

アンケート調査支援システム

有馬 昌宏・井内 善臣・辻 新六

1. はじめに

何らかの問題が生じたとき、そして、その問題を解決するために必要とされる情報を、問題に関係する多くの個人や組織から直接的に収集しなければならないとき、アンケート調査の必要性が生じる。しかし、実際には、アンケート調査に不可欠な2つの条件、すなわち、

①標本調査や統計解析の基本知識の習得

②大量の調査データの解析の手段

が不十分なため、調査の実施が断念されたり、調査が実施されたとしても問題解決のために有用な情報を得ることができなかったということになることが多いのではないだろうか。

今回は、アンケート調査を問題解決のために積極的に活用しようとしているときに役立つアンケート調査支援システム QUEST・PACK を紹介する。

2. システム開発の背景と目的

2.1 データ解析用プログラムパッケージの制約

大量の調査データを解析するためには、コンピ

ュータの利用が不可欠である。しかし、大型コンピュータ用に開発されたデータ解析用プログラムパッケージには次のような利用上の制約がある。

(1) 大型コンピュータを利用できる場所と時間が限られている。

(2) 利用する大型コンピュータとプログラムパッケージについて、ある程度の基礎知識を必要とする。

また、パソコンを利用するデータ解析用プログラムパッケージもいくつか開発されてはいるが、大量の調査データを効率的に解析できるレベルには未だ達していないように思われる。このため、自前で調査するとしても、調査データの解析作業だけは第三者に委託せざるをえないというのが現状ではないだろうか。

2.2 調査の基本の理解の困難性

コンピュータは、アンケート調査の必要条件ではあるが、十分条件ではない。調査に要した時間、費用、労力に見合うだけの価値をもつ情報を得るためには、標本設計、調査票作成、実査、調査データの解析の基本を理解したうえで、その技術を実際の調査で生かしていかなければならない。また、調査の企画から報告書の作成までのすべての調査プロセスを外部の機関に委託する場合でも、調査から得られた結論や提言を問題解決のために活用するためには、調査の基本の理解が不可欠であるということに変わりはない。ただし、 Σ や σ 等の記号が出てきただけで拒絶反応を示す人が多

ありま まさひろ 神戸商科大学 管理科学科
いのうち よしみ " 情報処理教育センター
つじ しんろく " 経済研究所
〒655 神戸市垂水区星陵台4-3-3

いので、調査の基本を理解してもらうのは容易ではない。

2.3 システムの開発目的

QUEST・PACK は、次のことをめざして開発した、対話型のアンケート調査支援システムである。

- (1) 大型コンピュータのデータ解析用プログラムパッケージの利用上の制約を解消する。
- (2) 調査データの解析以外の作業も支援する。
- (3) アンケート調査の基本を理解させるための教育コースでも利用できる。

開発に当っては以下に示す機能をもたせるように特に留意した。

- (a) パソコンに触れたことがない人でもすぐにご利用することができるよう、キーボードのテンキー部分だけで必要最小限の処理を行ない、誤ったキー操作は受けつけない。
- (b) マニュアルを見なくとも、ディスプレイに表示される指示や問合せに回答していくことにより、作業に必要な操作を自然に行なえる。
- (c) できるかぎり大量の調査データをできるかぎり高速で処理する。
- (d) 調査データの入力フォーマットの指定や調査データのチェックなどの煩雑な手続きを自動的に行なう。
- (e) 調査データの探索的な解析が可能のように、質的データのカテゴリの統合を行なえる。
- (f) 質問の回答パターンによって被調査者（サンプル）をいくつかのグループに分類し、分類されたグループ間の比較を行なえる。
- (g) 調査データの解析結果の出力にはグラフを活用し、プリンターに印刷される図表はそのままの形で報告書の図表としても利用できる。

3. システムの概要

3.1 システムの構成

表 1 QUEST・PACK の構成と機能

サブシステム	機 能
①標本の抽出	1. 標本の大きさの決定 2. 乱数の発生
②調査票の作成	1. 調査票の新規入力 2. 調査票の修正 3. 調査票の出力
③データファイルの作成	1. 調査データの入力 2. 調査データのチェック 3. 調査データの修正 4. 調査データの出力 5. 調査データの削除
④サンプルの選別	
⑤調査データの解析	1. 単純集計 2. クロス集計 3. 相関・回帰分析
⑥ファイルユーティリティ	1. ファイルの一覧表示 2. ファイルの名前変更 3. ファイルの複写 4. ファイルの保護および保護解除 5. ファイルの削除
⑦標本抽出実験	1. 標本分布 2. 信頼区間

QUEST・PACK は、表 1 に示す 7 つのサブシステムから構成されており、メニュー選択方式により実行したい作業を選択していく。作業が選択されると、

①実行する作業内容

②これから実行する作業に関連する情報や入力した情報を確認するための一覧表

③システムからの指示または問合せ

がディスプレイに表示されるので、一覧表を参照しながらシステムの指示や問合せに答えていけばよい。図 1 は、質問形式の入力を選択したときの表示の例であり、8 番目の質問項目の質問形式を入力することを指示している。

3.2 標本の抽出

「標本の大きさの決定」では、①単純無作為抽出法、②層化抽出法（同数割当法、比例割当法）、③2 段階抽出法（等確率副次抽出、確率比例抽出）について、目標精度と信頼度から必要サンプル数を計算する。

「乱発の発生」では、乱数の最大値と発生個数

調査名：「緑」の満足度調査

質問数：12

番号	質問項目名	回答形式	選択肢数	回答数	回答形式一覧
1	質問項目 1	単一	3	1	
2	質問項目 2	無制限複数	6	6	単一回答——1
3	質問項目 3	無制限複数	5	6	制限複数回答——2
4	質問項目 4	単一	7	1	無制限複数回答——3
5	質問項目 5	単一	4	1	一部順位回答——4
6	質問項目 6	単一	3	1	完全順位回答——5
7	質問項目 7	単一	7	1	数値記入回答——6
8	質問項目 8				文字記入回答——7

次画面：ROLL UP 前画面：ROLL DOWN

質問形式を入力して下さい。回答形式の番号は、上表右側を参照して下さい。

質問番号 = 8 回答形式 =

図 1 質問形式の入力

を与えることにより、指定された数だけの乱数を昇順に並べ換えてプリンターに出力する。

3.3 調査票の作成

QUEST・PACK を特徴づけるのがこのサブシステムであり、調査票に記載される①調査名、②質問の数、③質問の回答形式(④選択肢数、⑤回答数)、⑥質問項目名、⑦質問文(⑧選択肢文、⑨データの測定単位)とデータファイル名および調査概要メモを定義する。

ここで、質問項目名と選択肢文は、調査データの解析結果のタイトルやカテゴリー名として出力され、「質問番号××」と「選択肢番号××」がデフォルト名として用意されている。また、質問の回答形式は、

- (1) 単一回答(「該当するものを1つ選べ」)
- (2) 制限複数回答(「該当するものを3つ(まで)選べ」)
- (3) 無制限複数回答(「該当するものをすべて選べ」)
- (4) 一部順位回答(「望ましい順に3つ選べ」)
- (5) 完全順位回答(「望ましい順に順番をつけよ」)
- (6) 数値記入回答(「……の数はいくつ?」)

(7) 文字記入回答(「……について自由に答えよ」)

の7つの中から選択する(図1)。評定法や一対比較法の質問は単一回答形式の質問とみなす。

QUEST・PACK では、「調査票の作成」サブシステムを実行した後でなければ調査データの入力を行なうことはできない。これは、質問の回答形式に応じて調査データの入力が進められるからである。したがって、本サブシステムは、データ解析用プログラムパッケージでの変数名とデータ入力フォーマットの指定に対応していると考えられることもできる。

なお、調査概要メモには、調査実施日や調査実施方法などの任意の関連情報を記録しておくことができる。

3.4 データファイルの作成

QUEST・PACK では、データの入力フォーマットを指定する必要はなく、次の入力規則にしたがって、回収された調査票に記入されている回答のとおり調査データを入力していけばよい。

- (1) リターンキーで1つの回答の入力が完了する。
- (2) 複数個の回答が許される質問(マルチ質問)

番号	質問項目名	回答形式	回答
1	「緑」に対する満足度	1— 3— 1	1
2	「緑」に対する満足の理由	3— 6— 6	*
3	「緑」に対する不満の理由	3— 5— 5	*
4	緑化促進の重点場所	1— 7— 1	2
5	総合運動公園建設計画の知名	1— 4— 1	1
6	総合運動公園建設の賛否	1— 3— 1	1
7	居住地域	1— 7— 1	3
8	住居形態	1— 6— 1	2
9	居住年数	1— 5— 1	2
10	性別	1— 2— 1	

サンプル番号: 3

選択肢文表示・入力ミス: HELP 入力終了: /

質問番号 = 10

回答形式 = 単一

選択肢数 = 2

回答数 = 1

回答 =

図 2 調査データの入力

で、指定した回答数よりも実際の回答が少ない場合には、その質問のすべての回答の入力が完了したことをリターンキーで示す。

(3) 無回答は*で示す。

したがって次のように調査データの入力が簡単に行える。

- (a) 欠測値を指定する必要がある。
- (b) 選択肢の数が2桁である質問の回答を入力するさいに入力カラムに注意しなくてよい。たとえば回答が2であるとき、02を入力しなくてよい。
- (c) 無制限複数回答形式の質問では、各選択肢に回答しているかないかを0-1変数で示す必要はなく、選択肢の数だけデータを入力すればよい。たとえば7つの選択肢の中から2と3を回答しているとき0110000と入力し

ないで、2と3だけ入力する。

また、定義された選択肢数を超える値の回答はチェックされ、その入力が受けつけられないようになっている。

図2はデータを入力するときの表示の例であり、10番目の質問項目の回答を入力することを指示している。

入力されたデータは、1サンプルごとにディスプレイに表示されるとともにプリンターにも出力され(図3)、50サンプルごとにディスクに記録されていく。なお、すべてのサンプルの調査データを一括して入力する必要はない。

ところで、入力された調査データをそのまま解析することはできない。なぜならば、無回答や回答内容に論理的な矛盾のあるサンプルが存在するからである。このため、QUEST・PACKには、

サンプル番号

```

1: 1 010010 ***** 7 1 2 3 3 4 1 35.00 3
2: 2 ***** 01000 3 1 2 2 4 1 2 39.00 8
3: 1 ***** ***** 2 1 1 3 2 2 1 28.00 5
4: 1 001010 ***** 3 2 1 1 2 2 1 40.00 3
5: 1 000001 ***** 2 2 3 3 * 1 2 33.00 8

```

図 3 調査データのプリンターへの出力

サンプル番号	3		
質問番号	1	-->	1
質問番号	2	-->	*****
サンプル番号	5		
質問番号	8	-->	*
サンプル番号	13		
質問番号	1	-->	1
質問番号	3	-->	00100
質問番号	9	-->	*

ろ過的質問無回答
無回答
ろ過的質問矛盾回答
無回答

図 4 チェックされた調査データのプリンターへの出力

「調査データのチェック」が用意されている。これを選択すると、

- ①質問に無回答があるサンプル
- ②制限複数回答形式と一部順位回答形式の質問で、入力された回答の数が指定された回答数よりも少ないサンプル
- ③ろ過的質問(「前の質問で「1.賛成」と答えた人に聞きます」というタイプの質問)に矛盾する回答をしているサンプル

をサーチし、そのサンプル番号とチェック理由の一覧をプリンターに出力する(図4)。

3.5 サンプルの選別

アンケート調査では、男女・年齢別にテレビの視聴時間に違いがあるかどうかを分析するというように、質問に対する回答パターンによってサンプルをいくつかのグループに分類し、グループ間での平均や比率の比較を行なうことがある。「サンプルの選別」では、このような分析が可能のように、指定した質問項目で特定の回答内容を示すサンプルだけを選別し、選別されたサンプルの調査データから構成される新しいデータファイルを作

成する。選別の条件は、複数の質問に対して設定することができる。この「サンプルの選別」と次に説明する「クロス集計」を組み合わせれば、三重クロス集計表や四重クロス集計表などによる分析も可能になる。

3.6 調査データの解析

QUEST・PACK で利用できる調査データの解析機能は、①単純集計、②クロス集計、③相関・回帰分析、の3つであり、多変量解析手法やデータ加工機能は提供されていない。ただし、QUEST・PACK のデータファイルを大型コンピュータのデータ解析用プログラムパッケージのデータファイルに変換することは可能である。

各機能の出力内容は以下に示すとおりであるが、いずれの場合にも、指定した質問項目の質問に無回答のサンプルは解析の対象から自動的に除外される。また、カテゴリーの統合を行ない、統合化された新カテゴリーのもとで単純・クロス集計結果を出力させることができる。ただし、解析結果をディスプレイに表示できるのは、カテゴリーの数が9までの場合に限定される。

質問番号： 1 「「緑」に対する満足度」の棒グラフ

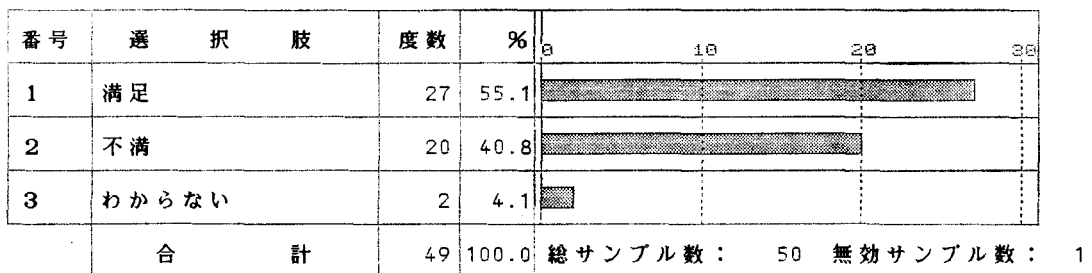


図 5 度数分布表と棒グラフのディスプレイ表示

質問番号： 4 「緑化推進の重点場所」 の円グラフ

総サンプル数： 50
無効サンプル数： 1

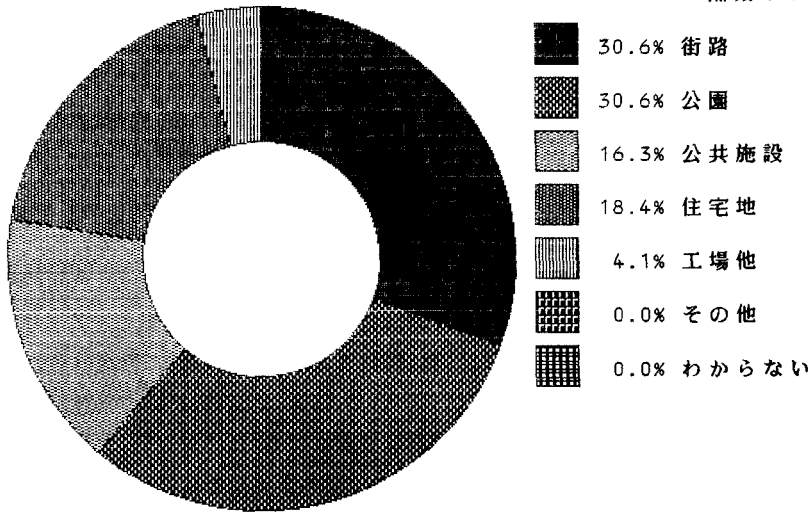


図 6 円グラフの出力

1) 単純集計

調査データの質問項目ごとの集計結果を、

- ①棒グラフと組み合わせられた度数分布表(図5)
- ②円グラフ(図6)
- ③帯グラフ(図7)

のいずれかの形式で出力する。数値記入回答形式の質問項目については、データを区分してカテゴリー化する必要があるが、平均、分散、歪度、尖度などの記述統計量も出力される。

2) クロス集計

2つの質問項目の調査データのクロス集計結果を

- ①クロス集計表(図8)
- ②ステレオグラム(図9)

のいずれかの形式で出力する。また、クロス集計表のプリンターへの出力では、クラマーのV係数、コンティンジェンシー係数と質問項目間に関連性があるかどうかの χ^2 検定結果も印刷される。ただし、数値記入回答形式の質問項目を集計の対象として選択した場合には、データのカテゴリー化が

質問番号： 8 「住居形態」 の帯グラフ

総サンプル数： 50
無効サンプル数： 0

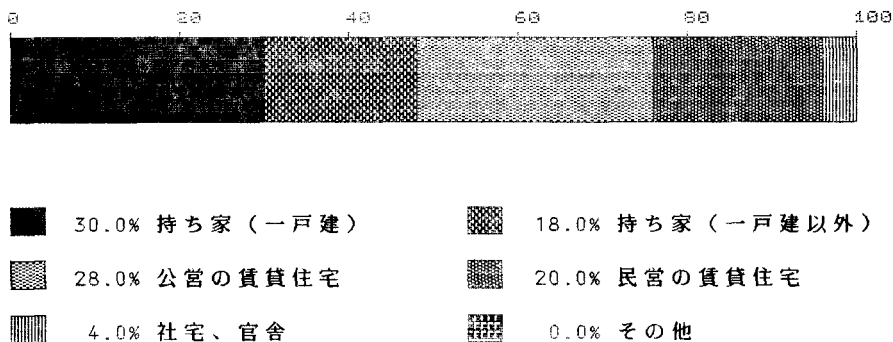


図 7 帯グラフの出力

*** (縦) 質問番号 1 と (横) 質問番号 8 のクロス集計 ***

住居形態	持ち家（一戸建）							
	持ち家（一戸建以外）							
	公営の賃貸住宅							
	民営の賃貸住宅							
	社宅、官舎							
「緑」に対する満足度	その他							
	合 計							
無効サンプル数	0							
無効サンプル数	1							
満足	9	5	5	6	2	0	27	
不満	5	4	8	3	0	0	20	
わからない	0	0	1	1	0	0	2	
合 計	14	9	14	10	2	0	49	

図 8 クロス集計表のディスプレイ表示

必要である。

3) 相関・回帰分析

独立変数および従属変数となる質問項目を指定することにより、相関係数と回帰式を計算し、

① スキャターダイアグラム

② 分散分析表

を出力する。

3.7 ファイルユーティリティ

作成されたデータファイルを管理するために用

意されているのが「ファイルユーティリティ」であり、作成したデータファイルのバックアップを取ったり、データファイルに書き込み防止の保護をかけることなどができる。

3.8 標本抽出実験

QUEST・PACK には、

① 標本分布と正規分布との間の関係

② 区間推定での信頼度の意味と標本の大きさと信頼区間の幅の関係

*** (左) 質問番号 3 と (右) 質問番号 4 のステレオグラム ***

「緑」に対する不満の理由

緑化推進の重点場所

- 1: 住宅の庭に緑がない
- 2: 公園の緑が少ない
- 3: 街路樹が貧弱
- 4: 家屋が密集
- 5: その他

- 1: 街路
- 2: 公園
- 3: 公共施設
- 4: 住宅地
- 5: 工場他
- 6: その他
- 7: わからない

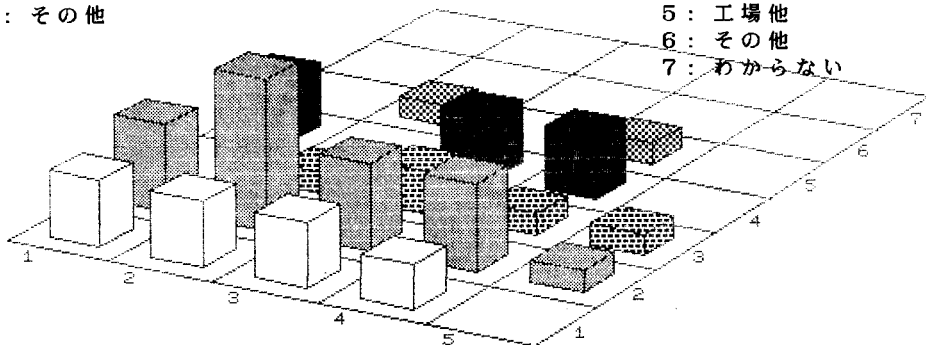
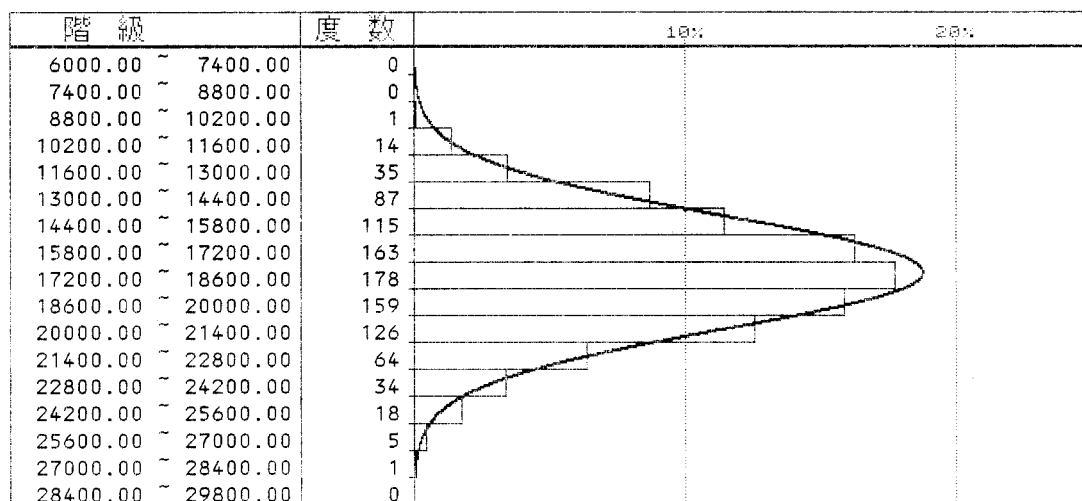


図 9 ステレオグラムのディスプレイ表示

(注) 「緑化促進の重点場所」の「6.その他」と「7.わからない」をカテゴリーの統合機能を利用して「5.工場地」と統合すれば、より見やすい図になる。

母平均 17750.00 分布の平均 17830.10 組合せの数 658008
母標準偏差 7017.83 標準誤差 3477.49 標本の大きさ 5 抽出回数 1000



度数表示 → *, 表示変更 → / , スケール変更 → + , 終 → J (実現度数)

図 10 標本抽出実験(標本分布)

の理解を助けるシミュレーションプログラムが用意されている。すなわち、入力したN個のデータからn個のデータを無作為に抽出して、

- 1) 標本平均または標本比率を計算するという実験をm回行ない、その結果得られる分布と、入力したN個のデータの平均と分散から理論的に計算される正規曲線を重ねて表示する。
- 2) 標本平均または標本比率の $(1-\alpha) \times 100\%$ 信頼区間を求めるという実験をm回行ない、その結果を表示する

というシミュレーションを行なう(図10, 図11)。

実験できる標本抽出法としては、①単純無作為抽出法、②系統抽出法、③層化抽出法(同数割当法, 比例割当法)、④2段抽出法(等確率副次抽出, 確率比例抽出)が選択できる。

4. システムの制約と利用可能性

4.1 システムの制約

QUEST・PACK では、調査データの記憶媒体としてフロッピーディスクを利用するため、処

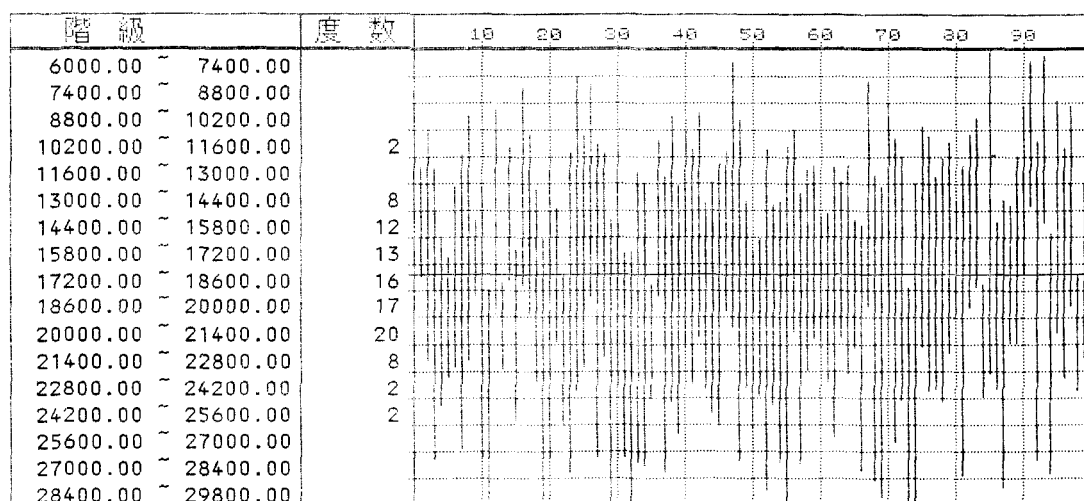
理することのできる調査票の大きさとサンプル数には次の制約がある。

- (1) 調査票の質問の数は最大99までである。
- (2) ディスクの容量、調査票の質問の数、各質問の回答形式や質問文の長さなどに依存するが、単一回答形式の質問20問と数値記入回答形式の質問10問からなる調査票で1MBのディスクを利用する場合には、処理できるサンプルの数は約10,000である。

また、調査データの解析に要する時間は、次のとおりである(使用機種PC-9800)。

- (1) 調査票の質問の数や処理の対象となる質問項目の回答形式に依存するが、サンプル数が1000の場合に単純集計を行なうと、数値記入回答形式以外の質問項目で約1分、数値記入回答形式の質問項目では記述統計量を計算するため約2分を要する。
- (2) クロス集計では、単純集計に要する時間にクロスの分類をとるための比較演算の時間が加算され、サンプル数1000、単一回答形式の

母平均 17750.00 分布の平均 18188.00 組合せの数 658008
母標準偏差 7017.83 信頼度 95.00% 標本の大きさ 5 抽出回数 100



実験による信頼度 = 90.00 %

終 → J

図 11 標本抽出実験(信頼区間)

2つの質問項目の場合には、クロス集計結果を得るまでに約90秒を要する。

4.2 QUEST・PACK の利用可能性

QUEST・PACK は、以下に示す場合に特に利用の効果があるものと期待される。

(1) 予備調査による調査票の修正

実査のやり直しがきかないアンケート調査では収集されるデータの内容を規定する調査票が調査の成否の鍵を握る重要な役割を果たす。このため予備調査によって質問の妥当性の検討を行なうことが望ましい。QUEST・PACK による予備調査データの解析結果から、問題解決のために必要な情報を導出できるかどうかを検討することができるであろう。

また、予備調査を実施しない場合でも、50サンプル程度の仮想データを入力して解析し、質問の妥当性をチェックすることに QUEST・PACK を利用することもできる。質問の修正は、「調査票の作成」サブシステムで行なうことができる。

(2) 大型コンピュータの利用に制約がある場合
QUEST・PACK は、大型コンピュータのデータ解析用プログラムパッケージの制約を解消することを目的として開発されたシステムであり、サンプル数が1000程度までの場合に最大の効果を発揮する。また、解析結果の作表・作図の手間を考えると、サンプル数が2000~3000程度までの調査でも、QUEST・PACK の利用効果は十分あると思われる。

(3) 調査データの入力手段

QUEST・PACK は、初めてキーボードに触れるという人でも容易にデータ入力作業を行なうことができるように設計されている。したがって、QUEST・PACK を利用して作成したデータファイルを大型コンピュータのデータ解析用プログラムパッケージのデータファイルに変換し、調査データの解析は大型コンピュータで行なうという利用の仕方もある。

(4) アンケート調査の教育コースでの利用
われわれは、

①「標本抽出実験」を実行させたり「調査データの解析」の出力結果を利用しながらアンケート調査の基本を説明し、

②サンプル数50程度の簡単なアンケート調査を実際に行なってもらう

という形で QUEST・PACK をアンケート調査の教育コースに利用しているが、単にテキストと黒板を使って説明していたときと比べて、受講生に興味をもたせながらアンケート調査の基本と実際を理解してもらうことができるようになったとわれわれなりに評価している。

5. おわりに

今回紹介した QUEST・PACK は、昨年の秋季研究発表会の OR ソフトウェアショーで発表したシステムに改良を加えたものである。発表会において貴重なコメントをくださった学会員の方々に感謝いたします。

参考文献

有馬昌宏, 井内善臣, 辻新六: アンケート調査の基本と実際, 朝倉書店, 1986



研究部会報告

●交通問題●

●第12回

日時: 昭和61年 2月26日(水) 18:00~20:00

場所: 東洋経済新報社ビル (日本橋) 出席者: 12名

テーマ: 「鉄道技師の見た国際復興銀行の活動」

講師: 島 隆氏 (日立製作所)

世界銀行グループの中の1つである国際復興銀行で鉄道プロジェクトに対する貸し付け承認の仕事に従事していた経験から最近の同銀行の活動状況, 認可に対する条件等について説明があった。

●第13回

日時: 昭和61年 3月19日(水) 18:00~20:00

場所: 東洋経済新報社ビル (日本橋) 出席者: 12名

テーマ: 「自由討論」

昨年度からのひきつづきで、国鉄に関する諸問題について検討するとともに、次のテーマの中から研究を進めたいという提案がなされた。1. タクシー会社, 2. 徒歩からの交通の見直し, 3. めまい現象 (運賃体系を旅客と通勤の2本立とする。), 4. 環状道路・東京湾横断道路, 5. 通勤電車のクラス分け, 6. 見学会 (物流会社), 7. 列車の等級の多様化など。

●D P●

日時: 4月22日(火) 18:00~20:00

場所: 日科技連 出席者: 7名

発表: 安田正実(千葉大) “多変量停止問題に対する変化する支払の変形”

プレイヤーへの支払に対する無作為ベクトル別の2つの種類が存在するとする。系は個々の停止戦略を集計することによって停止する時、プレイヤーに与えられる支払は彼が“stop”または“continue”と宣言するにしたがって決定される。2人のプレイヤーの零和と簡単な優位則の問題の場合はいわゆる Dynkin のゲームに帰着する。最適方程式が導かれ Nash 平衡戦略が得られた。また他に“ポテンシャル理論法による OLA 最適停止問題に対する値について”の発表も同時に行なわれた。

●日本のシステム科学(研究グループ)●

●第13回

日時: 昭和61年 4月19日(土) 14:00~17:00

場所: 八丁堀 東京都勤労福祉会館 出席者: 6名

テーマ: 「平和: 共存共栄のオペレーションズリサーチ」

発表者: 上田亀之助 (上田イノベーション研究所・杉野女子大学)

平和とは共存共栄が実現した場合の状態を示すコトバです。平和とは相手や関係者もその気になってくれないと実現できない状態です。ですから、自分が平和を保つ方法論と相手がこちらに対して平和を保ってくれる、双方のオペレーションズをリサーチする必要があります。