

Excel

在统计分析中的应用

■ 基础知识
典型范例
综合实战

■ 专业作者：

由精通Excel的统计分析教学专家结合多年教学与工作经验精心编著，详细讲述如何在Excel中实现各种统计分析功能，符合统计工作者的应用需求

■ 涵盖面广：

涵盖统计学的绝大部分知识，包括：统计基础与数据描述、描述性统计、数据库统计函数与数据透视表、统计指数、概率分布与概率分布图、抽样与抽样分布、参数估计、假设检验、非参数检验、方差分析、相关分析、回归分析与预测和时间序列分析等

■ 结合实践：

提供160多个来自实际工作中的应用案例，并给出了详尽的操作步骤和解说，读者可根据案例学习如何将Excel应用到自己的工作中，或者根据统计需求直接套用

钟晓鸣 万小笠 编著

1CD

超值视频教学演示光盘

■ 60节Excel视频动画教学演示

■ 9大类254个办公模板

■ 167个范例源文件

 **科学出版社**
北京科海电子出版社
www.khp.com.cn

钟晓鸣 万小笠 编著

Excel

在统计分析中的应用

■ 基础知识
典型范例
综合实战

内 容 提 要

本书由浅入深系统地介绍了如何在 Excel 2003 中实现各种统计分析功能,帮助读者在巩固统计学知识的基础上,将 Excel 作为进行统计分析的一门有效工具熟练地运用于日常的工作、学习及科研中。本书共分 14 章,涵盖基础统计学的绝大部分内容,主要包括:统计基础与数据描述、描述性统计、数据库统计函数与数据透视表、统计指数、概率分布与概率分布图、抽样与抽样分布、参数估计、假设检验、非参数检验、方差分析、相关分析、回归分析与预测、时间序列分析和综合案例分析。

随书配套光盘提供了 254 个 Excel 常用模板和本书案例源文件,并配有 Excel 软件教学演示,方便读者学习使用。

本书既可作为统计初学者或有一定统计学基础的用户学习 Excel 的辅导书,也可作为高校经济和金融相关专业师生和各类培训班进行统计学及 Excel 学习的教材和参考用书。

图书在版编目(CIP)数据

Excel 在统计分析中的应用:基础知识、典型范例、综合
实战 / 钟晓鸣, 万小笠编著. —北京: 科学出版社, 2008
ISBN 978-7-03-023774-3

I. E… II. ①钟… ②万… III. 电子表格系统, Excel—
应用—统计分析 IV. C819

中国版本图书馆 CIP 数据核字(2008)第 201777 号

责任编辑: 刘志燕 / 责任校对: 杨慧芳

责任印刷: 科 海 / 封面设计: 洪文婕

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京市鑫山源印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2009 年 2 月第 一 版

开本: 16 开

2009 年 2 月第一次印刷

印张: 25.75

印数: 0 001~4 000

字数: 626 千字

定价: 45.00 元(含 1CD 价格)

(如有印装质量问题, 我社负责调换)

前言

Excel 是 Microsoft 公司开发的 Office 办公软件包中的最重要的部分之一, 由于其采用电子表格技术, 从诞生一开始便与数据统计之间有着天然的联系。随着其版本的逐渐提高, 统计分析功能也日渐强大, 其中专为统计设计的各类统计函数简化了函数的计算, 而其中通过加载宏添加的数据分析工具更是使复杂的统计分析过程变得快捷和易实现。

为了帮助众多希望运用统计软件在工作和科研中进行统计分析的读者能够掌握并熟练应用 Excel, 笔者精心编著了本书。本书按照基础统计学的架构, 在简单介绍统计知识的基础上结合具体例子分析如何应用 Excel 实现各种类型的统计分析, 对每一个例子给出了详细的操作步骤, 通过对统计理论的回顾和操作实例的讲解实现运用统计学的原理和以 Excel 为工具解决实际中的问题。

本书特色

1. 内容充实, 涵盖面广

本书涵盖了基础统计学的绝大部分内容, 除了检验和回归等统计问题外, 书中还介绍了统计指数、非参数检验和时间序列分析等内容, 而这些内容正有着越来越多的应用。

2. 实例丰富、典型

在简单回顾统计知识的基础上, 本书重点通过大量现实中的例子来讲解, 一方面使读者理解面对此类问题该选用何种统计方法, 另一方面也领悟如何通过 Excel 解决现实问题。

3. 编排科学, 注重图解

本书在内容的编排和操作步骤说明上反复比较, 通过添加指示文本框和中间说明提示, 尽量使得读者快速理解作者的意图, 并在按照操作步骤进行实现的基础上, 达到融会贯通和熟练运用的目的。在此基础上, 本书在每章都给出了有针对性的习题, 读者可以通过习题中的实例进行进一步的练习。

4. 讲解通俗, 步骤详细

在本文实例的讲解中, 给出了具体分析和详尽的步骤。在此基础上, 本书在相应的图中运用文本框进行了标注, 使得读者可以马上抓住操作步骤的要点。

本书内容

第 1 章首先在介绍几个统计学概念基础上, 讲解如何通过 Excel 绘制频率分布图、相对频率分布图、累计频率分布图和直方图。

第 2 章主要介绍如何利用 Excel 进行描述性统计, 从中心趋势、离中趋势、偏度、峰度 4 个方面对数据进行全方位的描述。

第3章首先介绍 Excel 内部数据库及其基本概念, 然后通过举例说明如何创建数据库并实现查询、筛选和分类汇总, 最后介绍了 Excel 中的数据透视表和数据透视图。

第4章主要介绍如何利用 Excel 计算统计指数来描述单个变量或一组变量的相对变化。

第5章介绍了几种常见的离散型概率分布和连续型概率分布, 并运用 Excel 绘制不同分布的概率密度函数图和分布函数图。

第6章主要介绍如何在 Excel 中实现最常用的三种随机抽样: 简单随机抽样、等距抽样、分层抽样, 并对比分析了比较特殊的一类抽样: 周期抽样, 最后研究了不同情况下的抽样分布。

第7章介绍如何通过抽样后的样本对总体的参数进行估计, 介绍了点估计和区间估计两种主要类型。

第8章主要介绍在 Excel 中如何进行各种类型的假设检验, 此外还重点介绍了相关的检验函数以及在分析工具中的4种检验宏工具: z 检验和3类 t 检验。

第9章介绍了6种常用的非参数检验方法: χ^2 、符号检验、Wilcoxon 带符号的等级检验、Mann-Whitney U 检验、Kruskal-Wallis H 检验和等级相关检验。

第10章介绍3种方差分析: 单因素方差分析、无重复的双因素分析和有重复的双因素分析。本章重点在于掌握方差分析原理的基础上, 熟练运用分析工具中的3种方差分析工具。

第11章介绍变量之间的相关关系的分析, 可分为简单线性相关和多元相关两类, 而多元相关的度量又分为多元相关系数和偏相关系数, 重点在于熟练运用相关函数和相关分析工具求得各类相关系数。

第12章介绍回归分析, 讲解利用 Excel 的趋势线以及各类回归函数和回归分析工具实现各种类型的回归分析, 并在此基础上进行预测。

第13章简要介绍了时间序列分析的概念, 以及如何运用移动平均、趋势回归和指数平滑3种方法对趋势变动进行分析和预测, 采用长期趋势消除法和季节虚拟变量回归法进行季节变动调整。

第14章介绍了如何使用 Excel 对具体复杂的例子进行综合分析。在本章中, 综合应用了前面章节中所讲解的各种统计分析方法和工具, 结合实际应用介绍了 Excel 进行统计分析的方法。

本书适合的读者

本书既可作为统计初学者或有一定统计学基础的用户学习 Excel 的辅导书, 也可作为高校经济、管理和工程等专业相关师生和各类培训班进行统计学及 Excel 学习的教材和参考用书。

本书作者

本书由钟晓鸣、万小笠主编, 参与编写的还有王俊标、陈晨、高守传、郭瑞、周宇炜、蔡雪燕、陈杰、荣飞、郑林、张路平、陈刚、陈冠军、陈杰、罗皓茵、赵正坤、公芳亮、程明雷, 在此一并表示感谢。

由于编者水平有限, 且编著时间仓促, 书中难免存在疏漏之处, 敬请广大读者朋友批评指正, 并提出宝贵意见。

编者
2009.1



目 录

第 1 章 统计基础与数据描述	1
1.1 总体和样本	1
1.2 数据类型	2
1.2.1 连续数据与离散数据	2
1.2.2 横截面数据和时间序列数据	2
1.2.3 组数据和非组数据	3
1.3 数据描述	11
1.3.1 频率分布图	12
1.3.2 相对频率分布图	16
1.3.3 累积频率分布图或累积曲线	18
1.3.4 直方图	20
1.4 小结	23
1.5 习题	23
第 2 章 描述性统计	25
2.1 中心趋势	25
2.1.1 算术平均值	25
2.1.2 几何平均值	28
2.1.3 众数	31
2.1.4 中位数	32
2.1.5 调和平均数	35
2.2 离中趋势	35
2.2.1 方差	36
2.2.2 标准差	38
2.2.3 四分位数偏差	40
2.2.4 方差系数	43
2.3 偏度	44
2.3.1 偏斜度	45
2.3.2 矩偏度系数	46
2.3.3 四分位数偏度系数	47

2.3.4 Spearman 偏度系数	48
2.4 峰度	49
2.4.1 峰值	50
2.4.2 矩峰度系数	51
2.5 利用分析工具进行描述性统计	52
2.6 小结	55
2.7 习题	55
第3章 数据库统计函数与数据透视表	56
3.1 数据库	56
3.1.1 数据库的创建	56
3.1.2 数据的查询与筛选	59
3.1.3 数据的分类汇总	63
3.2 数据库统计函数	70
3.2.1 计数函数	71
3.2.2 求和、乘积函数	73
3.2.3 最值函数	75
3.2.4 均值、方差、标准差函数	77
3.2.5 其他数据库函数	81
3.3 数据透视表	82
3.3.1 创建数据透视表	83
3.3.2 编辑数据透视表	88
3.4 小结	91
3.5 习题	91
第4章 统计指数	93
4.1 综合指数	93
4.1.1 同等加权指数	93
4.1.2 基期加权指数 (拉氏指数)	95
4.1.3 现期加权指数 (帕氏指数)	97
4.1.4 资本加权指数	101
4.2 平均指数	103
4.2.1 算术平均指数	103
4.2.2 几何平均指数	105
4.2.3 调和平均指数	107
4.3 小结	108

4.4 习题	108
第5章 概率分布及概率分布图	111
5.1 离散型概率分布	111
5.1.1 二项分布	112
5.1.2 超几何分布	120
5.1.3 泊松分布	124
5.2 连续型概率分布	129
5.2.1 正态分布	130
5.2.2 标准正态分布	140
5.2.3 对数正态分布	144
5.2.4 其他连续型分布	146
5.3 小结	150
5.4 习题	150
第6章 抽样与抽样分布	151
6.1 随机抽样	151
6.1.1 采用随机数函数进行简单抽样	152
6.1.2 采用随机数发生器进行简单随机抽样	157
6.1.3 采用抽样宏进行简单随机抽样	160
6.1.4 采用随机数函数进行等距抽样	163
6.1.5 分层抽样	165
6.2 周期抽样	167
6.3 选择抽样样本大小	169
6.3.1 估计均值时样本大小的选择	169
6.3.2 估计比例时样本大小的选择	172
6.4 抽样分布	173
6.4.1 大样本的抽样分布	174
6.4.2 小样本的抽样分布	177
6.5 小结	189
6.6 习题	189
第7章 参数估计	191
7.1 总体均值的估计	191
7.1.1 总体方差已知下估计	191
7.1.2 总体方差未知且为小样本下估计	193

7.1.3 总体方差未知且为大样本下估计	194
7.2 总体均值之差估计	196
7.3 总体比例和比例之差的估计	197
7.3.1 总体比例估计	197
7.3.2 总体比例之差估计	199
7.4 总体方差和方差比估计	200
7.4.1 总体方差估计	200
7.4.2 总体方差比的估计	201
7.5 小结	202
7.6 习题	203
第8章 假设检验	204
8.1 总体均值的假设检验	206
8.1.1 总体方差已知下均值的假设检验	206
8.1.2 方差未知且为小样本下总体均值检验	210
8.1.3 总体方差未知且为大样本下的检验	213
8.2 总体均值之差假设检验	215
8.2.1 总体方差已知下的检验	215
8.2.2 总体方差未知且为大样本下检验	219
8.2.3 总体方差未知且为小样本下检验	221
8.3 总体比例和比例之差的假设检验	230
8.3.1 总体比例的假设检验	230
8.3.2 总体比例之差假设检验	231
8.4 总体方差的假设检验	235
8.4.1 单个正态总体方差的假设检验	235
8.4.2 两总体方差的检验	236
8.5 小结	241
8.6 习题	241
第9章 非参数检验	243
9.1 χ^2 检验	243
9.1.1 拟合优度检验	244
9.1.2 独立性检验	248
9.2 符号检验	253
9.2.1 简单符号检验	253
9.2.2 Wilcoxon 带符号的等级检验	255

9.3	Mann-Whitney U 检验	259
9.4	Kruskal-Wallis H 检验	262
9.5	等级相关检验	264
9.6	小结	268
9.7	习题	269
第 10 章 方差分析		271
10.1	单因素方差分析	271
10.1.1	单因素方差分析方法	272
10.1.2	方差分析表	275
10.1.3	单因素方差分析工具的运用	278
10.2	双因素方差分析	280
10.2.1	无重复的双因素方差分析	280
10.2.2	有重复的双因素方差分析	286
10.3	小结	291
10.4	习题	291
第 11 章 相关分析		293
11.1	简单线性相关	293
11.1.1	散点图法	294
11.1.2	相关函数法	297
11.1.3	相关分析工具	299
11.1.4	相关系数的检验	301
11.2	多元相关	305
11.2.1	多元相关系数	305
11.2.2	偏相关系数	307
11.3	小结	309
11.4	习题	309
第 12 章 回归分析与预测		312
12.1	一元线性回归分析与预测	312
12.1.1	应用散点图和趋势线进行回归分析	313
12.1.2	应用回归函数进行回归分析	317
12.1.3	应用回归分析工具进行回归分析	323
12.1.4	一元线性预测	326
12.2	多元线性回归分析与预测	329

12.2.1 运用 LINEST 函数进行多元线性回归分析	330
12.2.2 运用回归分析工具进行多元线性回归分析	332
12.2.3 多元线性预测	335
12.3 一元非线性回归分析与预测	338
12.3.1 指数回归与预测	338
12.3.2 对数回归与预测	345
12.3.3 幂函数回归与预测	349
12.3.4 多项式回归与预测	353
12.4 小结	357
12.5 习题	357
第 13 章 时间序列分析	359
13.1 移动平均法测定长期趋势	360
13.1.1 中心移动平均法	360
13.1.2 历史移动平均法	363
13.2 趋势回归法测定长期趋势	367
13.2.1 线性趋势方程	367
13.2.2 二次多项式趋势方程	370
13.2.3 指数趋势方程	373
13.3 指数平滑法测定长期趋势	375
13.3.1 确定最佳阻尼系数 $(1-\alpha)$	376
13.3.2 指数平滑分析工具	379
13.4 季节变动的测定	381
13.4.1 长期趋势消除法	381
13.4.2 季节虚拟变量回归法	384
13.5 小结	387
13.6 习题	387
第 14 章 综合实例分析	389
14.1 利用分析工具绘制直方图	389
14.2 描述性统计	391
14.3 数据库的创建及数据库统计函数	392
14.4 参数估计	397
14.5 假设检验	398
14.6 相关分析	399
14.7 回归分析	402

统计基础与数据描述

统计学是一门关于用科学的方法收集、整理、汇总、描述和分析数据资料，并在此基础上进行推断和决策的科学。狭义的统计用来统指数据或者从数据中得到的一些数字。

从统计的定义可以看出，统计的关键在于对数据的分析与加工，而 Excel 强大的数据分析功能则恰恰与统计所要处理的问题相适应，因此从 Excel 产生之初便被广泛地应用于统计中，而专为统计分析所开发的各种宏更使得 Excel 成为统计分析中一种实用而高效的工具。虽然 SPSS、SAS 等专业统计软件在某些方面具有更为强大的统计分析功能，但其或者需要专业的编程，或者需要高昂的价格，因此普及性远远不如 Excel。

在开始 Excel 的统计之旅前，本章首先介绍统计的几个基本概念和统计中最简单的数据描述，这是掌握用 Excel 进行统计分析的必备知识。

1.1 总体和样本

当收集分析一组反映人或物的特征数据时，如果对组中的每一个元素进行观察，则得到的数据即为总体。但在实际工作中，由于各种条件的限制或者从经济的角度考虑，往往仅对组中一部分数据进行观测，这一部分即为样本。

如果样本能很好地反映总体的特性，那么就可以通过样本的分析来对总体下结论，在这种情况下进行的统计工作称为归纳统计学或统计推断。如某学校辅导员需要对毕业生找工作时的服装消费状况进行调查，一种方法是该辅导员通过动员所有院系的辅导员将所有毕业生的花费进行逐一统计，最终得到毕业生服装花费的总体。但实际中往往无法也没必要统计所有同学的消费金额，因此可以采用另一种方法，即通过在每个院系中抽取一定数量的学生进行调查，调查所得的数据即为样本，根据抽取的样本即可推断出所有毕业生服装花费的总体。

如果仅仅描述和分析特定的对象而不下结论或者对较大的群体不进行推断，而用一个数来概括一组数据的特征，这种情况下的统计称为描述性统计或演绎统计学。

1.2 数据类型

数据有各种各样的形式，数据类型往往决定或影响统计分析使用方法的选择，因此在使用 Excel 进行统计分析之前，必须明确数据的不同类型。

1.2.1 连续数据与离散数据

按照是否连续，数据可以分为连续（continuous）数据和离散（discrete）数据。连续数据可以取区间内的任何值，也就是说数据可用一个连续的标准度量，度量值只受精度的限制。典型的连续数据的例子是投资收益率，收益可以为 5%、5.001%或 4.9999%，同样的还有时间、速度等。

离散数据则仅能取区间内的有限个值，例如由于股票受最小变动单位的限制，如果股票的最小变动单位为 0.01 元，则对应股票的价格只能为 5、5.01 或 4.99，而不能为 5.005，同样的还有班级人数、股票的成交量等。

后面章节的例子将会说明连续数据和离散数据对描述统计量的计算产生的差别。

1.2.2 横截面数据和时间序列数据

按照数据反映的时间，数据可分为横截面数据和时间序列数据。横截面数据描述的是某一特定时间内一组变量的状况，例如表 1-1 中描述的 2005 年 3 月 21 日上证 50 指数样本股的 20 种股票的收盘价格的一组数据，即为横截面数据；时间序列数据则反映某一特定变量随时间的变化，最为典型的就是股票价格的时间序列，例如表 1-2 中深发展（000001）在 2004 年 12 月中 22 个交易日的价格组成的数据，即为时间序列数据。

表 1-1 上证 50 指数样本股横截面数据

代码	证券名	价格	代码	证券名	价格
600000	浦发银行	7.24	600026	中海发展	9.65
600004	白云机场	8.54	600028	中国石化	4.20
600006	东风汽车	2.94	600029	南方航空	3.75
600008	首创股份	8.20	600030	中信证券	5.00
600009	上海机场	16.65	600033	福建高速	8.38
600011	华能国际	6.97	600036	招商银行	8.51
600015	华夏银行	4.05	600050	中国联通	2.90
600016	民生银行	5.69	600100	清华同方	12.92
600018	上港集箱	16.37	600104	上海汽车	4.66
600019	宝钢股份	6.17	600171	上海贝岭	8.42

表 1-2 深发展 2004 年 12 月价格

日期	价格	日期	价格
2004-12-1	6.94	2004-12-12	6.91
2004-12-2	7	2004-12-13	6.86
2004-12-3	7.02	2004-12-14	6.86
2004-12-4	6.98	2004-12-15	6.97
2004-12-5	7	2004-12-16	6.88
2004-12-6	6.96	2004-12-17	6.82
2004-12-7	7.09	2004-12-18	6.75
2004-12-8	6.94	2004-12-19	6.69
2004-12-9	6.87	2004-12-20	6.66
2004-12-10	6.96	2004-12-21	6.58
2004-12-11	6.99	2004-12-22	6.59

1.2.3 组数据和非组数据

按照数据是否分组,数据又可分为组数据和非组数据。当只有较少部分数据被处理时,这些数据可能是未被加工的或者处于非组状态(ungrouped),这样的数据称为非组数据。由于信息量较少,读者仍然能理解这些数据,例如表 1-3 中是上证 180 指数 2004 年 12 个月度观测值。

表 1-3 上证 180 指数 2004 年月度值

月	2004-1	2004-2	2004-3	2004-4	2004-5	2004-6
值	3 019.04	3 140.53	3 213.48	2 912.81	2 819.49	2 528.44
月	2004-7	2004-8	2004-9	2004-10	2004-11	2004-12
值	2 548.49	2 479.16	2 591.00	2 452.73	2 490.28	2 362.07

当数据量相当大时,例如,如果需要描述 2002 年到 2004 年 3 年间上证 180 指数的月度值和月度收益,就需要将其加工汇总,将非组数据转化为一个频率表(frequency),以方便分析人员掌握其中的规律,此时便将非组数据转化为了组(grouped)数据。

下面通过举例说明如何在 Excel 中实现将非组数据转化为组数据。

例 1.1 求解上证 180 指数的月收益率

上证 180 指数 2002 年到 2004 年对应的月度观测值如表 1-4 所示,试求指数对应的月收益率,其中月收益率=(本月指数值-上月指数值)/上月指数值。

表 1-4 上证 180 指数 2002 年至 2004 年月度值

时间	指数值	时间	指数值	时间	指数值
2002-1	2 828.95	2003-1	2 794.97	2004-1	3 019.04
2002-2	2 869.65	2003-2	2 793.41	2004-2	3 140.53
2002-3	2 941.43	2003-3	2 810.75	2004-3	3 213.48
2002-4	3 028.49	2003-4	2 882.80	2004-4	2 912.81
2002-5	2 773.89	2003-5	3 010.19	2004-5	2 819.49
2002-6	3 299.06	2003-6	2 846.85	2004-6	2 528.44
2002-7	3 138.92	2003-7	2 861.76	2004-7	2 548.49
2002-8	3 124.83	2003-8	2 741.57	2004-8	2 479.16
2002-9	2 953.91	2003-9	2 619.23	2004-9	2 591.00
2002-10	2 817.64	2003-10	2 592.19	2004-10	2 452.73
2002-11	2 658.88	2003-11	2 681.86	2004-11	2 490.28
2002-12	2 561.45	2003-12	2 828.80	2004-12	2 362.07

具体操作步骤如下：

Step 01 新建一个空白工作表，单击 A1 单元格，输入表头“上证 180 指数”。

Step 02 首先合并单元格以设置表头。选择 A1:D1 单元格，右击选中的区域，在弹出的快捷菜单中选择【设置单元格格式】命令，弹出【单元格格式】对话框，在【数字】选项卡的【分类】列表框中选择【文本】选项，如图 1-1 所示。

Step 03 选择【单元格格式】对话框中的【对齐】选项卡，选中其中的【合并单元格】复选框，如图 1-2 所示。

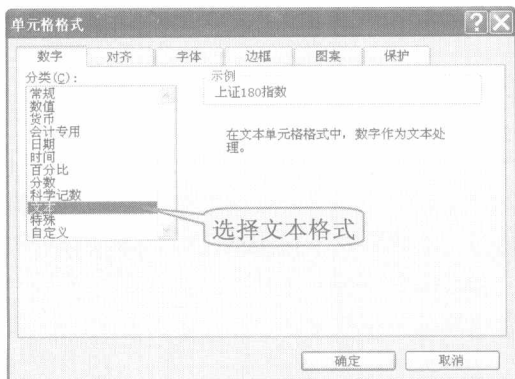


图 1-1 【数字】选项卡



图 1-2 【对齐】选项卡

Step 04 选择【单元格格式】对话框中的【字体】选项卡，在【字体】列表框中选择【宋体】选项，在【字号】列表框中选择【12】选项，如图 1-3 所示，完成后单击【确定】按钮。

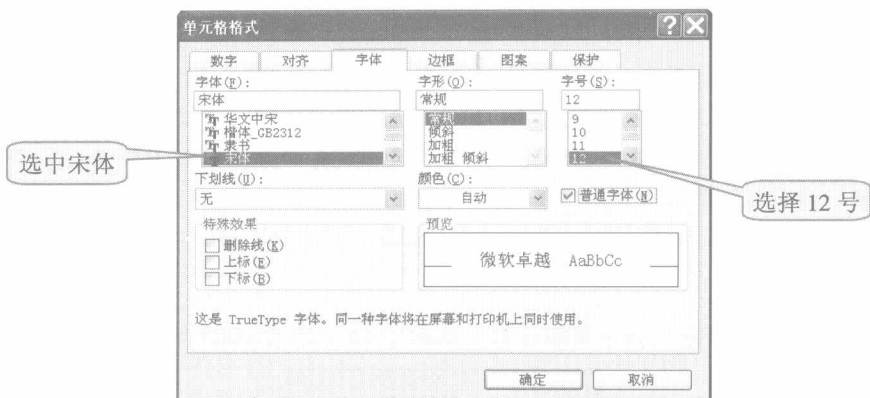


图 1-3 【字体】选项卡

Step 05 单击单元格 A2，输入“时间”；单击单元格 B2，输入“指数”；单击单元格 C2，输入“收益率”。

Step 06 单击单元格 A3，输入“2002-1”，并右击，在弹出的快捷菜单中选择【设置单元格格式】命令，然后在弹出的【单元格格式】对话框中选择【数字】选项卡，在【分类】列表框中选择【自定义】选项，在【类型】文本框中输入“yyyy-m”，如图 1-4 所示，完成后单击【确定】按钮。

Step 07 运用自动填充单元格命令给出所有的日期。单击 A3 单元格，将光标置于 A3 单元格右下角，此时光标变为小黑十字，如图 1-5 所示，拖动鼠标至 A38 单元格，即可实现所有日期的自动填充。

说明：自动填充功能是 Excel 的一大特色，熟练地使用自动填充功能将大大简化某些重复性输入工作。

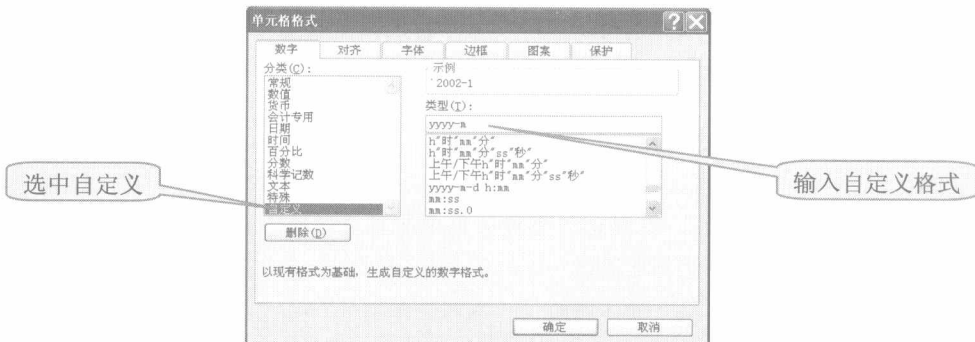


图 1-4 自定义数字格式

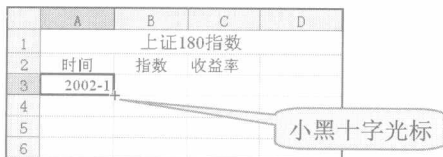


图 1-5 自动填充

Step 08 在 B3~B38 单元格中输入对应每月的指数观测值。

Step 09 求解每月收益率，月收益率=（本月指数-上月指数）/上月指数。

单击 C4 单元格，在编辑栏中输入“=（”，单击 B4 单元格，在编辑栏中输入“-”，单击 B3 单元格，在编辑栏中输入“）/”，再单击 B3 单元格，完成后按回车键。

Step 10 使用自动填充功能求出其他月份的收益率。单击 C4 单元格，将鼠标置于 C4 单元格右下角，待出现小黑十字光标后，拖动鼠标至 C38 单元格，即可将所有月份的收益率全部求出。

最终结果如图 1-6 所示，由于篇幅限制，隐藏了 15~26 行 2003 年间的记录。

	A	B	C	D
1	上证180指数			
2	时间	指数	收益率	
3	2002-1	2,828.95		
4	2002-2	2,869.65	0.014387	
5	2002-3	2,941.43	0.025014	
6	2002-4	3,028.49	0.029598	
7	2002-5	2,773.89	-0.08407	
8	2002-6	3,299.06	0.189326	
9	2002-7	3,138.92	-0.04854	
10	2002-8	3,124.83	-0.00449	
11	2002-9	2,953.91	-0.0547	
12	2002-10	2,817.64	-0.04613	
13	2002-11	2,658.88	-0.05635	
14	2002-12	2,561.45	-0.03664	
27	2004-1	3,019.04	0.067251	
28	2004-2	3,140.53	0.040241	
29	2004-3	3,213.48	0.023229	
30	2004-4	2,912.81	-0.09357	
31	2004-5	2,819.49	-0.03204	
32	2004-6	2,528.44	-0.10323	
33	2004-7	2,548.49	0.00793	
34	2004-8	2,479.16	-0.0272	
35	2004-9	2,591.00	0.045112	
36	2004-10	2,452.73	-0.05337	
37	2004-11	2,490.28	0.015309	
38	2004-12	2,362.07	-0.05148	
39				

图 1-6 上证 180 指数月收益率

从图 1-6 中可以看出，虽然求得了三年中每个月的指数收益率，但由于数据太多，无法直观地观察出其中的规律，甚至无法对这一系列的数据进行说明，因此有必要对其加工汇总，将非组数据转化为组数据。

说明：Excel 的隐藏功能也经常用到的功能之一，可以隐藏选中的行和列。选中需要隐藏的行（或列），右击，在弹出的快捷菜单中选择【隐藏】命令，即可将选中的行（或列）隐藏。当选中包含隐藏行（或列）的区域后，右击，在弹出的快捷菜单中选择【取消隐藏】命令，即可将隐藏的行（或列）重新显示。

对于例 1.1 中如此多的数据，可以将其分为不同的组区间或等级，并采用不同区间的观测值数目（或称频率）来描述，给出对应的频率表。非组数据的分组可以通过 COUNTIF 计算不同组区间内满足条件的单元格的数目间接求解，也可以运用频率分布函数 FREQUENCY 直接返回频率数组，下面对这两种方法分别举例说明。

对数据进行分组时，应首先判断数据大体所在的区域，找出最大值和最小值，然后合理地确定出对应的分组区间（又称组距），且区间不能有重叠部分。