

Week2 確認テスト 解答・解説

以下に、世帯の消費の実態を調べたデータ（ダミーデータ）がある。

このデータを使って分析してみよう。

※課題に必要なとなる分析の Excel でのやり方については補講を参照。

【問題 1】 利用データ：dummydata_A.xlsx

消費支出のうち、食費の平均値、分散、標準偏差の組み合わせとして正しいものを、次の(1)～(4)のうちから一つ選びなさい。（平均値、分散、標準偏差は Excel の分析ツールで算出、計算値は小数第二位で四捨五入した。）

- | | | | | | | |
|-----|-----|---------|----|---------------|------|----------|
| (1) | 平均値 | 75594.2 | 分散 | 867,413,901.7 | 標準偏差 | 24,213.2 |
| (2) | 平均値 | 65594.2 | 分散 | 853,413,901.7 | 標準偏差 | 29,213.2 |
| (3) | 平均値 | 85594.2 | 分散 | 766,413,901.7 | 標準偏差 | 27,213.2 |
| (4) | 平均値 | 55594.2 | 分散 | 466,413,901.7 | 標準偏差 | 22,213.2 |

【解答 1】

(2)

【解説 1】

Excel の分析ツール→基本統計量で、平均値等の代表値や散布度（基本統計量）は算出できる。

この機能を使って食費の基本統計量を算出すると右表となる。

よって、正解は(2)の平均値 65594.2、分散 853,413,901.7、標準偏差 29,213.2。

なお、これらの値は、以下の Excel 関数でも算出できる。

平均値 =AVERAGE(データの範囲)

分散 =VAR(データの範囲)

標準偏差 =STDEV(データの範囲)

食費	
平均	65594.2042
標準誤差	292.1324874
中央値（メジアン）	57522.5
最頻値（モード）	45092
標準偏差	29213.24874
分散	853413901.7
尖度	1.207859899
歪度	1.043183373
範囲	217758
最小	9094
最大	226852
合計	655942042
標本数	10000

【問題 2】 利用データ：dummydata_A.xlsx

階級間隔を 10,000 円として食費の度数分布表を作成した時、20,001～30,000 円以下の階級の度数として正しいものを、次の(1)～(4)のうちから一つ選びなさい。

- (1) 464
- (2) 564
- (3) 664
- (4) 764

【解答 2】

(1)

【解説 2】

Excel の分析ツール→ヒストグラムで、度数分布表は作成できる。この機能を使って、食費の度数分布表を作成する。

まず、階級間隔を指定する表を作成する。間隔は 10,000 円であり、また先の基本統計量から最小値が 9,094、最大値が 226,852 であることが分かっているので、右下のような表とする。(Excel の分析ツールでは、階級の境界値が「～以下」を示すことに注意。例：階級値 20,000 とした場合「10,000 より大きく 20,000 以下」の値が算出される)

この表を使って分析ツールにより度数分布表を作成すると左下の表となる。

表により 20,001～30,000 円以下の階級の度数は 464 で、(1)が正解となる。

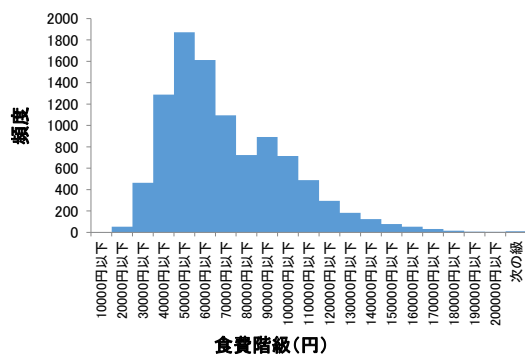
食費階級 (円)
10000
20000
30000
40000
50000
60000
70000
80000
90000
100000
110000
120000
130000
140000
150000
160000
170000
180000
190000
200000

食費階級 (円)	頻度
10000	1
20000	53
30000	464
40000	1288
50000	1871
60000	1611
70000	1095
80000	723
90000	892
100000	714
110000	488
120000	295
130000	183
140000	123
150000	78
160000	52
170000	32
180000	16
190000	6
200000	5
次の級	10

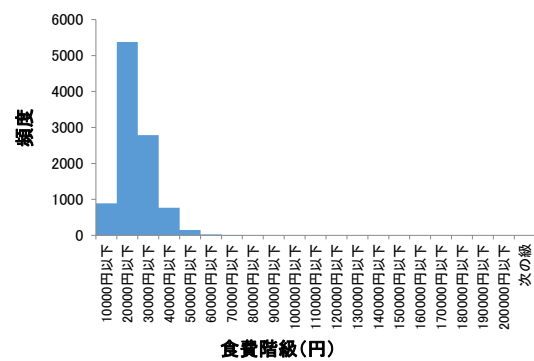
【問題 3】 利用データ：dummydata_A.xlsx

階級間隔を 10,000 円とした食費のヒストグラムとして、最も近いものを次の(1)～(4)のうちから一つ選びなさい。

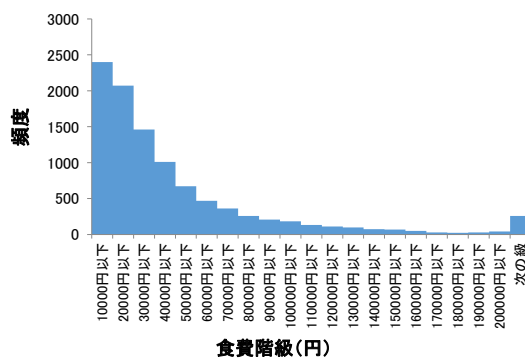
(1)



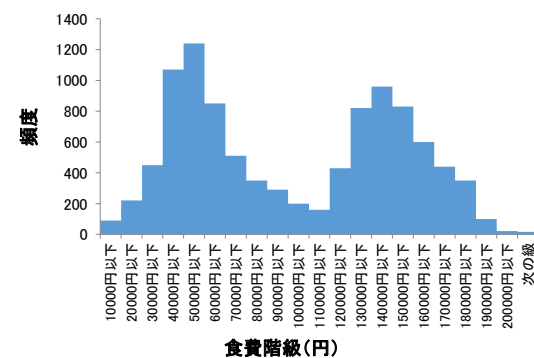
(2)



(3)



(4)



【解答 3】

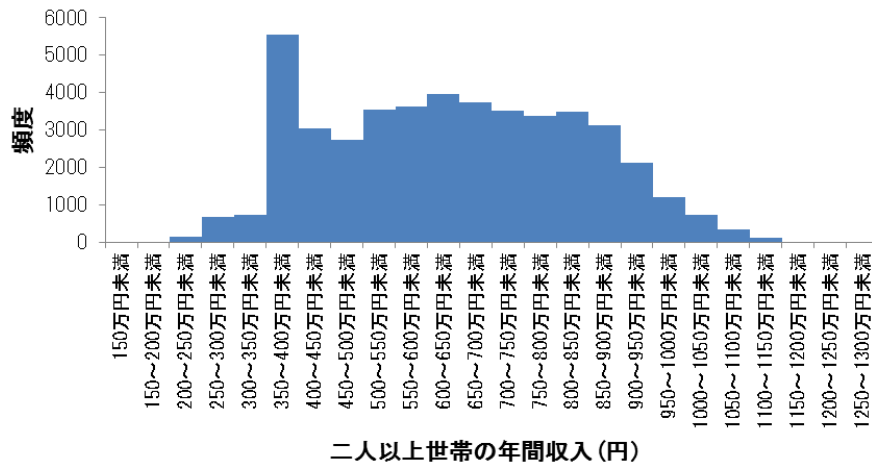
(1)

【解説 3】

Excel の分析ツール→ヒストグラムによってグラフを作成すると (1)となる。

【問題 4】

下図はある地域の二人以上世帯の年間収入を示したヒストグラムです。このヒストグラムの解釈として、最も適切でないものを次の(1)～(4)のうちから一つ選びなさい。



- (1) データは、150 万円未満～1300 万円未満に分布している
- (2) 最もサンプルが集中しているのは、350～400 万円未満の階級である
- (3) 年間収入の最小は 200～250 万円未満の階級である
- (4) 年間収入の最大は 1100～1150 万円未満の階級である

【解答 4】

(1)

【解説 4】

ヒストグラムから最小は 200～250 万円未満の階級、最大は 1100～1150 万円未満の階級であることがわかるため、データの分布は 200～1150 万円未満である。よって(1)が回答となる。

【問題 5】 利用データ：dummydata_A.xlsx

「その他の消費支出」を除く消費支出の中で最もばらつきの大きい項目を、次の(1)～(4)のうちから一つ選びなさい。(ばらつきの指標は標準偏差とする)

- (1) 食費
- (2) 住居費
- (3) 保健医療費
- (4) 交通・通信費

【解答 5】

(4)

【解説 5】

Excel の分析ツール→基本統計量で消費支出 9 項目の標準偏差を算出すると以下となる。(小数点第一位で四捨五入)

食費	29,213
住居費	38,626
光熱・水道費	8,009
家具・家事用品費	11,957
被服及び履物費	14,293
保健医療費	19,456
交通・通信費	71,950
教育費	47,476
教養娯楽費	32,800

標準偏差の値が大きい方がばらつきは大きいため、上記のうち最も値が大きい交通・通信費が回答となる。

なお、今回のように複数の消費支出を対象に標準偏差を算出する場合は、関数の利用が便利である。

標準偏差 =STDEV(データの範囲)

※ このダミーデータは、独立行政法人 統計センターが提供している一般用マイクロデータを加工し作成したものです。