 - データのばらつきを表す値
 - 平均偏差と分散
 - **標準偏差**
 - 分散の意味
 - 平均値・分散・標準偏差の変換

標準偏差

標準偏差 (standard deviation) =



某国民的アイドル集団 (77 人) の身長

- 平均値 = $\frac{148+148.5+\dots+172.3}{77} = 158\text{cm}$
- 分散 = $\frac{(148-158)^2+(148.5-158)^2+\dots+(172.3-158)^2}{77} = 26.0 \text{ cm}^2$
- 標準偏差 = $\sqrt{\frac{(148-158)^2+(148.5-158)^2+\dots+(172.3-158)^2}{77}} = 5.1 \text{ cm.}$

ここまで来たよ

- 1 復習:データの位置を代表する値
 - 代表値
- 2 データのばらつきを表す値
 - データのばらつきを表す値
 - 平均偏差と分散
 - 標準偏差
 - 分散の意味
 - 平均値・分散・標準偏差の変換

分散の意味

Q3

Quiz(分散の意味)

あるクラスで行われたテストで、英語の平均点は 60 点、標準偏差 10 点.
数学の平均点は 60 点、標準偏差 20 点.
英語の 70 点と数学の 70 点、どちらのほうが価値ある?

- 1 たぶん英語のほうが価値ある
- 2 たぶん数学のほうが価値ある
- 3 どちらも同じ
- 4 これだけの情報ではまったくわからない
- 5 平均点が 60 点だと再テストがあるだろう

Q4

Quiz(分散の意味)

野球のレギュラー選手 A,B のシーズンデータから、1 試合あたりのヒットの本数を調べた。

A は平均 1.5 本, 標準偏差 1.5 本.

B は平均 1.5 本, 標準偏差 0.5 本.

よく「確実性がある」「むらがない」などといわれるのはどちらの選手?

- ① A
- ② B
- ③ どちらも同じ
- ④ これだけの情報ではまったくわからない
- ⑤ 阪神に来てほしい

Q5

Quiz(分散の意味)

同じ太さの毛で作った2個のブラシ A,B の毛の長さを調べた.

A は平均 3cm, 標準偏差 0.2mm.

B は平均 2cm, 標準偏差 0.1mm.

触ったときになめらかなのはどっち?

- ① A
- ② B
- ③ どちらも同じ
- ④ これだけの情報ではまったくわからない
- ⑤ ブラシの色による

ここまで来たよ

- 1 復習:データの位置を代表する値
 - 代表値
- 2 データのばらつきを表す値
 - データのばらつきを表す値
 - 平均偏差と分散
 - 標準偏差
 - 分散の意味
 - 平均値・分散・標準偏差の変換

平均値・分散・標準偏差の変換

$a \times X, X + b$ の平均値・分散・標準偏差は、 X の平均値・分散・標準偏差からわかっちゃう!

| データ | 平均値 | | 分散 | 標準偏差 | |
|------------------|----------|----------------|---------|--------------|----------------------|
| X | X の平均値 | | X の分散 | | X の標準偏差 |
| $X + b$ | X の平均値 | $+b$ | X の分散 | | X の標準偏差 |
| $a \times X$ | X の平均値 | $\times a$ | X の分散 | $\times a^2$ | X の標準偏差 $\times a$ |
| $a \times X + b$ | X の平均値 | $\times a + b$ | X の分散 | $\times a^2$ | X の標準偏差 $\times a$ |

定数を加える

100cm を除いた部分だけで考えると? $b = +100$.

$$(X + b) \text{ の平均値} = X \text{ の平均値} + b$$

$$\begin{aligned} & \frac{[48 + 100] + [48.5 + 100] + \cdots + [72.3 + 100]}{77} \\ &= \frac{48 + 48.5 + \cdots + 72.3}{77} + 100 = (58 + 100)\text{cm} \end{aligned}$$

$$(X + b) \text{ の分散} = X \text{ の分散}$$

$$\begin{aligned} & \frac{[(48 + 100) - (58 + 100)]^2 + \cdots + [(72.3 + 100) - (58 + 100)]^2}{77} \\ &= \frac{[48 - 58]^2 + \cdots + [72.3 - 58]^2}{77} = 26.0\text{cm}^2 \end{aligned}$$

$$(X + b) \text{ の標準偏差} = X \text{ の標準偏差}$$

定数をかける

cm じゃなく m で考えると? $a = 0.01$.

$$(a \times X) \text{ の平均値} = X \text{ の平均値} \times a$$

$$\begin{aligned} & \frac{(148 \times 0.01) + (148 \times 0.01) + \cdots + (172.3 \times 0.01)}{77} \\ &= \frac{148 + 148.5 + \cdots + 172.3}{77} \times 0.01 = 158 \times 0.01\text{m} \end{aligned}$$

$$(a \times X) \text{ の分散} = X \text{ の分散} \times a^2$$

$$\begin{aligned} & \frac{[(148 \times 0.01) - (158 \times 0.01)]^2 + \cdots + [(172.3 \times 0.01) - (158 \times 0.01)]^2}{77} \\ &= \frac{[148 - 158]^2 + \cdots + [172.3 - 158]^2}{77} \times 0.01^2 = 26.0 \times 0.01^2\text{m}^2 \end{aligned}$$

$$(a \times X) \text{ の標準偏差} = X \text{ の標準偏差} \times a$$

$$\sqrt{26.0 \times 0.01^2} = \sqrt{26.0} \times 0.01 = 5.1 \times 0.01\text{m}$$

Q6

Quiz(平均値と分散)

次のデータを考える: 1009 1010 1012 1012 1012

- ① 平均値を求めよう
- ② 標準偏差を求めよう

連絡

- 今週は授業内で紙を 1 枚提出
- 加減乗除と平方根 (ルート) の使える電卓持ってきてね. 関数電卓でなくてもいいです. 携帯電話の機能・アプリでもかまいません.
- 各追加 2 ピーナッツ=計 4 ピーナッツになる課題.

提出: 2013-10-16 水 or 2013-10-23 水の授業

① 龍谷大学 e ラーニングシステム

<https://moodle.media.ryukoku.ac.jp/> → リメディアルコース統計学 → 第 1 章終了テスト.

② 龍谷大学 e ラーニングシステム

<https://moodle.media.ryukoku.ac.jp/> → リメディアルコース統計学 → 第 2 章終了テスト.

このサイトには, <http://hig3.net> → 龍大 Moodle, や Info Seta → e ラーニングサイト → 新 e ラーニングシステム でも到達できます. すべてを送信して終了する → レビューを終了する の後に出る, 「あなたの前回受験の要約」 ページ (下) を印刷して, 紙で提出. (スクリーンショットを課題にアップロードしてもいい)

あなたは樋口 三郎としてログインしています (ログ

フォーメーション

読み取れるように 修了テスト 第2章

評価方法: 最高評点

あなたの前回受験の要約

| 受験 | 状態 | 評点 / 100.00 | レビュー |
|----|------------------------------------|-------------|------|
| 1 | 終了 送信日時 2013年 10月 8日(火曜日) 17:39 | 0.00 | レビュー |

最高評点: 0.00 / 100.00

もう一度受験する