

1 はじめに

前回の復習

- ・ 分散、標準偏差
 - データのバラツキを表す指標
 - 平均値と同じくらい重要
 - ◇ メロンパンの重さの分散が大きい → パン屋としては問題
 - テストの得点の分散が大きい → できる人もいればできない人もいる
 - 毎日の売上高の分散が大きい → 売上高の予測は簡単ではない
 - 分散 = $\frac{\text{(各データの平均値からの差の二乗の和)}}{\text{データの個数}}$
 - 標準偏差 = $\sqrt{\text{分散}}$

EXCEL 関数

- 最小値 =min(範囲)
- 最大値 =max(範囲)
- 平均値 =average(範囲)
- 中央値 =median(範囲)
- 最頻値 =mode(範囲)
- 分散 =varp(範囲) 又は var.p(範囲)
- 標準偏差 =stdevp(範囲) 又は stdev.p(範囲)

演習 1 次の問の○×を考えてください。

- (1) 分散は、「各データの平均値からの差」を用いて計算する。
- (2) あるデータの標準偏差の単位は、あるデータの単位と同じである。

演習 2

(1) 日本人の男性の身長の変動幅が一番大きいのは何歳の時ですか？

(2) 日本人の女性では何歳の時ですか？

日本人の年齢別・性別別の身長と体重

※ SD : 標準偏差

年齢(歳)	男身長平均(cm)	男身長SD(cm)	女身長平均(cm)	女身長SD(cm)	男体重平均(kg)	男体重SD(kg)	女体重平均(kg)	女体重SD(kg)
6	116.78	5.00	115.95	4.96	21.29	2.90	21.02	3.11
7	122.99	4.93	121.72	5.08	24.00	3.37	23.29	3.41
8	128.12	5.36	127.71	5.59	26.64	4.27	26.32	4.27
9	133.54	5.64	133.52	6.13	29.90	5.41	29.48	5.24
10	138.78	5.92	140.28	6.66	33.54	6.58	33.89	6.32
11	145.05	7.14	147.26	6.65	37.92	8.05	39.18	7.32
12	152.69	7.88	152.06	6.01	43.52	8.83	43.21	7.63
13	160.11	7.52	155.06	5.11	49.04	8.80	46.49	6.69
14	165.37	6.72	156.79	5.23	53.41	8.24	49.16	6.55
15	168.61	5.77	157.09	5.34	59.24	9.56	51.37	6.93
16	170.14	5.79	157.51	5.45	60.50	8.72	51.73	6.97
17	171.01	5.77	157.73	5.35	62.48	8.69	51.91	6.80
18	171.12	5.55	158.03	5.37	62.64	9.01	51.68	7.02
19	171.58	5.63	158.23	5.56	63.08	8.16	51.32	6.19
20	171.66	5.60	158.32	5.52	63.95	8.41	51.08	6.10
21	171.74	5.57	158.42	5.48	64.82	8.66	50.84	6.01
22	171.83	5.54	158.51	5.44	65.70	8.91	50.60	5.92
23	171.89	5.53	158.56	5.43	65.89	8.93	50.56	5.92
24	171.95	5.51	158.61	5.41	66.07	8.94	50.52	5.92
25	172.01	5.50	158.65	5.40	66.26	8.96	50.47	5.92

(出典) 2009 年度体力・運動能力調査 (厚生労働省) を基に konisimple 氏が作成

演習 3 (1)～(3)で分散が一番大きいのはどれであることを予想してから、実際に分散と標準偏差を求めてください。

(1) {10, 10, 30, 30}

					合計	平均
データ	10	10	30	30		
偏差						
(偏差) ²						

分散＝

標準偏差＝

(2) {20, 20, 20, 20}

					合計	平均
データ	20	20	20	20		
偏差						
(偏差) ²						

分散＝

標準偏差＝

(3) {10, 15, 25, 30}

					合計	平均
データ	10	15	25	30		
偏差						
(偏差) ²						

分散＝

標準偏差＝

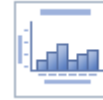
演習 4 横浜市のスターバックス、ドトールの店舗データを用いて以下の作業を行ってください。

- (1) スターバックス店舗の最寄駅からの距離について、度数分布図を作成して下さい。
- (2) ドトール店舗の最寄駅からの距離について、度数分布図を作成して下さい。
- (3) スターバックス店舗の最寄駅からの距離について、平均値、標準偏差を求めて下さい。
- (4) ドトール店舗の最寄駅からの距離について、平均値、標準偏差を求めて下さい。
- (5) スターバックスとドトールの店舗の立地の違いについてわかったことを述べて下さい。

主な手順

(1) スタバ店舗の最寄り駅からの距離の度数分布図の作成

- ① (見出し行を含めスターボックスの店舗データを選択) → 「挿入」
→ 「ピボットテーブル」 → 「OK」
- ② フィールドリストの「駅からの距離」を「行ラベル」のボックスにドラッグ
- ③ 同じ「駅からの距離」を「 Σ 値」のボックスにドラッグ
→ 「合計/駅から距離」と表示されている右側の▼
→ 「値フィールドの設定」 → 「合計」を「データの個数」に変更 → 「OK」
- ④ (「行ラベル」の列の任意のセルを選択) → (右クリック) → 「グループ化」
→ 「先頭の値」に「0」を入力 → 「OK」 → (度数分布表の完成)
- ⑤ (度数分布表(「見出し行」を含め、「総計の行」は含めない)を選択)
→ 「挿入」 → 「グラフ」 → 「縦棒」 → (2D縦棒の一番左のものを選択)
→ 「グラフのレイアウト」で次のスタイルを選択。



⑥ タイトルとラベルの入力

- ✧ グラフタイトル: 「スターボックスの立地 (横浜市)」
- ✧ 横軸ラベル: 最寄り駅からの距離 (m)
- ✧ 縦軸ラベル: 店舗数

(2) ドトール店舗の最寄り駅からの距離の度数分布図の作成

(略)

(3) スターバックスの最寄駅からの距離の平均値と標準偏差の計算

- (最初のシートに戻る) →
- スターバックス店舗について、
最寄駅からの距離の平均値を =average 関数で計算
最寄駅からの距離の標準偏差を =stdevp 関数で計算

(4) ドトール店舗の最寄駅からの距離の平均値と標準偏差の計算

(略)

(5) 提出シートの作成と印刷

- ① 「提出」シートを開く
- ② ヘッダーに学籍番号と氏名を入力
- ③ 計算した「最寄駅からの距離」の平均値、標準偏差を「提出」シートの表にコピー
- ④ 作成した2つの「度数分布図」を「提出」シートの下部にコピー
- ⑤ 「スターバックスとドトールの立地の違い」を記入
- ⑥ 印刷が1ページに収まるように設定する
「ページレイアウト」→「拡大縮小印刷」→「横：1ページ、縦：1ページ」
- ⑦ 印刷 → (提出)

(参考) 印刷範囲を指定する方法 (必要な場合)

(印刷範囲の選択) → 「ページレイアウト」→「印刷範囲」→「印刷範囲の設定」