



简要说明

EpiData

EpiData 是一个用于数据输入及相关信息的计算机程序。

当你已经将数据收集在纸上，并且要对数据进行分析或制表时，需要使用 EpiData 这个软件，利用它可以产生简单的频数和变量清单，但 EpiData 软件主要是用于数据输入。

在数据录入过程中，可以对变量定义范围或做一些计算。你还可以从一个清单上选择一些数字码代表相应的意义，并加以保存，如 (1 = No 2 = Yes), 这些文本清单可以输出成为 "数值标记"。对于日期的输入也很容易，如将日期设置为 "dd/mm/yyyy" 格式，2003 年 10 月 9 日可以输入为 "09/10/2003"。

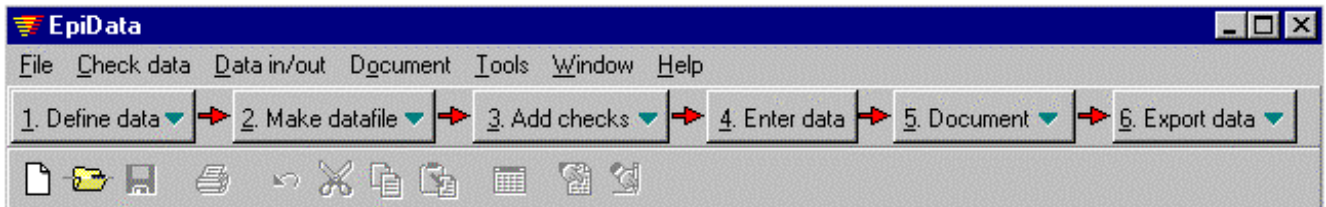
EpiData 既适合应用于简单的数据集即只有一个问卷或一个实验室数据表，也可以应用于很多问卷，数据集，很多实验室部门的数据。这里只介绍最简单的情况。它的基本原理是基于 DOS 环境下工作的 Epi Info v6。你可以先书写一些简单的文本行，然后将它转换成数据输入格式，然后再添加数据控制功能，条件跳转和字段的计算功能。

程序下载和软件安装 EpiData 是一个免费软件。可以从 www.epidata.dk 网站下载这个软件，然后按照给出的指令去做即可完成安装工作。EpiData 的安装过程对于你的计算机没有任何影响。EpiData 由一个程序文件和帮助文件组成。(从技术上讲，EpiData 由几个独立文件组成，都存在同一个目录下，所有选择项都存在一个.ini 文件中)。

局限性

记录数 (观察值) 原则上没有限制。(已经实验过，记录数可以超过十万条)。对于八万条记录的查询在使用 Pentium I 200Mhz 处理器时不超过 1 秒。数据文件结构不能超过 999 文本行。

如何运行 EpiData



EpiData 的屏幕具有标准的视窗形式，它包括一个菜单和两个工具条。

"工作过程工具条" 可以指导用户从 "1. 定义数据" 到 "6. 数据转出"。

1. 定义数据

对于每个变量需要给出 3 种信息:

- A.. 输入字段名 (变量, 如. v1).
- B.. 变量描述文本. (如 性别或 "出生日期")
- C.. 字段定义, 如. ## 代表两位数字.

其他数据类型可以是布尔函数 (yes-no), 或声索引字段.

选择项设置 (文件菜单) 由用户决定使用变量命名方法。

- | | |
|--|---------------------------|
| a. v1sex (10 first characters in sentence) | 选择“自动命名”字段名是 v1sex |
| b. v1 (first word of sentence). | 选择“第一个单词”字段名是 v1 |

其他选择项: 背景颜色, 字段颜色, 行高度等.使用 Stata 或 SPSS 的用户应当使用 "第一个单词"命名方法; 字段名将是变量名. Stata 的用户选择小写字段名.

2. 数据文件的产生.

当完成数据定义后，你可以预览数据输入格式，产生实际的数据文件。

```

我的第一个数据录入表
id          (自动编号) <idnum >
v1          性别 #
v2          身高 (米)  #.##
v3          体重 (公斤)  ###.#
bmi         体重指数  ##.##
v4          出生日期 <dd/mm/yyyy>
age        年龄      ###
s1          居住国  _____
s2          城市 (现在地址) <a >
t1          今天日期 <Today-dmy>

```

3. 添加/修改 核查功能

EpiData 软件的最强大的功能是在数据输入过程中可以指定输入范围和进行计算。

- 严格限制输入为某些数值并且为输入的数字给出文本描述。
- 指定数据输入顺序，如某些问题只限于对男性个体，(jumps)
- 在数据输入过程中使用计算功能。例如根据访问日期和出生日期计算年龄。但是一般计算是在分析过程中进行。
- 具有帮助功能和一些计算扩展功能如, if .. then ...endif 结构。
(可以参考安装给出的实例，还可以访问网站 www.epidata.dk/examples.php).

当你开始使用“添加/修改”部分时，出现一个窗口：在上部显示变量名 (v8) . 标记和变量类型 (number). 下面是定义块：

Range, Legal 定义允许输入的数据。

Jumps 在输入数据后决定是否跳转。

(本例中数值为 1 跳至 v10)

Must enter: 如果设置为 Yes 必须给出数值。(否则可以为空)

Repeat: 重复上个记录的数值，即如果不改变仍然维持前一个记录的数值，这个数值可以改变。

Value label: 对于分类数据它定义了数值的具体意义。如 1=男 2=女。

Edit: 可以定义其它相关的核查命令。具体参考帮助文件。

Save: 保存当前的定义。

在本例中定义了 **Range,Legal** 和 **Value labels** . 实际上一般只需使用一个即可。

Variable	V8
Rigidfix:	Number
Range, Legal	1-8,99
Jumps	1>v10
Must enter	Yes
Repeat	No
Value label	label_v8rigidfix

4. 数据输入

打开数据文件并且输入数据，增加或查找数据。字段 V1 右面的蓝色文本 Female 是根据核查文件，在数据输入后显示的。本例子中体重指数 Body mass index 和年龄 age 是计算字段。

文件类型：

- A. 数据格式定义文件. **first.qes**
- B. 真正的数据文件. **first.rec**.
- C. 核查文件. **first.chk**
- D. 辅助文件. **first.not** 和运行记录文件 **first.log** .

ID	(automatic id number)	1
U1	sex	2 Female
U2	Height (meter)	1.75
U3	Weight (kilo)	67.0
BMI	Body Mass Index	21.88
U4	Date of birth	12/12/1956
AGE	Age today	44
S1	Country of Residence	Denmark
S2	City (Current address)	00 Odense
T1	Today's Date	27/01/2001

5. 与数据相关的文件

当完成一些字段定义后可以显示数据文件结构及相关信息。例如 (*first.rec*):

```

数据文件: C:\data\first.rec
文件标记: My first test datafile is an example

文件大小:          612 bytes
最后更新日期:     28. Jan 2001 12:14
字段数: 7
记录数: 0
核查功能: Yes (Last revision 28. Jan 2001 12:02)

Fields in datafile:

```

No.	变量名	标记	字段类型	宽度	核查内容	数值标记
1	id		ID-number	6		
2	v1	sex	Integer	1		sex 1: Male 2: Female 9: Unknown
3	v2	Height (meter)	Fixed number	4:2	Legal: 0.0-2.30,9	
6	v4	Date of birth	Date (dmy)	10		

(other fields omitted)

当输入一些数据后，你可以对输入的部分或全部数据进行显示，也可以显示标记。

```

Observation 1

```

id	1	v1	Male	v2	1.92
v4	12/12/1945	s1	denmark	s2	Copenhagen

可以显示 "编码本", 包括简单的频数表

```

v2 ----- Sex
      type: Integer
      value labels: sex
      range/legal: 1-2,2
      missing: 0/25
      range: [1,2]
      unique values: 2
      tabulation:
          Freq.   Pct.   Value  Label
          11    44.0     1    Male
          14    56.0     2    Female
v3 ----- Temp
      type: Floating point
      range/legal: 36.00-40.00
      missing: 0/25
      range: [36.00,37.50]
      unique values: 12
      mean: 36,84
      std. dev: 0,37

```

6. 数据转出和数据加密.

作为“工具”的一部分，可以产生加密或不加密的压缩文件。

注意如果忘记密码不能再恢复这个文件.

工具和其它功能

EpiData 包括其它的数据处理功能如, 比较两个文件, 并以字段为单位给出不同之处; 更改调查表文件后, 将原输入数据转入新调查表数据文件而不丢失原来输入的数据. 建立树形结构编码, 产生关联数据输入, 等. 以后还要添加, 用户定义菜单, 检查文件使用的语言等。

分析

EpiData 包括几乎所有函数. 开发基本分析软件包的工作已经开始. 请及时查看相关网站 www.epidata.dk. 可以将输入数据转出为各种数据格式, 如 ascii 文件, dbaseIII 或 Excel Stata, SAS 和 SPSS).

其它资源

帮助文件及其它补充文件可以参考 <http://www.epidata.dk>

讨论组及用户支持将在百忙之中 2003 年秋建立.

关于 EpiData

EpiData 是基于 Windows 95/98/NT/2000 (32 bit) 的一个数据输入程序. EpiData 的开发首先是由 Jens M. Lauritsen, MD. PhD, Denmark. 程序发布由: The EpiData Association, Odense Denmark. 程序设计者为: Jens M.Lauritsen & Michael Bruus.

引证:

Lauritsen JM 和 Bruus M. EpiData (第 3 版). 一个为保证数据输入可靠的强大工具. The EpiData Association, Odense, Denmark, 2003.

第一版 EpiData 开发者是 Lauritsen JM, Bruus M., Myatt MA, EpiData, 的版本 1.0-1.5. 数据输入的一个有效的工具. County of Funen Denmark and Brixton Health UK. 2001.

资助与感谢.

资助者清单见 [Http://www.epidata.dk/funding.htm](http://www.epidata.dk/funding.htm) 未来赞助和感谢:

[Http://www.epidata.dk/credit.htm](http://www.epidata.dk/credit.htm) . 各种语言的版本, 见 [Http://www.epidata.dk](http://www.epidata.dk)

如对未来发展表示捐助可发电子邮件 info@epidata.dk

声明

EpiData 软件已经被开发并经检查, 以确保数据可靠输入. 我们为此作了极大的努力, 但是不能对任何使用者承担, 由于使用该软件引起数据丢失的责任.