

Current Index to Statistics

Technical Reference Manual
CIS EXTENDED DATABASE

1997 Edition

Sponsored by
The American Statistical Association and
The Institute of Mathematical Statistics

Ronald A. Thisted

Department of Statistics, The University of Chicago

Michael J. Wichura

Department of Statistics, The University of Chicago

Klaus Hinkelmann

Department of Statistics, Virginia Tech

Bruce E. Trumbo

Department of Statistics, California State University, Hayward

November 1997

Editor, *Current Index to Statistics*

RONALD A. THISTED 1995–

Editors, *CIS Printed Volume*

| | | |
|---------------------|---------------|---------|
| KLAUS HINKELMANN | Volumes 21– | 1995– |
| Edward E. Gbur, Jr. | Volumes 16–20 | 1990–94 |
| Richard K. Burdick | Volumes 11–15 | 1985–89 |
| James E. Gentle | Volumes 6–10 | 1980–84 |
| Brian L. Joiner | Volumes 1–5 | 1975–79 |

Editors, *CIS/ED*

| | |
|--------------------|---------|
| MICHAEL J. WICHURA | 1995– |
| Ronald A. Thisted | 1994–95 |
| Bruce E. Trumbo | 1990–93 |

CIS Management Committee, 1997

| | |
|--------------------|-------------------|
| Paul Shaman, CHAIR | |
| Douglas M. Bates | Michael J. Meyer |
| Julia A. Norton | Alan M. Zaslavsky |

Copyright © 1997

American Statistical Association & Institute of Mathematical Statistics
All rights reserved.

Licensees of the 1997 *CIS/ED* are granted a non-exclusive, royalty-free license to copy and to distribute this *CIS/ED Technical Reference Manual* to authorized users of the database. Each such copy must contain the title page and this page in its entirety.

No bibliography, index, or other work (whether in computer-readable, printed or other form) derived from more than 250 Database records contained in the Database may be distributed without the written permission of the authorized representative of the *CIS/ED*. Any publication prepared with the substantial assistance of the Database should acknowledge its use and cite it as a reference: “Michael J. Wichura, Editor (1997). *Current Index to Statistics Extended Database*. American Statistical Association (Alexandria, VA) and Institute of Mathematical Statistics (Hayward, CA).”

Certain programs, files, and documents on CD-ROM distribution of the *CIS/ED* are copyright by their individual authors. Each such item contains a notice of copyright, and may also contain other restrictions on redistribution described herein.

Information concerning current prices, licensing, and availability can be obtained from our page on the World Wide Web, (URL: <http://www.stat.uchicago.edu/~cis>), or by contacting either the ASA (+1 703 684 1221) or IMS (+1 510 783 8141) offices.

Contents

| | | |
|----------|---|-----------|
| 1 | CD-ROM Organization | 1 |
| 2 | Technical Specifications and Contents | 2 |
| 2.1 | Macintosh Files | 2 |
| 2.2 | DOS/Windows Files | 2 |
| 2.3 | Unix™ Files | 2 |
| 3 | Field 1 Specifications | 4 |
| 3.1 | Books | 4 |
| 3.2 | Articles in journals and other periodicals, <i>excluding</i> those distributed primarily in electronic form | 5 |
| 3.3 | Articles in proceedings and edited books | 6 |
| 3.4 | Computer readable material (programs, databases etc.) <i>excluding</i> contents of electronic journals or periodicals | 7 |
| 3.5 | Articles in electronic journals and periodicals | 8 |
| 3.6 | Administrative records | 10 |
| 3.7 | Symbols used in Column 27 | 11 |
| 3.8 | Future types of entries | 11 |
| 3.9 | Sample records and information for programmers | 11 |
| 4 | Abbreviation File Formats | 13 |
| 4.1 | ABBRMST.V97 | 13 |
| 4.2 | ABBR.V97 | 15 |
| 4.3 | ABB.V97 | 16 |
| 4.4 | AB.V97 | 16 |
| 4.5 | ABAMS.V97 | 17 |
| 4.6 | ABBRAMS.V97 | 17 |
| 5 | Notes on Fields 5, 6 and 7 | 18 |
| 5.1 | Accent marks | 18 |
| 5.2 | Transliterations | 19 |
| 5.3 | U.S. spelling | 19 |
| 5.4 | Field 7 administrative information | 20 |
| 6 | Names Transliterated from Cyrillic | 21 |

| | | |
|----------|--|-----------|
| 7 | Abstracts and Book Reviews | 26 |
| 7.1 | Organization of Abstract/Book Review Files | 26 |
| 7.2 | The Data Files | 27 |
| 7.3 | The Index Files | 33 |
| 7.4 | The Catalog File | 33 |
| 7.5 | Printing Abstracts and Book Reviews | 34 |

1 CD-ROM Organization

CIS/ED is encoded in ASCII files organized using the ISO-9660 format for CD-ROMs. For the 1997 edition, the CD-ROM is prepared in a “hybrid” format that will make installation and use easier for Macintosh users, while maintaining ISO-9660 compatibility for users of other operating systems.

To deal with operating system differences, most files on the CD-ROM are present in three different versions — one for DOS/Windows, one for Macintosh, and one for UnixTM. Users of a particular operating system will typically be interested in only one of the versions. Each of the versions contains the entire database, as well as auxiliary indexes and other files.

The CD-ROM has a directory for each major system. These areas contain

1. The DOS directory contains MS-DOS and Windows readable files.
2. UnixTM readable files, containing the same files as in (1), except with UnixTM newlines substituted for the CR/LF end-of-line convention used by MS-DOS. These files will also have file names (possibly appearing to the user in lower-case letters) corresponding to the file names used by MS-DOS. The UnixTM collection of files is contained in a directory named UNIX, to “hide” it from MS-DOS.
3. Macintosh readable files, containing the same files as in (1), except with Macintosh newlines substituted for the CR/LF end-of-line convention used by MS-DOS. File names correspond to the file names used by MS-DOS, but may appear more “Mac-like.” For instance, program names (that is, applications) may not end in “.exe” as they do in MS-DOS, and Macintosh-specific files and programs may have names that are not constrained by the “8.3” requirements of DOS.

The CD-ROM contains a Macintosh HFS file system, which is not directly accessible to DOS, Windows, and UnixTM users.

Within each platform-specific directory, the CD-ROM contains directories containing license agreements, abstracts, indexes, abbreviations, etc.

2 Technical Specifications and Contents

The CD-ROM version of the *CIS/ED* is distributed on a single CD-ROM in hybrid HFS/ISO-9660 format.

The HFS partition contains a standard Macintosh file system. This file system contains Macintosh-specific files, indexes, and programs.

The ISO-9660 partition contains DOS and Unix™-readable files. On this partition, directory and file names are restricted to have at most 8 alphanumeric characters, followed optionally by a period and an extension consisting of no more than three alphanumeric characters. This “8.3” format guarantees compatibility with DOS systems.

2.1 Macintosh Files

Files in the HFS (Macintosh) partition of the CD-ROM are either

- Executable binary files (applications) that run under System 7.0 and later on Apple Macintosh computers. Some applications are included that may run properly only on the Motorola 68000 series of computers or only on the PowerPC family of computers.
- ASCII files, with individual lines terminated by the CR character (Hexadecimal 0D).

2.2 DOS/Windows Files

Files in the DOS directory of the CD-ROM are either

- Executable binary files that run under MS-DOS 6.0 and later on machines compatible with the Intel 80286 cpu instruction set, or
- ASCII files, with individual lines terminated by a CR/LF pair of characters (Hexadecimal 0D0A).

IBM-compatible use of the CD-ROM requires Microsoft CD-ROM Extensions, version 2.0 or higher, or equivalent.

2.3 Unix™ Files

Files in the UNIX directory of the CD-ROM are

- ASCII files, with individual lines terminated by the LF character (Hexadecimal 0A),
- TAR- or SHAR-format files, which require de-archiving using shell commands.

3 Field 1 Specifications for the CIS/ED 1997 Edition

Unique Identifiers

These pages give the current specifications for Field 1 of *CIS/ED*. Field 1 is a 28 character field which for historical reasons has various fixed-width subfields. The entries in the database have been sorted by a case-insensitive sort on Field 1.

Unique Identifiers

Field 1 is a unique identifier for entries in the database. When Field 1 for two or more indexed items would otherwise be identical, column 27 is used to assign a unique character, typically a lower-case letter, for each entry. This occurs most frequently for book reviews, since reviews of more than one book often appear on a single journal page. In some instances, the database records the *range* of pages devoted to book reviews in the particular issue of the journal. In this case, the same page range has sometimes been assigned to all reviews in that issue. Similar considerations apply to discussion papers. In some circumstances the individual discussants' remarks are separately indexed, but the page range refers to the entire collection of discussants' remarks. A third common situation in which Field 1 would not be unique involves books by the same publisher, either with the same number of pages (sometimes estimated), or an unknown number of pages (denoted in each case by a question mark). The (unique) entry in Field 1 is used to link database entries with abstracts.

3.1 Books

| | | |
|-------|------|---|
| 1- 2 | (2) | Year |
| 3-18 | (16) | Publisher abbreviation |
| 19-20 | (2) | Spaces |
| 21-25 | (5) | Number of pages Right justified, if numeric '?' in column 24 if unknown |
| 26 | (1) | The letter p |
| 27 | (1) | Optional symbol (used for uniqueness if needed) |
| 28 | (1) | The letter B if no summary contained in <i>CIS/ED</i> |

The letter b (lower case), if summary available

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| (a) | 9 | 1 | P | u | b | l | i | s | h | e | r | A | b | b | r | e | v | i |
| (b) | 9 | 1 | P | u | b | l | i | s | h | e | r | A | b | b | r | e | v | i |
| | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |

Notes: Example (b) illustrates an entry for a book with an unknown number of pages.

3.2 Articles in journals and other periodicals, *excluding* those distributed primarily in electronic form

- 1- 2 (2) Year
- 3-10 (8) Journal Abbreviation (NOT including vol/issue)
- 11-13 (3) Volume number
- 14-16 (3) Issue number (if any)
- 17-21 (5) Beginning page number:
Right justified, if numeric or roman numeral
'?' in column 20, if unknown
- 22 (1) Hyphen
- 23-26 (4) Ending page number:
Right justified (rarely truncated), if numeric
Left justified, if roman numeral
'?' in column 24, if unknown
- 27 (1) Optional symbol (+, ?, or Capital Letter; if
needed, lower case letter is used for uniqueness)
- 28 (1) The letter J if no abstract contained in *CIS/ED*
The letter j (lower case), if abstract available

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| (a) | 9 | 1 | J | o | u | r | n | l | A | b | | 1 | 2 | | 1 | 2 | 3 | 4 |
| (a1) | 9 | 1 | J | o | u | r | n | l | A | b | | 1 | 2 | | 1 | 2 | 3 | 4 |
| (b) | 9 | 1 | J | o | u | r | n | l | A | b | | 1 | 2 | | 1 | 2 | 3 | 4 |
| (b1) | 9 | 1 | J | o | u | r | n | l | A | b | | 1 | 2 | | ? | - | ? | J |
| (c) | 9 | 1 | J | o | u | r | n | l | A | b | | 1 | 2 | | 1 | 2 | 3 | 4 |
| (d) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | 3 | | 1 | 2 | 3 | 4 |
| (e) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | / | | 1 | 1 | 2 | 3 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (e1) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | / | 1 | 2 | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | J | | | | | |
| (e2) | 9 | 1 | J | o | u | r | n | l | A | b | | | / | 8 | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | J | | | | | |
| (f) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | / | A | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | J | | | | | |
| (f1) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | / | ? | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | J | | | | | |
| (f2) | 9 | 1 | J | o | u | r | n | l | A | b | | | 1 | 2 | 3 | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | J | | | | | |
| (g) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | 3 | | | 1 | 2 | 3 | 4 | 5 | - | . | . | 1 | 2 | J | | | | |
| (h) | 9 | 1 | J | o | u | r | n | l | A | b | 1 | 2 | | | | | | | | | | | | i | v | - | v | i | i | J |
| | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | | | | |

Notes: (a) and (a1) differ only in that record (a1) corresponds to an article whose abstract is supplied on the *CIS/ED* CD-ROM distribution. See the *User Guide* section on abstracts for additional information.

(e)–(e2) Issue numbers 1–9 are written with a slash; issue numbers above 99 (very rare — (f2) shows an example) are written without a slash. [In the printed volumes and printouts of retrieved records, software should insert a slash.] A few periodicals have only issue numbers. Also, a few journals have neither volume nor issue numbers; their periodicity is indicated only by the year of publication. These have spaces in columns 11–16, and can be distinguished from proceedings by looking at column 28.

(g) It would be very unusual to have a five-digit page number in a journal that starts renumbering with each issue, so either column 16 or 17 is ordinarily blank for easy readability of raw database records.

For special issues of journals, the format allows designations such as 34/ A, 11/ S, etc. without causing confusion. In some cases this is superior to putting a capital letter in column 27. (For example, sometimes there is an entire “supplemental issue” in the volume. However, sometimes there are supplemental pages or appendixes at the end of an issue or one particular issue is split into two parts, A and B; then a symbol in column 27 could be used, with an issue number if needed.) Of course, journals with regular A and B series, have separate journal abbreviations for the two series.

3.3 Articles in proceedings and edited books

- 1- 2 (2) Year
- 3-12 (10) Abbreviated title of edited book or proceedings
 (Rarely include digits.)
- 13-16 (4) Spaces (or “volume” number, right justified)
- 17-21 (5) Beginning page number:

- Right justified, if numeric or roman numeral;
'?' in column 20, if unknown
- 22 (1) Hyphen
- 23-26 (4) Ending page number:
Right justified (rarely truncated), if numeric
Left justified, if roman numeral
'?' in column 24, if unknown
- 27 (1) Optional symbol (used for uniqueness if needed)
- 28 (1) The letter P if no abstract contained in *CIS/ED*
The letter p (lower case), if abstract available

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| (a) | 9 | 1 | P | r | o | c | e | e | d | A | b | b | | | | | | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | P |
| (b) | 9 | 1 | P | r | o | c | e | e | d | A | b | b | (| 1 |) | | | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | P |
| (b1) | 9 | 1 | P | r | o | c | e | e | d | A | b | b | (| 1 | 2 |) | | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | P |
| (c) | 9 | 1 | P | r | o | c | e | e | d | A | b | b | (| S |) | | | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | P |
| (d) | 9 | 1 | P | r | o | c | e | e | d | A | b | b | 1 | / | 2 | | | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | P |
| (e) | 9 | 1 | P | r | o | c | e | e | d | A | b | b | 1 | 2 | 3 | | | | 1 | 2 | 3 | 4 | - | 1 | 2 | 3 | 4 | P |
| | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |

Notes: (b)–(b1) “Volume” number is in parentheses unless (very rarely) the number of digits exceeds two, as in example (e). Except where space constraints dictate, volume numbers appear in parentheses. When they must be omitted (as in examples (d) and (e), search software should act as if the volume field were enclosed in parentheses. If parentheses are present in Field 1, column 16 is guaranteed to contain the character ‘)’.

3.4 Computer readable material (programs, databases etc.) *excluding* contents of electronic journals or periodicals

- 1- 2 (2) Year
- 3-18 (16) Publisher, Producer, or Distributor abbreviation
- 19-20 (2) Spaces
- 21-24 (4) Number of kilobytes, megabytes, or gigabytes
If unknown, the character ? is in column 23
- 25-26 (2) The letters kb, mb, or gb
- 27 (1) Optional symbol (used for uniqueness if needed)

- 28 (1) The letter C if no summary contained in *CIS/ED*
 The letter c (lower case), if summary available

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (a) | 9 | 1 | P | u | b | P | r | o | d | o | r | D | i | s | t | r | i | b | | | | | 1 | 2 | 3 | m | b | C |
| (b) | 9 | 1 | P | u | b | P | r | o | d | o | r | D | i | s | t | r | i | b | | | | | | | ? | m | b | C |

3.5 Articles in electronic journals and periodicals

This specification was used for the first time in 1994. Because electronic journals and other publications are still evolving, it is likely that these specifications will change over the next few years to accommodate evolving standards of bibliographic documentation. It is even possible that electronic journals will revise their own means of self-reference and will make these changes retroactive. The *CIS/ED* will reflect these changes as they occur in the future.

As of 1994, electronic journals have adopted different schemes for citing references to articles. Column 26 of an entry's *CIS/ED* record uses a code to indicate the scheme used by the particular publication. The codes in current use are

- p This code is used if the source assigns virtual page numbers to its contents, and articles are sequentially paginated by volume or volume and issue. Such publications can be thought of as an electronic image of a printed publication, and page numbers are well defined. They identify both position of an article within the publication and position within individual articles. In this case the "size" subfield of Field 1 contains the ending page number of the article.
- f This code is used if the source identifies articles by a file number or reference number and articles are individually paginated within articles. This scheme uses the file number to locate the article within an issue, and a page number to identify locations within an article. In this case the "size" subfield contains the number of pages.
- s This code is used if the source identifies articles by a file number but articles are not paginated. In this case, the file number locates

articles within an issue, but there are no assigned pages within articles to locate position within articles. For such entries, the “size” subfield contains the number of 1024-byte blocks (rounded upward), to give a rough indication of the size of the item.

- t This code is used when the source identifies articles by specifying a unique article name by which the material can be retrieved, rather than by an identifying page or file number. In this case, the file name is included parenthetically as part of the title field (Field 2) in the main database entry, and columns 17–20 of the corresponding index record contain an integer assigned by the editors of the *CIS/ED*.

NOTE. The identifying number used in record of type ‘t’ are *not* assigned by the editors of the electronic journal, and should *not* be used in bibliographic citations to articles contained in such a journal. The CIS-assigned number is used in order to link database entries with abstracts, when the latter are present, and provide for a unique identifier in Field 1 for each entry in the database.

- 1- 2 (2) Year
- 3-10 (8) Journal Abbreviation (**not** including vol/issue)
- 11-13 (3) Volume number
- 14-16 (3) Issue number (if any)
- 17-20 (4) Identification field
Depends on contents of column 26; see table below
- 21 (1) ‘-’ (if source sequentially paginated), ‘␣’ otherwise
- 22-25 (4) Size field
Depends on contents of column 26; see table below
- 26 (1) A code letter indicating the contents of columns 17–25, according to the following table. (See notes above.)

| Col 26 | Identification Field contains | Col 21 | Size Field contains |
|-----------|----------------------------------|-----------|------------------------|
| p | Starting page number | - | Ending page number |
| f | Article sequence number | | Number of pages |
| s | Article sequence number | | Number of 1Kb blocks |
| t | CIS ID number | | Number of 1Kb blocks |

- 27 (1) Reserved (may be used if needed for uniqueness)
- 28 (1) The letter E, if no abstract is available
The letter e, if an abstract is available

```

      1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8
      - - - - -
(a)  9 3 J o u r n l A b   1 2           1 2 3 -   1 2 3 p   e
(b)  9 3 J o u r n l A b   1 2           4 5 6     1 2 3 f   E
(c)  9 3 J o u r n l A b   1 2           4 5 6     1 2 3 s   e
(d)  9 3 J o u r n l A b   1 2           1 2 3     1 2 3 t   e
      - - - - -

```

Notes: (a) represents an entry from a periodical distributed primarily in electronic form, which assigns sequential (virtual) page numbers to its contents. The abstract for this entry is available on the *CIS/ED* CD-ROM. (b) represents an entry which is identified by a file number, and which contains a certain number of pages. This approach is sometimes used by journals that publish \TeX , postscript, or dvi files, and which assign to each article a unique number. (c) represents an article which is assigned a number, and whose size is indicated only by a number of bytes. The examples in (c) and (d) correspond to entries for which an abstract is available.

3.6 Administrative records

A small number of records in the database do not report bibliographic information. These are administrative records which credit those who have worked on the database, show copyright notices, and allow for internal editorial record keeping. These records have the same number of fields as other records. The contents of these fields are undefined from the user's perspective, except that column 28 of Field 1 contains the letter Z.

Search software ordinarily should not print such records in response to a query of the database.

- 1-26 (26) Reserved
- 27 (1) Optional symbol (used for uniqueness if needed)
- 28 (1) The letter Z

```

      1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8
      - - - - -
(a)  9 3 C I S : A / Z                                     Z
      - - - - -

```

3.7 Symbols used in Column 27

Column 27 of Field 1 is reserved for special purposes. Table 3.7 contains all current uses of this column. When a non-blank character is encountered in column 27, search software should take appropriate actions.

Table 1: Non-blank codes used in Column 27

| | |
|-------------------|--|
| lower-case letter | Sequentially-assigned to make Field 1 unique |
| upper-case letter | Indicates pages in a supplement or addendum. The following have been used: A, B, E, R, S, T. |
| * | Non-standard coding of remainder of Field 1 [A fuller (or more correct) citation will appear in Field 7] |
| + | Article continued on unspecified pages of a magazine or journal |
| ? | Exact page numbers are in doubt |

3.8 Future types of entries

In the future, the character in column 28 will also be used to distinguish among types of records not now in the database, for example, dissertations (D), technical reports (T), and videos or motion pictures (V). In future years, book reviews and software reviews (currently of type J) will be distinguished from journal articles by a separate code (probably R).

3.9 Sample records and information for programmers

A dialogue with software developers and users has been important to *CIS/ED* since its inception. For example, input has been requested via notices in society bulletins, technical sessions at meetings, and surveys. Files of sample records have been provided on disk and, more recently (1993 and 1994) have been posted on Statlib.

Beginning in 1996, sample records, calls for discussion on changes in format, and production schedules will be posted on the World Wide Web at <http://www.stat.uchicago.edu/~cis>.

We wish to formalize the connection with potential software developers and to foster an even closer and more direct working relationship, including the opportunity to comment in advance on proposed changes in format and contents. In the future, developers who wish to receive early notification of changes to the database format should contact the Editor via electronic mail (cised@galton.uchicago.edu) to register as a developer. Those who do not have access to electronic mail should contact the Editor by mail (*Current Index to Statistics*, Department of Statistics, The University of Chicago, 5734 University Avenue, Chicago, IL 60637 USA).

4 Abbreviation File Formats

This section describes the files contained in the directory **ABBREV**.

Those files whose names start with “**AB**” contain variants of the tables which associate the keys used in Field 1 of the database to particular journal, book, proceedings, and publisher names¹.

In the documentation below, we have used the same publisher example, journal example, and edited book example for each of the files.

4.1 **ABBRMST.V97**

The fullest list (with most complete information about sources) is in the file **ABBRMST.V97**, the master abbreviation file. Each record in this file consists of a “record type” indicator, followed by either four or five fields delimited by the **#** symbol. The record-type indicator (P=publisher, J=Journal, E=Edited book or proceedings, Z=Administrative) is not preceded by the **#** symbol.

For P-type records (Publishers), Field 1 contains the CIS publisher key (16 characters, padded on right with blanks). Field 2 contains the full name of the publisher, and Field 3 contains the location of publisher. This location is typically a city and/or country. This field may contain a cross-reference to another abbreviation denoted by **@L()**; see the subsection entitled “**@-codes**” below. Field 4 contains administrative history codes used by the editors.

Publisher example

```
P#ACM:NY          #Association for Computing Machinery
                  #New York#BT94;94A#
```

For J-type records (Journals), Field 1 contains the *CIS* key (16 characters, padded on right with blanks). These are the codes that are used to identify journal titles in Field 1 of *CIS/ED*. Field 2 contains the journal title, and Field 3 contains the location of publisher. For almost 90% of the entries, the contents of the location field is a cross-reference to another

¹Because publishers frequently change their names, street addresses, and even cities and countries without notice to compilers of indexes, we have eliminated most street addresses for publishers. Any addresses that remain in the abbreviation files should not be trusted; many of these addresses are twenty years out of date.

abbreviation denoted by @L() [See “@- codes” below]. Field 4 contains administrative history codes used by the editors.

Journal example

```
J#ACMMaSfw          #ACM Transactions on Mathematical
                    Software#@L(ACM:NY)#BT94#
```

For E-type records (Proceedings and Edited Books), Field 1 contains the *CIS* key (16 characters, padded on right with blanks). These are the codes that are used in Field 1 of *CIS/ED* to identify book or proceedings titles. Field 2 contains the title. Field 3 contains the editors of the volume, if any, with names in inverted order (Smith, John), separated by the ; symbol. Field 4 contains the location of publisher. This will typically be a cross-reference to another abbreviation denoted by @L() [See “@-codes” below]. Field 5 contains administrative history codes used by the editors.

Edited book example

```
E#ACmplxSrvy       #Analysis of Complex Surveys
                    #Skinner, C. J.;Holt, D.;Smith, T. M. F.
                    #@L(Wiley:NY:UK)#BT94;94a#
```

@-codes

The codes below are used to associate (parsable) information about publishers and publications.

@L() “Published by”

@S() “Distributed by” [This is equivalent to “See @L()”]

@D() “A division of”

@J() Used to link a journal which has changed its name to the variant name; this code is used in conjunction with the terms “Formerly” or “Became”.

[Square brackets]

Comments and notations that may be helpful but are not part of a publisher's name or location may be placed in square brackets ([]). Here are some examples of the use of square brackets:

```
P#SovRadio:Moscw #Soviet Radio Moscow [Sovietskoye Radio]
                  #Moscow#BT94;94U;94A#
J#ZeitWahr       #Zeitschrift f\"ur Wahrscheinlichkeitstheorie
                  und Verwandte Gebiete [Became: @J(ProbTher)]
                  #@L(Springer:Brln:NY)#BT94;94U#
E#StudHSilVn    #Studies in Honour of Silvio Vianelli
                  [Facolta di Economia e Commercio]
                  #Istituto di Statistica (Palermo)
                  #BT94;94U;94A#
```

4.2 ABBR.V97

Each element of ABBR.V97 consists of a three- or four-field record delimited by # symbols. Within each field, the bracketed material from ABBRMST.V97 has been deleted, and all remaining cross-references (@-codes) have been expanded.

For Publishers:

1. 16-character *CIS* key
2. Full name of publisher
(with "Publishing" "Co" "Inc" "Ltd", etc. deleted)
3. City and/or country of publisher

Example:

```
#ACM:NY          #Association for Computing Machinery
                  #New York#
```

For Journals:

1. 16-character *CIS* key
2. Full name of publication
3. Publisher and address

Example:

```
#ACMMaSfw          #ACM Transactions on Mathematical Software
                   #Association for Computing Machinery
                   (New York)#
```

For Edited Books and Proceedings:

1. 16-character *CIS* key
2. Full name of publication
3. Editors, in inverted order
4. Publisher with city and/or country enclosed in parentheses

Example:

```
#ACmplxSrvy        #Analysis of Complex Surveys
                   #Skinner, C. J.;Holt, D.;Smith, T. M. F.
                   #Wiley (New York)#
```

4.3 ABB.V97

This file is identical to `ABBR.V97`, except that Field 2 and the last field contain abbreviated versions of the journal and publisher, and all publisher information is omitted for journals. This format may be the most convenient one to use for automatic bibliography generation.

Examples:

```
#ACM:NY            #Assoc. Comput. Mach.#New York#
#ACMMaSfw          #ACM Trans. Math. Software##
#ACmplxSrvy        #Anal. Complex Surveys
                   #Skinner, C. J.;Holt, D.;Smith, T. M. F.
                   #Wiley (New York)#
```

4.4 AB.V97

This file is a human-readable file containing the same information as `ABB.V97`, but in a form much less amenable to automated parsing than the. For edited books and proceedings, the editor field is separated from the title field by the three-character separator `;` (that is, semicolon-blank-underscore). Editors names are listed in inverted order and are separated by semicolons. The last editor is followed by the notation (Ed), provided that there is at least one editor. The editor field is separated from the location field by the `;` separator. Thus, the editor field is

empty when there are no editors listed, as in the last example below. This format was chosen to maintain maximal compatibility with the format used in 1993.

Examples:

| | |
|------------|--|
| ACM:NY | Assoc. Comput. Mach. (New York) |
| ACMMaSfw | ACM Trans. Math. Software |
| ACmplxSrvy | Anal. Complex Surveys; _Skinner, C. J.;Holt, D.;Smith,T. M. F. (Ed); _Wiley (New York) |
| AlgorCmplx | Algorithms \& Complexity: New Dir. \& Recent Results; _; _Academic (New York; London) |

4.5 ABAMS.V97

This file contains one record for each publication or publisher that is represented *both* in the *CIS/ED* and in the master list of *Mathematical Reviews*, published by the American Mathematical Society. Each record contains two fields (surrounded by # symbols). The first field contains the *CIS* key, without trailing blanks. The second field contains the shortened name corresponding to the key as used in *Math Reviews*. The records in this file are sorted on the *CIS* key in a case-insensitive manner.

Examples:

```
#ACMMaSfw#ACM Trans. Math. Software#
#ATTTJ#AT\&T Tech. J.#
#EcomtrRv#Econometric Rev.#
```

4.6 ABBRAMS.V97

This file is a variant of *ABAMS.V97* and is in the same format as that file. Field 2 of this file contains the full *long* name corresponding to the *CIS* key, according to *Math Reviews*. The records in this file are also sorted on the *CIS* key in a case-insensitive manner.

Examples:

```
#ACMMaSfw#Association for Computing Machinery. Transactions
on Mathematical Software#
#ATTTJ#AT\&T Technical Journal#
#EcomtrRv#Econometric Reviews#
```

5 Notes on Fields 5, 6 and 7

This chapter contains technical details on the preparation of Fields 5 and 6, and brief comments on Field 7. These fields were added to *CIS/ED* to make computer searches easier and more productive. However, software that prints out the results of searches will ordinarily omit Fields 5, 6, and 7, since they do not contribute additional bibliographic information.

In the preparation of Fields 5 and 6, the guiding principle has been the overall usefulness for searches, rather than correctness in any linguistic, orthographic, or other sense. (There is ample evidence that some copy editors of journals are not linguists; some users (and editors!) of *CIS/ED* may not be linguists either.)

5.1 Accent marks

Here we use the word *accent* to refer to any mark that modifies a letter or letters of the “English” alphabet to spell non-English proper nouns. In most cases Fields 5 and 6 simply repeat accented names with all accent codes removed. Such instances are referred to as having been “STRIPPED.” Here are some examples:

| | | | |
|-----------------------------|--------------|--------|-------------|
| <code>Cram\’er</code> | (Cramér) | yields | Cramer |
| <code>\c{C}inlar</code> | (Çinlar) | yields | Cinlar |
| <code>Pr{\ae}stgaard</code> | (Præstgaard) | yields | Praestgaard |
| <code>Stanis{\l}aw</code> | (Stanisław) | yields | Stanislaw |

There are three systematic exceptions:

- (a) **Umlaut/dieresis.** Prior to 1988, umlauted German vowels would usually be rendered with the letter *e* following the vowel (whether or not the word was capitalized). Thus, *für* would be denoted as *fuer*. French, Spanish, etc. diereses are typographically indistinguishable from umlauts and so they often received the same treatment, even though this makes no linguistic sense. To conform with this convention, Field 5 or 6 always gives two alternative spellings when umlauts/diereses are removed — one with the extra *e* and one without. The purpose is to provide a uniform means for searching, and no regard is given to the correctness of these conventions from a linguistic standpoint.
- (b) Scandinavian *Ø, ø*. Two versions, one with *o* and one with *oe*, are provided.

- (c) Scandinavian $\text{Å}, \text{å}$. Two versions, one with **a** and one with **aa**, are provided.

Here are some examples:

| | | | | | |
|--------------------------|------------|--------|-----------|-----|----------|
| <code>K\unsch</code> | (Künsch) | yields | Kuensch | and | Kunsch |
| <code>Cs\org\H{o}</code> | (Csörgő) | yields | Csoergo | and | Csorgo |
| <code>No\el</code> | (Noël) | yields | Noel | and | Noel |
| <code>S{o}rensen</code> | (Sørensen) | yields | Soerensen | and | Sorensen |
| <code>{AA}ke</code> | (Åke) | yields | Aake | and | Ake |

When there is more than one of these cases per name, minimal and maximal versions appear together, but are not mixed. For example, both **Aake Soerensen** and **Ake Sorensen**, but neither *Aake Sorensen* nor *Ake Soerensen* would appear in Field 5.

5.2 Transliterations

Russian names such as Chebyshev, Gikhman, Markov, Prokhorov, and Skorokhod have a standardized entry in Field 5 or 6 if a different spelling is used in Fields 2, 3, or 4: A full list is given in Chapter 6.

Certain names from non-Roman alphabets (mainly Russian, but also Hebrew, Arabic, etc.) are sometimes transliterated into the Roman alphabet using apostrophes (to represent a vast variety of soft signs, glottal stops, breathing sounds, etc.). Unfortunately, journals and translators have widely differing practices in the use of such apostrophes (*Dynkin* and *Dyn'kin*, *Silvestrov* and *Sil'vestrov*, etc.), making searches difficult. Wherever a transliterated name with an apostrophe has been found (several hundred records), it has been repeated in Field 5 or 6 without the apostrophe.

5.3 U.S. spelling

A systematic search was made on the sequence *-our*. Instances in which the “American” spelling has *-or* yielded entries in Field 6 when the word in question might reasonably be used as a search word. *Behaviour* is seldom a sensible search word (e.g., *asymptotic behaviour* would be found by looking for *asymptotic*) and so has not usually been respelled. However, *behavioural science*, *nearest neighbour*, *Prohorov neighbourhood*, *labour market*, *tumour growth*, *coloured noise*, etc. have been respelled. On an

unsystematic basis, other words, when found, have been respelled in the U.S. fashion in Field 6, for example, *anesthesia*, *fetus*, etc.

5.4 Field 7 administrative information

For the curious, here are examples of some of the types of codes that may have been inserted into Field 7 to show the source of a record or later modifications to it. **PV18** indicates that the record appeared in printed Volume 18 of the printed version of *Current Index to Statistics*. For the first five printed volumes, it is not feasible to show the exact volume number: the designation **JMBT** is used (for B. Joiner, R. Milton, R. Burdick, and B. Trumbo, each of whom have had something to do with these records as they currently appear.) **CILM** designates pre-1980 records that appeared in the *Cumulative Index to Linear Models*, **IMSC** refers to records prepared or verified for the IMS index, and **JASI** to records prepared or verified for the recently published JASA Index. A notation such as **BB94** denotes a record that was changed or added to *CIS/ED* in this edition by someone with initials **BB**.

Records for the period 1975–1979 that carry the code **JMBT** have been retrieved from deteriorated media and may contain errors; the presence of another code for these records (other than **CILM**) indicates subsequent comparison with original sources, and hence greater reliability. Of all the codes used, only **CILM** carries any indication of subject matter.

6 Names Transliterated from Cyrillic

Several journals indexed by the *Current Index to Statistics* are published in Slavic languages, and many authors are residents of Slavic countries, that use Cyrillic rather than Roman characters for writing the language. Since 1983, the American Mathematical Society has been using a single method for transliterating (romanizing) Cyrillic characters; at least five other methods are in common use, however.

Consequently, depending on the source of the information and the individual or organization translating words and transliterating names, a single author may appear in different places with quite different romanized versions of his or her name. This fact makes searching for authors such as R. Z. Khas'minskiĭ difficult, since this name (ХАСЬМИНСКИЙ) can be — and is! — transliterated in many different ways, many of which are not adjacent in Roman alphabetical order.

We have instituted a systematic search for names which could appear in more than one transliteration, and have found that the *CIS/ED* contains each of the following names in more than one variant transliteration. In each such case, the transliteration given in the table below, using the current American Mathematical Society system, is contained in Field 5 (after removing any accents, apostrophes, and other TEX codes), regardless of the transliteration used in Field 3.

| | | | |
|-----------------|---------------|---------------|--------------|
| Aĭvazyan | АЙВАЗЯН | Belyaev | БЕЛЯЕВ |
| Alekseev | АЛЕКСЕЕВ | Bikyalis | БИКЯЛИС |
| Aleshkyavichene | АЛЕШКЯВИЧЕНЕ | Bloznyalis | БЛОЗНЯЛИС |
| Alyushina | АЛЮШИНА | Bogdanovich | БОГДАНОВИЧ |
| Ambartsumyan | АМБАРЦУМЯН | Vokuchava | БОКУЧАВА |
| Anan'evskiiĭ | АНАНЬЕВСКИЙ | Bol'shev | БОЛЬШЕВ |
| Antonets | АНТОНЕЦ | Borovskikh | БОРОВСКИХ |
| Antyufeev | АНТЮФЕЕВ | Bratiĭchuk | БРАТИЙЧУК |
| Aref'eva | АРЕФЬЕВА | Brodskiiĭ | БРОДСКИЙ |
| Arsenishvili | АРСЕНИШВИЛИ | Brusilovskiiĭ | БРУСИЛОВСКИЙ |
| Artamonovskiiĭ | АРТАМОНОВСКИЙ | Bukharaev | БУХАРАЕВ |
| Asriev | АСРИЕВ | Bulinskiiĭ | БУЛИНСКИЙ |
| Ayupov | АЮПОВ | Burnashev | БУРНАШЕВ |
| Badalbaev | БАДАЛБАЕВ | Butsan | БУЦАН |
| Bagdonavichus | БАГДОНАВИЧУС | Vyambazhav | БЯМБАЖАВ |
| Belopol'skaya | БЕЛОПОЛЬСКАЯ | Chebotarev | ЧЕБОТАРЕВ |

| | | | |
|------------------|---------------|---------------|--------------|
| Chentsov | ЧЕНЦОВ | Ginovyan | ГИНОВЯН |
| Cherenkov | ЧЕРЕНКОВ | Godovanchuk | ГОДОВАНЧУК |
| Cherkashin | ЧЕРКАШИН | Golinskiĭ | ГОЛИНСКИЙ |
| Cherkasov | ЧЕРКАСОВ | Gorodetskiĭ | ГОРОДЕЦКИЙ |
| Chistyakov | ЧИСТЯКОВ | Granovskiĭ | ГРАНОВСКИЙ |
| Chitashvili | ЧИТАШВИЛИ | Grintsevichus | ГРИНЦЕВИЧУС |
| Chobanyan | ЧОБАНЯН | Grinyuvene | ГРИНЮВЕНЕ |
| Chuprunov | ЧУПРУНОВ | Gurevich | ГУРЕВИЧ |
| Daletskiĭ | ДАЛЕЦКИЙ | Gurskiĭ | ГУРСКИЙ |
| Darkhovskiĭ | ДАРХОВСКИЙ | Gushchin | ГУЩИН |
| Dashevskiĭ | ДАШЕВСКИЙ | Gvantseladze | ГВАНЦЕЛАДЗЕ |
| Davidovich | ДАВИДОВИЧ | Ibramkhalilov | ИБРАМХАЛИЛОВ |
| Davydov | ДАВЫДОВ | Ignatyuk | ИГНАТЮК |
| Demurishvili | ДЕМУРИШВИЛИ | P'inskiĭ | ИЛЬИНСКИЙ |
| Dem'yanov | ДЕМЬЯНОВ | Inzhevito | ИНЖЕВИТОВ |
| Didkovskiĭ | ДИДКОВСКИЙ | Ishchenko | ИЩЕНКО |
| Diesperova | ДИЕСПЕРОВА | Ivinitiĭ | ИВИНИЦКИЙ |
| Dochviri | ДОЧВИРИ | Ivkovich | ИВКОВИЧ |
| Donchenko | ДОНЧЕНКО | Kaĭmanovich | КАЙМАНОВИЧ |
| Dorogovtsev | ДОРОГОВЦЕВ | Kabalevskiĭ | КАБАЛЕВСКИЙ |
| Drozhhzina | ДРОЖЖИНА | Kakosyan | КАКОСЯН |
| Dunaev | ДУНАЕВ | Kalenskiĭ | КАЛЕНСКИЙ |
| Dzhafarov | ДЖАФАРОВ | Kanevskiĭ | КАНЕВСКИЙ |
| Dzhamirzaev | ДЖАМИРЗАЕВ | Kaniovskiĭ | КАНИОВСКИЙ |
| Dzhvarsheishvili | ДЖВАРШЕЙШВИЛИ | Kartashev | КАРТАШЕВ |
| Eidukyavichyus | ЕИДУКЯВИЧЮС | Kartashov | КАРТАШОВ |
| Ekushov | ЕКУШОВ | Khadzhiev | ХАДЖИЕВ |
| Enyukov | ЕНУКОВ | Khakhubiya | ХАХУБИЯ |
| Ershov | ЕРШОВ | Khalidov | ХАЛИДОВ |
| Evstigneev | ЕВСТИГНЕЕВ | Khalikulov | ХАЛИКУЛОВ |
| Ezhov | ЕЖОВ | Kharin | ХАРИН |
| Gabovich | ГАБОВИЧ | Kharlamov | ХАРЛАМОВ |
| Gadzhiev | ГАДЖИЕВ | Khashimov | ХАШИМОВ |
| Gaivoronskiĭ | ГАИВОРОНСКИЙ | Khas'minskiĭ | ХАСЬМИНСКИЙ |
| Galstyan | ГАЛСТЯН | Khmaladze | ХМАЛАДЗЕ |
| Gal'chuk | ГАЛЬЧУК | Khokhlov | ХОХЛОВ |
| Ganikhodzhaev | ГАНИХОДЖАЕВ | Kholevo | ХОЛЕВО |
| Gaposhkin | ГАПОШКИН | Khusanbaev | ХУСАНБАЕВ |
| Gikhman | ГИХМАН | Kitaev | КИТАЕВ |

| | | | |
|----------------|--------------|-----------------|---------------|
| Kolchinskiĭ | КОЛЧИНСКИЙ | Malinovskiĭ | МАЛИНОВСКИЙ |
| Kolmanovskiĭ | КОЛМАНОВСКИЙ | Malyarenko | МАЛЯРЕНКО |
| Kolodii | КОЛОДИЙ | Malyshev | МАЛЬШЕВ |
| Kolomiets | КОЛОМИЕЦ | Malyukyavichyus | МАЛЮКЯВИЧЮС |
| Kordonskiĭ | КОРДОНСКИЙ | Malyutov | МАЛЮТОВ |
| Korenevskiĭ | КОРЕНЕВСКИЙ | Maniya | МАНИЯ |
| Korolyuk | КОРОЛЮК | Manstavichyus | МАНСТАВИЧЮС |
| Korshunov | КОРШУНОВ | Marushin | МАРУШИН |
| Kozachenko | КОЗАЧЕНКО | Matskyavichyus | МАЦКЯВИЧЮС |
| Kozarovitskiĭ | КОЗАРОВИЦКИЙ | Matveev | МАТВЕЕВ |
| Krasnitskiĭ | КРАСНИЦКИЙ | Men'shikov | МЕНЬШИКОВ |
| Kravtsov | КРАВЦОВ | Mikhaïlovskaya | МИХАЙЛОВСКАЯ |
| Krizhyus | КРИЖЮС | Mikhaïlov | МИХАЙЛОВ |
| Kuchkarov | КУЧКАРОВ | Mikulyavichus | МИКУЛЯВИЧУС |
| Kudlaev | КУДЛАЕВ | Mil'shtein | МИЛЬШТЕЙН |
| Kudzhma | КУДЖМА | Mirakhmedov | МИРАХМЕДОВ |
| Kulinich | КУЛИНИЧ | Miroshin | МИРОШИН |
| Kul'chitskiĭ | КУЛЬЧИЦКИЙ | Mirzakhmedov | МИРЗАХМЕДОВ |
| Kurchenko | КУРЧЕНКО | Misheikis | МИШЕИКИС |
| Kuritsyn | КУРИЦЫН | Mitropol'skiĭ | МИТРОПОЛЬСКИЙ |
| Kurzanskiĭ | КУРЖАНСКИЙ | Mkrтчyan | МКРТЧЯН |
| Kutoyants | КУТОЯНЦ | Mnatsakanov | МНАЦАКАНОВ |
| Kuznetsova | КУЗНЕЦОВА | Mogul'skiĭ | МОГУЛЬСКИЙ |
| Kuznetsov | КУЗНЕЦОВ | Moklyachuk | МОКЛЯЧУК |
| Kvaratskheliya | КВАРАЦХЕЛИЯ | Molchanov | МОЛЧАНОВ |
| Labkovskiĭ | ЛАБКОВСКИЙ | Mosyagin | МОСЯГИН |
| Laurinchikas | ЛАУРИНЧИКАС | Mukhin | МУХИН |
| Lavrent'ev | ЛАВРЕНТЬЕВ | Mukhomor | МУХОМОР |
| Lepskiĭ | ЛЕПСКИЙ | Mutafchiev | МУТАФЧИЕВ |
| Levitskaya | ЛЕВИЦКАЯ | Nadaraya | НАДАРАЯ |
| Lifshits | ЛИФШИЦ | Nagaev | НАГАЕВ |
| Linetskiĭ | ЛИНЕЦКИЙ | Nakhapetyan | НАХАПЕТЯН |
| Liptser | ЛИПЦЕР | Nakonechnyi | НАКОНЕЧНЫЙ |
| Livshits | ЛИВШИЦ | Nemirovskiĭ | НЕМИРОВСКИЙ |
| Lumel'skiĭ | ЛУМЕЛЬСКИЙ | Nikolaev | НИКОЛАЕВ |
| Lyubich | ЛЮБИЧ | Nikol'skiĭ | НИКОЛЬСКИЙ |
| Lyubinskas | ЛЮБИНСКАС | Norvaisha | НОРВАЙША |
| Makhno | МАХНО | Nosovskiĭ | НОСОВСКИЙ |
| Malevich | МАЛЕВИЧ | Oseledets | ОСЕЛЕДЕЦ |

| | | | |
|------------------|---------------|----------------|--------------|
| Ostrovskii | ОСТРОВСКИЙ | Shervashidze | ШЕРВАШИДЗЕ |
| Pachkauskas | ПАЧКАУСКАС | Shevlyakov | ШЕВЛЯКОВ |
| Pecherskii | ПЕЧЕРСКИЙ | Shirokov | ШИРОКОВ |
| Pechinkin | ПЕЧИНКИН | Shiryaev | ШИРЯЕВ |
| Pergamenshchikov | ПЕРГАМЕНЩИКОВ | Shkol'nik | ШКОЛЬНИК |
| Petrosyan | ПЕТРОСЯН | Shtar'kov | ШТАРЬКОВ |
| Piranashvili | ПИРАНАШВИЛИ | Shtatland | ШТАТЛАНД |
| Pirdzhanov | ПИРДЖАНОВ | Shurenkov | ШУРЕНКОВ |
| Pisanets | ПИСАНЕЦ | Sidoravichyus | СИДОРАВИЧЮС |
| Piunovskii | ПИУНОВСКИЙ | Siluyanova | СИЛУЯНОВА |
| Polishchuk | ПОЛИЩУК | Sirazhdinov | СИРАЖДИНОВ |
| Poznyak | ПОЗНЯК | Skorokhod | СКОРОХОД |
| Prokhorov | ПРОХОРОВ | Skvortsov | СКВОРЦОВ |
| Pugachev | ПУГАЧЕВ | Slivnyak | СЛИВНЯК |
| Pukhal'skii | ПУХАЛЬСКИЙ | Slobodenyuk | СЛОБОДЕНЮК |
| Purtukhiya | ПУРТУХИЯ | Smyshlyaeva | СМЫШЛЯЕВА |
| Rachkauskas | РАЧКАУСКАС | Solov'ev | СОЛОВЬЕВ |
| Radavichyus | РАДАВИЧЮС | Startsev | СТАРЦЕВ |
| Rakhimov | РАХИМОВ | Statulyavichus | СТАТУЛЯВИЧУС |
| Romanovskii | РОМАНОВСКИЙ | Stepakhno | СТЕПАХНО |
| Rozovskii | РОЗОВСКИЙ | Sukhov | СУХОВ |
| Rubshtein | РУБШТЕЙН | Sveshnikov | СВЕШНИКОВ |
| Rukhin | РУХИН | Sytaya | СЫТАЯ |
| Ryzhov | РЫЖОВ | Tarashchanskii | ТАРАЩАНСКИЙ |
| Sachkov | САЧКОВ | Tartakovskii | ТАРТАКОВСКИЙ |
| Sadovskii | САДОВСКИЙ | Tikhomirov | ТИХОМИРОВ |
| Safaryan | САФАРЯН | Timofeev | ТИМОФЕЕВ |
| Sakhanenko | САХАНЕНКО | Tkachuk | ТКАЧУК |
| Sakhobov | САХОБОВ | Tomashevskii | ТОМАШЕВСКИЙ |
| Sakovich | САКОВИЧ | Topchi | ТОПЧИЙ |
| Sarafyan | САРАФЯН | Torondzhadze | ТОРОНДЖАДЗЕ |
| Saulis | САУЛИС | Tret'yakov | ТРЕТЬЯКОВ |
| Semovskii | СЕМОВСКИЙ | Troitskii | ТРОИЦКИЙ |
| Serdobol'skii | СЕРДОБОЛЬСКИЙ | Tsirel'son | ЦИРЕЛЬСОН |
| Sevast'yanov | СЕВАСТЬЯНОВ | Tsitritskii | ЦИТРИЦКИЙ |
| Shaikhet | ШАИХЕТ | Tsvetkov | ЦВЕТКОВ |
| Shakhbazov | ШАХБАЗОВ | Tyurin | ТЮРИН |
| Shatashvili | ШАТАШВИЛИ | Ulanovskii | УЛАНОВСКИЙ |
| Shergin | ШЕРГИН | Ul'yanov | УЛЬЯНОВ |

| | | | |
|------------------|--------------|---------------|-------------|
| Ushakov | УШАКОВ | Yudin | ЮДИН |
| Vakhaniya | ВАХАНИЯ | Yurachkovskii | ЮРАЧКОВСКИЙ |
| Vaninskiĭ | ВАНИНСКИЙ | Yurinskiĭ | ЮРИНСКИЙ |
| Vasil'ev | ВАСИЛЬЕВ | Yushkevich | ЮШКЕВИЧ |
| Venttsel' | ВЕНТЦЕЛЬ | Zaigraev | ЗАИГРАЕВ |
| Vershik | ВЕРШИК | Zaitsev | ЗАИЦЕВ |
| Vil'chevskii | ВИЛЬЧЕВСКИЙ | Zakharin | ЗАХАРИН |
| Vishnevskii | ВИШНЕВСКИЙ | Zakharov | ЗАХАРОВ |
| Vladimirskii | ВЛАДИМИРСКИЙ | Zhiglyavskii | ЖИГЛЯВСКИЙ |
| Volkovinskiĭ | ВОЛКОВИНСКИЙ | Zhilinskas | ЖИЛИНСКАС |
| Vorobeychikov | ВОРОБЕЫЧИКОВ | Zhurbenko | ЖУРБЕНКО |
| Vysochanskiĭ | ВЫСОЧАНСКИЙ | Zinchenko | ЗИНЧЕНКО |
| V'yugin | ВЬЮГИН | Zolotukhina | ЗОЛОТУХИНА |
| Yadrenko | ЯДРЕНКО | Zuev | ЗУЕВ |
| Yakovleva | ЯКОВЛЕВА | Zyablov | ЗЯБЛОВ |
| Yanushkyavichene | ЯНУШКЯВИЧЕНЕ | | |
| Yanushkyavichyus | ЯНУШКЯВИЧЮС | | |

7 Abstracts and Book Reviews

CIS/ED has abstracts for all articles published in the following journals during the years indicated. (Years in italics denote abstracts from the first issue of the journal.)

- *The American Statistician*, 1996–
- *Journal of Agricultural, Biological, and Environmental Statistics*, 1996–
- *Journal of the American Statistical Association*, 1996–
- *Journal of Business and Economic Statistics*, 1996–
- *Journal of Computational and Graphical Statistics*, 1992–
- *Journal of Statistics Education*, 1993–
- *Technometrics*, 1996–

We hope to be able to add abstracts from other publications as they become available. The presence of an abstract in the *CIS/ED* database is signalled by a lower-case letter in column 28 of Field 1 of an article's main entry.

CIS/ED also has the text for book reviews published in the following journals during the years indicated.

- *The American Statistician*, 1996–
- *Journal of the American Statistical Association*, 1996–
- *Technometrics*, 1996–

The presence of the text of a book review in the *CIS/ED* database is signalled by a lower-case letter in column 28 of Field 1 of the review's main entry.

This document outlines the format for the abstract/book review files and related data structures. As we gain more experience with abstracts, book reviews, and related material, and with integrating them into the database, we expect to refine the data structures still further.

7.1 Organization of Abstract/Book Review Files

Abstracts and reviews supplied with the database are linked to individual *CIS/ED* records through a collection of three types of files:

1. A set of one or more *data files* which contain the actual abstract or review, together with related information. These files are in ASCII format, typically with \TeX typesetting codes. They contain the actual copies of the abstracts. In 1997, there is one file, named `ABSDAT.A97`.
2. A set of one or more *index files*. Each record in an index file contains a field that identifies an article entry in the database, another field that identifies the data file in which the abstract or review resides, using an abbreviated file code, and a third field that identifies the location within the data file where the abstract or review begins. Other fields in the index record provide additional information that may be of use to writers of search software, but which are not necessary for locating an abstract or review. Index files are in ASCII format for 1997, but may be encoded in another format in the future. In 1997, there is one index file, named `ABSNDX.A97`.
3. A single *catalog file*. This file contains two kinds of information. First, it contains a list of the index files to be searched. Second, it contains a list of links from the abbreviated file code used by the index files to the physical file names where the data files are located. This file can be edited to make configuration of search software easier and to provide a degree of flexibility in configuring a search system. The catalog file will be in ASCII format. The catalog file is named `ABSCAT.A97` for 1997.

The internal format of each file is described in the sections below.

7.2 The Data Files

The data files consist of abstracts and reviews, one after another in free format. Each abstract consists of the following elements:

- Exactly one Line B (Begin abstract)
- Zero or more Line D (Define environment)
- Zero or more Line T (Title)
- Zero or more Line A (Author)
- Zero or more Line W (Where the author is located)
- Exactly one Line X (Indicating start of abstract teXt)
- One or more lines with the text of the abstract (Data Lines)

- Exactly one Line Y (Indicating end of abstract text)
- Zero or more Line K (Key Words)
- Exactly one Line E (End of abstract)

Text of book reviews consist of similar elements:

- Exactly one Line C (Begin book review)
- Zero or more Line D (Define environment)
- Zero or more Line T (Title or work reviewed)
- Zero or more Line A (Author)
- Zero or more Line I (Publisher information)
- Zero or more Line W (Where the author/reviewer is located)
- Exactly one Line X (Indicating start of review teXt)
- One or more lines with the text of the review (Data Lines)
- Exactly one Line Y (Indicating end of review text)

If the review contains references, then the following lines are included.

- Exactly one Line F (Begin book review references)
- One or more lines containing the reference list
- Exactly one Line G (End book review references)

The review concludes with

- Zero or more line R (Reviewer)
- Exactly one Line E (End of abstract)

The contents of the various types of lines are described in detail below.

7.2.1 Lines B, C

Each abstract begins with a line containing the characters “\B” in columns 1 and 2; abstracts begin with “\C.” These characters are always followed by a comment (possibly empty) enclosed by braces. If the abstract/review corresponds to an entry in the *CIS/ED* (which is almost always the case), then the first 28 characters of the comment will be identical to the contents of Field 1 of the *CIS/ED* entry. This establishes a link between the abstract and a specific article citation in the database.

7.2.2 Line T

These lines contain the title of the article being abstracted. Such lines begin with the characters “\T{”, continue on one or more lines with the text of the article’s title, and are concluded by the single character “}”.

7.2.3 Line A

These lines contain the author or authors of the article being abstracted or reviewed. Such lines begin with the characters “\A{”, continue on one or more lines with the names of the article’s author(s), and are concluded by the single character “}”.

7.2.4 Line R

These lines contain the author(s) *of the review*. Such lines begin with the characters “\R{”, continue on one or more lines with the names of the review’s author(s), and are concluded by the single character “}”.

7.2.5 Line W

These optional lines contain the address or affiliation of the author(s) and/or reviewer(s) of the article. Such lines begin with the characters “\W{”, continue on one or more lines with the names of the author(s)’ address(es), and are concluded by the single character “}”.

7.2.6 Lines X, Y, and Data lines

The data lines contain the text of the abstract or the review itself. These lines are preceded by a line containing only “\X”, and followed by a single line containing only “\Y”. While the contents of Lines B, T, A, and K are typically adapted by the editors of the *CIS/ED*, the data lines are typically **not** edited, but are reproduced as supplied by the original producer of the computer-readable abstract. Such abstracts are often recorded using \TeX or \LaTeX to encode mathematical symbols and constructions such as accents, superscripts, and the like. However there is no standard that all journals have adopted, and as a consequence, abstracts from different sources may not be coded consistently.

In those cases wherein the abstracts as supplied cause serious problems with \TeX or \LaTeX , they have been edited by the editors of *CIS/ED* to the smallest extent possible to make processing by \TeX or \LaTeX possible.

This often involves (a) the introduction of $\text{T}_{\text{E}}\text{X}$ control sequences and/or (b) the definition of $\text{T}_{\text{E}}\text{X}$ control sequences used in the abstract. The latter is accomplished through defining D Lines as described below.

7.2.7 Line D

These lines contain necessary definition or redefinition of $\text{T}_{\text{E}}\text{X}$ or $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ control sequences. Such lines begin with the characters “ $\backslash\text{D}\{$ ”, continue on one or more lines with an appropriate $\text{T}_{\text{E}}\text{X}$ $\backslash\text{def}$ command, and are concluded by the single character “ $\}$ ”.

7.2.8 Line K

These lines contain any key words of the article being abstracted that were directly supplied by the provider of the abstract. Such lines begin with the characters “ $\backslash\text{K}\{$ ”, continue on one or more lines with the article’s key words, and are concluded by the single character “ $\}$ ”.

7.2.9 Line F

A line that begins with “ $\backslash\text{F}$ ” denotes the beginning of the reference list for a book review.

7.2.10 Line G

A line that begins with “ $\backslash\text{G}$ ” closes the reference list for a book review.

7.2.11 Line I

Lines with “ $\backslash\text{I}$ ” contain publisher information, enclosed in braces($\{\}$).

7.2.12 Line E

These lines mark the end of the abstract. Such lines contain only the characters “ $\backslash\text{E}$ ”.

7.2.13 Line Z

Information within the braces following “ $\backslash\text{Z}$ ” is administrative information that is not part of the review or abstract, and should be ignored.

7.2.14 Limitations on data files

1. Data files must not to exceed 9,999,999 bytes in length.
2. Individual abstracts in data files are not to exceed 99,999 bytes in length.
3. No line in an abstract is permitted to contain user-defined T_EX control sequences that consist of a single upper-case roman character. These are used within the abstract file; other codes may be used in the future. This restriction does not apply to control sequences *pre-defined* in T_EX: \H, \L, \O, \P, and \S (which are used to produce ö, L, Ø, ¶, and §, respectively).
4. No T_EX or L^AT_EX control sequence (other than those defined in plain T_EX or standard L^AT_EX) may be used without first being preceded by a D Line within which the control sequence is defined.

7.2.15 An Example of an Abstract

```
\B{92JCmpGrSt 1 197- 211 j}
\T{Ridge Finding From Noisy Data}
\A{Peter Hall, Wei Qian, and D.M. Titterington}
\X
Methods are presented for detecting ridges and/or antiridges
using noisy data. Several alternative criteria are proposed
for identifying points on a ridge and procedures for
following a ridge line are discussed. The methods are
illustrated by examples.
\Y
\K{Image processing; Ridge finding; Smoothing.}
\E
```

7.2.16 An Example of a Book Review

```
\C{96JASA      91      434- 435 j}
\Z{ID433-50-BR}
\T{Measurement, Regression, and Calibration}
\A{Philip J. \lastname{Brown}}
\I{New York: Oxford University Press, 1993. ix $$ 201 pp.
  \ $45.}
```

```
\X
```

As its title suggests, this book is a compendium of ideas from the areas of regression and calibration; I am not quite sure why the word ‘measurement’ appears in the title. The book does a good job of bringing together many of the latest developments in these areas, in which the author is a recognized expert.

```
\dots
```

In this age of information, calibration is increasingly important, but this topic often has not been given enough attention in regression texts. *{it Measurement, Regression, and Calibration}* is a welcome attempt to bring together state-of-the-art developments in this area and should be a valuable addition to the library of anyone interested in calibration.

```
\Y
```

```
\F
```

Denham, M. C., and Brown, P. J. (1993), ‘Calibration in Many Variables,’ *{it Applied Statistics}*, 42, 515--528.

Oman, S. D. (1991), ‘Random Calibration With Many Measurements: An Application of Stein Estimation,’ *{it Technometrics}*, 33(2), 187--195.

```
\G
```

```
\R{Esteban \lastname{Walker}}
```

```
\E
```

7.3 The Index Files

Each record in an index file is in fixed format:

- 1-28 (28) Field 1. The contents of this field must be identical to the corresponding field for the article in the *CIS/ED* database, which is used as a unique article identifier for those entries that have an abstract or review.
- 29 (1) Space
- 30 (1) Data file code. This is a single ASCII character, typically a letter or a digit. This character is case-sensitive. It represents the data file in which the abstract is to be found, according to the lookup table in the Catalog File, as specified in section 7.4.4 below.
- 31-39 (9) Offset from the beginning of the data file, in bytes.
- 40-45 (6) Length of the entry, in bytes.
- 46 (1) Format code, possibly blank.
Codes in current use are:
 - P Plain \TeX
 - L \LaTeX
 - T Both plain \TeX and \LaTeX
 - W Microsoft RTF format
 - A ASCII (coding incompatible with \TeX)

7.4 The Catalog File

The catalog file resides in a file named `abstract.cat` by default. It consists of records in ASCII format, terminated by a newline character.

7.4.1 Record 1: Number of index files (I)

This record contains a positive integer (I) indicating the number of index files to be used. Note: search software must take account of the possibility that there is more than one index file ($I > 1$).

7.4.2 Records 2 through $I+1$: Index file names

Each record contains the full pathname of an index file. Any search software will search each file listed, in the order listed, when searching for abstracts.

7.4.3 Record $I + 2$: Number of data files (F)

This record contains a single positive integer (F) indicating the number of data files that are referred to by the index files.

7.4.4 Record $I + 3$ through $I + F + 3$: Data codes and file names

Each record consists of a single-character ASCII code, followed by a single space, followed by the full pathname of the data file to which that code is assigned. The ASCII code character in column 1 of this record corresponds to the codes used in column 30 of index file records.

7.5 Printing Abstracts and Book Reviews

The codes embedded in the data file can be used for printing by defining suitable \TeX control sequences. For our editorial use, we proofread the abstracts on the database by typesetting them using the following \LaTeX code, which places each abstract on a separate page in a structured way:

```
%  
% Driver.tex (LaTeX 2.09)  
%  
\documentstyle[cisabs]{article}  
\begin{document}  
\input abstract  
\end{document}
```

For those who use \TeX instead of \LaTeX , the following codes will produce a similar result, placing each abstract on a separate page, but with a different choice of margins and typestyle than \LaTeX uses:

```
%  
% Plain-TeX variant of driver  
%  
\input cisabs.sty  
\input abstract  
\bye
```

The \LaTeX style file `cisabs.sty` referred to in the drivers above, is included in the CD-ROM distribution of *CIS/ED*, in the ABSTRACT directory. This file can also be input to plain \TeX , as it uses no constructions unique to \LaTeX .