

## 2010 高教社杯全国大学生数学建模竞赛

### 承 诺 书

我们仔细阅读了中国大学生数学建模竞赛的竞赛规则.

我们完全明白, 在竞赛开始后参赛队员不能以任何方式(包括电话、电子邮件、网上咨询等)与队外的任何人(包括指导教师)研究、讨论与赛题有关的问题。

我们知道, 抄袭别人的成果是违反竞赛规则的, 如果引用别人的成果或其他公开的资料(包括网上查到的资料), 必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺, 严格遵守竞赛规则, 以保证竞赛的公正、公平性。如有违反竞赛规则的行为, 我们将受到严肃处理。

我们参赛选择的题号是(从 A/B/C/D 中选择一项填写):     B    

我们的参赛报名号为(如果赛区设置报名号的话):     07050    

所属学校(请填写完整的全名):     哈尔滨工业大学    

参赛队员(打印并签名): 1.     孙阳    

2.     张兴起    

3.     雷博    

指导教师或指导教师组负责人(打印并签名):     尚寿亭    

日期:     2010     年     9     月     13     日

---

赛区评阅编号:

## 2010 高教社杯全国大学生数学建模竞赛

### 编 号 专 用 页

赛区评阅编号：

赛区评阅记录：

评 阅 人										
评 分										
备 注										

全国统一编号：

全国评阅编号：

# 对上海世博会文化影响力的定量评估

## 摘要

世界博览会是人类文明的驿站，是多元文化跨国界交流的平台。2010 年上海世博会是首次在中国举办的世界博览会。上海世博会的举办，无论对上海还是整个中国都有着深远的影响。它将给经济和文化等多个领域带来前所未有的推动力。本文从上海世博会对文化的影响这个侧面来定量评估 2010 年上海世博会的影响力。

本文对世博会的文化影响力的研究是从文化市场影响力、文化资源影响力、文化环境影响力 3 个方面进行分析，对文化商品贸易份额等 9 个角度加以量化。

我国承办过的与本届世博会相似的世界性博览会有 1999 昆明世界园艺博览会、2006 沈阳世界园艺博览会、2008 年奥运会。但考虑到奥运会的赛会性质和举办目的显著不同，且人们对其的关注程度差异较大，所以本文以昆明和沈阳的园艺博览会为样本对本届世博会进行文化影响力评估。我们采用中国统计局《年度统计年鉴》中的客观数据，通过因子分析和 BP 神经网络两种方法评估出世博的文化影响力。

为了权衡众多影响因素对文化影响力方面评价时的相关性，本文首先运用了线性关系因素的因子分析法，得出了各影响因素对文化影响力的影响权重。再通过已公布的 2010 年 6 月的统计数据对尚在进行中的上海世博会的文化影响进行了量化预测。

然而，影响文化影响力指数的因素之间有一定的非线性，所以要求求解算法最好有一定的非线性映射能力，因此我们又利用了 BP 神经网络的方法建立了神经网络模型。利用 BP 神经网络算法的泛化能力和容错能力建立出理论上更加完善的解算方案。

最后我们对两种方法做了系统性比较评价，可行性和误差分析。

**关键字：上海世博会 文化影响力 BP 神经网络模型 因子分析法 非线性回归**

## 引言

世界博览会是人类文明的驿站，是各种新思想、新科技、新发明、新创造、新理念充分展示的盛会，也是多元文化跨国界交流的平台。从 1851 年英国伦敦举办首届现代意义的世界博览会以来，世博会以创新为基石，以科技展示和文化展现作为其实现教育功能的两大手段，以科技的光芒指引人类前行，用文化的力量震撼人类的心灵。

## 一、问题重述

2010 年上海世博会是首次在中国举办的世界博览会。从 1851 年伦敦的“万国工业博览会”开始，世博会正日益成为各国人民交流历史文化、展示科技成果、体现合作精神、展望未来发展等的重要舞台。世博会,在为本国经济创造效益的同时,其文化方面的贡献也是意义深远的. 众所周知,世博会,本身就起着文化方面的宣传和推广作用.同时,由于它的举办,而间接发挥的文化作用更是不可估量的.这些统合起来,即是世博文化方面的影响力.

## 二、模型假设

- 2.1 文化影响力是相对的，可以通过与其它地区的比较来衡量。
- 2.2 文化影响力的总和为 100，地区文化影响力等于它占整体文化影响力总和的比例。
- 2.3 假定在世博会期间，每个月份文化市场影响力数值相同。
- 2.4 假定在世博会期间，只有世博会影响文化影响力。

## 三、符号说明

符号	符号说明
$Q$	调整激励函数形式的 Sigmoid 参数
$O_i$	节点 $i$ 的输出
$N$	样本数
$x_i$	网络输入
$y_i$	网络输出
$O_{ik}$	网络输出为 $y_k$ 节点 $i$ 的输出
$net_{jk}$	网络输出为 $y_k$ 节点 $j$ 的输入
$E$	误差函数
$\hat{y}_k$	网络实际输出
$M$	层数
$W_{ij}$	权值
$z_i$	测试变量
$U_i$	独特因子
$F_p$	公共因子
$A$	因子负荷矩阵
$a_{ij}$	第 $i$ 个变量( $z_i$ )在第 $j$ 个公共因子 $F_j$ 上的负荷

## 四、问题分析

世界博览会是人类文明的驿站，是各种新思想、新科技、新发明、新创造、新理念充分展示的盛会，也是多元文化跨国界交流的平台。从 1851 年英国伦敦举办首届现代

意义的世界博览会以来，世博会以创新为基石，以科技展示和文化展现作为其实现教育功能的两大手段，以科技的光芒指引人类前行，用文化的力量震撼人类的心灵。

上海世博会的举办无论对上海还是整个中国都有着深远的影响。它将给经济和文化等多个领域前所未有的推动力，本文我们将从上海世博会对文化市场影响力这个侧面来定量评估 2010 年上海世博会的影响力。

上海世博会上，各参展国家的展馆及其展示，都是各国历史和文化的缩影。其间的 2 万场文化演艺娱乐活动中，将有许多是参展国组织的具有本国文化特色的活动。通过建筑、音乐、戏曲、舞蹈等人类共同的语言，中华文明的精髓将在与世界各种文明相互交流、碰撞甚至摩擦中得以延绵，中华文化的内涵将得以丰富，中华文化的实体将得以发展。举办此次世博会将全面提升中国的文化“软实力”。

一方面，上海世博会中国国家馆集中展现中华文化的魅力瑰宝，成为中国这个千年文明古国在 2010 年世博会上的“国家名片”。另一方面，上海世博会期间的文化活动也将成为展现中华文化源远流长、根深叶茂的又一重要形式。上海世博会的举办，将为中华文化与世界多元文化的交融提供平台，从而为中华文化吸引力的提升提供沃土。

据介绍，上海世博会期间，预计在园区内举办的各类文化演艺娱乐活动将达到 2 万场，而中国 56 个民族的特色文化、300 多个地方传统剧种、500 多种国家级的非物质文化遗产的展现，将成为上海世博会 2 万场活动中的重要组成部分，以喜闻乐见的形式向来自世界各地的参观者展现中华文化。上海世博会特别鼓励具有中国元素的原创作品在世博会上进行首演，人们共同期待着上海世博会能够成为中华文化继往开来的一个崭新起点。

4.1 文化影响力概念

文化影响力是一个地区对整个文化市场和文化生活的客观影响的总和。对文化市场的影响可以分为直接影响和间接影响；直接影响来源于文化市场竞争，间接影响与文化资源竞争相关。对文化生活的影响同样可以分为直接影响和间接影响；直接影响与文化贸易（文化市场影响）相关，间接影响与文化资源和文化环境相关。文化资源影响文化贸易和文化生活。文化环境包括政治环境、经济环境、社会环境、生态环境和国民素质等；文化环境不仅影响文化生活，而且影响文化贸易。如果把与文化市场竞争相关的影响力简称文化市场影响力、把与文化资源竞争相关的影响力简称为文化资源影响力、把与文化环境相关的影响力简称为文化环境影响力，那么，文化影响力可以大致分解为三个部分：文化市场影响力、文化资源影响力和文化环境影响力。文化影响力评价衡量一个地区对整个文化市场和文化生活的三个基本方面的影响大小——文化市场、文化资源和文化环境影响。

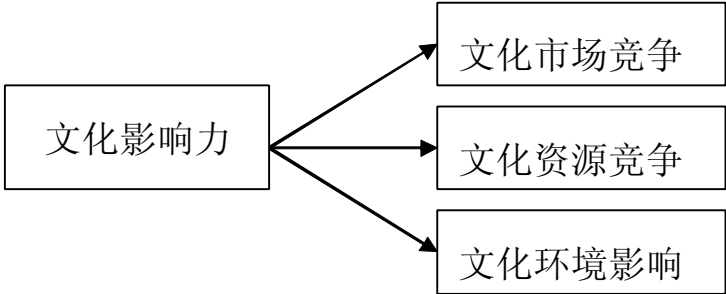


图 4-1 文化影响力结构示意图

## 4.2 文化影响力的评价原理

文化影响力评价的原理是：文化影响力是一个地区对整个文化市场和文化生活的客观影响的大小；文化影响力指数等于文化市场影响力、文化资源影响力和文化环境影响力的相对水平的加权平均值，等于参与评价的文化指标的影响指数的加权平均值（表 4-2）。文化影响力评价是一种相对水平评价。

假设和推论	主要的内容
主要假设	
假设一	文化影响力是一个地区对整个文化市场和文化生活的客观影响的大小
假设二	文化影响力与文化市场影响力、文化资源影响力和文化环境影响力正相关
假设三	整体文化影响力的总和为 100，地区的文化影响力等于它占整体影响力总和的比例
假设四	存在一个标杆，它的文化影响力的各项指标位居第一
假设五	标杆的文化影响力等于它的影响力占整体影响力总和的比例
假设六	标杆的各项文化指标的影响指数为 100
假设七	标杆的文化影响力指数等于各项指标的影响指数的加权平均值，等于 100
假设八	文化影响力是相对的，可以通过与标杆的比较来衡量
主要推论	
模型一	某个地区的文化影响力的单个指标的影响指数等于它的数值与标杆数值相比的相对值
模型二	某个地区的文化影响力指数等于参与评价的文化指标的单个指数的加权平均值

表 4-2 文化影响力评价的概念模型

## 4.3 文化影响力的评价指标

为了客观、准确、完整地并且定量的评价 2010 上海世博会文化的影响力，必须建立一个能够全方位、多角度反映文化影响力的评价指标体系。表 4-3 是文化影响力的标准恒定指标。

一级指标	二级指标 (S)	三级指标 (T)	指标含义	单位
文化影	文化市场影响力	文化商品贸易份额	文化商品贸易/文化商品总贸易	%
		文化服务贸易份额	文化服务贸易/文化服务总贸易	%
		国际旅游收支份额	国际旅游收支/国际旅游收支总和	%
		国际旅游人次份额	出入境 旅游人次/出入境旅游总人次	%
	文化资源影响力	图书出版种类份额	图书出版种类/世界图书出版种类	%



$$\begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \end{pmatrix} = (a_{ij})_{m \times p} \cdot \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_p \end{pmatrix} + \begin{pmatrix} c_1 U_1 \\ c_2 U_2 \\ \vdots \\ c_m U_m \end{pmatrix} \quad (\text{式5-2})$$

$$\text{简记为 } Z = A \cdot F + C U \quad (\text{式5-3})$$

$\begin{matrix} (m \times 1) & (m \times p) & (p \times 1) & (m \times m) & (m \times 1) \\ & & & \text{(对角阵)} & \end{matrix}$

且满足: (I)  $P \leq m$

(II)  $\text{COV}(F, U) = 0$  (即  $F$  与  $U$  是不相关的)

(III)  $E(F) = 0$   $\text{COV}(F) = (1 \cdot \cdot \cdot 1)_{p \times p} = I_p$

即  $F_1, \dots, F_p$  不相关, 且方差皆为 1, 均值皆为 0

(IV)  $E(U) = 0$   $\text{COV}(U) = I_m$  即  $U_1, \dots, U_m$  不相关, 且都是标准化的变量, 假定  $z_1, \dots, z_m$  也是标准化的, 但并不相互独立。

式中  $A$  称为因子负荷矩阵, 其元素(即(5.2.2-1)中各方程的系数)  $a_{ij}$  表示第  $i$  个变量( $z_i$ ) 在第  $j$  个公共因子  $F_j$  上的负荷, 简称因子负荷, 如果把  $z_i$  看成  $P$  维因子空间的一个向量, 则  $a_{ij}$  表示  $z_i$  在坐标轴  $F_j$  上的投影。

### 5.1.3 因子算法模型的应用

为了综合考虑文化商品贸易份额、文化服务贸易份额、国际旅游收支份额、国际旅游人次份额、图书出版种类份额、电影产量份额、劳动生产率、国际移民份额、森林覆盖率这 9 个因素对文化影响力的影响, 下面对其进行因子分析。通过数据判断 9 个因素之间的关系, 可以使用因子分析法。

设存在 1999 年中国文化影响力、2006 年中国文化影响力两个变量,  $m=2$ , 记  $z_1$  = “1999 年中国文化影响力”,  $z_2$  = “2006 年中国文化影响力”。公共因子  $F_1$  = “文化商品贸易份额”,  $F_2$  = “文化服务贸易份额”,  $F_3$  = “国际旅游收支份额”,  $F_4$  = “国际旅游人次份额”,  $F_5$  = “图书出版种类份额”,  $F_6$  = “电影产量份额”,  $F_7$  = “劳动生产率”,  $F_8$  = “国际移民份额”,  $F_9$  = “森林覆盖率”, 由此可得线性方程:

$$\begin{cases} z_1 = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + \dots + a_{19}F_9 \\ z_2 = a_{21}F_1 + a_{22}F_2 + a_{23}F_3 + \dots + a_{29}F_9 \end{cases} \quad (\text{式5-4})$$

矩阵:



$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = (a_{ij})_{2 \times 9} \cdot \begin{pmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \\ F_5 \\ F_6 \\ F_7 \\ F_8 \\ F_9 \end{pmatrix} \quad (\text{式5-5})$$

设对变量  $Z_i$  进行测试得容量为  $n=2$  的观测值 ( $i=1, 2$ )

$$Z_{i1}, Z_{i2}, \dots, Z_{in} \quad (i=1 \sim m)$$

$$\text{记 } r_{ij} = \frac{L_{ij}}{\sqrt{L_{ii}L_{jj}}}$$

$$\text{其中 } l_{ij} = \sum_{k=1}^n z_{ik} z_{jk} - \frac{1}{n} (\sum_k z_{ik}) (\sum_k z_{jk})$$

$$l_{ii} = \sum_{k=1}^n z_{ik}^2 - \frac{1}{n} (\sum_{k=1}^n z_{ik})^2 \quad (\text{式 5-6})$$

称  $r_{ij}$  为  $Z_i, Z_j$  的样本相关系数

$$\text{记 } R = (r_{ij}) = \begin{pmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1m} \\ r_{21} & 1 & r_{23} & \cdots & r_{2m} \\ \vdots & & & & \\ r_{m1} & r_{m2} & r_{m3} & \cdots & 1 \end{pmatrix}_{m \times m} \quad \text{此为 } Z \text{ 的样本相关矩阵, 是一个 } m \text{ 阶对称}$$

阵。

得到测试变量  $Z$  的样本相关矩阵  $R$  之后, 现求主因子解。

(1) 求  $R$  的特征根, 即解方程:

$$|\lambda E - R| = \begin{vmatrix} \lambda - 1 & -r_{12} & \cdots & -r_{1m} \\ -r_{21} & \lambda - 1 & \cdots & -r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ -r_{m1} & -r_{m2} & \cdots & \lambda - 1 \end{vmatrix} = 0 \quad (\text{式5-7})$$

由  $R$  是非负定阵, 解出的特征值都是非负的, 将其非零特征值按从大到小排序并重新编码:  $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ 。

(2) 按预先规定所取的  $P=9$  个公共因子的累计方差贡献率达到的百分比 (85%), 使



电影产量份额	电影产量/世界电影产量（部）	182	2935	6.20%
劳动生产率	经济环境，GDP/劳动力（亿/万人）	80579.4	70586	11416
国际移民份额	社会环境，国际移民/世界移民总数（万人）	91	10022	0.91%
森林覆盖率	生态环境，森林面积/国土面积	17491	95998.90231	18.22%

表 5-2 1999 昆明园艺博览会文化影响力各决定因素统计 II

项目	算法	第三产业 GDP 数额（亿元）	占第三产业 GDP 比例	数额（亿元）	世界文化商品总贸易（亿元）	比例
文化商品贸易份额	文化商品贸易/文化商品总贸易	84769.4	0.0911	7722.49234	77070.78184	10.02%
文化服务贸易份额	文化服务贸易/文化服务总贸易	84769.4	0.0801	6790.02894	79976.78375	8.49%

表 5-3 2006 沈阳园艺博览会文化影响力各决定因素统计 I

项目	算法	分子量	分母量	比值
国际旅游收支份额	国际旅游收支/国际旅游收支总和（百万美元）	33949	271374.9001	12.51%
国际旅游人次份额	入境旅游人次/入境旅游总人次（万人）	4221	5673	74.41%
图书出版种类份额	图书出版种类/世界图书出版种类（种）	141831	778863	18.21%
电影产量份额	电影产量/世界电影产量（部）	424	3475	12.20%
劳动生产率	GDP/劳动力（亿/千万人）	211923.5	76.400	27.739
国际移民份额	社会环境，国际移民/世界移民总数（万人）	291	20949	1.39%
森林覆盖率	生态环境，森林面积/国土面积	26591.202	160671.9154	16.55%

表 5-4 2006 沈阳园艺博览会文化影响力各决定因素统计 II

## 2) 文化影响力指数的数据评估

因为文化影响力的评估有很强的主观因素，我们在评估世博会文化影响力时采用专

家评估法。专家评估法也称专家调查法，该方法以专家为索取未来信息的对象，组织各领域的专家运用专业方面的知识和经验，通过直观的归纳，对预测对象过去和现在的状况、发展变化过程进行综合分析与研究，找出预测对象变化、发展规律、从而对预测对象未来的发展区实际状况做出判断。

所以我们在专家评分过程中选择了来自不同地区的 5 名同学，并为他们提供 1999 年昆明世界园艺博览会、2006 年沈阳世界园艺博览会的相关资料，5 名同学依据文化影响力评价标准中的 9 个指标分别对其进行评分，满分为 5 分。

统计后的表格如表 5-5：

	1	2	3	4	5	平均值
沈阳世博会	3	2.7	2.5	2.8	3.1	2.8
昆明世博会	2.8	2.8	2	1.6	1.9	2.3

表 5-5 博览会文化影响力专家评估统计

5.1.5 采用因子分析模型对数据处理结果与分析

将文化商品贸易份额、文化服务贸易份额、国际旅游收支份额、国际旅游人次份额、图书出版种类份额、电影产量份额、劳动生产率、国际移民份额、森林覆盖率这9个影响文化影响力的相关数据代入模型，最终得出权向量矩阵。

对其列表统计如表5-6：

公共因子	权重 %	总和 %
文化商品贸易份额	17	100
文化服务贸易份额	15	
国际旅游收支份额	23	
国际旅游人次份额	19	
图书出版种类份额	8	
电影产量份额	4	
劳动生产率	3	
国际移民份额	2	
森林覆盖率	8	

表 5-6 利用因子分析法所得各影响因素权重

可以得出：文化影响力与文化商品贸易份额、文化服务贸易份额、国际旅游收支份额、国际旅游人次份额、图书出版种类份额、电影产量份额、劳动生产率、国际移民份额、森林覆盖率的函数关系式确定为：

$$y=0.17F_1+0.15F_2+0.23F_3+0.19F_4+0.08F_5+0.04F_6+0.00003F_7+0.02F_8+0.08F_9$$

（式5-9）

其中  $F_1$  = “文化商品贸易份额”，  $F_2$  = “文化服务贸易份额”，  $F_3$  = “国际旅游收支份额”，  $F_4$  = “国际旅游人次份额”，  $F_5$  = “图书出版种类份额”，  $F_6$  = “电影产量份额”，  $F_7$  = “劳动生产率”，  $F_8$  = “国际移民份额”，  $F_9$  = “森林覆盖率”  
 带入数据相应数据得：

$$1999\text{ 年昆明园艺博览会 } y_1=2.3$$

说明可通过因子分子模型解算，得出正确结论。

## 5.2 BP 神经网络算法

### 5.2.1 BP 神经网络算法的简介

目前，已发展了几十种神经网络，如 Hopfield 模型、Feld-mann 等连接网络模型、Hinton 等的波尔茨曼机模型以及 Rumel-hart 等的多层感知模型和 Kohonen 的自组织网络模型等等。在众多神经网络模型中，应用最广泛的是多层感知机神经网络。多层感知机神经网络的研究始于 20 世纪 50 年代，但一直进展不大。直到 1985 年，Rumelhart 等人提出了误差反向传递学习算法（即 BP 算法），实现了 Minsky 的多层网络设想，如图 5-7 所示。

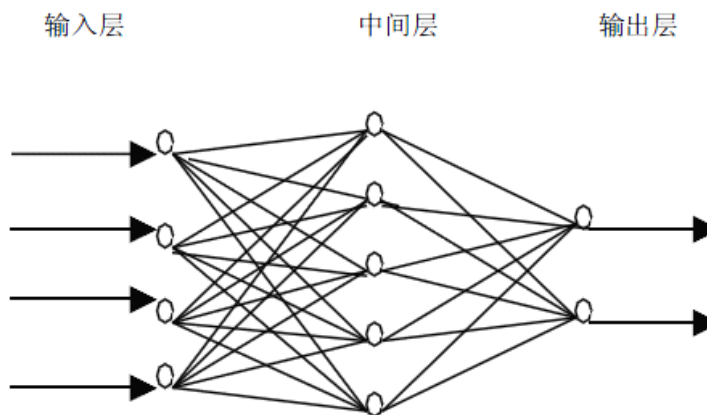


图 5-7 BP 神经网络模型

BP 神经网络是由输入层、中间层、输出层组成的阶层型神经网络。中间层可扩展为多层，相邻层之间的各神经元进行全连接，而每层各神经元之间无连接。当一对学习模式提供给网络后，各神经元获得网络的输入响应产生连接权值，然后按减小希望输出与实际输出误差的方向，从输出层经各中间层逐层修正各连接权，回到输入层。此过程反复交替进行，直至网络的全局误差趋向给定的极小值，即完成学习的过程。运用神经网络评价企业文化影响力的优点表现在以下几个方面：①不需要过程非常复杂的人工演算，减少出错率；②不需要确定权数（一般是根据经验确定权数，人为性比较强），提高精确度；③只需要输入已有的参考数值，计算机便能很快算出数值结果节省时间。

### 5.2.2 BP 神经网络算法的模型

BP 算法不仅有输入层节点、输出层节点，还可有 1 个或多个隐含层节点。对于输入信号，要先向前传播到隐含层节点，经作用函数后，再把隐节点的输出信号传播到输出节点，最后给出输出结果。节点的作用的激励函数通常选取 S 型函数，如

$$f(x) = \frac{1}{1 + e^{-x/Q}} \quad (\text{式 5-10})$$

式中 Q 为调整激励函数形式的 Sigmoid 参数。该算法的学习过程正向传播和反向传播组成。在正向传播过程中，输入信息从输入层经隐含层逐层处理，并传向输出层。每一层神经元的状态只影响下一层神经元的状态。如果输出层得不到期望的输出，则转入反向传播，将误差信号沿原来的连接通道返回，通过修改各层神经元的权值，使得误差信号最小。

设含有 n 个节点的任意网络，各节点之特性为 Sigmoid 型。为简便起见，指定网络

只有一个输出  $y$ , 任意节点  $i$  的输出为  $O_i$ , 并设有  $N$  个样本  $(x_k, y_k)$  ( $k=1, 2, 3 \dots N$ ), 对某一输入  $x_k$ , 网络输出为  $O_{jk}$  的节点  $j$  的输出为  $y_k$ , 节点  $j$  的输入为

$$net_{jk} = \sum_i W_{ij} O_{ik} \quad (\text{式 5-11})$$

并将误差函数定义为

$$E = \frac{1}{2} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \quad (\text{式 5-12})$$

其中  $\hat{y}_k$  为网络实际输出, 定义  $E_k = (y_k - \hat{y}_k)^2$ ,  $\delta_{jk} = \frac{\partial E_k}{\partial net_{jk}}$

且  $O_{jk} = f(net_{jk})$  于是

$$\frac{\partial E_k}{\partial W_{jk}} = \frac{\partial E_k}{\partial net_{jk}} \frac{\partial net_{jk}}{\partial W_{jk}} = \frac{\partial E_k}{\partial net_{jk}} O_{jk} = \delta_{jk} O_{jk} \quad (\text{式 5-13})$$

当  $j$  为输出节点时  $O_{jk} = \hat{y}_k$

$$\delta_{jk} = \frac{\partial E_k}{\partial \hat{y}_k} \frac{\partial \hat{y}_k}{\partial net_{jk}} = -(y_k - \hat{y}_k) f'(net_{jk}) \quad (\text{式 5-14})$$

当  $j$  不是输出节点, 则有

$$\delta_{jk} = \frac{\partial E_k}{\partial net_{jk}} = \frac{E_k}{O_{jk}} \frac{O_{jk}}{net_{jk}} = \frac{E_k}{O_{jk}} f'(net_{jk}) \quad (\text{式 5-15})$$

$$\begin{aligned} \frac{\partial E_k}{\partial O_{jk}} &= \sum_m \frac{\partial E_k}{\partial net_{mk}} \frac{\partial net_{mk}}{\partial O_{jk}} \\ &= \sum_m \frac{\partial E_k}{\partial net_{mk}} \frac{\partial}{\partial O_{jk}} \sum_i W_{mi} O_{ik} \\ &= \sum_m \frac{\partial E_k}{\partial net_{mk}} \sum_i W_{mj} = \sum_m \delta_{mk} W_{mj} \end{aligned} \quad (\text{式 5-16})$$

因此

$$\begin{cases} \delta_{jk} = f'(net_{jk}) \sum_m \delta_{mk} W_{mj} \\ \frac{\partial E_k}{\partial W_{ij}} = \delta_{mk} O_{ik} \end{cases} \quad (\text{式 5-17})$$

如果有 M 层，而 M 层含输入节点，第一层为输入节点，则 BP 算法为；

第一步，选取初始权值  $W$

第二步，重下述过程直至收敛

a. 对于  $k=1$  到 N

(1) 计算  $O_{ik}$ ,  $net_{jk}$  和  $\hat{y}_k$  的值(正向过程);

(2) 对各层从 M 到 2 反向计算(反向过程)

b. 对同一节点  $j \in M$ , 由式(34.1)和(34.2)计算  $\delta_{jk}$

第三步，修正权值， $W_{ij} = W_{ij} - \mu \frac{\partial E}{\partial W_{jk}}$ ,  $\mu > 0$ , 其中

$$\frac{\partial E}{\partial W_{jk}} = \sum_k^N \frac{\partial E}{\partial W_{jk}} \quad (\text{式 5-18})$$

从上述 BP 算法可以看出，BP 模型把一组样本 I/O 问题变成一个非线性优化问题，它使用的是优化中最普遍的梯度下降法。如果把神经网络看成输入到输出的映射，则这个映射是一个高度非线性映射。设计一个神经网络专家系统重点在于模型的构成和学习算法的选择。一般来说，结构是根据所研究领域及要解决的问题确定的。通过对所研究问题的大量历史资料数据的分析及目前的神经网络理论发展水平，建立合适的模型，并针对所选的模型采用相应的学习算法，在网络学习的过程中，不断的调整网络参数，直到输出结果满足要求。

### 5.2.3 采用 BP 神经网络模型的解算

对于 BP 神经网络模型的解算，我们没有采用传统的自编程序的方法，我们利用 DPS 软件,不但应用简单，而且与自编软件相比，可以使时间和空间代价更少。

DPS 软件提供多功能数理统计和数学模型处理软件系统，它将数值计算、统计分析、模型模拟以及画线制表等功能融为一体，兼有如 Excel 等流行电子表格软件系统和若干专业统计分析软件系统的功能。其中的神经网络工具箱可以进行 BP 网络算法等多种算法的解算。

## 六、模型应用

### 6.1 上海世博会文化影响力指数预测

利用建立的模型对本次上海世博会的文化影响力的量化指数进行计算。因此我们需要得到上海世博会期间的个因素的统计结果，我们通过中国旅游指导机构公布的 2010 年 6 月国际旅游收支情况见表 6-1 可以得到国际旅游方面的指标。但是由于世博会并未结束，所以有些长期统计因素的统计值并未公布，所以它们的值只能靠预测。

	计量单位	6 月	同比增长(%)	1-6 月	同比增长(%)
<b>旅游人数</b>	<b>万人次</b>	<b>1 088.30</b>	<b>9.41</b>	<b>6 551.99</b>	<b>5.57</b>
# 外 国 人	万人次	223.51	30.67	1 255.09	22.97
香港同胞	万人次	630.89	4.00	3 923.67	3.06
澳门同胞	万人次	189.84	3.11	1 125.80	-3.53
台湾同胞	万人次	44.06	33.64	247.43	17.13
<b>旅游人数</b>	<b>万人次</b>	<b>458.51</b>	<b>15.87</b>	<b>2 716.74</b>	<b>10.63</b>
# 外 国 人	万人次	182.30	32.67	1 022.58	24.97
香港同胞	万人次	206.56	3.86	1 293.06	2.72
澳门同胞	万人次	32.18	3.22	190.81	-3.49
台湾同胞	万人次	37.47	32.70	210.30	16.18
<b>旅游（外汇）收入</b>	<b>亿美元</b>	<b>37.94</b>	<b>23.36</b>	<b>218.62</b>	<b>16.05</b>
# 外 国 人	亿美元	22.91	32.56	128.52	24.86
香港同胞	亿美元	7.92	3.91	49.43	2.84
澳门同胞	亿美元	2.30	3.17	13.62	-3.51
台湾同胞	亿美元	4.82	32.76	27.05	16.23

表 6-1 2010 年 6 月国际旅游收支情况

对长期统计量的预测，我们采用曲线拟合的方法，外推插值出 2010 年的统计数据。以下我们以电影产量份额的因素作为算例进行说明。

我们通过过去 9 年的电影产量如表 6-2，通过插值的方法得出图 6-3，从图中可以直观的得到 2010 年的电影产量为 572.9 部

年 份	故事 片	动画 片	科学教育 片	纪录 片	合计（部）
2001	88	1	56	9	154
2002	100	2	60	7	169
2003	140	2	53	6	201
2004	212	4	30	10	256
2005	260	7	33	2	302
2006	330	13	36	13	392
2007	402	6	34	9	451
2008	406	16	39	16	477
2009	412	24	49	21	506

表 6-2 国内电影产量统计



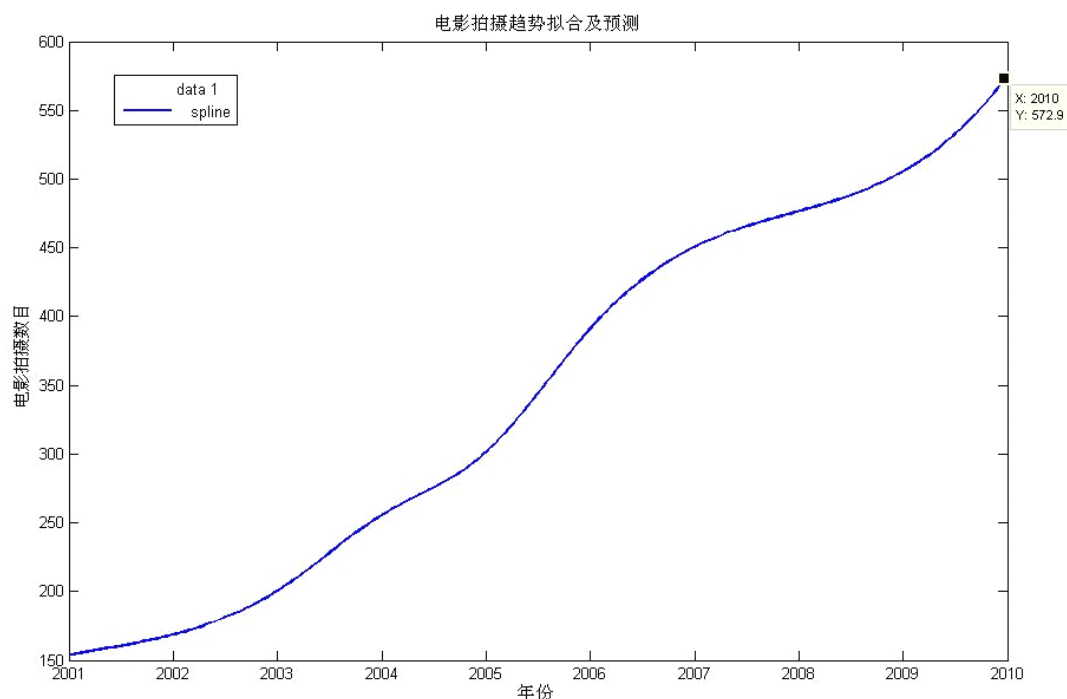


图 6-3 电影产量插值结果

其他未得到年度统计数据的因素也依照相同方法得到，结果如表 6-4

项目	算法	第三产业 GDP 数额（亿元）	占第三产业 GDP 比例	数额（亿元）	世界文化商品总贸易（亿元）	比例
文化商品贸易份额	文化商品贸易/文化商品总贸易	32591.202	0.0246	561.0743622	55007.29041	1.35%
文化服务贸易份额	文化服务贸易/文化服务总贸易	32591.202	0.0151	268.5711402	54810.43678	0.84%

图 6-4 2010 上海世博会文化影响力各决定因素预测 I

项目	算法	分子量	分母量	比值
国际旅游收支份额	国际旅游收支/国际旅游收支总和（百万美元）	15120	561713.1474	2.72%
国际旅游人次份额	入境旅游人次/出入境旅游总人次（万人）	960	1766	48.2%
图书出版种类份额	图书出版种类/世界图书出版种类（种）	183831	4418411	5.21%
电影产量份额	电影产量/世界电影产量（部）	573	2935	6.90%
劳动生产率	经济环境，GDP/劳动力（亿/万人）	32591.2	70432	11416
国际移民份额	社会环境，国际移民/世界移民总数（万人）	95.2	10022	1.23%

森林覆盖率	生态环境，森林面积/ 国土面积	17491	95998.90231	18.92%
-------	--------------------	-------	-------------	--------

表 6-5 2010 上海世博会文化影响力各决定因素预测 II

通过以上数据，分别利用两种模型算法，解得上海世博会的文化影响力指数，见表 6-6

	因子分析法	BP 神经网络
上海世博会文化影响力指数	3.1	3.2

表 6-6 上海世博会文化影响力预测

以上结果与沈阳园艺博览会和昆明园艺博览会的文化影响力指数对比如图 6-7、图 6-8

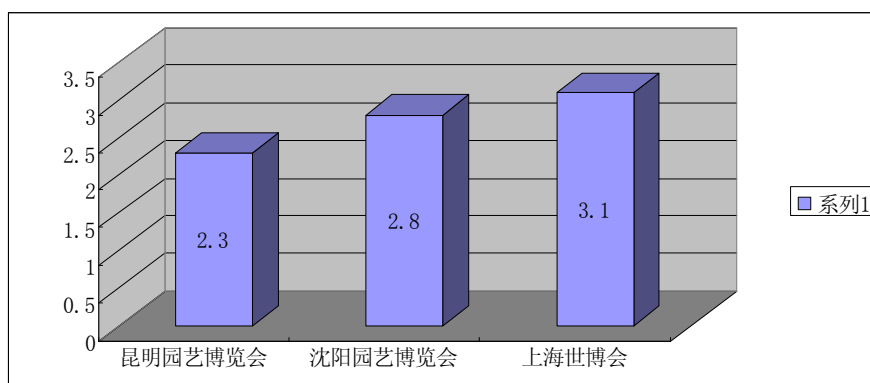


图 6-7 三次博览会文化影响力指数对比（因子分析法）

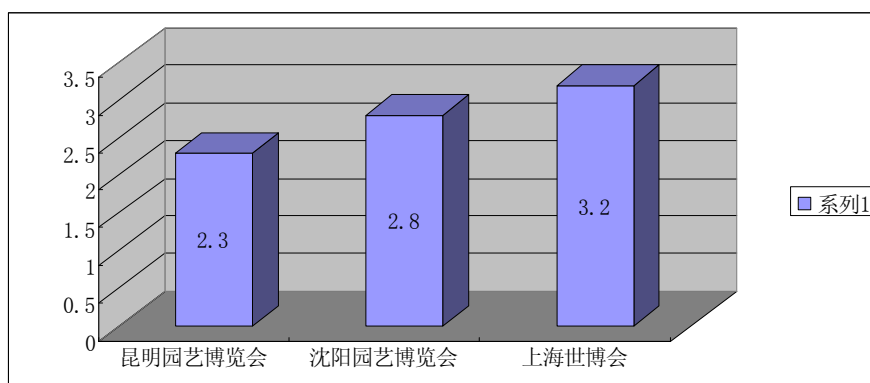


图 6-8 三次博览会文化影响力指数对比（BP 神经网络法）

## 6.2 BP 神经网络的解算残差分析

其中对 BP 神经网络得出拟合残差，可见拟合的近似程度很高，足以满足要求，如下图所示：

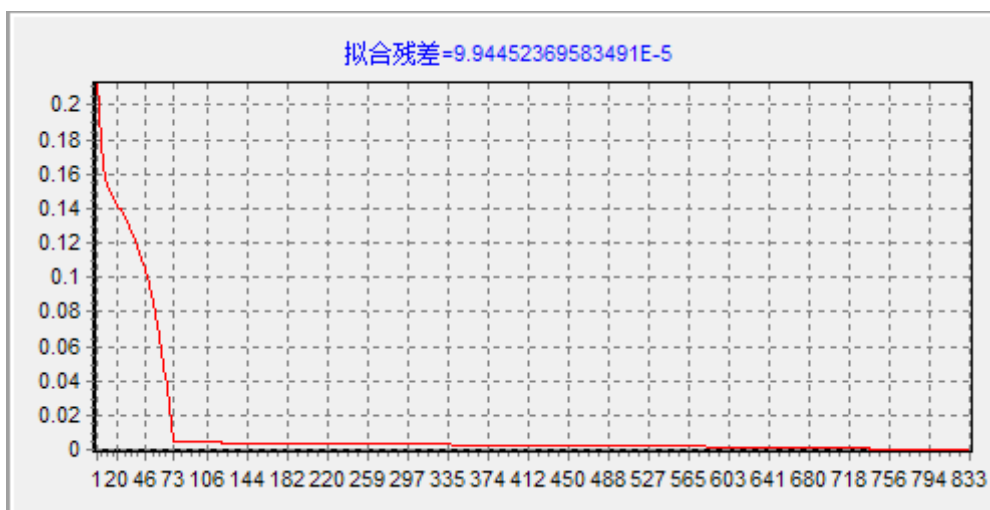


图 6-9 对上海世博会文化影响力 BP 算法残差分析

## 七、算法复杂性分析及模型的评价

### 7.1 复杂性分析

通过推导我们可解得第 $k+1$ 步的最优学习参数 $\alpha$ 、 $\beta$ 、 $\gamma$ ，由(式5-18)更新权值得到 $W(k+1)$ 使得误差函数 $E(\alpha, \beta, \gamma)$ 达到最小。从解的结果看，最优学习参数的表达式比较繁琐，但仔细观察(式5-17)式会发现，求解学习参数的计算量只与输出层的神经元个数有关，与网络的结构无关。而且(式5-15)——(式5-17)中 $\frac{\partial E_k}{\partial w}$ 的值在求 $\nabla E(k)$ 时可以直接得到，并不需要增加计算量。

由于学习参数对网络的训练有着很大的影响，许多研究者对此作了很多的研究。但是较多文献所提出的自适应学习率是根据学习误差通过一个放缩因子来进行调节，该调节方法是很粗糙的，有时会引起震荡现象。而通过本文提出的算法所得到的最优学习参数的估计是比较精确的，每一步迭代都是使误差函数达到最小。

### 7.2 模型评价

#### 7.2.1 因子分析模型的缺点

对因子分析模型和理论特别是小样本因子分析模型漏掉了方差贡献小的公共因子、没有除去零方差的特殊因子等情形进行讨论和说明。

#### 7.2.2 BP 神经网络模型的优点

运用 BP 神经网络评价文化影响力的优点主要表现在以下几个方面：

- ①避开权数、确定问题。如上述操作，在整个 BP 神经网络学习过程中没有复杂的人工计算过程，不需要计算衡量权数。因为在权数的定值上，有很大模糊性，BP 神经网络很好地避开了这点；
- ②指出症结所在。在一般的评价方法的结果中，只能见到一个数据或一个笼统的评价，很难或者不能发现企业文化中哪方面存在问题。在图 6-9 中可以直接明了地看出文化指标中的短板和需要改进的地方；
- ③运用 DPS 软件操作过程简单易行。整个操作过程不需要人工计算，只需要在 DPS 中输入数据。

#### 7.2.3 BP 神经网络模型的缺点

对初始权重非常敏感，极易收敛于局部极小，BP 算法本身就是一个优秀的局部搜索算法，加上 BP 对初始网络权重非常敏感，以不同的权重初始化网络，BP 算法会收敛于不同的局部极小。

其次是它往往停滞于误差梯度曲面的平坦区，收敛缓慢甚至不能收敛在误差梯度曲面的平坦区，误差梯度信息极小，每次对权重的改变量也极小，使得网络收敛及其缓慢，甚至网络不能收

敛。

### 7.3 模型可行性分析

虽然模型是完全客观准确的，但在应用于实际问题上也存在一些可行性问题。在实际问题中，各因素对文化影响力的贡献权重往往不是一成不变，而应该是时变的值，所以本文中利用已有观测（两届世界性博览盛会）对上海世博会的文化影响力中各指标的权重进行评定存在一定的制约。

此外，由于上海世博会未结束，导致一些评定文化影响力的指标数据不完善，并且只能用上海世博会期间某个月份的数据来代替整体上海世博会的影响，定量分析上会存在一定的片面性。

#### 【参考文献】

- [1] 姜启源，数学模型（第三版）[M]，北京：高等教育出版社，2003；
- [2] 韩中庚，数学建模竞赛，北京：科学出版社，2007；
- [3] 陈华溢，基于 BP 神经网络的企业文化影响力评价[J]，江苏，2008；
- [4] 中华人民共和国国家统计局，<http://www.stats.gov.cn/>，2010.9.11；
- [5] 中国发展门户网，中国文化影响力的评价、原理和指标，[http://cn.chinagate.cn/reports/whxdh/2009-01/24/content\\_17182808.htm](http://cn.chinagate.cn/reports/whxdh/2009-01/24/content_17182808.htm)，2010.9.11。
- [6] 世界银行 WDI 数据库

## 附录

### 3-6 第三产业增加值构成

本表按当年价格计算。

单位：%

行 业	1991	1992	1993	1994	1995	1996	1997	1998	1999
总 计	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
农、林、牧、渔服务业	0.7	0.7	0.7	0.7	0.6	0.6	0.8	0.8	0.8
地质勘查业水利管理业	1.1	1.1	1.2	1.3	1.4	1.4	1.3	1.2	1.2
交通运输、仓储及邮电									
通信业	19.5	18.4	18.7	18.0	17.0	17.1	16.5	16.4	16.5
交通运输和仓储业	17.5	16.3	16.1	14.8	13.2	12.9	11.7	11.5	11.3
邮电通信业	2.0	2.1	2.6	3.2	3.8	4.2	4.8	4.9	5.2
批发和零售贸易餐饮业	28.9	29.9	27.3	27.1	27.5	27.2	26.7	26.1	25.6
金融、保险业	17.8	17.5	18.2	18.5	19.4	19.7	19.7	18.6	17.9
房地产业	5.1	5.7	5.7	5.8	5.9	5.6	5.5	5.8	5.7
社会服务业	6.2	6.6	7.9	8.0	8.6	8.4	9.5	10.5	10.7
卫生体育和社会福利业	3.0	2.9	2.9	2.9	2.7	2.8	2.7	2.7	2.7
教育、文化艺术及广播									
电影电视业	6.3	6.0	6.3	6.5	6.3	6.6	6.8	7.2	7.8
科学研究和综合技术服									
务业	1.3	1.4	1.3	1.4	1.5	1.6	1.9	1.9	2.1
国家机关、政党机关和									
社会团体	9.2	8.9	8.7	8.6	8.0	7.9	7.7	7.8	8.1
其他行业	0.9	1.0	1.1	1.1	1.0	1.0	1.0	1.0	1.0

## 18-5 国际旅游外汇收入及构成

单位：百万美元

指 标	1999		2000	
	数 额	比 重	数 额	比 重
总计	14099	100.0	16224	100.0
长途交通	4165	29.6	4880	30.1
民航	3140	22.3	3502	21.6
铁路	445	3.2	596	3.7
汽车	382	2.7	596	3.7

轮船	198	1.4	186	1.1
游览	749	5.3	751	4.6
住宿	2034	14.4	2205	13.6
餐饮	1528	10.8	1523	9.4
商品销售	2771	19.7	3231	19.9
娱乐	845	6.0	1193	7.3
邮电通讯	411	2.9	503	3.1
市内交通	533	3.8	534	3.3
其他服务	1063	7.5	1404	8.7

### 18-1 旅 游 事 业 发 展 情 况

指 标	1996	1997	1998	1999	2000
旅行社总数 (个)	<b>4252</b>	<b>4986</b>	<b>6222</b>	<b>7326</b>	<b>8993</b>
国际旅行社	977	991	1312	1256	1268
国内旅行社	3275	3995	4910	6070	7725
旅行社职工人数(人)	<b>87555</b>	<b>94829</b>	<b>100448</b>	<b>108830</b>	<b>164336</b>
国际旅行社	53093	48881	52290	47153	68093
国内旅行社	34462	45948	48158	61677	96243
涉外饭店总数 (个)	<b>4418</b>	<b>5201</b>	<b>5782</b>	<b>7035</b>	<b>10481</b>
入境旅游人数 (万人)	<b>5112.75</b>	<b>5758.79</b>	<b>6347.84</b>	<b>7279.56</b>	<b>8344.39</b>
外国人	674.43	742.80	710.77	843.23	1016.04
华 侨	15.46	9.90	12.07	10.81	7.55
港澳同胞	4249.47	4794.33	5407.54	6167.06	7009.94
台湾同胞	173.39	211.76	217.46	258.46	310.86
三大旅行社接待					
入境旅游人数(万人)	<b>171.46</b>	<b>189.68</b>	<b>167.00</b>	<b>205.11</b>	<b>223.30</b>
中国国际旅行社	80.80	88.45	77.45	100.25	91.16
中国旅行社	59.25	66.94	53.79	65.10	88.32
中国青年旅行社	31.41	34.29	35.76	39.76	43.82
国内居民出境总人数 (万人)	<b>758.82</b>	<b>817.54</b>	<b>842.56</b>	<b>923.24</b>	<b>1047.26</b>
#因私出境人数	241.39	243.96	319.02	426.61	563.09
国内旅游人数(万人)	<b>63900</b>	<b>64400</b>	<b>69450</b>	<b>71900</b>	<b>74400</b>
旅游收入					

国际旅游收入	(亿美元)	102.00	120.74	126.02	140.99	162.24
国内旅游收入	(亿元)	1638.38	2112.70	2391.18	2831.92	3175.54

21-13 全国图书、杂志和报纸出版情况

年份 地区	图 书				杂 志				报 纸			
	种 数		总印数 (亿册、 张)	总印张数 (亿印张)	种 数 (种)	每期平均 印 数 (万册)	总印数 (亿册)	总印张数 (亿印 张)	种 数 (种)	每期平均 印 数 (万份)	总印数 (亿份)	总印张数 (亿印张)
	(种)	#新出版										
1978	14987	11888	37.7	135.4	930	6200	7.6	22.7	186	4280	127.8	113.5
1979	17212	14007	40.7	172.5	1470	7960	11.8	30.1		4762	130.8	123.0
1980	21621	17660	45.9	195.7	2191	10298	11.3	36.7	188	6236	140.4	141.7
1981	25601	19854	55.8	217.7	2801	13096	14.6	45.4	242	7152	140.7	133.6
1982	31784	23445	58.8	222.0	3100	13885	15.1	15.1	277	8074	140.0	129.1
1983	35700	25826	58.0	232.4	3415	15995	17.7	52.5	340	9611	155.1	142.7
1984	40072	28794	62.5	260.6	3907	20440	21.8	64.3	458	16246	180.8	162.3
1985	45603	33743	66.7	282.8	4705	23952	25.6	77.3	1445	19107	246.8	202.8
1986	51798	39426	52.0	220.3	5248	21980	24.0	68.1	1574	18387	242.7	203.5
1987	60213	42854	62.5	261.3	5687	24375	25.9	72.7	1611	20275	264.5	223.6
1988	65962	46774	62.3	269.0	5865	23275	25.5	71.2	1537	19788	267.8	231.3
1989	74973	55475	58.6	243.6	6078	17145	18.4	50.7	1576	15288	207.0	179.5
1990	80224	55245	56.4	232.1	5751	16156	17.9	48.1	1444	14670	211.3	182.8
1991	89615	58467	61.4	266.1	6056	18216	20.6	54.4	1524	16393	236.5	205.8
1992	92148	58169	63.4	280.4	6486	20506	23.6	62.7	1657	18031	257.9	238.8
1993	96761	66313	59.3	282.3	7011	20780	23.5	64.2	1788	18478	263.8	287.1
1994	103836	69779	60.1	297.2	7325	19763	22.1	63.9	1953	17736	253.2	310.8
1995	101381	59159	63.2	316.8	7583	19794	23.4	67.0	2089	17644	263.3	359.6
1996	112813	63647	71.6	360.5	7916	19300	23.1	68.1	2163	17877	274.3	392.4
1997	120106	66585	73.1	364.0	7918	20046	24.4	73.3	2149	18259	287.6	459.8
1998	130613	74719	72.4	373.6	7999	20928	25.4	79.9	2053	18211	300.4	540.0
1999	141831	83095	73.2	391.4	8187	21845	28.5	96.8	2038	18632	318.4	636.7
2000	143376	84235	62.7	376.2	8725	21544	29.4	100.0	2007	17914	329.3	799.8

21-8 摄 制 电 影 片 产 量

年份	电影故事片厂 (个)	故 事 片 (部)	美 术 片 (本)	科学教育片 (本)	纪 录 片 (本)
1952	4	4	2	41	157
1957	11	40	5	84	272
1962	16	34	17	94	133

1965	16	52	21	240	378
1975	15	27	11	214	313
1978	12	46	26	289	202
1979	17	65	25	349	317
1980	17	82	32	337	242
1981	19	105	33	277	276
1982	19	112	33	284	259
1983	19	127	37	343	299
1984	20	144	37	387	337
1985	20	127	45	357	419
1986	20	134	46	383	417
1987	22	146	45	353	347
1988	22	158	38	344	350
1989	22	136	53	334	259
1990	22	134	51	326	296
1991	22	130	46	351	283
1992	22	170	56	354	307
1993	22	154	47	252	300
1994	22	148	32	182	22
1995	30	146	37	40	111
1996	30	110	58	33	39
1997	31	88	28	34	95
1998	31	82	9	30	54
2000	31	91	1	49	10

注：本表电影故事片厂只包括国务院批准的厂。

## 5-1 就 业 基 本 情 况

项 目	1996	1997	1998	1999	2000
从业人员合计 (万人)	<b>68850</b>	<b>69600</b>	<b>69957</b>	<b>70586</b>	<b>71150</b>
第一产业	34769	34730	34838	35364	35575
第二产业	16180	16495	16440	16235	16009
第三产业	17901	18375	18679	18987	19566
从业人员构成(合计=100)					
第一产业	50.5	49.9	49.8	50.1	50.0
第二产业	23.5	23.7	23.5	23.0	22.5
第三产业	26.0	26.4	26.7	26.9	27.5
按城乡分从业人员(万人)					
城镇从业人员	19815	20207	20678	21014	21274
#国有单位	11244	11044	9058	8572	8102



城镇集体单位	3016	2883	1963	1712	1499
股份合作单位	-	-	136	144	155
联营单位	49	43	48	46	42
有限责任公司	-	-	484	603	687
股份有限公司	363	468	410	420	457
私营企业	620	750	973	1053	1268
港澳台商投资单位	265	281	294	306	310
外商投资单位	275	300	293	306	332
个体	1709	1919	2259	2414	2136
乡村从业人员	49035	49393	49279	49572	49876
#乡镇企业	13508	13050	12537	12704	12820
私营企业	551	600	737	969	1139
个体	3308	3522	3855	3827	2934
职工人数 (万人)	<b>14845</b>	<b>14668</b>	<b>12337</b>	<b>11773</b>	<b>11259</b>
国有单位	10949	10766	8809	8336	7878
城镇集体单位	2954	2817	1900	1652	1447
其他单位	942	1085	1628	1785	1935
城镇单位女性	-	-	-	<b>4613</b>	<b>4411</b>
从业人员 (万人)					
城镇登记失业人数 (万人)	<b>553</b>	<b>570</b>	<b>571</b>	<b>575</b>	<b>595</b>
城镇登记失业率 (%)	<b>3.0</b>	<b>3.1</b>	<b>3.1</b>	<b>3.1</b>	<b>3.1</b>

注：1.1990 年以后，经济活动人口、从业人员总计、城镇和乡村从业人员小计资料根据人口变动情况抽样调查调整，因此分地区、分经济类型、分行业的资料相加不等于总计。（下表同）。

2.1998 年及以后城镇单位从业人员、职工人数统计口径有调整，详见本篇末指标解释。