王后雄学案

教材完全解读

选修·专题

6大特色引发学考热潮
推动学习模式全面升级

独具创意 ☑ 同步突破 ☑
考向指引 ☑ 典例导思 ☑
考试工具 ☑ 核心预测 ☑

高中数学 选修1-2

丛书主编：王后雄
本册主编：童亚新 赵志培

小熊图书
接力出版社 Jieli Publishing House
全国百佳图书出版单位 Top 100 publishing house in China

王后雄学案

教材完全解读

总策划：熊 辉

选修·专题

高中数学 选修1-2

丛书主编：王后雄
本册主编：童亚新 赵志培
编 委：胡建平 丁仁贵
王 涛 汪 飞
林菊芳

小熊图书
接力出版社 Jieli Publishing House
全国百佳图书出版单位 Top 100 publishing house in China

图书在版编目（CIP）数据

教材完全解读. 高中数学. 1-2：选修 / 童亚新主编.—5版.—南宁：接力出版社，2011.9
ISBN 978-7-80732-471-3

Ⅰ.①教… Ⅱ.①童… Ⅲ.①中学数学课—高中—教学参考资料 Ⅳ.①G634

中国版本图书馆CIP数据核字（2011）第192199号

总 策 划：熊　辉
责任编辑：吴惠娟
责任校对：陈　娟
封面设计：木头羊

JIAOCAI WANQUAN JIEDU
GAOZHONG SHUXUE
教材完全解读
高中数学　选修1-2
丛书主编：王后雄　本册主编：童亚新　赵志培
*
社　长：黄　俭　总编辑：白　冰
接力出版社出版发行
广西南宁市园湖南路9号　邮编：530022
E-mail：jielipub@public.nn.gx.cn
河南省瑞光印务股份有限公司印刷　全国新华书店经销
*
开本：889毫米×1194毫米　1/16　印张：10　字数：264千
2011年10月第5版　2011年10月第6次印刷
ISBN 978-7-80732-471-3
定价：16.30元

教材完全解读

本书特点

1. 以《课程标准》、《考试大纲》为编写依据，完全解读知识、方法、能力、考试题型，全面提高学习成绩。

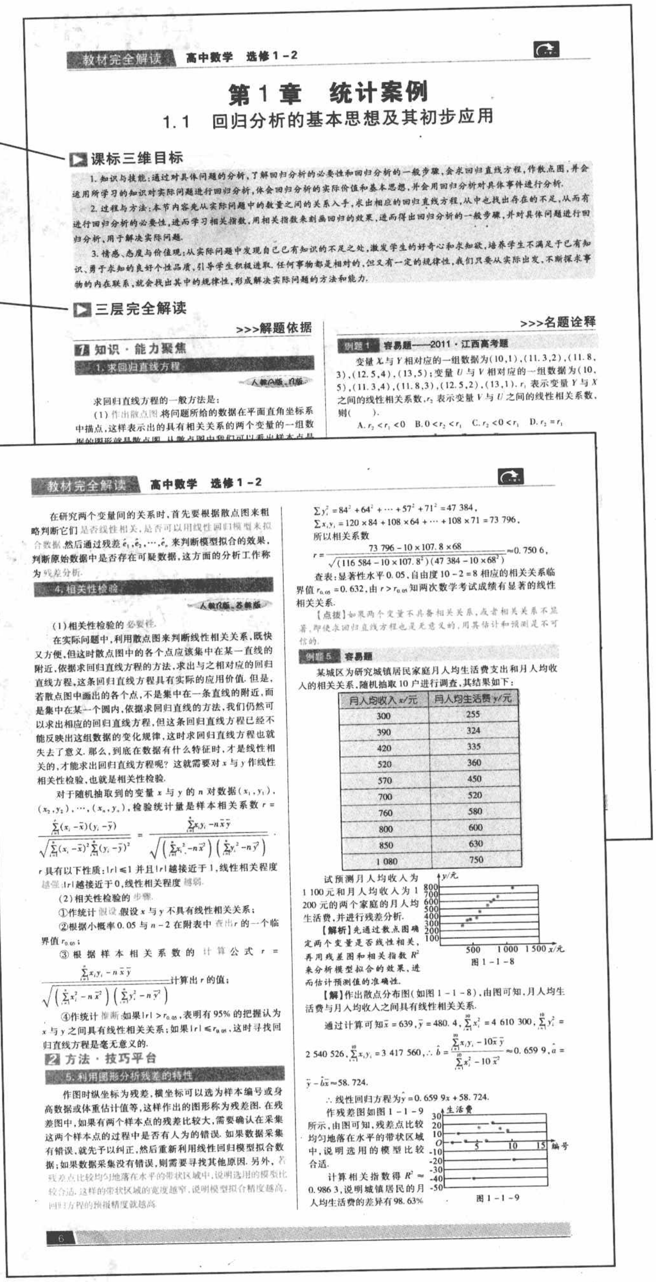
2. 采用国际流行的双栏对照案例编写方式，左栏对教材全解全析，在学科层次上力求讲深、讲透、讲出特色；右栏用案例诠释考点，对各个考点各个击破。

明确每课学习要求

以课标为依据，三维目标全解教材学习要求，提供总体的学习策略，提出具体的学习要诀，体现目标控制学习规则。

三层完全解读

从知识、方法、思维三个方面诠释教材知识点和方法点，帮您形成答题要点、解题思维，理清解题思路，揭示考点实质和内涵。


教材完全解读　高中数学　选修1－2

第1章　统计案例

1.1　回归分析的基本思想及其初步应用

课标三维目标

三层完全解读

>>>解题依据　>>>名题诠释

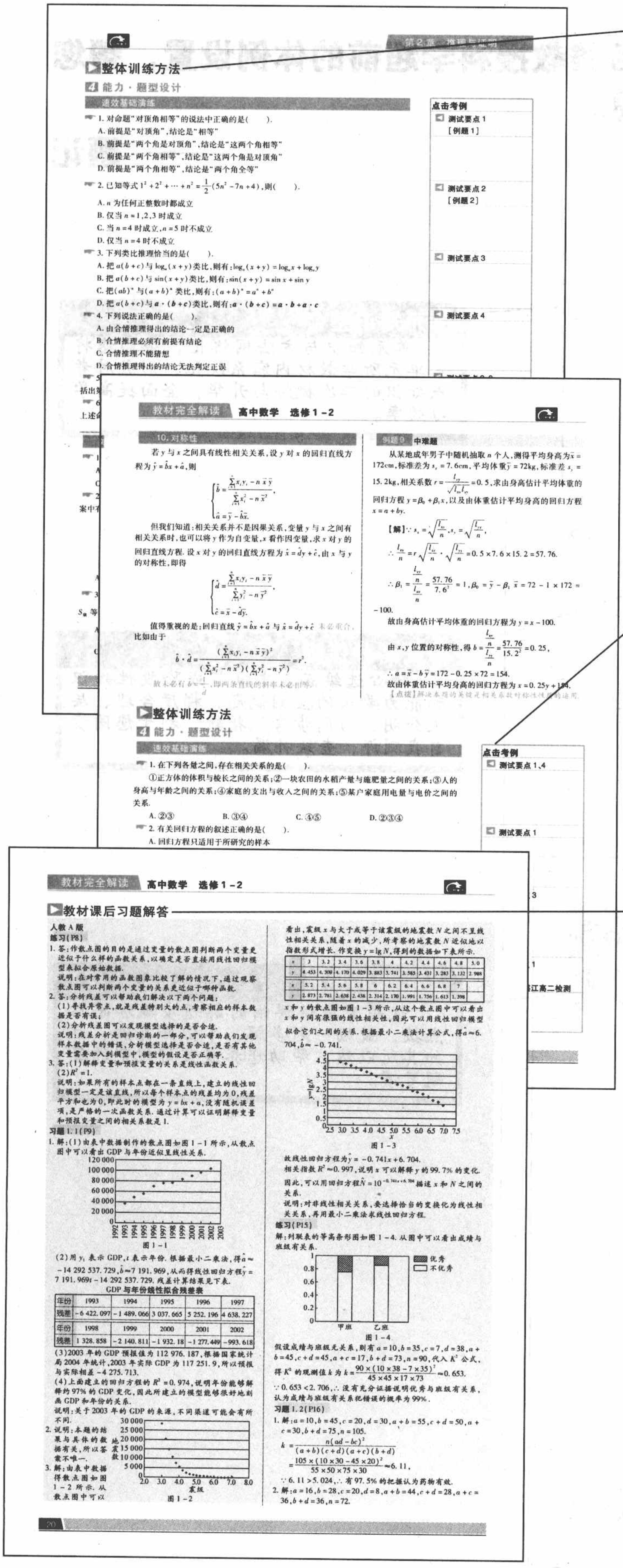
1 知识·能力聚焦

1. 求回归直线方程

4. 相关性检验

2 方法·技巧平台

5. 利用图形分析残差的特性



整体训练方法

针对本节重点、难点、考点及考试能力达标所设计的题目。题目难度适中,是形成能力、考试取得高分的必经阶梯。

解题错因导引

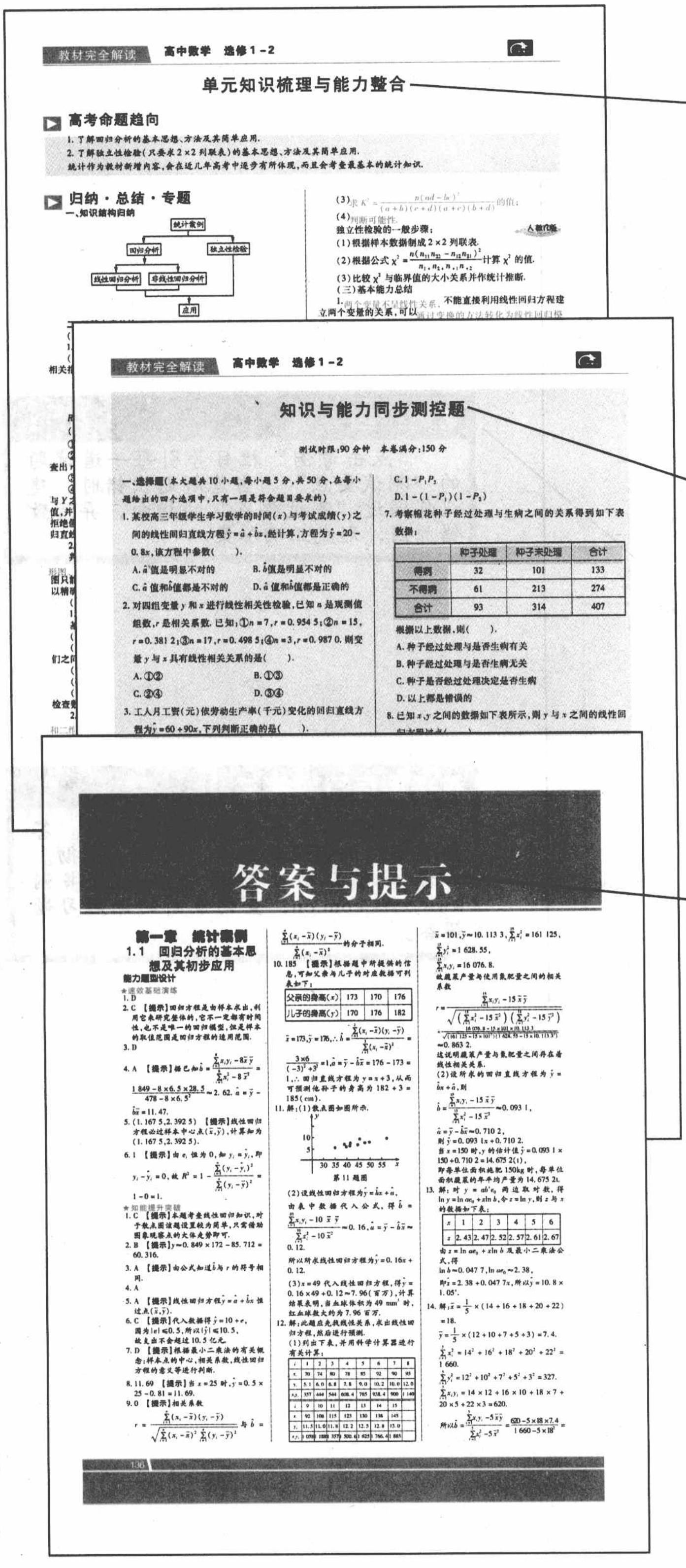
“点击考例”栏目导引每一道试题的“测试要点”。当您解题出错时，建议您通过“测试要点”的指向，弄清致错原因，形成正确答案。

教材课后习题解答

帮助您弥补课堂上听课的疏漏。答案准确，讲解繁简适度、到位、透彻。如课本有变动，请登陆小熊图书网www.xxts.com.cn，查询最新的课后习题解答。

教辅大师、特级教师王后雄教授科学超前的体例设置，帮您赢在学习起点，成就人生夙愿。

——题记



单元知识整合

单元知识与方法网络化，帮助您将本单元所学教材内容系统化，形成对考点知识的二次提炼与升华，全面提高学习效率。

考试高分保障

精心选编涵盖本章节或阶段性知识和能力要求的检测试题，梯度合理、层次分明，与同步考试接轨，利于您同步自我测评，查缺补漏。

点拨解题思路

试题皆提供详细的解题步骤和思路点拨，鼓励一题多解。不但知其然，且知其所以然，帮助您养成良好规范的答题习惯。

目录

目录

知识与方法 阅读索引

模块学习指南

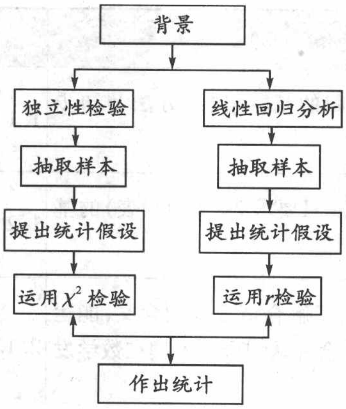
“课程标准”与“完全解读”内容对照表

序号	课程标准(高中数学选修 1－2)	完全解读内容＊页码
1	通过对典型案例的探究,进一步了解回归的基本思想、方法及初步应用	1.1　回归分析的基本思想及其初步应用＊P_4
2	通过对典型案例的探究,了解独立性检验(只要求 2×2 列联表)的基本思想、方法及初步应用	1.2　独立性检验的基本思想及其初步应用＊P_{13}
3	结合已学过的数学实例和生活中的实例,了解合情推理的含义,能运用归纳和类比等方法进行简单的推理,体会并认识合情推理在数学发展中的作用	2.1.1　合情推理＊P_{28}
4	结合已学过的数学实例和生活中的实例,体会演绎推理的重要性,掌握演绎推理的基本方法,并能运用它们进行一些简单推理. 通过具体实例,了解合情推理和演绎推理之间的联系和差异	2.1.2　演绎推理＊P_{39}
5	结合已经学过的数学实例,了解直接证明的两种基本方法:分析法和综合法,并了解分析法和综合法的思考过程和特点	2.2.1　综合法和分析法＊P_{50}
6	结合已经学过的数学实例,了解间接证明的一种基本方法——反证法,并了解反证法的思考过程和特点	2.2.2　反证法＊P_{62}
7	在问题情境中了解数系的扩充过程,体会实际需求与数学内部的矛盾在数系扩充过程中的作用,感受人类理性思维的作用及数与现实世界的联系. 理解复数的基本概念以及复数相等的充要条件	3.1.1　数系的扩充和复数的概念＊P_{79}
8	了解复数的几何意义	3.1.2　复数的几何意义＊P_{84}
9	(1)了解复数的代数表示法及其几何意义 (2)能进行复数代数形式的四则运算,了解复数代数形式的加、减运算的几何意义	3.2.1　复数代数形式的加减运算及其几何意义＊P_{90} 3.2.2　复数代数形式的乘除运算＊P_{97}
10	(1)通过具体实例,进一步认识程序框图 (2)通过具体实例,了解工序流程图 (3)能绘制简单实际问题的流程图,体会流程图在解决实际问题中的作用	4.1　流程图＊P_{113}
11	(1)通过实例,了解结构图,运用结构图梳理已学过的知识、收集整理资料信息 (2)结合作出的结构图与他人进行交流,体会结构图在揭示事物联系中的作用	4.2　结构图＊P_{120}

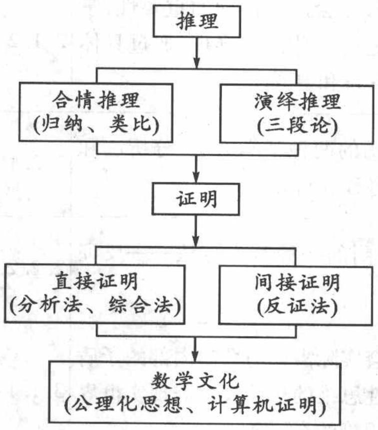
全书知识结构图解·名师学法指津

一　全书知识结构图解

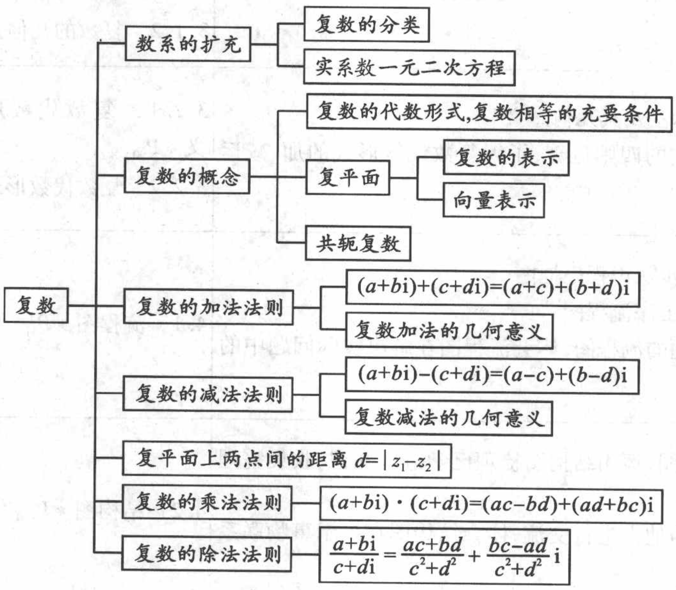
第一章　统计案例



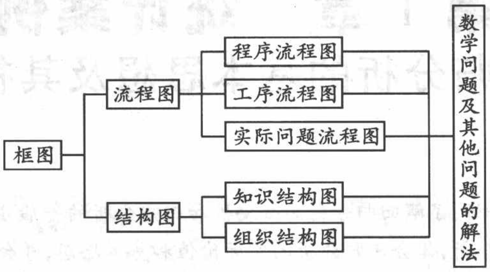
第二章　推理与证明



第三章　数系的扩充与复数的引入



第四章　框图



二　名师学法指津

第一章“统计案例”中学习的统计方法是在前面学习统计和回归直线的基础上，通过分析诸多典型案例来进行的，在学习中要注意以下三点：

1. 同学们需要亲自实践，参与解决具体的实际问题，学习处理数据的一些常用方法，会用必要的现代技术手段来处理有关数据，如计算器、计算机等.

2. 学习回归分析要在复习、回顾两个变量的相关性的基础上进行，做到前后联系，逐步推进.

3. 在独立性检验中要特别关注方法的应用，不要过多探究方法的理论依据；要注意不能单纯记忆公式和套用公式，要多探寻规律，合理建模，形成方法.

第二章“推理与证明”中内容的实用性较强，同时又是数学逻辑性的重要体现. 在学习中要做到以下两点：

1. 多联系生活实际和其他学科的结论，体会推理的应用性.

2. 通过数学实例去理解合情推理和演绎推理，不必刻意追求概念的抽象表述.

第三章“数系的扩充与复数的引入”，新教材降低了对此内容的要求，只要求学习复数的概念、复数的代数形式及几何意义、加减乘除运算及加减法的几何意义. 学习时要做到以下三点：

1. 在学习复数的概念和运算时，要注意避免繁琐的计算，要学会灵活运用一些基本技巧.

2. 在学习复数的几何意义时，要学会运用数形结合的思想解决有关问题.

3. 多阅读一些引申的内容，扩大自己的知识面及数学视野.

第四章“框图”，本章是体现数学应用性的重要内容，学习时要做到以下三点：

1. 学习框图，要从分析实际入手，学会运用框图表示数学计算和证明过程中的主要思路与步骤、实际问题中的工序流程、某一知识系统的结构关系等.

2. 在运用框图的过程中理解流程图和结构图的特征.

3. 积极实践，掌握框图的用法，体验用框图解决问题的优越性.

第1章 统计案例

1.1 回归分析的基本思想及其初步应用

课标三维目标

1. 知识与技能：通过对具体问题的分析，了解回归分析的必要性和回归分析的一般步骤，会求回归直线方程，作散点图，并会运用所学习的知识对实际问题进行回归分析，体会回归分析的实际价值和基本思想，并会用回归分析对具体事件进行分析.

2. 过程与方法：本节内容先从实际问题中的数量之间的关系入手，求出相应的回归直线方程，从中也找出存在的不足，从而有进行回归分析的必要性，进而学习相关指数，用相关指数来刻画回归的效果，进而得出回归分析的一般步骤，并对具体问题进行回归分析，用于解决实际问题.

3. 情感、态度与价值观：从实际问题中发现自己已有知识的不足之处，激发学生的好奇心和求知欲，培养学生不满足于已有知识、勇于求知的良好个性品质，引导学生积极进取. 任何事物都是相对的，但又有一定的规律性，我们只要从实际出发，不断探求事物的内在联系，就会找出其中的规律性，形成解决实际问题的方法和能力.

三层完全解读

>>>解题依据

1 知识·能力聚焦

1. 求回归直线方程

人教A版、B版

求回归直线方程的一般方法是：

(1)作出散点图，将问题所给的数据在平面直角坐标系中描点，这样表示出的具有相关关系的两个变量的一组数据的图形就是散点图. 从散点图中我们可以看出样本点是否呈条状分布，从而判断两个量是否具有线性相关关系.

(2)求回归系数 $\hat{a},\hat{b}$，其中

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

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

将所给的数据 x,y 列成相应的表格，如下表所示：

序号	x	y	x^2	y^2	xy
1	x_1	y_1	x_1^2	y_1^2	x_1y_1
2	x_2	y_2	x_2^2	y_2^2	x_2y_2
…	…	…	…	…	…
n	x_n	y_n	x_n^2	y_n^2	x_ny_n
Σ	$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_iy_i$

由此可知 $\bar{x}=\frac{\sum x_i}{n},\bar{y}=\frac{\sum y_i}{n}$，其中 $(\bar{x},\bar{y})$ 称为样本点的中心.

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

(3)写出回归直线方程 $\hat{y}=\hat{b}x+\hat{a}$，并用回归直线方程进行预测说明：当 x 取 x_0 时，由线性回归方程可得 $\hat{y}_0$ 的值，从而可进行相应的判断.

>>>名题诠释

例题1 容易题——2011·江西高考题

变量 X 与 Y 相对应的一组数据为(10,1)，(11.3,2)，(11.8,3)，(12.5,4)，(13,5)；变量 U 与 V 相对应的一组数据为(10,5)，(11.3,4)，(11.8,3)，(12.5,2)，(13,1). r_1 表示变量 Y 与 X 之间的线性相关系数，r_2 表示变量 V 与 U 之间的线性相关系数，则(　　).

A. $r_2<r_1<0$　B. $0<r_2<r_1$　C. $r_2<0<r_1$　D. $r_2=r_1$

【解析】$\because r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$，∴通过计算得第一组变量正相关，第二组变量负相关. 故选C.

【答案】C

例题2 容易题

某工厂1~8月份某种产品的产量与成本的统计数据见下表：

月份	1	2	3	4	5	6	7	8
产量(吨)	5.6	6.0	6.1	6.4	7.0	7.5	8.0	8.2
成本(万元)	130	136	143	149	157	172	183	188

以产量为 x，成本为 y.

(1)画出散点图；

(2)y 与 x 是否具有线性相关关系？若有，求出其回归方程.

【解析】本题主要考查对两个变量的回归分析，画出散点图，代入回归系数公式即可得.

【解】(1)由表画出散点图，如图1-1-7所示.

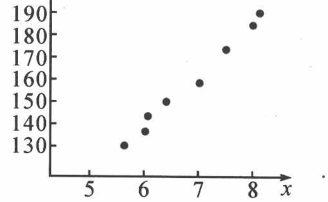


图1-1-7

2. 随机误差

人教A版

对于某一个 x_i，由回归直线方程：$\hat{y}=\hat{b}x+\hat{a}$ 可以确定一个 $\hat{y}_i$ 的值，但是由于测量本身存在误差，或者测量受其他因素的影响，或者回归直线方程本身存在误差，或者方程受某些随机因素的影响，使得 $\hat{y}_i$ 与测得的实际数据 y_i 之间存在误差，并不相等，出现 $y_i=\hat{y}_i+\sum_i$，其中 $\sum_i$ 为随机误差.

从散点图中我们可以看到，样本点分布在某一条直线附近，而不是在一条直线上，所以不能用一次函数 $y=bx+a$ 来描述它们之间的关系，我们用下面的线性回归模型来表示：$y=bx+a+e$，其中 a,b 为模型的未知参数，e 称为随机误差. 我们选用的线性模型往往只是一种近似的模型，除解释变量外，还有其他因素也会导致随机误差 e 的产生.

随机误差 e 的主要来源：(1)用线性回归模型近似真实模型(真实模型是客观存在的，通常我们并不知道真实模型是什么)所引起的误差. 可能存在非线性的函数能够更好地描述 y 与 x 之间的关系，但是现在却用线性函数来表述这种关系，结果会产生误差. 这种由模型近似所引起的误差包含在 e 中. (2)忽略了某些因素的影响. 影响变量 y 的因素不只是变量 x，可能还包括其他许多因素(例如在描述身高和体重关系的模型中，体重不仅受身高的影响，还会受遗传基因、饮食习惯、生长环境等其他因素的影响)，它们的影响都体现在 e 中. (3)观测误差. 由于测量工具(比如一个人的体重是确定的数，不同的秤可能会得到不同的观测值)的原因，导致 y 的观测值产生误差，这样的误差也包含在 e 中.

3. 回归分析

人教A版

(1)线性相关关系强弱的描述，样本相关系数的计算公式：

$$r=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$，当 $r>0$ 时，表明两个变量正相关；当 $r<0$ 时，表明两个变量负相关. r 的绝对值越接近1，表明两个变量的线性相关性越强；r 的绝对值接近于0时，表明两个变量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.

(2) $\sum_{i=1}^{n}(y_i-\bar{y})^2$ 表示总的效应，称为总偏差平方和；数据点和它在回归直线上相应位置的差异 $(y_i-\hat{y}_i)$ 是随机误差的效应，称 $\hat{e}_i=y_i-\hat{y}_i$ 为残差；$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 称为残差平方和，$\sum_{i=1}^{n}(y_i-\bar{y})^2-\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$ 为两者之差，它代表了随机误差的效应，解释变量效应的这个值称为回归平方和. 我们用相关指数 R^2 来刻画回归的效果，其计算公式是 $R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$，$R^2$ 的值越大，说明残差平方和越小，也就是说模型的拟合效果越好. 在线性回归模型中，R^2 表示解释变量对预报变量变化的贡献率，R^2 越接近于1，表示回归效果越好(因为 R^2 越接近于1，表示解释变量和预报变量的线性相关性越强). 如果某组数据可能采取几种不同回归方程进行回归分析，则可以通过比较 R^2 的值来作出选择，即选择 R^2 大的模型作为这组数据的模型.

(2)从图上可看出，这些点基本上散布在一条直线附近，可以认为 x 和 y 线性相关关系显著，下面求其回归方程，首先列出下表.

序号	x_i	y_i	x_i^2	y_i^2	x_iy_i
1	5.6	130	31.36	16 900	728.0
2	6.0	136	36.00	18 496	816.0
3	6.1	143	37.21	20 449	872.3
4	6.4	149	40.96	22 201	953.6
5	7.0	157	49.00	24 649	1 099.0
6	7.5	172	56.25	29 584	1 290.0
7	8.0	183	64.00	33 489	1 464.0
8	8.2	188	67.24	35 344	1 541.6
Σ	54.8	1 258	382.02	201 112	8 764.5

$\bar{x}=6.85$，$\bar{y}=157.25$.

$$\therefore \hat{b}=\frac{\sum_{i=1}^{8}x_iy_i-8\bar{x}\bar{y}}{\sum_{i=1}^{8}x_i^2-8\bar{x}^2}=\frac{8\ 764.5-8\times6.85\times157.25}{382.02-8\times6.85^2}\approx22.17,$$

$\hat{a}=\bar{y}-\hat{b}\bar{x}=157.25-22.17\times6.85\approx5.39$，

故线性回归方程为 $\hat{y}=22.17x+5.39$.

【点拨】求回归直线方程时，首先应注意到，只有在散点图大致呈线性分布时，求出的回归直线方程才有实际意义，否则，求出的回归直线方程毫无意义.

例题3　容易题——2011·山东青岛模拟

为了考察两个变量 y 与 x 的线性相关性，测得 x、y 的13对数据，若 y 与 x 具有线性相关关系，则相关系数 $|r|$ 的取值范围是________.

【解析】查表得显著性水平0.05，自由度 $13-2=11$ 相应的相关系数临界值 $r_{0.05}=0.553$. 所以 y 与 x 若具有线性相关关系，则相关系数 $|r|$ 的取值范围是 $(0.553,1]$.

【答案】$(0.553,1]$

例题4　容易题

现随机抽取了我校10名学生在入学考试中的数学成绩(x)与入学后的第一次考试数学成绩(y)，数据如下表：

学生号	1	2	3	4	5	6	7	8	9	10
x_i	120	108	117	104	103	110	104	105	99	108
y_i	84	64	84	68	69	68	69	46	57	71

请问：这10名学生的两次数学考试成绩是否具有显著的线性相关关系？

【解析】若 x 与 y 呈线性相关关系，就无需进行相关性检验，否则需进行相关性检验，可由公式 $r=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sqrt{(\sum_{i=1}^{n}x_i^2-n\bar{x}^2)(\sum_{i=1}^{n}y_i^2-n\bar{y}^2)}}$ 判断.

【解】$\bar{x}=\frac{1}{10}(120+108+\cdots+99+108)=107.8$，

$\bar{y}=\frac{1}{10}(84+64+\cdots+57+71)=68$，

$\sum x_i^2=120^2+108^2+\cdots+99^2+108^2=116\ 584$，

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，这方面的分析工作称为残差分析.

4. 相关性检验

人教B版、苏教版

(1)相关性检验的必要性.

在实际问题中，利用散点图来判断线性相关关系，既快又方便，但这时散点图中的各个点应该集中在某一直线的附近，依据求回归直线方程的方法，求出与之相对应的回归直线方程，这条回归直线方程具有实际的应用价值. 但是，若散点图中画出的各个点，不是集中在一条直线的附近，而是集中在某一个圆内，依据求回归直线的方法，我们仍然可以求出相应的回归直线方程，但这条回归直线方程已经不能反映出这组数据的变化规律，这时求回归直线方程也就失去了意义. 那么，到底在数据有什么特征时，才是线性相关的，才能求出回归直线方程呢？这就需要对 x 与 y 作线性相关性检验，也就是相关性检验.

对于随机抽取到的变量 x 与 y 的 n 对数据 (x_1,y_1)，(x_2,y_2)，…，(x_n,y_n)，检验统计量是样本相关系数 $r=\dfrac{\sum\limits_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i-\bar{x})^2\sum\limits_{i=1}^{n}(y_i-\bar{y})^2}}=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sqrt{\left(\sum\limits_{i=1}^{n}x_i^2-n\bar{x}^2\right)\left(\sum\limits_{i=1}^{n}y_i^2-n\bar{y}^2\right)}}$.

r 具有以下性质：$|r|\leqslant 1$ 并且 $|r|$ 越接近于1，线性相关程度越强；$|r|$ 越接近于0，线性相关程度越弱.

(2)相关性检验的步骤.

①作统计假设，假设 x 与 y 不具有线性相关关系；

②根据小概率0.05与 $n-2$ 在附表中查出 r 的一个临界值 $r_{0.05}$；

③根据样本相关系数的计算公式 $r=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sqrt{\left(\sum\limits_{i=1}^{n}x_i^2-n\bar{x}^2\right)\left(\sum\limits_{i=1}^{n}y_i^2-n\bar{y}^2\right)}}$ 计算出 r 的值；

④作统计推断：如果 $|r|>r_{0.05}$，表明有95%的把握认为 x 与 y 之间具有线性相关关系；如果 $|r|\leqslant r_{0.05}$，这时寻找回归直线方程是毫无意义的.

2 方法·技巧平台

5. 利用图形分析残差的特性

作图时纵坐标为残差，横坐标可以选为样本编号或身高数据或体重估计值等，这样作出的图形称为残差图. 在残差图中，如果有两个样本点的残差比较大，需要确认在采集这两个样本点的过程中是否有人为的错误. 如果数据采集有错误，就先予以纠正，然后重新利用线性回归模型拟合数据；如果数据采集没有错误，则需要寻找其他原因. 另外，若残差点比较均匀地落在水平的带状区域中，说明选用的模型比较合适. 这样的带状区域的宽度越窄，说明模型拟合精度越高，回归方程的预报精度就越高.

$\sum y_i^2=84^2+64^2+\cdots+57^2+71^2=47\ 384$，

$\sum x_iy_i=120\times 84+108\times 64+\cdots+108\times 71=73\ 796$，

所以相关系数

$$r=\frac{73\ 796-10\times 107.8\times 68}{\sqrt{(116\ 584-10\times 107.8^2)(47\ 384-10\times 68^2)}}\approx 0.750\ 6,$$

查表：显著性水平0.05，自由度 $10-2=8$ 相应的相关关系临界值 $r_{0.05}=0.632$，由 $r>r_{0.05}$ 知两次数学考试成绩有显著的线性相关关系.

【点拨】如果两个变量不具备相关关系，或者相关关系不显著，即使求回归直线方程也是无意义的，用其估计和预测是不可信的.

例题5 容易题

某城区为研究城镇居民家庭月人均生活费支出和月人均收入的相关关系，随机抽取10户进行调查，其结果如下：

月人均收入 x/元	月人均生活费 y/元
300	255
390	324
420	335
520	360
570	450
700	520
760	580
800	600
850	630
1 080	750

试预测月人均收入为1 100元和月人均收入为1 200元的两个家庭的月人均生活费，并进行残差分析.

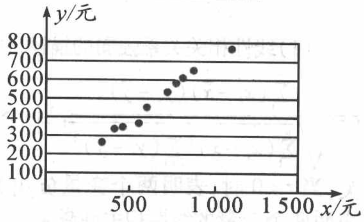


图1－1－8

【解析】先通过散点图确定两个变量是否线性相关，再用残差图和相关指数 R^2 来分析模型拟合的效果，进而估计预测值的准确性.

【解】作出散点分布图(如图1－1－8)，由图可知，月人均生活费与月人均收入之间具有线性相关关系.

通过计算可知 $\bar{x}=639$，$\bar{y}=480.4$，$\sum\limits_{i=1}^{10}x_i^2=4\ 610\ 300$，$\sum\limits_{i=1}^{10}y_i^2=2\ 540\ 526$，$\sum\limits_{i=1}^{10}x_iy_i=3\ 417\ 560$，$\therefore \hat{b}=\dfrac{\sum\limits_{i=1}^{10}x_iy_i-10\bar{x}\bar{y}}{\sum\limits_{i=1}^{10}x_i^2-10\bar{x}^2}\approx 0.659\ 9$，$\hat{a}=\bar{y}-\hat{b}\bar{x}\approx 58.724$.

$\therefore$ 线性回归方程为 $\hat{y}=0.659\ 9x+58.724$.

作残差图如图1－1－9所示，由图可知，残差点比较均匀地落在水平的带状区域中，说明选用的模型比较合适.

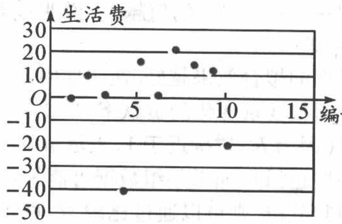


图1－1－9

计算相关指数得 $R^2\approx 0.986\ 3$，说明城镇居民的月人均生活费的差异有98.63%

几种常见的残差图：

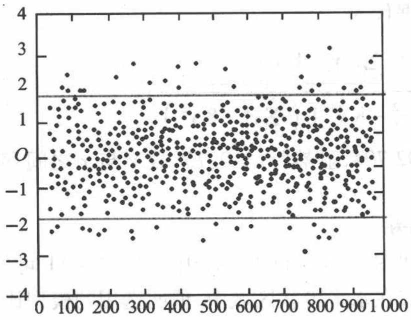


图 1－1－1

图 1－1－1：残差图中的点分布在以原点为中心的水平带形区域上，并且沿水平方向散点的分布规律相同，说明残差是随机的，所选择的回归模型是合理的.

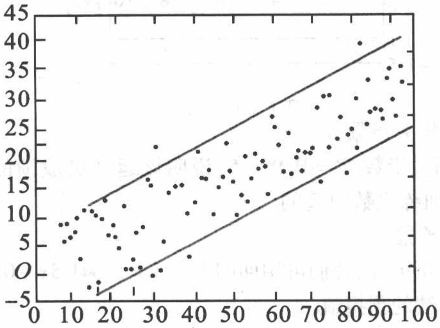


图 1－1－2

图 1－1－2：残差图中的点分布在一条倾斜的带形区域上，并且沿带形区域方向散点的分布规律相同，说明残差与横坐标变量间有线性关系，此时所选择的回归模型的效果不是最好的，有改进的余地.

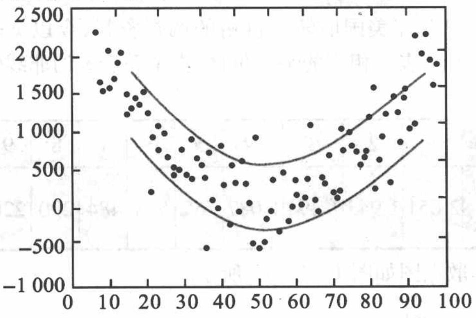


图 1－1－3

图 1－1－3：残差图中的点分布在一条二次曲线的弯曲带形区域上，说明残差与横坐标变量间有二次关系，此时所选用的回归模型的效果不是最好的，有改进的余地.

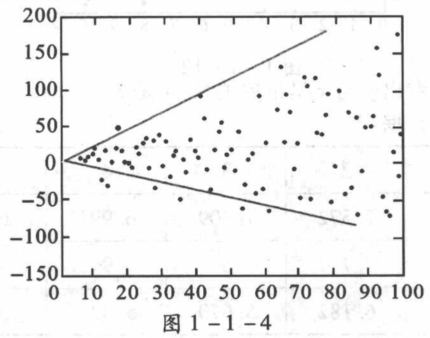


图 1－1－4

是由月人均收入引起的.

由以上分析可知，我们可以利用回归方程 $\hat{y}=0.659\,9x+58.724$ 来作为月生活费的预报值.

将 $x=1\,100$ 代入回归方程得 $y=784.61$（元）；

将 $x=1\,200$ 代入回归方程得 $y=850.60$（元）.

故预测月人均收入分别为 1 100 元和 1 200 元的两个家庭的月人均生活费分别为 784.61 元和 850.60 元.

例题 6　中难题

某运动员训练次数与运动成绩的数据如下表：

次数（x）	30	33	35	37	39	44	46	50
成绩（y）	30	34	37	39	42	46	48	51

（1）作出散点图；

（2）求出回归方程；

（3）计算相关系数，并利用其检验两变量间线性相关关系的显著性；

（4）作出残差图；

（5）计算相关指数 R^2；

（6）试预测该运动员训练 47 次及 55 次的成绩.

【解】（1）作出该运动员训练次数（x）与成绩（y）的散点图，如图 1－1－10 所示，由散点图可知，它们之间具有线性相关关系.

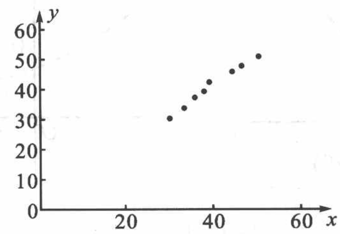


图 1－1－10

（2）列表计算：

次数 x_i	成绩 y_i	x_i^2	y_i^2	x_iy_i
30	30	900	900	900
33	34	1 089	1 156	1 122
35	37	1 225	1 369	1 295
37	39	1 369	1 521	1 443
39	42	1 521	1 764	1 638
44	46	1 936	2 116	2 024
46	48	2 116	2 304	2 208
50	51	2 500	2 601	2 550

由上表可求得 $\overline{x}=39.25$，$\overline{y}=40.875$，$\sum\limits_{i=1}^{8}x_i^2=12\,656$，

$\sum\limits_{i=1}^{8}y_i^2=13\,731$，$\sum\limits_{i=1}^{8}x_iy_i=13\,180$，

$$\therefore \hat{b}=\frac{\sum\limits_{i=1}^{8}(x_i-\overline{x})(y_i-\overline{y})}{\sum\limits_{i=1}^{8}(x_i-\overline{x})^2}=\frac{\sum\limits_{i=1}^{8}x_iy_i-8\,\overline{x}\,\overline{y}}{\sum\limits_{i=1}^{8}x_i^2-8\,\overline{x}^2}\approx 1.041\,5,$$

$\hat{a}=\overline{y}-\hat{b}\overline{x}=-0.003\,875$，

$\therefore$ 回归方程为 $\hat{y}=1.041\,5x-0.003\,875$.