万向思维

特别合作 sina 新浪教育

倍速™

100+100+100 ?= 1000000

学习法

学习策略+漫画释义+综合应用+课后解答

高中数学 选修1-2

人教A版

总主编 刘增利

打造学科状元

北京出版社出版集团 BEIJING PUBLISHING HOUSE(GROUP)

北京教育出版社 BEIJING EDUCATION PUBLISHING HOUSE

特别合作 sina 新浪教育

倍速

100+100+100=1000000

学习法

高中数学 选修1－2

人教A版

总主编 刘增利

学科主编 杨文彬

本册主编 陈勤 韩廷蕴

编者 陈勤 李丽丽

北京出版社出版集团
BEIJING PUBLISHING HOUSE(GROUP)

北京教育出版社
BEIJING EDUCATION PUBLISHING HOUSE

编读交流平台

主编邮箱:zhubian@ wxsw. cn（任何疑问、意见或建议,皆请提出,我们是很虚心的。）

投稿邮箱:tougao@ wxsw. cn(想让大家分享你的学习心得和人生体验吗？快投稿吧!）

求购邮箱:qiugou@ wxsw. cn(什么书适合自己,在哪能买到？我们的选书顾问为你量身选择。）

图书质量监督电话:010－62380997　010－58572393　010－82378880(含图书内容咨询)

传真:010－62340468

销售服务短信:

中国移动用户发至　625551001

中国联通用户发至　725551001

小灵通用户发至　9255551001

建议咨询短信:

中国移动用户发至　625556018

中国联通用户发至　725556018

小灵通用户发至　9255556018

想知道更多的图书信息,更多的学习资源,请编辑手机短信"万向思维"发送至50120;想知道更多的考试信息,更多的学习方法,请编辑相应的手机短信"小学学习方法""初中学习方法"或"高中学习方法"发送至50120。

通信地址:北京市海淀区王庄路1号清华同方科技广场B座11层万向思维(邮编100083)。

最新"万向思维金点子"奖学金获奖名单

2006年12月10日

一等奖:

狄　欢(江苏溧阳)

二等奖:

秦文莉(安徽宿州)　周文颖(河北迁西)

熊秋艳(云南墨江)　方　莱(安徽蚌埠)

李　昊(河南潢川)　马建明(安徽阜南)

王晓楠(辽宁本溪)　常思佳(黑龙江明水)

樊昕阳(河南安阳)　陈佳莹(浙江慈溪)

……

2007年7月10日

一等奖:

周　政(甘肃庆阳)　李贵兵(陕西石泉)

二等奖:

张　雪(安徽寿县)　尹寒梅(四川岳池)　夏佳志(湖北孝感)　李文霞(青海湟中)

宁年宝(福建三明)　雷裕鹏(福建福安)　谭进艳(广东廉江)　郑　慧(海南儋州)

李莹莹(黑龙江嫩江)　司晗广(河南许昌)　卢建英(云南绿春)　伍冬林(四川南充)

吴栩莹(浙江上虞)　黄洁仪(广东大朗)　郭　磊(陕西咸阳)　何　攀(甘肃庆阳)

陈斯文(福建龙海)　铖东东(内蒙古赤峰)　胡承贤(江西宜春)　倪　燕(四川成都)

……

倍速学习法　高中数学选修1－2　人教A版

策划设计　北京万向思维基础教育教学研究中心数学教研组

总主编　刘增利

学科主编　杨文彬

本册主编　陈　勤　韩廷蕴

责任编辑　毕　伶

责任审读　陈　勤

责任校对　晁　鲁　曹　攀　彭凤珠

责任录排　李翠翠

封面设计　魏　晋

版式设计　廉　赢

插图作者　范金凤

出　　版　北京出版社出版集团
　　　　　北京教育出版社

发　　行　北京出版社出版集团

印　　刷　陕西思维印务有限公司

经　　销　各地书店

开　　本　890×1240　1/32

印　　张　8

字　　数　224千字

版　　次　2007年11月第1版

印　　次　2007年11月第1次印刷

书　　号　ISBN 978－7－5303－6146－7/G·6065

定　　价　10.80元

版权所有　翻印必究

目录

第一章 统计案例

第二章 推理与证明

CONTENTS

目录

CONTENTS

此为试读,需要完整PDF请访问: www.ertongbook.com

目录

第三章 数系的扩充与复数的引入

CONTENTS

目录

CONTENTS

第一章　统计案例

化归的思想方法

在解决数学问题时，常遇到一些问题直接求解较为困难，需将原问题转化为一个（相对来说，自己较熟悉的、相对原问题较简单的）新问题，通过新问题的求解达到解决问题的目的，这一思想方法称为“化归”.

本章整体感知

数理统计学是收集、分析、研究数据，并对所研究的问题作出结论的科学.

本章是在学习过抽样、样本估计总体、相互独立事件同时发生的概率、线性回归等基本知识的基础上提出来的. 本章中我们将通过对案例的讨论，进一步讨论线性回归分析方法及其应用，并初步了解独立检验的基本思想，认识统计方法在决策中的作用.

本章的主要内容是回归直线方程的分析及相关检验、独立性检验两部分. 其重点是回归分析的初步应用，难点是独立性检验的应用. 我们要注重对案例的分析，对数据的收集、整理、分析，要注意理论联系实际. 本章所使用的数学思想方法有统计、推理、分析等.

1.1　回归分析的基本思想及其初步应用

知识结构 · 理清知识脉络

- 回归分析
  - 随机误差
  - 样本点的中心
  - 样本相关系数
  - 残差分析 — 相关指数 R^2
  - 建立回归模型

自主学习 · 享受探究乐趣

一、新知导入

忆旧（知识回顾）

1. 抽样方法

一般地，设一个总体含有 N 个个体，从中逐个不放回地抽取 n 个个体作为样本（$n<N$），如果每次抽取时总体内各个个体被抽取的机会都相等，就把这种抽样方法叫做简单随机抽样.

2. 用样本估计总体

一般地，用样本估计总体可分成两种：一种是用样本的频率分布估计总体的频率分布；另一种是用样本的数字特征（如平均数、标准差）估计总体的数字特征.

3. 回归直线方程

如果散点图中点的分布从整体上看大致在一条直线附近，我们称这两个变量之间具有线性相关关系，这条直线叫做回归直线，这条回归直线的方程叫做回归直线方程.

迎新（问题引入）

在实际问题中，我们常常会遇到多个变量同处于一个过程之中，它们相互联系、互相制约. 有的变量之间有完全确定的函数关系，还有一些变量，它们之间也有一定的关系，然而这种关系并不完全确定，它们的关系就不能用一个确定的函数关系式表达出来.

一般地，把两个变量分为解释变量 x 与预报变量 y，作出散点图，从点的分布特征来判断它们是否线性相关. 若线性相关，则可能用回归直线方程来解决相关的实际问题. 然而，由于存在误差，有时很难说这些点是否分布在同一条直线的附近，这时就很难判断这两个变量之间是否具有相关关系，因此单纯地用散点图判断主观性太强. 在这种情况下，我们通常在回归分析时用相关系数来检验两个变量相关关系的强弱.

二、教材详析

知识点 1. 回归方程

(1)线性回归模型和随机误差

对于一组具有线性相关关系的数据$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$,其散点图中,样本点分布在某一条直线附近,但又不是在一条直线上,所以不能用一次函数$y=bx+a$来描述它们之间的关系,这时我们采用线性回归模型

$$y=bx+a+e$$

来描述它们之间的关系.其中,a和b为模型的未知参数,e称为随机误差.

预报变量y由解释变量x和随机误差e共同确定.

(2)回归直线方程的系数估计和样本点的中心

这组线性相关的数据$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$的回归直线方程$\tilde{y}=bx+a$的斜率和截距的最小二乘法估计公式分别为:

$$\hat{b}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2},$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}.$$

其中$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i,(\bar{x},\bar{y})$称为样本点的中心.

我们用直线$\hat{y}=\hat{b}x+\hat{a}$来近似地估计回归直线$\tilde{y}=bx+a$.

特别提示:

(1)回归直线过样本点的中心$(\bar{x},\bar{y})$.

(2)在线性回归模型中,随机误差e的方差σ^2越小,通过回归直线$\tilde{y}=bx+a$预报真实值y的精确度越高.

(3)引起预报值$\hat{y}$与真实值y之间的误差的原因:一是随机误差e,它引起真实值y和回归直线上的点$\tilde{y}$之间的误差;另一方面,由于$\hat{a}$和$\hat{b}$为截距和斜率的估计值,它们与真实值a和b之间也存在着误差,它们引起了预报值$\hat{y}$和真实回归直线上的点的$\tilde{y}$之间的差异.

(4)回归系数$\hat{b}$也可以用公式$\hat{b}=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}$来计算,其推导公式如下:

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}(x_iy_i-x_i\bar{y}-\bar{x}y_i+\bar{x}\bar{y})$$

$$=\sum_{i=1}^{n}x_iy_i-\bar{y}\sum_{i=1}^{n}x_i-\bar{x}\sum_{i=1}^{n}y_i+n\bar{x}\bar{y}$$

$$=\sum_{i=1}^{n}x_iy_i-\bar{y}(n\bar{x})-\bar{x}(n\bar{y})+n\bar{x}\bar{y}=\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y},$$

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}(x_i^2-2x_i\bar{x}+\bar{x}^2)=\sum_{i=1}^{n}x_i^2-2\bar{x}\sum_{i=1}^{n}x_i+n\bar{x}^2$$

$$=\sum_{i=1}^{n}x_i^2-2\bar{x}(n\bar{x})+n\bar{x}^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2,$$

$$\therefore \hat{b}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}.$$

(5)回归方程只适用于我们所研究的样本的总体；样本的取值范围会影响回归方程的适用范围，一般不能超过这个适用范围，否则将没有实用价值；回归方程一般都具有时间性.

知识点2. 衡量线性相关关系的强弱

样本相关系数 r 可以用来衡量两个变量之间的线性相关关系，其计算公式为

$$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}.$$

$r>0$ 表明两个变量正相关，$r<0$ 表明它们负相关. $|r|$越接近1，表明两个变量的线性相关性越强；$|r|$越接近0，表明它们之间几乎不存在线性相关关系. 通常当$|r|>0.75$时，认为两个变量有很强的线性相关关系.

特别提示：

r 的公式也可以写成 $r=\dfrac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sqrt{(\sum_{i=1}^{n}x_i^2-n\bar{x}^2)(\sum_{i=1}^{n}y_i^2-n\bar{y}^2)}}$. 这由 $\hat{b}$ 的推导过程可以得出.

知识点3. 偏差平方和分解

(1)总偏差平方和

解释变量和随机误差的组合效应造成了观测值 y_i 和总的平均值 $\bar{y}$ 之间的差异，这个总效应用总偏差平方和

$$\sum_{i=1}^{n}(y_i-\bar{y})^2$$

来表示.

(2)残差

数据点和它在回归直线 $\hat{y}=\hat{b}x+\hat{a}$ 上相应位置的差异$(y_i-\hat{y}_i)$是随机误差的效应，称

$$\hat{e}_i=y_i-\hat{y}_i=y_i-\hat{b}x_i-\hat{a},i=1,2,\cdots,n$$

此为试读，需要完整PDF请访问：www.ertongbook.com

为相应于点(x_i,y_i)的残差.

(3)残差平方和

残差平方和

$$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$$

代表了随机误差的效应.

(4)回归平方和

解释变量的效应用回归平方和

$$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$$

来表示.

(5)偏差平方和分解公式

总偏差平方和 = 回归平方和 + 残差平方和.

即
$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2+\sum_{i=1}^{n}(y_i-\hat{y}_i)^2.$$

特别提示：

(1)总偏差平方和刻画了预报变量y的变化程度.

(2)在回归平方和中，所有预测值$\hat{y}_i(i=1,2,\cdots,n)$的平均值也等于$\bar{y}$，即$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i$. 所以回归平方和又可以写成$\sum_{i=1}^{n}\left(\hat{y}_i-\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i\right)^2$，它刻画了估计量$\hat{y}=\hat{a}+\hat{b}x$的变化程度. 因为估计量由解释变量$x$决定，所以回归平方和刻画了预报变量的变化中由解释变量通过线性回归模型所引起的那部分变化程度.

(3)残差平方和刻画了残差变量变化的程度.

(4)预报变量的变化程度可以分解为由解释变量引起的变化程度与残差变量的变化程度之和.

知识点4. 残差分析

在研究两个变量间的关系时，首先要根据散点图来粗略判断它们是否线性相关，是否可以用线性回归模型来拟合数据. 然后，可以通过残差

$$\hat{e}_1,\hat{e}_2,\cdots,\hat{e}_n$$

来判断模型拟合的效果，判断原始数据中是否存在可疑数据. 这方面的分析工作就称为残差分析.

(1)残差分析的一般方法

①作残差图. 作图时，纵坐标为残差，横坐标可以选为样本编号，或有关数据(如身高数据、体重的估计值等)，这样作出的图形称为残差图.

若残差点比较均匀地落在水平带状区域中，说明选用的模型比较合适，这样的带状区域越窄，说明模型的拟合精度越高，回归方程的预报精度也越高；若残差点分布

不均匀,首先应确认采集的标本是否有误,如果有误,就予以纠正,然后再重新利用线性回归模型来拟合数据,如果数据采集没有错误,则寻找其他原因.

②利用相关指数 R^2 来刻画回归效果.

R^2 的计算公式:$R^2=1-\dfrac{\sum\limits_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i-\bar{y})^2}$.

即:$R^2=1-\dfrac{\text{残差平方和}}{\text{总偏差平方和}}=\dfrac{\text{回归平方和}}{\text{总偏差平方和}}$.

R^2 的取值越大,说明残差的平方和越小,即随机误差产生的效应越小,说明模型的拟合效果越好. 在线性回归模型中,R^2 表示解释变量对于预报变量的贡献率,可以说预报变量 y 的变化中的 $100R^2\%$ 是由解释变量 x 引起的,或者说解释变量 x 可以解释预报变量 y 的 $100R^2\%$ 的变化. 因此,R^2 越接近 1,表示回归的效果越好.

在线性回归模型中,R^2 越接近 1,表示解释变量和预报变量的线性相关性越强.

(2)在含有一个解释变量的线性回归模型中,相关指数 R^2 恰好等于相关系数 r 的平方,即 $R^2=r^2$.

$\because -1\leqslant r\leqslant 1,\therefore 0\leqslant R^2\leqslant 1$.

在一元线性回归模型中,相关指数 R^2 和相关系数 r 都能刻画用线性回归模型拟合数据的效果.

当 $r=\pm 0.9$ 时,$R^2=0.81$. 通常当 $R^2>0.80$ 时,认为线性回归模型对该数据是很有效的,这时两个变量的相关系数的绝对值几乎超过 0.9.

特别提示:

(1)残差 e 受许多条件的影响,也受我们所选用的线性模型的影响.

(2)作残差图有时不够精确,也难于区分拟合效果的好坏,因此多数情况下,选用计算相关指数 R^2 来说明拟合效果.

知识点 5. 建立回归模型

(1)建立回归模型的基本步骤:

①确定研究对象,明确哪个变量是解释变量,哪个变量是预报变量;

②画出确定好的解释变量和预报变量的散点图,观察它们之间的关系(如是否存在线性关系等);

③由经验确定回归方程的类型(如数据呈线性关系,则选用线性回归方程 $y=bx+a$);

④按照一定规则估计回归方程中的参数(如最小二乘法);

⑤得出结果后,分析残差图是否有异常(个别数据对应残差过大,或残差呈现不随机的规律性等),若存在异常,则检查数据是否有误,或模型是否合适等.

(2)比较多个模型的拟合效果

①用模型的残差的平方和的大小来判断模型的拟合效果. 残差平方和越小的模型,拟合的效果越好.

②用相关指数 R^2 来比较模型的拟合效果，R^2 越大，模型的拟合效果越好.

特别提示：

由回归方程得到的预报值 $\hat{y}$ 是预报变量的可能取值的平均值，而不是预报变量的精确值.

解题方法 · 乘坐智慧快车

一、基础经典全析

题型 1　回归分析

例 1 为了了解某一地区母亲身高 x 与女儿身高 y 的相关关系，随机测得 10 对母女的身高如下表所示：

母亲身高 x(cm)	159	160	160	162	159	154	158	159	157	158
女儿身高 y(cm)	158	159	160	161	161	156	157	160	157	158

试对 x 与 y 进行一元线性回归分析，并预报当母亲身高为 161 cm 时，女儿的身高为多少.

分析：为了寻找两个随机变量间的线性关系，一般先作散点图. 把这 10 对数据画出散点图，如图 1－1－1 所示，由图可以看出，x 与 y 之间有近似的线性相关关系，据此图用回归直线方程解决问题.

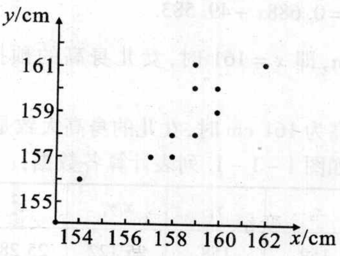

图 1－1－1

解法一：画出散点图如图 1－1－1，可以看出样本点呈条状分布，母亲的身高和女儿的身高有较好的线性相关关系.

$\bar{x}=\frac{1}{10}\sum_{i=1}^{10}x_i=158.6$，

$\bar{y}=\frac{1}{10}\sum_{i=1}^{10}y_i=158.7$.

计算数据如下表：

x_i	159	160	160	162	159	154	158	159	157	158
$x_i-\bar{x}$	0.4	1.4	1.4	3.4	0.4	-4.6	-0.6	0.4	-1.6	-0.6
$(x_i-\bar{x})^2$	0.16	1.96	1.96	11.56	0.16	21.16	0.36	0.16	2.56	0.36
y_i	158	159	160	161	161	156	157	160	157	158
$y_i-\bar{y}$	-0.7	0.3	1.3	2.3	2.3	-2.7	-1.7	1.3	-1.7	-0.7
$(y_i-\bar{y})^2$	0.49	0.09	1.69	5.29	5.29	7.29	2.89	1.69	2.89	0.49

$\therefore \sum_{i=1}^{10}(x_i-\bar{x})^2=40.4,\sum_{i=1}^{10}(y_i-\bar{y})^2=28.1,\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y})=27.8.$

$$\therefore r=\frac{\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{10}(x_i-\bar{x})^2\sum_{i=1}^{10}(y_i-\bar{y})^2}}=\frac{27.8}{\sqrt{40.4\times28.1}}\approx0.825.$$

$\because 0.825>0.75$，$\therefore$ 可以认为 x 和 y 之间有很强的线性相关关系.

$$\hat{b}=\frac{\sum_{i=1}^{10}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{10}(x_i-\bar{x})^2}=\frac{27.8}{40.4}\approx0.688,$$

$$\hat{a}=\bar{y}-\hat{b}\bar{x}\approx158.7-0.688\times158.6\approx49.583.$$

$\therefore$ 回归直线方程为 $\hat{y}=0.688x+49.583$.

当母亲身高为 161 cm，即 $x=161$ 时，女儿身高的预报值为 $\hat{y}\approx0.688\times161+49.583=160.351$(cm).

这就是说，当母亲身高为 161 cm 时，女儿的身高大致是 160.4 cm.

解法二：画出散点图如图 1-1-1. 列表计算各数据：

序号	x_i	y_i	x_iy_i	x_i^2	y_i^2
1	159	158	25 122	25 281	24 964
2	160	159	25 440	25 600	25 281
3	160	160	25 600	25 600	25 600
4	162	161	26 082	26 244	25 921
5	159	161	25 599	25 281	25 921
6	154	156	24 024	23 716	24 336
7	158	157	24 806	24 964	24 649
8	159	160	25 440	25 281	25 600

续表

10	158	158	24 964	24 964	24 964
Σ	1 586	1 587	251 726	251 580	251 885

$\therefore \bar{x}=\frac{1}{10}\times 1\ 586=158.6, \bar{y}=\frac{1}{10}\times 1\ 587=158.7, 10\bar{x}\bar{y}=10\times 158.6\times 158.7$ $=251\ 698.2, 10\bar{x}^2=10\times 158.6^2=251\ 539.6, 10\bar{y}^2=10\times 158.7^2=251\ 856.9.$

$$\therefore r=\frac{\sum_{i=1}^{10}x_iy_i-10\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{10}x_i^2-10\bar{x}^2\right)\left(\sum_{i=1}^{10}y_i^2-10\bar{y}^2\right)}}$$

$$=\frac{251\ 726-251\ 698.2}{\sqrt{(251\ 580-251\ 539.6)(251\ 885-251\ 856.9)}}\approx 0.825.$$

$\because 0.825>0.75, \therefore$ 可以认为 x 和 y 有很强的线性相关关系.

$$\hat{b}=\frac{\sum_{i=1}^{10}x_iy_i-10\bar{x}\bar{y}}{\sum_{i=1}^{10}x_i^2-10\bar{x}^2}=\frac{251\ 726-251\ 698.2}{251\ 580-251\ 539.6}\approx 0.688,$$

以下同解法一.

点拨:线性回归分析的步骤方法为:①首先作出统计假设;②求出线性相关系数;③由相关系数确定回归直线方程是否有意义;④写出线性回归方程,解决有关问题.

例 2 在彩色显影中,由经验可知:形成染料光学密度 y 与析出银的光学密度 x 由公式 $y=Ae^{\frac{b}{x}}(b<0)$ 表示. 现测得试验数据如下表:

x_i	0.05	0.06	0.25	0.31	0.07	0.10	0.38	0.43	0.14	0.20	0.47
y_i	0.10	0.14	1.00	1.12	0.23	0.37	1.19	1.25	0.59	0.79	1.29

试求 y 对 x 的回归方程.

分析:从题意可知,这不是一个线性回归分析问题,而是一个非线性回归分析问题,由于题目中已给定了要求的曲线为 $y=Ae^{\frac{b}{x}}$ 类型,我们只要将它转化为线性方程,再通过所给的 11 对样本数据,求出 A 和 b 的值即可确定 x 与 y 的相关关系的曲线方程.

解:由题意知,对于给定的公式 $y=Ae^{\frac{b}{x}}(b<0)$ 两边取自然对数,得 $\ln y=\ln A+\frac{b}{x}$.

与线性回归方程相对照可以看出,只要取 $u=\frac{1}{x}, v=\ln y, a=\ln A$,就有 $v=a+bu$. 这是 v 对 u 的线性回归直线方程,对此我们再套用相关性检验,求回归系数 b 和 a. 题目中所给的数据由变量置换 $u=\frac{1}{x}, v=\ln y$,变为如下表所示的数据:

u_i	20.000	16.667	4.000	3.226	14.286	10.000
v_i	−2.303	−1.966	0	0.113	−1.470	−0.994
u_i	2.632	2.326	7.143	5.000	2.128	
v_i	0.174	0.223	−0.528	−0.236	0.255	

可以求得 $r\approx-0.998$. 由于 $|r|\approx0.998>0.75$，可知，u 与 v 具有很强的线性相关关系.

计算得 $\hat{b}\approx-0.146$，$\hat{a}\approx0.548$，所以 $\hat{v}=0.548-0.146u$.

把 u 和 v 置换回来可得，$\ln\hat{y}=0.548-\dfrac{0.146}{x}$.

所以 $\hat{y}=e^{0.548-\frac{0.146}{x}}=e^{0.548}\cdot e^{-\frac{0.146}{x}}\approx1.730e^{-\frac{0.146}{x}}$.

所以回归曲线方程为 $\hat{y}=1.730e^{-\frac{0.146}{x}}$.

点拨：由于我们已经掌握了线性回归分析的思想方法，所以解决本题的思路是通过适当的变量置换，把非线性回归方程转化为线性回归方程，然后再套用线性回归分析的解题步骤即可.

题型 2　残差分析

例 3 在一段时间内，某种商品的价格 x（元）和需求量 y（件）之间的一组数据为：

价格 x	14	16	18	20	22
需求量 y	56	50	43	41	37

求出 y 对 x 的回归直线方程，并说明拟合效果的好坏.

分析：依据题中数据求出 y 对 x 的回归直线方程，从而列出残差表，根据公式 $R^2=1-\dfrac{\sum\limits_{i=1}^{5}(y_i-\hat{y}_i)^2}{\sum\limits_{i=1}^{5}(y_i-\bar{y})^2}$，计算出相关指数 R^2，结合 R^2 的大小即可得出结论.

解：$\bar{x}=\dfrac{1}{5}\times(14+16+18+20+22)=18$，

$\bar{y}=\dfrac{1}{5}\times(56+50+43+41+37)=45.4$，

$\sum\limits_{i=1}^{5}x_i^2=14^2+16^2+18^2+20^2+22^2=1\ 660$，

$\sum\limits_{i=1}^{5}x_iy_i=14\times56+16\times50+18\times43+20\times41+22\times37=3\ 992$.

所以 $\hat{b}=\dfrac{\sum\limits_{i=1}^{5}x_iy_i-5\bar{x}\bar{y}}{\sum\limits_{i=1}^{5}x_i^2-5\bar{x}^2}=\dfrac{3\ 992-5\times18\times45.4}{1\ 660-5\times18^2}=-2.35$，

所以 $\hat{a}=\bar{y}-\hat{b}\bar{x}=45.4-(-2.35)\times18=87.7$.

此为试读，需要完整PDF请访问：www.ertongbook.com