

出題範囲—第0章及び第1章

注意事項—解答レポートは原則手書きで作成すること。レポート用紙のサイズ・種類は問わない。解答のみで良い。ただし、ヒストグラム・散布図を作成する問題は、Excel でこれらを作成したものをWordにコピーしそれを印刷したものをよい。切り貼りをしても良い。度数分布表は、手書き(手計算)で作成しても良いし、Excel を用いて印刷してもどちらでも良い。

問題の解き方も説明すべき問題は、解答までの道筋をきちんと述べよ。また、計算は結果のみならず途中の計算式も分かる様に示せ。ただし、問題文中に計算結果のみ示す旨の指示がある場合は計算結果のみで良い。

第1問(準備)

(1) $\sum_{k=4}^6 3x_{j,k} y_{l,k}$ を Σ を用いない式に直せ。

(2) 値10.4と値17.7の間を4等分したときの10.4に近い側の4分の1の位置にある値が $\frac{3}{4} \times 10.4 + \frac{1}{4} \times 17.7$

であることを次の手順で証明しなさい。まず、値10.4と値17.7の間を2等分したときの真ん中の値aは

$$\frac{\square \times 10.4 + \square \times 17.7}{2}$$

であり、値10.4と値aの間を更に2等分したときの真ん中の値は

$$\frac{\frac{\square \times 10.4}{2} + \frac{\square \times 10.4 + \square \times 17.7}{2}}{2}$$

となり、これを簡潔にすれば求めるべき式が得られる。

(3) 名義尺度データの例を(プリントにあるもの以外で)1つ挙げよ。また、その解答した例のデータの型を述べよ。

(4) $x_1=0.6$, $x_2=0.1$, $\sum_{i=1}^3 x_i = 1$ のとき x_3 を求めよ。

第2問(母集団・標本・データ)

日本人男性の年齢別のエコノミークラス症候群(旅行血栓症)の実態を把握するために次の様な調査を行った。日本人男性20歳~69歳の全874名にこの調査に協力してもらい、それぞれの男性の最低血圧と血液粘度を、飛行機搭乗30分前・搭乗開始1時間後・搭乗開始3時間後の計3回測定した。調査に協力して戴いた男性の年齢別の内訳は下表の様になった。

20歳代	30歳代	40歳代	50歳代	60歳代
125名	186名	不明	213名	163名

エコノミークラス症候群とは、飛行機のように非常に狭い席に長時間座っている際の下肢の圧迫による下肢静脈のうっ滞と、水分不足による血液粘度の上昇により、血栓が生じ血管壁に付着する現象から起因する一群の症状である。飛行機が目的地に着陸し席を立つと長時間圧迫されていた足の静脈に付着していた血栓が血管壁からはがれ、静脈流に乗って肺にとび、肺の血管を閉塞させると、急性肺動脈血栓塞栓症が起こります。またこの血栓が脳に移動して血管を閉塞させると脳塞栓、心臓の血管を閉塞させると急性心筋梗塞となり危険です。

- (1) この調査では母集団が5個あります。その母集団5個とは何か?各母集団に対する標本は何か?また各標本の大きさも答えよ。
- (2) この調査ではデータが6項目あります。そのデータ6項目とは何か?またそれらのデータの種類・型・単位を述べよ。(単位は必修問題ではなく分かる者だけで良しとする。)

第3問(医療系データの特性値)

ある病院の一部の患者の体温(°C)・血圧(mmHg)・朝食の摂取カロリー(kcal)をまとめると下表のようになった。

患者番号	1	2	3	4
体温	37.5	37.6	37.5	37.9
最大血圧	112	112	127	121
最小血圧	85	81	87	81
カロリー	1760	1820	2670	670

各データの平均値・中央値・最頻値・最大値・最小値・範囲・第1四分位値・第3四分位値・四分位範囲・分散・標準偏差・歪度を求めよ。

注意 解答に際して、400人の患者のデータを扱っている状況を想定し、各特性値がそれぞれ意味があるのかどうかを検討しながら解答せよ。400人分のデータを手で計算するのは困難なので、この問題自体は4名だけのデータにしてある。

第4問(アンケート調査の特性値と分布)

ある病院の病室環境改善計画の一環として一部の患者17名にアンケート調査を行った。そのアンケートの中の1つの質問とその回答結果が以下のようになった。

質問 この病院のベッドの寝心地は入院前の自宅での状態に比べていかがですか？

- 回答
1. 自宅より寝心地はかなり悪い。
 2. 自宅より寝心地は悪い。
 3. 自宅と比べて寝心地に違いはない。
 4. 自宅より寝心地は良い。
 5. 自宅より寝心地はかなり良い。

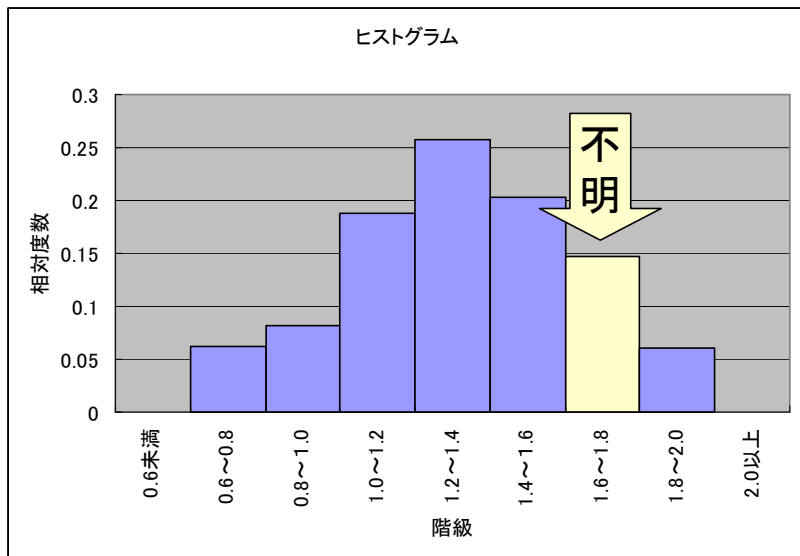
結果 1. 3. 3. 2. 3. 4. 3. 2. 4. 3. 3. 3. 1. 2. 2. 1. 3.

- (1) このデータの平均値・中央値・最頻値・範囲・標準偏差を求めよ。
- (2) 度数分布表とヒストグラムを作成せよ。ただし、度数分布表の項目は度数・累積度数・相対度数にせよ。また、ヒストグラムの縦軸は相対度数のみにせよ。

第5問(応用)

ある大学のある学部の全学生4000名の左の視力の検査結果をまとめた度数分布表とヒストグラムが下のようになった。

階級	相対度数
0.6未満	0.0000
0.6~0.8	0.0625
0.8~1.0	0.0825
1.0~1.2	0.1875
1.2~1.4	0.2575
1.4~1.6	0.2025
1.6~1.8	不明
1.8~2.0	0.0600
2.0以上	0.0000



- (1) 左の視力が1.6以上1.8未満の学生の人数を求めよ。
- (2) このデータの第1四分位値は幾つか？ただし、第1四分位値の値そのものを解答するのではなく、第1四分位値がどの階級に属するかを答えよ。
- (3) このデータの範囲はどの程度になり得るか？ただし、解答は幾つより大きく幾つより小さいの形で、可能な限りその幅を狭めなさい。

- (4) このデータの平均値はどの程度になり得るか？ただし、解答は幾つ以上幾つ未満の形で、可能な限りその幅を狭めなさい。

第6問(Excel 統計)

ある高校のあるクラスの一部の生徒10名の英語と数学の期末テストの結果が下の様にExcelシートにまとめられている。以下の特性値をExcel関数で計算する場合にセルに入力すべきものを省略なしで(セル範囲もきちんと)述べよ。

- (1) 英語の平均値
- (2) 数学の中央値
- (3) 数学の第3四分位値
- (4) 英語の分散
- (5) 数学の歪度
- (6) 英語の標準偏差
- (7) 英語と数学の相関係数

	A	B	C	D
1	学籍番号	英語(点)	数学(点)	
2	1002	76	79	
3	1005	23	40	
4	1007	79	81	
5	1010	94	98	
6	1014	56	99	
7	1018	27	28	
8	1020	38	33	
9	1021	20	24	
10	1027	79	81	
11	1032	64	72	
12				

第7問(Excel 統計)

第6問の英語と数学のデータを用いて以下の間に答えよ。データは各自でExcelに入力せよ。このPDFファイルからExcelにコピーすれば良い。

- (1) 英語と数学の散布図を作成せよ。
- (2) 英語と数学の平均値と標準偏差を求めよ。この問題はExcel関数を用いて計算し、解答は計算結果のみ答えること。
- (3) 英語と数学の相関係数を求めるための計算式に必要な数値を代入した完全な式を記述せよ。この問題は計算結果を答えるのではなく、計算式に値を代入した式のみを示せ。

第8問(実践)

水尾のホームページの第1回レポート用標本(Excel)のシート1にあるデータを用いて以下の間に解答せよ。必要な計算は全てExcelで行い、計算結果の数値のみ解答せよ。

- (1) 統計学・疫学・基礎看護学Iの中でどの講義の分布が最も集中していないばらついた分布と言えるかを必要な特性値を用いて客観的に説明せよ。同様に、どの講義の分布が最もひずみの少ない対称的な綺麗な分布と言えるかを必要な特性値を用いて客観的に説明せよ。
- (2) 統計学のデータを用いて度数分布表とヒストグラムを作成しなさい。また作成したヒストグラムを分析して統計学の成績状況に関して把握できる特徴を考察・議論しなさい。

注意 階級の分割個数(あるいは階級の幅)は最も適切だと判断できるものを自分で決めなさい。

- (3) 統計学のデータを用いて、(2)で決めた階級の分割より分割個数の少ないヒストグラムと分割個数の多いヒストグラムを作成しなさい。そして(2)のヒストグラムと合わせた3つのヒストグラムを比較して、階級分割の仕方がヒストグラムを見る人間に与える印象の違いについて考察・議論しなさい。
- (4) 『統計学と疫学』・『統計学と基礎看護学Ⅰ』の組み合わせに対して散布図を作成しなさい。
- (5) (4)の組み合わせに対して相関係数を求めなさい。
- (6) (4), (5)の結果から、『統計学の成績と疫学の成績』の関係の強さと『統計学の成績と基礎看護学Ⅰの成績』の関係の強さに関して分析できることを考察・議論しなさい。

第9問(実践)

水尾のホームページの第1回レポート用標本(Excel)のシート2にあるデータを用いて以下の問に解答しなさい。必要な計算は全てExcelで行い、計算結果の数値のみ解答せよ。

- (1) アンケートの回答データに対して特性値の中で意味のあるものを全て求めなさい。
- (2) アンケートの回答データを用いて度数分布表とヒストグラムを作成しなさい。
- (3) (1), (2)の結果を分析して自由に考察・議論しなさい。