

統計検定2級

2018年6月

問1

一番優しい、統計学の教本

吉田

2018年6月の問題

問1 次の表は、2017年度のサッカーJリーグのチーム年間総得点のデータである。

J1				J2				J3			
チーム名	年間総得点	チーム名	年間総得点	チーム名	年間総得点	チーム名	年間総得点	チーム名	年間総得点		
札幌	39	清水	36	山形	45	岐阜	56	盛岡	32	北九州	44
仙台	44	磐田	50	水戸	45	京都	55	秋田	53	鹿児島	49
鹿島	53	G大阪	48	群馬	32	岡山	44	福島	39	琉球	44
浦和	64	C大阪	65	千葉	70	山口	48	栃木	44	F東京	23
大宮	28	神戸	40	東京V	64	讃岐	41	YS横浜	41	G大阪	23
柏	49	広島	32	町田	53	徳島	71	相模原	34	C大阪	23
FC東京	37	鳥栖	41	横浜FC	60	愛媛	54	長野	34		
川崎F	71			湘南	58	福岡	54	富山	37		
横浜FM	45			松本	61	長崎	59	藤枝	50		
甲府	23			金沢	49	熊本	36	沼津	60		
新潟	28			名古屋	85	大分	58	鳥取	31		
平均	44.06			平均	54.45			平均	41.06		
標準偏差	12.89			標準偏差	11.77			標準偏差	8.06		

資料：J.LEAGUE Data Site

[1] 下の図I～IIIは、「チーム年間総得点（総得点）」、「チーム年間総得点の平均からの偏差（偏差）」および「チーム年間総得点の標準化得点（標準化得点）」をJ1, J2, J3のカテゴリーごとに分けて示した箱ひげ図である。

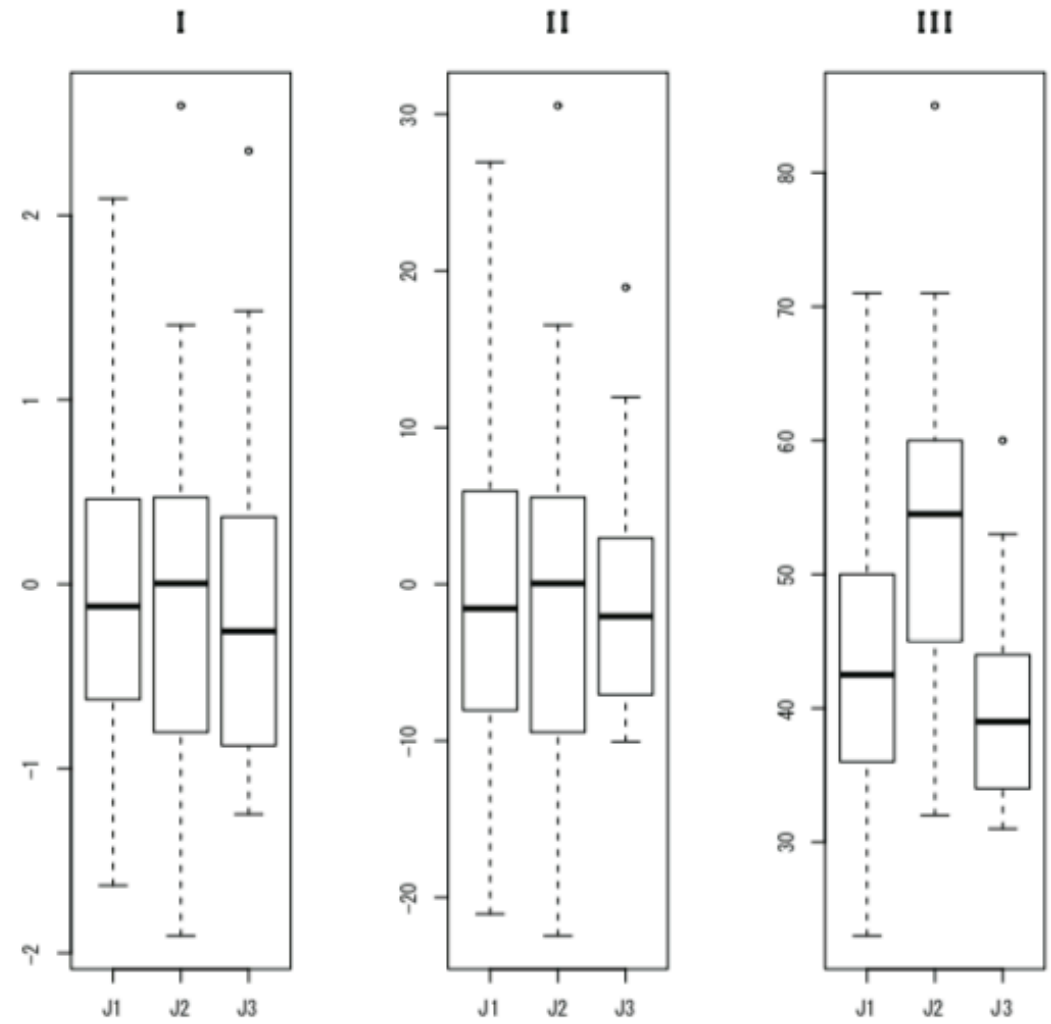
なお、これらの箱ひげ図では、“「第1四分位数」－「四分位範囲」×1.5”以上の値をとるデータの最小値、および“「第3四分位数」＋「四分位範囲」×1.5”以下の値をとるデータの最大値までひげを引き、これらよりも遠い値を外れ値として○で示している。

図I～IIIに対応する変数の組合せとして、次の①～⑤のうちから最も適切なものを一つ選べ。 1

- ① 総得点：I, 偏差：II, 標準化得点：III
- ② 総得点：II, 偏差：I, 標準化得点：III
- ③ 総得点：II, 偏差：III, 標準化得点：I
- ④ 総得点：III, 偏差：I, 標準化得点：II
- ⑤ 総得点：III, 偏差：II, 標準化得点：I

[2] J2のチーム年間総得点において、平均から標準偏差の2倍以上離れた観測値は何個あるか。次の①～⑤のうちから適切なものを一つ選べ。 2

- ① 0個 ② 1個 ③ 2個 ④ 3個 ⑤ 4個以上



正解

■問 1 (1) : 5

■問 1 (2) : 2

この問題で知っておくべき3つの知識

■箱ひげ図

■偏差

■標準化

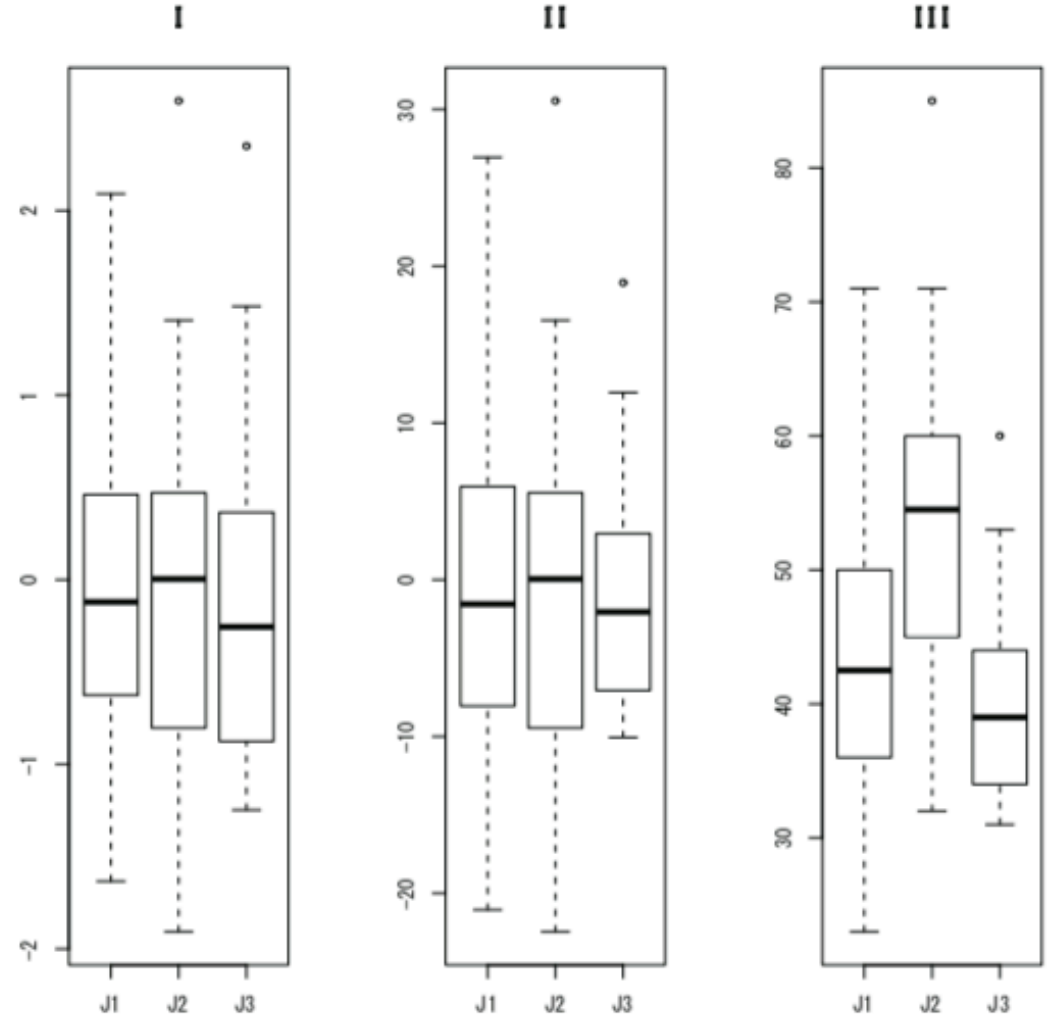
問2 (1) で気になる用語は？

[1] 下の図I～IIIは、「チーム年間総得点（総得点）」、「チーム年間総得点の平均からの偏差（偏差）」および「チーム年間総得点の標準化得点（標準化得点）」をJ1, J2, J3のカテゴリごとに分けて示した箱ひげ図である。

なお、これらの箱ひげ図では、「第1四分位数」-「四分位範囲」×1.5以上の値をとるデータの最小値、および「第3四分位数」+「四分位範囲」×1.5以下の値をとるデータの最大値までひげを引き、これらよりも遠い値を外れ値として○で示している。

図I～IIIに対応する変数の組合せとして、次の①～⑤のうちから最も適切なもの一つ選べ。 1

- ① 総得点：I, 偏差：II, 標準化得点：III
- ② 総得点：II, 偏差：I, 標準化得点：III
- ③ 総得点：II, 偏差：III, 標準化得点：I
- ④ 総得点：III, 偏差：I, 標準化得点：II
- ⑤ 総得点：III, 偏差：II, 標準化得点：I



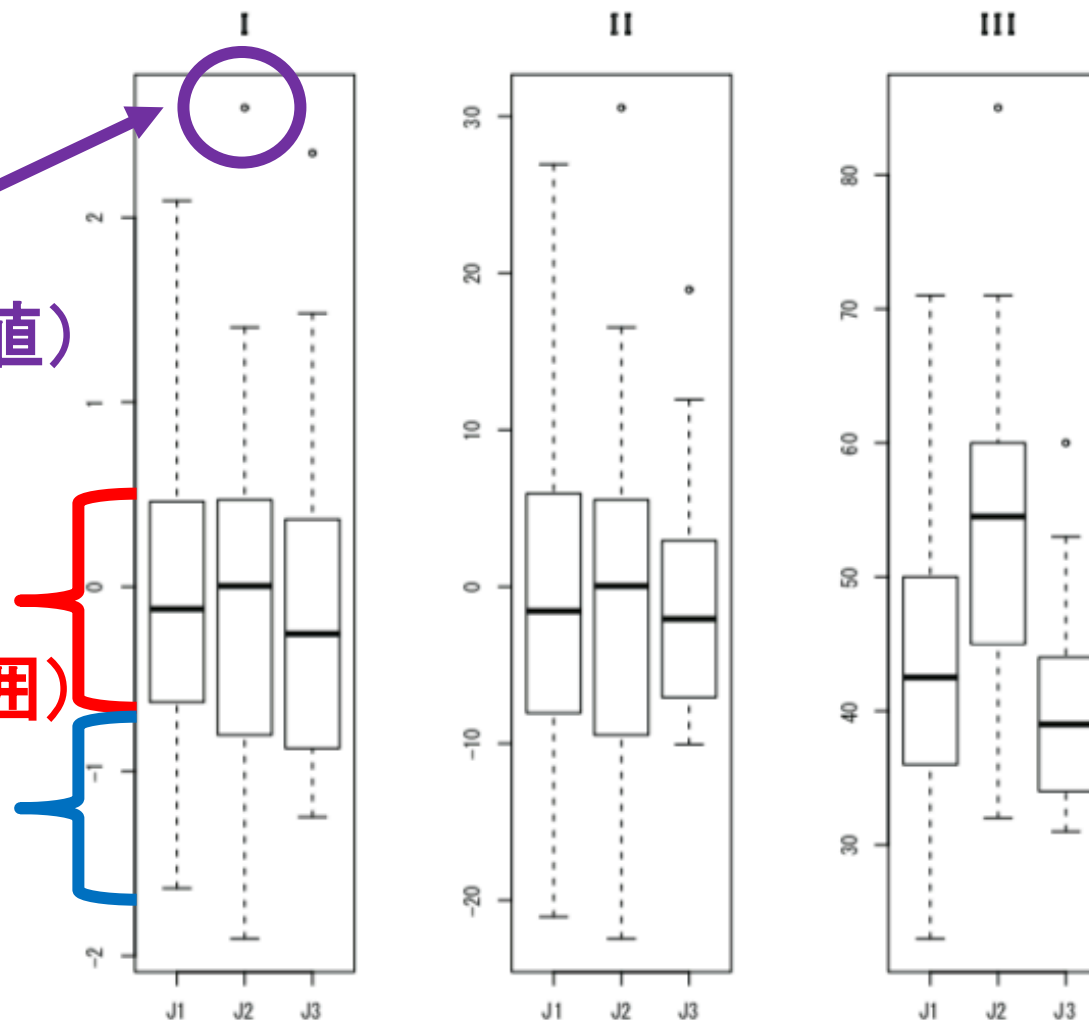
必ず知っておくべきグラフ

■箱ひげ図

外れ値
(ひげの範囲より離れた値)

箱
(四分位範囲:
第1四分位数と第3四分位の範囲)

ひげ
(四分位範囲*1.5)



偏差って何？

■偏差：平均値と各データの差

J1			
チーム名	年間総得点	チーム名	年間総得点
札幌	39	清水	36
仙台	44	磐田	50
鹿島	53	G 大阪	48
浦和	64	C 大阪	65
大宮	28	神戸	40
柏	49	広島	32
FC 東京	37	鳥栖	41
川崎 F	71		
横浜 FM	45		
甲府	23		
新潟	28		
平均	44.06		
標準偏差	12.89		

例) 札幌の偏差:

$$39 - 44.06 = -5.06$$

特徴:

生データ(総得点)よりも小さい値になる

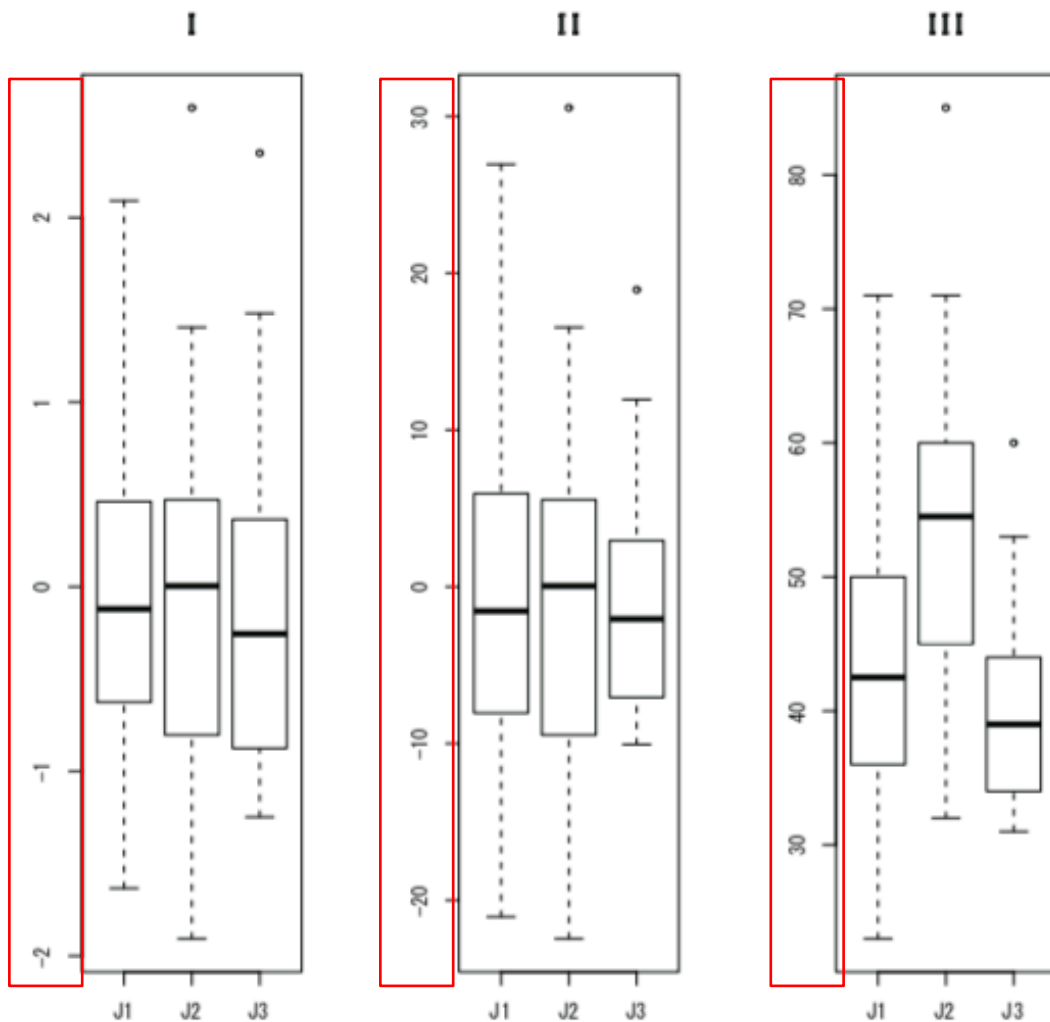
標準化とは？どんなメリットがあるの？

$$\frac{\text{データ} - \text{平均値}}{\text{標準偏差}}$$

標準化後、平均値=0、標準偏差=1の正規分布
(標準正規分布)に従うようになる

問 1 (1) の解説

縦軸に注目！



縦軸のスケール:

I < II < III

標準化

偏差

総得点

正解: 5

問 1 (2) で気になる用語

■標準偏差の2倍以上離れた観測値

[2] J2 のチーム年間総得点において、平均から標準偏差の2倍以上離れた観測値は何個あるか。次の①～⑤のうちから適切なもの一つ選べ。

① 0個

② 1個

③ 2個

④ 3個

⑤ 4個以上



標準化の知識が必要！

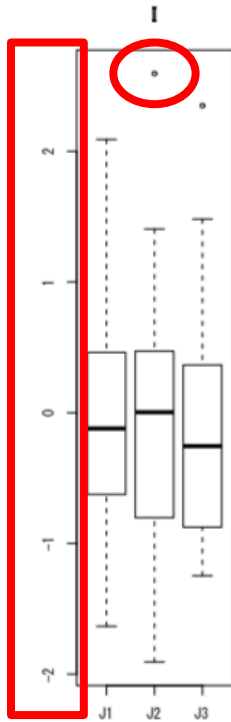
標準化とは？どんなメリットがあるの？（再掲）

$$\frac{\text{データ} - \text{平均値}}{\text{標準偏差}}$$

標準化後、**平均値=0**、**標準偏差=1**の正規分布
(標準正規分布)に従うようになる

問題文を読み替えるとどうなる？

■標準偏差の2倍以上離れた → 標準化後の分布で2以上離れた



外れ値は1つだけ：
正解：2

この動画で学んだ用語

■箱ひげ図

■四分位範囲

■偏差：平均値と各データの差

■標準化：
$$\frac{\text{データ} - \text{平均値}}{\text{標準偏差}}$$

正解

■問 1 (1) : 5

■問 1 (2) : 2