

RSS Higher Certificate in Statistics, Specimen A

Module 1: Data Collection and Interpretation

1. 異なる研究室での検査の手順が似たような結果を与えるかどうかを知るための計画において、金属疲労テストを3種類の張力レベル（レベル1が最も弱く、レベル3が最も強い）に関して行った。合成物質の標本は同じ束（batch）から得られ、金属疲労テストはヨーロッパの異なる9つの研究室で行われた。下表は、各研究室の結果（疲労のサイクル）とその平均を張力レベルごとに示したものである。

<i>laboratory</i>	<i>strain level 1 results</i>	<i>mean (1)</i>	<i>strain level 2 results</i>	<i>mean (2)</i>	<i>strain level 3 results</i>	<i>mean (3)</i>
A	7335, 6882, 8353	7523	1336, 1693	1515	690, 735	712.5
B	3512, 4000, 4300	3937	977, 1152	1065	428, 460, 473	453.7
C	3822, 5558, 6910	5430	1516, 1607, 1650	1591	630, 675	652.5
D			1740, 1852	1796	447, 718, 935	700
E	4382, 6239, 8930	6517	1488, 1501, 1516	1502	674, 681, 781	712
F	4030, 4138, 4202	4123	1095, 1205, 1290	1197	465, 380, 395	413.3
G	5000, 5245, 5452	5232	1031, 1156, 1238	1142	380, 408, 454	414
H	2510, 2604, 2811, 2900, 2986	2762	738, 771, 864, 883	814	303, 325, 329	319
J	2247, 3800, 5267	3771	957, 1156, 1202	1105	225, 487	356

次の2点に関心がある。

- (1) 異なる研究室で得られた測定値の間の本質的な違いはあるか？
- (2) ここでテストしたのは別の張力レベルでの疲労のサイクルを予測するモデルを見つけることはできるか？

表のデータをもとに、適切なグラフを描き、それを根拠として (1) と (2) について大まかな結論を予備報告書 (preliminary report) にまとめよ。(詳細な統計解析や計算の繰り返しにより、時間を過度に費やさないように注意すること。) この報告書は、専門知識を持たない者にも分かるように書くこと。

2. (i) 単純無作為抽出という用語について理解していることを述べよ。また、単純無作為抽出法が適切な抽出法ではなく、他の抽出法と組み合わせた方が望ましいと思われる条件について記述せよ。
- (ii) 非標本誤差について3つ異なるタイプをあげ、これらの誤差が生じる状況について簡単に記述せよ。
- (iii) 電話調査の主な欠点について概説せよ。

3. 抽出された世帯に送られた顧客調査のアンケート用紙の一部が次のページに示されている。それぞれの世帯にアンケートを記入してもらい、(無料の) 返信用の封筒に入れて郵便で返信してもらうようにしている。

このアンケートの長所と短所についてコメントせよ。また、改善できると思われる質問について代わりの表現 (phrasing) を提案せよ。

Question 3 のアンケートは次のページに示されている。

Question 3 のアンケート

市民顧客調査

説明

- 1) このアンケートはあなたの世帯で主に買い物をする人が記入してください。
- 2) あなたに当てはまる解答について適切な四角に「×」を記入してください。
- 3) 記入したアンケートを返信用の封筒で返信してください。

1. あなたと家族について

1. あなたの誕生日は？ |_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|
 日 月 年
2. あなたは？
 1 既婚 2 独身
3. 家庭には何人住んでいますか？ _____
4. 家族全体の収入に最も当てはまるものはどれですか？
 1 £ 5000 未満 4 £ 20000-£ 30000
 2 £ 5000-£ 10000 5 £ 30000-£ 50000
 3 £ 10000-£ 20000 6 £ 50000 以上

2. あなたの買い物

1. あなたは食事や食料雑貨の買い物でどのお店を利用しますか？（具体的に） _____
2. なぜそこで買い物をするのですか？
 1 距離 4 駐車設備
 2 利便性 5 価格
 3 商品の質 6 食料の品揃え
3. 食料雑貨にどれくらい使っていますか？
 1 £ 15 未満 4 £ 35-£ 44.99 7 £ 75-£ 89.99
 2 £ 15-£ 24.99 5 £ 45-£ 59.99 8 £ 90-£ 104.99
 3 £ 25-£ 34.99 6 £ 60-£ 74.99 9 £ 105 以上
4. 過去3カ月の間にどれくらいの頻度で次の商品を購入しましたか？
 環境にやさしい商品
 1 非常に頻繁に 2 多少頻繁に
 3 あまり購入していない 4 全く購入していない

リサイクル商品

- 1 非常に頻繁に 2 多少頻繁に
3 あまり購入していない 4 全く購入していない
5. あなたは朝食のシリアルを何袋買いますか？
 1 2 3 4
6. あなたがいつも買う朝食のシリアルのブランドはどこですか？（複数回答可）
- 1 Kellog
2 Nabisco
3 Jordan
4 お店のもの
5 その他（挙げてください）

4. (i) 調査における目標母集団 (target population) と調査母集団 (study population) との間に存在しうる違いを説明せよ。また、これら2つの用語を定義し、違いが存在しそうなときと存在しそうでないときの状況の例をあげよ。

調査のための優れた抽出枠 (sampling frame) を得ることの重要性について述べ、さらにそれがどのように得られるか例を2つあげよ。

(ii) 以下に与えられる記述統計量のコンピューターの出力を用いて、データセットを要約するのにどのようにこれらの統計量を用いればよいか、またデータのサブセット (Sector A, B, C) がどのように比較されるか説明せよ。この記述統計量と一緒に、元のデータから作成することができ、データの解釈の手助けとなるような他の情報や図表を示せ。適切な図を用いて解答を示すこと。なお、このデータはある地域での異なる3つの分野 (A, B, C) における賃金である。

Sector	N	Mean	Median	StDev	SEMean	Min	Max	Q1	Q3
A	188	9.609	8.500	5.920	0.432	1.75	44.5	5.250	12.120
B	48	9.492	9.245	4.244	0.613	3.00	20.4	5.890	11.658
C	14	9.507	10.375	3.633	0.971	3.75	15.0	6.500	11.808