

## RSS Higher Certificate in Statistics, Specimen

### Module 8: Survey sampling and estimation

1. 清涼飲料会社の商品マネージャーは主な銘柄の 1 つについて多くの缶が一括して生産されている時間の間、500 個の缶を無作為抽出するよう手配した。
  - (a) 抽出の主な目的は生産ラインの稼働中にラインを占めている缶の精度の評価と規定の目標平均値である 300 ml を十分満たしているかどうかをチェックすることである。この標本に含まれる缶の平均の容量は 302 ml であり、標本の標準偏差は 0.5 ml であった。
    - (i) 缶の平均容量の近似 95% および 99% 信頼区間を求めよ。
    - (ii) 政府の規定では缶の平均容量が少なくとも 300 ml になるように要求している。(i) における結果を用いて、充填設備によって現在生産されている容量を変えた方がいいかどうか、その理由とともに述べよ。
    - (iii) 会社内での日常的な使用のために、商品マネージャーは平均量を  $\pm 0.05$  ml で推定するように要求した。この要求を満たすような 95% 信頼区間を構成するためにどれだけの標本サイズが必要か。
  - (b) 500 個の抽出された缶の次の測定は缶のリングプルが正しく取り付けられているかどうかというものであった。500 個の缶のうち 12 個は欠陥のあるリングプルが取り付けられているということが指摘された。

欠陥のあるリングプルが取り付けられた缶の割合に関する 95% 信頼区間を求めよ。また、(統計学者でない) 商品マネージャーに分かるような言葉でこの信頼区間が示していることを説明せよ。
  - (c) 商品マネージャーは実務的な理由で系統抽出を用いたいと強く思うであろう。これを正当化するためにはどのような仮定が必要か。また商品マネージャーにどのようなアドバイスをするか簡単に説明せよ。
2. ある政府機関は職員による旅費の請求額について調査をしたいと思っている。機関長は費用削減に苦悩しており、財務担当チームに月ごとの請求費の標本を綿密に調べるように指示をした。これは 3 ヶ月間行われる予定である。

標本には 2 つの目的が含まれる： (1) 機関全体にわたる旅費の性質を財務担当チームに知らせる。(2) 職員が誤った額や過剰な額の請求をするのを抑止する役割を担う。

財務担当チーム長は抽出方法を開発するためにあなたのアドバイスを求めている。

以下の情報は抽出案を開発する上で助けになるかもしれない。

  - 機関は 5 つの部門からなっており、そのうちいくつかの部門は他の部門よりも出張が多い。
  - 各請求には特有の照会番号が付けられている。請求の詳細 (受領者、部門、日付、目的、請求額) が主要データベースに保存されている。

- (i) 財務担当チーム長は「母集団」、「抽出」、「フレーム」、「無作為抽出」、「層化無作為抽出」という用語を聞いたことはあるが、それらの正確な意味はあまり詳しくは理解していない。これらの用語が何を意味するか、今回の標本調査と関連づけながら簡潔な言葉で説明せよ。
  - (ii) 財務担当チームが行うことのできる調査方法について簡単に説明せよ。機関が要求していることとデータの入手可能性を考慮に入れること。
  - (iii) もし1ヶ月間の追跡調査を行うことが決まっているならば、推定に比推定法や回帰推定法を用いるべきかどうか、その理由とともに述べよ。
3. 2005年にある大学の研究者は、ある地域の公立学校 (state school) で働いている教員が様々な教育問題に対してどのような見解を持っているかを調べるために、教員の抽出を行った。同じ地域の私立学校 (independent school) で働いている教員についても同様に抽出を行った。重要な質問の1つは「あなたの努力は校長先生に認められていると思いますか?」というものであり、この質問に対する結果は以下の表にまとめられている。

<i>Education sector</i>	<i>Total number of teachers, N</i>	<i>Sample size, n</i>	<i>Number who answered "yes"</i>
State	5000	300	120
Independent	800	50	32

以下の各問いに答えよ。

- (i) 公立学校から抽出された教員が校長先生から努力が認められていると感じている母比率の近似 95%信頼区間を表せ。
- (ii) 研究者はこの質問で「はい」と答えた教員の割合が2つの学校区分で差があるどうかに関心がある。この差の近似 95%信頼区間を求め、それから導かれる結論を述べよ。
- (iii) 努力が校長先生に認められていると感じているこの地域の全ての教員の割合を近似 95%信頼区間の形で表せ。
- (iv) 層別化の主な目的を2つ述べよ。また、この調査ではこれらの目的がどのように適用されるか述べよ。
- (v) この調査は総サンプルサイズ 250 で次の年も実施されることになっている。2005年の調査結果を基準として、公立学校の教員と私立学校の教員との間の抽出の最適な配分を求めよ。

4. 中堅電機会社は本社勤務の従業員に対し IT 分野の近年の変化における彼らの見解を確かめるため調査を行った。

3つの部門のスタッフに質問を行った。調査に用いられる重要な変数の1つは「あなたはコンピュータのスクリーンの前に週に何時間座っていますか」という質問に対する解答である。結果は以下に示されている。

<i>Department</i>	<i>Number in department, N</i>	<i>Sample size, n</i>	<i>Mean weekly number of hours, <math>\bar{x}</math></i>	<i>Standard deviation of the number of hours, s</i>
Accounts	40	15	30	6
R&D	20	10	15	4
Marketing	10	3	20	3

- (i) (a) 有限修正, (b)  $t$  検定 は調査データの分析においていつ使われるべきか説明せよ。この調査から推測を行うときに、それらを用いないことでどのような影響があるだろうか。
- (ii) 各3部門に対し、コンピュータのスクリーンの前に従業員が座っている平均時間について95%信頼区間を求めよ。
- (iii) マーケティング部門の従業員がコンピュータのスクリーンの前に座っている平均時間の95%信頼区間が $\pm 3$ 時間以内となるためにはどれくらいの標本の大きさが必要であろうか。
- (iv) 全ての従業員のコンピュータのスクリーンの前に座っている平均時間に対する95%信頼区間を求めよ。