

(2019年6月28日版)

## 統計検定4級新出題範囲例題集（問題および略解）

1. これは2020年4月から変更される出題範囲に対応した統計検定4級の例題集です。
2. 本例題集は4級の全範囲を網羅しているのではなく、新出題範囲「統計的問題解決の方法」、「標本調査」、「データの要約 - 箱ひげ図」に関する問題のみ掲載しています。
3. 本例題集は必要に応じて改訂されます。

問1 問題解決のサイクルについて述べた記述として、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 客観性を保つため、分析する際は得られたデータのみに着目すればよく、問題意識を持って取り組むのはよくない。
- ② 調査票を用いて調査を行う場合、調査票の質問文は、できるだけ専門用語を用いた正確な表現を心がける必要がある。
- ③ はじめに実験や調査を行い、データを収集することによって問題を明確化し、その後、データの分析を行う。
- ④ はじめに問題を明確化し、実験や調査を行い、データを収集した後にデータの分析を行う。問題が解決しなかった場合は問題を明確化するところに戻る。
- ⑤ 分析結果をもとに問題解決を図ったところ、1回のサイクルでは問題が解決できなかった場合、実験や調査が失敗であったと結論付ける。

問2 A市の中学校のうち、年度当初に視力検査を実施している中学校の1年生男子の中から無作為に500名を抽出し、その結果を集めた。この標本調査の母集団と標本について、次の①～⑤のうちから最も適切な組合せを一つ選べ。

- ① 母集団：全国の中学生全体  
標本：A市の中学生全体
- ② 母集団：A市の中学生全体  
標本：年度当初に視力検査を行った中学生
- ③ 母集団：A市の中学校で視力検査を実施した中学校の1年生男子全体  
標本：上の母集団から無作為に抽出した500名
- ④ 母集団：A市の中学校で年度当初に視力検査を実施した中学校の1年生男子全体  
標本：上の母集団から無作為に抽出した500名
- ⑤ 母集団：A市の中学校で年度当初に視力検査を実施した中学校  
標本：無作為に抽出した500名の生徒が所属する中学校

問3 無作為抽出法について述べた記述として、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 母集団に含まれるすべての人や物に異なる番号を付けて、調査する人の好きな番号を選べば無作為抽出になる。
- ② 無作為に抽出するとは、調査する人が抽出方法を自由に決められることである。
- ③ 無作為に抽出する方法には、乱数さいを用いる方法や乱数表を用いる方法などがある。
- ④ 標本の大きさが適正であれば、母集団に含まれる人や物が同じ確率で選ばれなくてもよい。
- ⑤ どんな方法を用いても無作為に抽出することは不可能なので、調査する人の判断である程度の妥協は許される。

問4 壺の中に色以外では区別のつかない小さな白玉と黒玉が合計1000個入っている。この壺の中から20個を無作為に抽出したところ、黒玉は4個であった。このとき、壺の中に入っている黒玉の個数はおよそ何個と推定されるか、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 150個      ② 200個      ③ 250個      ④ 300個      ⑤ 500個

問5 ある会社では、健康診断の結果を利用して健康状況を測る指標のBMIを計算し、社員の健康管理を行っている。BMIは(体重 kg)÷(身長 m)<sup>2</sup>で計算される。たとえば、身長172cm、体重75kgの人のBMIは、 $75 \div 1.72^2 = 25.35 \approx 25.4$ である。この会社では男性社員についてBMIの値に基づき、次の表のような解釈を行っている。

BMI	健康状態
17.6未満	やせすぎ
17.6以上19.8未満	やせ気味
19.8以上24.2未満	理想体重
24.2以上26.4未満	過体重
26.4以上	肥満

企画部の17人、営業部の29人、人事部の11人の男性社員のBMIを計算して、小数第2位を四捨五入した値を使い、部署ごとにBMIの分布の5数要約をまとめたところ、次の表のようになった。

5数要約	企画部	営業部	人事部
最小値	19.3	17.0	17.0
第1四分位数	22.1	21.0	22.4
中央値	24.3	22.2	24.3
第3四分位数	26.4	26.0	25.7
最大値	31.0	27.1	30.9

〔1〕男性社員の健康状態に関して、3つの部署の状況を述べた記述として、次の①

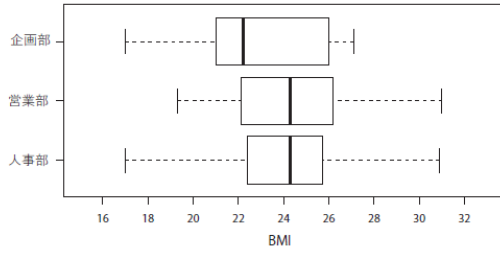
～⑤のうちから最も適切なものを一つ選べ。

- ① 中央値と最大値をみると、営業部は企画部や人事部に比べて低い。
- ② 3つの部署ともやせすぎの人がいる。
- ③ やせすぎと肥満の人がいるのは人事部だけである。
- ④ 企画部と人事部において、BMIが中央値以上の人は同じ人数である。
- ⑤ 四分位範囲から判断すると、BMIのばらつきが一番大きいのは人事部である。

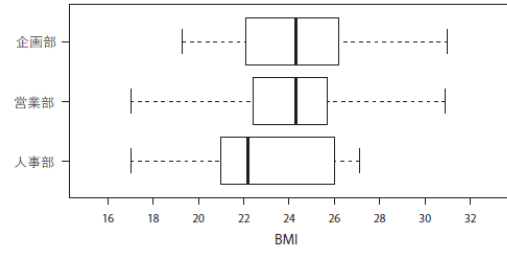
〔2〕3つの部署の箱ひげ図として正しいものを，次の①～④のうちから一つ選べ。

6

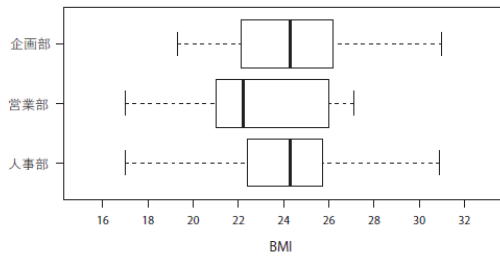
①



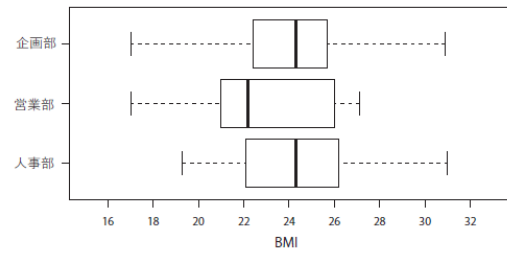
②



③



④



## 略解

※各問題の前に参考となる用語の解説があります。また、解説に示されたページは「統計検定4級対応 資料の活用」(東京図書)のページに対応します。

### <統計的問題解決の方法>

統計的問題解決の方法として **PPDAC サイクル**がある。PPDAC とは、Problem (問題：解決したい問題を整理する)、Plan (計画：調査や実験をデザインする)、Data (データ：データの収集と整理)、Analysis (分析：データの分析)、Conclusion (結論：結論をまとめ判断する) の頭文字をとったものである。これらのステップをサイクルとしてくり返して実行することを意識した、統計的な問題解決の活動が大切である。(p. 4)

### 問1

PPDAC サイクルに関する理解を問う問題である。

1

正解 ④

- ①：誤り。データの分析を行う際、最初に設定した問題を意識して分析方法を検討する必要がある。
- ②：誤り。実験・調査の計画の際、できるだけ対象者が質問を誤解しないように心掛けることが重要で、常に専門用語を用いる必要はない。
- ③：誤り。明確化された問題の解決に役立つように実験・調査の計画を立て、それに基づきデータを収集する。
- ④：正しい。最初の段階では問題そのものがそれほど明確でない場合が多いので、はじめに問題を明確化する。
- ⑤：誤り。問題の解決に至るプロセスは、必ずしも1回の実験や調査で完了するものではなく、何度も実験や調査をくり返す中でより良い結果を得ることもある。よって、正解は④である。

### <標本調査>

ある集団について何らかの興味ある特性や傾向を知りたいとき、この集団に含まれるすべての人や物全体を**母集団**、母集団から選ばれた人や物全体を**標本**という。母集団全てについて調査したとき**全数調査**といい、標本について調査をしたとき**標本調査**という。5年に

1 度実施される**国勢調査**は全数調査の一例で、新聞等でみられる**世論調査**は標本調査の一例である。

例えば、花子さんが自分の所属する中学の生徒（全員で 1000 人）の好みについて知りたいとき、中学の生徒全員（1000 人）が母集団である。一部の人を選んだなら、これが標本で、この調査が標本調査である。花子さんが 200 人を選んだとき、**標本の大きさ**は 200 という。(p. 65)

標本を選ぶための抽出方法として、**無作為抽出**が知られている。無作為抽出を行うためには、乱数さい、乱数表などが用いられる。この方法を用いると、母集団に含まれる人が選ばれる確率が等しくなる。(p. 65-66)

無作為抽出により選ばれた標本により、母集団の特性は偏りなく推定できる。例えば、花子さんが無作為抽出した 200 人の中で 30 人（15%）が「水泳が好き」と回答したとき、生徒全員の中でおよそ  $0.15 \times 1000 = 150$  [人] が「水泳が好き」と回答すると推定する。

## 問 2

標本調査における母集団と標本に関する問題である。

2

正解 ④

本問では、母集団は「A 市の中学校のうち、年度当初に視力検査を実施している中学校の 1 年生男子」であり、標本は「無作為に抽出した 500 名」である。

- ①：誤り。母集団、標本ともに誤り。全国の中学生全体は母集団ではない。
  - ②：誤り。母集団、標本ともに誤り。A 市の中学生全体は母集団ではない。
  - ③：誤り。母集団、標本ともに誤り。この母集団には年度当初に実施していない生徒も含まれる。
  - ④：正しい。母集団、標本ともに正しい。
  - ⑤：誤り。母集団、標本ともに誤り。この調査は中学校を対象としていない。
- よって、正解は④である。

## 問 3

無作為抽出法に関する理解を問う問題である。

3

正解 ③

- ①：誤り。異なる番号を付けても、調査する人の好きな番号を選ぶと、選ばれ方が同確率となるとは限らず偏りが生じる可能性がある。
- ②：誤り。無作為に抽出するとは、母集団に含まれる人が選ばれる確率が等しくなるように抽出することであり、抽出方法を選ぶことではない。
- ③：正しい。無作為に抽出する方法には、乱数さいや乱数表を用いる方法がある。これ以外にもコンピュータによって乱数を発生させる方法などがある。

- ④：誤り。標本の大きさが適正であっても、例えば女性を多く選ぶ、子どもを多く選ぶなど偏りが生じるような選び方をしてはいけない。
- ⑤：誤り。母集団に含まれる人や物にすべて異なる番号をつけて、その番号を乱数さい、乱数表、コンピュータによって乱数を発生させ抽出すれば無作為に抽出することができる。
- よって、正解は③である。

#### 問4

標本調査を利用して母集団の比率を推定する問題である。

4

正解 ②

無作為に抽出した20個の玉（標本）における白玉の比率は、 $\frac{4}{20} = 0.2$ である。このことから、壺の中（母集団）における白玉の比率はおよそ0.2と推定できる。したがって、壺の中に入っている白玉の個数は、およそ $1000 \times 0.2 = 200$  [個]と推定される。

よって、正解は②である。

#### <箱ひげ図>

観測値を小さい順に並べて四つに等しく分けたときの三つの区切りの値を小さい方から順に、**第1四分位数**、**第2四分位数**、**第3四分位数**という。第2四分位数は中央値である。これらに最小値と最大値を加えた五つの数（5数）によりデータの特徴を示すことを**5数要約**という。また、第3四分位数から第1四分位数を引いた値を**四分位範囲**とよび、データのばらつきの具合を表す指標として用いられる。

5数要約を視覚的に表現したものが**箱ひげ図**である。第1四分位数、第2四分位数、第3四分位数を用いて「箱」を作る。通常、最小値から第1四分位数までの部分と第3四分位数から最大値までの部分を「ひげ」として表す。

箱ひげ図と関連する用語についてはp.111-118、に詳しく書かれている。

[注] 第1四分位数と第3四分位数を求める方法はいくつかあり、使用する統計分析用ソフトウェアによって多少の差があるが、それぞれの値がどの程度であるかを理解することが重要である。

#### 問5

{1} 5

正解 ①

5数要約を用いて情報を正しく判断できるかを問う問題である。

①：正しい。



- ②：誤り。企画部の BMI の最小値は 19.3 であり、やせすぎの判断基準である 17.6 未満を超えている。したがって、企画部にはやせすぎの人がいない。
- ③：誤り。営業部の最小値は 17.0 であり、最大値は 27.1 である。したがって、営業部にもやせすぎ（17.6 未満）と肥満（26.4 以上）の人がいる。
- ④：誤り。企画部は 17 名なので、BMI の値が中央値以上の人は 9 人以上 13 人以下である。人事部は 11 名なので、BMI が中央値以上の人は 6 人以上 8 人以下である。したがって、同じ人数になることはない。
- ⑤：誤り。各部のレンジ（範囲）R と四分位範囲 IQR を求めると、次の通りである。

企画部	: R=31.0-19.3=11.7	IQR=26.4-22.1=4.3
営業部	: R=27.1-17.0=10.1	IQR=26.0-21.0=5.0
人事部	: R=30.9-17.0=13.9	IQR=25.7-22.4=3.3

したがって、人事部のレンジが一番大きいものの、四分位範囲から判断すると人事部のばらつきが一番大きいとはいえない。  
よって、正解は①である。

[2]

正解 ③

5 数要約から適切な箱ひげ図を選ぶ問題である。5 数のうちの 1 つでも箱ひげ図に矛盾があれば適切でない。

- ①：誤り。例えば、企画部の最小値は 19.3 であるが、箱ひげ図では約 17 と読み取れるので適切でない。
- ②：誤り。例えば、営業部の最大値は 27.1 であるが、箱ひげ図では約 31 と読み取れるので適切でない。
- ③：正しい。各部署の 5 数要約の 5 つの値のいずれとも矛盾はなく、適切な箱ひげ図であると判断できる。
- ④：誤り。例えば、人事部の最小値は 17.0 であるが、箱ひげ図では 19 以上と読み取れるので適切でない。  
よって、正解は③である。