高中

同步学程

TONG BU XUE CHENG

高中新课程

数学

选修1—2

明天出版社

高中

同步学程

高中新课程

数学

选修 1-2

明天出版社

同 步 学 程

数学

选修 1－2

※

明天出版社出版发行

（济南市经九路胜利大街 39 号）

http://www.sdpress.com.cn

http://www.tomorrowpub.com

各地新华书店经销　山东省无棣县教育实业公司印刷厂印刷

※

787×1092 毫米　16 开　　6 印张　　165 千字

2009 年 1 月第 1 版　　2009 年 1 月第 1 次印刷

ISBN　978－7－5332－5996－9

定价：5.50 元

前言

为了更好地贯彻素质教育要求，落实《山东省普通高中课程设置及教学指导意见(试行)》，帮助广大师生准确理解和把握实验教材的内容和要求，全面提高学生的自主学习能力，我们依据教育部颁布的《普通高中课程方案(实验)》、各学科课程标准和现行教材，组织部分一线骨干教师和教学研究人员编写了这套《同步学程》丛书，主要供高中学生同步学习使用。这套丛书对指导普通高中新课程实验，提高学生的综合素质，都将起到积极的促进作用。

这套丛书包括思想政治、语文、数学、英语、物理、化学、生物、历史、地理共九个学科的所有必修模块和部分选修模块，并根据教学进度同步发行。各模块根据新课程的内容特点按单元(节、课)编写，指导学生在规定的课时内完成学习任务，提高学习效率。

这套丛书有以下几个方面的特点：

1. 注重体现普通高中课程改革的理念和要求，帮助师生进行课程实验，用好用活教材；

2. 注重体现“知识和能力、过程和方法、情感态度和价值观”的三维目标要求，在帮助学生牢固掌握基础知识的前提下，努力提高学生的应用能力；

3. 注重设置问题情境，拓宽知识背景，指导学生掌握科学的学习方法，自主探求未知领域，培养学生的探索精神和创新能力；

4. 注重与新课程实验的同步性，紧密配合各学科的学习，按单元(节、课)分配学习课时，组织学习训练内容，既便于教师指导又便于学生自学。

参加《数学》(选修1—2)编写工作的老师及分工情况：王洪峰(第一章)、于清堂(第二章)、吕文彬(第三章)、徐西文、王春山(第四章)。李在功、高天祥、商金琳、王春山、马仕光参加了审稿。王文清老师负责统稿。

希望这套《同步学程》丛书能够帮助同学们学好新课程，打牢基础，提升素质，实现理想。

2009年1月

目录

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

第一章
统 计 案 例

§1.1　回归分析的基本思想及其初步应用

课标要求

1. 了解样本点的中心、随机误差、残差、残差的平方和等概念.

2. 会用残差分析、判断线性回归模型的拟合效果.

3. 掌握建立回归模型的基本步骤.

基础诊断

1. 如果散点图中的点的分布从整体上看大致在________________，我们就称这两个变量之间具有线性相关关系，____________叫做回归直线.

2. 回归分析是对具有____________________的一种常用方法.

3. 线性回归模型 $y=bx+a+e$ 中 e 称为__________.

4. 总偏差平方和的表达式是________________.

5. 通过残差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__________.

典型示例

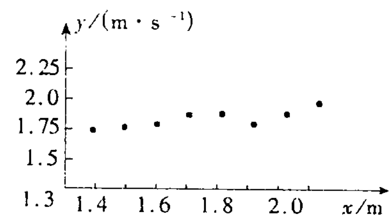
【例 1】研究某灌溉渠道水的流速 y 与水深 x 之间的关系，测得一组数据如下：

水深 x/m	1.40	1.50	1.60	1.70	1.80	1.90	2.00	2.10
流速 y/($\mathrm{m\cdot s^{-1}}$)	1.70	1.79	1.88	1.95	2.03	2.10	2.16	2.21

(1)求 y 对 x 的回归直线方程；

(2)预测水深为 1.95m 时水的流速是多少？

【思路分析】从散点图可以直观地看出变量 x 与 y 之间有无线性相关关系，为此把这 8 对数据描绘在平面直角坐标系中，得到平面上 8 个点，如图所示.



由图容易看出，x 与 y 之间有近似的线性相关关系，或者说，可以用一个回归直线方程

$$\hat{y}=bx+a$$

来反映这种关系，这些是我们在必修模块数学 3 中学过的知识.

进一步观察这 8 个点，容易发现它们并不是"严格地"在一条直线上. 对于某个 x_i，由上式能确定一个 $\hat{y}_i=bx_i+a$. 一般地说，由于测量流速可能存在误差，或者受某些随机因素的影响，或者上面的回归直线方程本身就不够精确，$\hat{y}_i$ 与测得的数据 y_i 很可能不相等，即 $y_i=\hat{y}_i+e_i(i=1,2,\cdots,8)$

其中 e_i 是随机误差项. 于是,就有

$y_i=bx_i+a+e_i(i=1,2,\cdots,8)$

这就是本题的线性模型.

从上述线性模型出发,我们可以求出 a 与回归系数 b 的估计值$\hat{a}$,$\hat{b}$,使得全部误差 $e_1,e_2,\cdots,e_8$ 的平方和达到最小,当然,这是一种很好的估计,最后得到求$\hat{a}$,$\hat{b}$的数学公式为

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

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

【解】(1)由题目所给数据可知:

$$\bar{x}=\frac{1}{8}\times(x_1+x_2+\cdots+x_8)=\frac{1}{8}\times14.00=1.75,$$

$$\bar{y}=\frac{1}{8}(y_1+y_2+\cdots+y_8)=\frac{1}{8}\times15.82=1.9775.$$

求得$\hat{b}=\frac{11}{15}\approx0.733$, $\hat{a}\approx0.694$.

y 对 x 的回归直线方程为

$\hat{y}=0.733x+0.694$.

回归系数$\hat{b}=0.733$ 的意思是,在此灌溉渠道中,水深每增加 0.1 米,水的流速平均增加 0.0733m/s,$\hat{a}=0.694$ 可以解释为水的流速中不受水深影响的部分.

(2)由(1)中求出的回归直线方程,把 $x=1.95$ 代入,易得

$\hat{y}=0.733\times1.95+0.694\approx2.12$(m/s).

计算结果表示,当水深为 1.95m 时可以预测渠水的流速为 2.12m/s.

【反思与小结】建立回归模型的一般步骤:(1)确定研究对象,明确两个变量即解释变量和预报变量;(2)画出散点图,观察它们之间的关系;(3)由经验确定回归方程类型(若呈线性关系,选用线性回归方程);(4)按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,或残差出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.

【例 2】假设关于某设备的使用年限 x 和所支出的维修费用 y(万元)有如下的统计资料:

使用年限 x	2	3	4	5	6
维修费用 y	2.2	3.8	5.5	6.5	7.0

若由资料知 y 对 x 呈线性相关关系.

(1)求相关系数 a,b;

(2)估计使用年限为 10 年时,维修费用是多少?

【解】直接利用公式求线性回归方程的系数 a,b.

(1)制表如下:

i	1	2	3	4	5	合计
x_i	2	3	4	5	6	20
y_i	2.2	3.8	5.5	6.5	7.0	25
x_iy_i	4.4	11.4	22.0	32.5	42.0	112.3
x_i^2	4	9	16	25	36	90
$\bar{x}=4;\bar{y}=5;$ $\sum_{i=1}^{5}x_i^2=90;\sum_{i=1}^{5}x_iy_i=112.3$						

于是有 $b=\frac{112.3-5\times4\times5}{90-5\times4^2}=\frac{12.3}{10}=1.23$,

$a=\bar{y}-b\bar{x}=5-1.23\times4=0.08$.

(2)回归直线方程为 $y=1.23x+0.08$,

当 $x=10$ 年时,

$y=1.23\times10+0.08=12.3+0.08=12.38$(万元),

即估计使用 10 年时,维修费用是 12.38 万元.

【误区警示】知道 x 与 y 是线性相关关系,无需进行相关性检验;否则,应首先进行相关性检验,如果本身两个变量不具备相关关系,或者说,它们之间相关关系不显著,即使求出回归直线方程也毫无意义,而且其估计和预测也是不可信的.

【反思与小结】因为 y 和 x 呈线性相关关

系，所以可以用一元线性相关的方法解决问题.

(1)利用公式：$b=\dfrac{\sum\limits_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sum\limits_{i=1}^{n}x_i^2-n\bar{x}^2}$，$a=\bar{y}-b\bar{x}$来计算回归系数，有时为了方便制表时，对应求出$x_iy_i$，$x_i^2$，以利于求和；

(2)获得线性回归方程后，取$x=10$，即得所求；

(3)本题应借用计算器计算，并列出表格，再按分析时的步骤进行.

归纳总结

比较两个模型的拟合效果的一般步骤：

1. 分别建立对应于两个模型的回归方程$\hat{y}^{(1)}=f(x,\hat{a})$与$\hat{y}^{(2)}=g(x,\hat{b})$，其中$\hat{a}$和$\hat{b}$分别是参数$a$和$b$的估计值.

2. 分别计算两个回归方程的残差平方和$\hat{Q}^{(1)}=\sum\limits_{i=1}^{n}(y_i-\hat{y}_i^{(1)})^2$与$\hat{Q}^{(2)}=\sum\limits_{i=1}^{n}(y_i-\hat{y}_i^{(2)})^2$.

3. 若$\hat{Q}^{(1)}<\hat{Q}^{(2)}$，则$\hat{y}^{(1)}=f(x,\hat{a})$的效果比$\hat{y}^{(2)}=g(x,\hat{b})$的好；反之，$\hat{y}^{(1)}=f(x,\hat{a})$的效果不如$\hat{y}^{(2)}=g(x,\hat{b})$的好.

拓展提高

1. 测得10对某国父子身高(单位：英寸)如下：

父亲身高(x)	60	62	64	65	66	67	68	70	72	74
儿子身高(y)	63.6	65.2	66	65.5	66.9	67.1	67.4	68.3	70.1	70

(1)如果y与x之间具有线性相关关系，求回归直线方程.

(2)如果父亲的身高为73英寸，估计儿子的身高.

【解】(1)由计算器易得：

$\bar{x}=66.8$，$\bar{y}=67.01$，$\sum\limits_{i=1}^{10}x_i^2=44794$，$\sum\limits_{i=1}^{10}y_i^2=44941.93$，

$\bar{x}\bar{y}=4476.27$，$\bar{x}^2=4462.24$，$\bar{y}^2=4490.34$，$\sum\limits_{i=1}^{10}x_iy_i=44842.4$.

设回归直线方程为$\hat{y}=bx+a$.

由$b=\dfrac{\sum\limits_{i=1}^{10}x_iy_i-10\bar{x}\bar{y}}{\sum\limits_{i=1}^{10}x_i^2-10\bar{x}^2}=\dfrac{44842.4-44726.7}{44794-44622.4}=\dfrac{79.7}{171.6}\approx0.4645$.

$a=\bar{y}-b\bar{x}=67.01-0.4645\times66.8\approx35.98$.

故所求的回归直线方程为$\hat{y}=0.4645x+35.98$.

(2)当$x=73$时，$\hat{y}=0.4645\times73+35.98\approx69.9$，所以当父亲身高为73英寸时，估计儿子的身高约为69.9英寸.

【反思与小结】借助于计算器应先求出各量，然后运用公式，计算.

展示平台

基础训练

1. 观察两个相关变量得如下数据

x	-1	-2	-3	-4	-5	5	4	3	2	1
y	-1	-2	-3	-4	-5	5	4	3	2	1

则两个变量间的回归直线方程为　(　　)

A. $y=\dfrac{1}{2}x-1$　　B. $y=x$

C. $y=2x+\dfrac{1}{3}$　　D. $y=x+1$

2. 预报变量是由________确定的.　(　　)

A. 解释变量

B. 随机误差

C. 解释变量和随机误差

D. 都不能确定

3. 回归直线方程参数a，b的最好估计是最小二乘估计$\hat{a}$和$\hat{b}$，则$\hat{a}=$________，$\hat{b}=$________.

4. 变量x，y之间的相关系数$r=$________.

5. 已知一个回归直线方程为$\hat{y}=2.5-3x$，则当变量x增加一个单位时，变量y的变化情况是________.

6. 在一段时间内,某种商品价格 x(万元)和需求量 y(t)之间的一组数据为:

价格 x	1.4	1.6	1.8	2	2.2
需求量 y	12	10	7	5	3

(1)画出散点图;

(2)求出 y 对 x 的回归直线方程,并在(1)的图形中画出它的图象;

(3)如果价格定为 1.9 万元,预测需求量大约是多少(精确到 0.01t)?

7. 弹簧长度 y(cm)随所挂物体质量 x(g)不同而变化的情况如下:

物质质量 x	5	10	15	20	25	30
弹簧长度 y	7.25	8.12	8.95	9.90	10.96	11.80

(1)画出散点图;

(2)求 y 对 x 的回归直线方程;

(3)预测所挂物体质量为 27g 时弹簧长度(精确到 0.01cm).

能力训练

8. 一个车间为了规定工时定额，需要确定加工零件所花费的时间，为此进行了10次试验，测得的数据如下：

零件数 x/个	10	20	30	40	50	60	70	80	90	100
加工时间 y/分	62	68	75	81	89	95	102	108	115	122

(1) y 与 x 是否具有线性相关关系？

(2) 如果 y 与 x 具有线性相关关系，求回归直线方程；

(3) 根据求出的回归直线方程，预测加工200个零件所用的时间为多少？

9. 一机器可以按各种不同速度运转，其生产物件有一些会有缺点. 每小时生产有缺点物件的多少，随机器运转的速度而变化，下列为其试验结果.

速度(转/秒)	每小时生产有缺点的物件数
8	5
12	8
14	9
16	11

(1) 求出机器速度影响每小时生产有缺点物件数的回归直线方程，并进行相关性检验；

(2) 若实际生产中所容许的每小时最大缺点物件数为10，那么，机器的速度每秒不得超过多少转？

反思提高

§1.2 独立性检验的基本思想及其初步应用

课标要求

1. 知道什么是 2×2 列联表及独立性检验.

2. 会根据题目所给列联表判断结论的可能性.

3. 了解独立性检验的基本思想.

基础诊断

1. 对于性别变量，其取值为男和女两种，这种变量的不同“值”表示个体所属的不同类别，像这样的变量称为____________.

2. 为了研究两个分类变量而列出的两个分类变量的频数表，称为________.

3. 独立性检验________________________独立性检验的基本思想类似于数学上的________法.

4. 独立性检验的一般步骤：

(1)根据样本数据制成____________；

(2)根据公式 $K^2=$__________________，计算 K^2 的值；

(3)比较 K^2 与临界值的大小关系作统计推断.

典型示例

【例 1】为了探究患慢性气管炎是否与吸烟有关，调查了 339 名 50 岁以上的人，调查结果如下表所示：

	患病	不患病	合计
吸烟	43	162	205
不吸烟	13	121	134
合计	56	283	339

试问：50 岁以上的人患慢性气管炎与吸烟习惯有关吗？

【分析】最理想的解决办法是向所有 50 岁以上的人作调查，然后对所得到的数据进行统计处理，但这花费的代价太大，实际上是行不通的，339 个人相对于全体 50 岁以上的人，只是一小部分，已学过总体和样本的关系，当用样本平均数、样本方差去估计总体相应的数字特征时，由于抽样的随机性，结果并不唯一. 现在情况类似，我们用部分对全体作推断，推断可能正确，也可能错误. 如果抽取的 339 个调查对象中很多人是吸烟但没患慢性气管炎，而虽不吸烟因身体体质差而患慢性气管炎，能够得出什么结论呢？我们有 95%(或 99%)的把握说事件 A 与事件 B 有关，是指推断犯错误的可能性为 5%(或 1%)，这也常常说成是“以 95%(或 99%)的概率”是一样的.

【解】根据列联表中的数据，得

$$K^2=\frac{339(43\times121-162\times13)^2}{205\times134\times56\times283}=7.469.$$

因为 $7.469>6.635$，所以我们有 99% 的把握说：50 岁以上的人患慢性气管炎与吸烟习惯有关.

【反思与小结】对两个分类变量进行独立性检验，要对样本的选取背景、时间等因素进行分析.

【例 2】在一次恶劣气候的飞行航程中调查男女乘客在机上晕机的情况如下表所示，根据此资料您是否认为在恶劣气候中男人比女人更容易晕机？

	晕机	不晕机	合计
男人	24	31	55
女人	8	26	34
合计	32	57	89

【解】这是一个 2×2 列联表的独立性检验问

题，根据列联表中的数据，得到

$$K^2=\frac{89(24\times26-31\times8)^2}{55\times34\times32\times57}=3.689.$$

因为 $3.689<3.841$，所以我们没有理由说晕机是否跟男女性别有关，尽管这次航班中男人晕机的比例 $\frac{24}{55}$ 比女人晕机的比例 $\frac{8}{34}$ 高，但我们不能认为在恶劣气候飞行中男人比女人更容易晕机.

【反思与小结】在使用 K^2 统计量作 2×2 列联表的独立性检验时，要求表中的4个数据大于等于5，为此，在选取样本的容量时一定要注意这一点，本例中的4个数据都大于5，是满足这一要求的.

归纳总结

一般地，假设有两个分类变量 x 和 y，它们的值域分别为 $\{x_1,x_2\}$ 和 $\{y_1,y_2\}$，其样本频数列联表(称为 2×2 列联表)为：

	y_1	y_2	总计
x_1	a	b	$a+b$
x_2	c	d	$c+d$
总计	$a+c$	$b+d$	$a+b+c+d$

若要推断的论述为

H_1:“X 与 Y 有关系”，

可以按如下步骤判断 H_1 成立的可能性：

1. 通过三维柱形图和二维条形图，可以粗略地判断两个分类变量是否有关系，但是这种判断无法精确地给出所得结论的可靠程度.

(1)在三维柱形图中，主对角线上两个柱形高度的乘积 ad 与副对角线上两个柱形高度的乘积 bc 相差越大，H_1 成立的可能性就越大；

(2)在二维条形图中，可以估计满足条件 $X=x_1$ 的个体中具有 $Y=y_1$ 的个体所占的比例 $\frac{a}{a+b}$，也可以估计满足条件 $X=x_2$ 的个体中具有 $Y=y_1$ 的个体所占比例 $\frac{c}{c+d}$，两个比例的值相差越大，H_1 成立的可能性就越大.

2. 可以利用独立性检验来考察两个分类变量是否有关系，并且能较精确地给出这种判断的可靠程度. 具体做法是：根据列联表所提供数据计算出随机变量 K^2 的值 k，其值越大，说明“X 与 Y 有关系”成立的可能性越大.

拓展提高

1. 某大型企业人力资源部为了研究企业员工工作积极性和对待企业改革态度的关系，随机抽取了189名员工进行调查，所得数据如下表所示：

	积极支持改革	不太赞成改革	合计
工作积极	54	40	94
工作一般	32	63	95
合计	86	103	189

对于人力资源部的研究项目，根据上述数据能得出什么结论?

【解】根据列联表中的数据，得到：

$$K^2=\frac{189(54\times63-40\times32)^2}{94\times95\times86\times103}=10.76$$

因为 $10.76>6.635$，所以有99%的把握说“员工工作积极”与“积极支持企业改革”是有关的，可以认为企业的全体员工对待企业改革的态度与其工作积极性是有关的.

【变式1】利用三维柱形图和二维条形图说明两者之间的关系.

【变式 2】比较用三维、二维图形与进行独立性检验的方法哪一种更精确，为什么？

2. 打鼾不仅影响别人休息，而且可能与患某种疾病有关，下表是一次调查所得的数据，试问：每一晚都打鼾与患心脏病有关吗？

	患心脏病	未患心脏病	合计
每晚都打鼾	30	224	254
不打鼾	24	1355	1379
合计	54	1579	1633

【解】根据列联表中的数据，得到

$$K^2=\frac{1633(30\times1355-224\times24)^2}{254\times1379\times54\times1579}$$

$=68.033$.

因为 $68.033>6.635$，所以有 99% 的把握说，每一晚都打鼾与患心脏病有关.

【变式 1】用三维柱形图和二维条形图说明两者之间的关系.

【变式 2】思考“打鼾与心脏病”从上述关系中，是否为“因果关系”？

展示平台

基础训练

1. 在一次独立性检验中，其把握性超过了 99%，则随机变量 K^2 的可能值为 ( )

A. 6.635　B. 5.024　C. 7.879　D. 3.841

2. 分类变量 X 和 Y 的列联表如下

	y_1	y_2	总计
x_1	a	b	$a+b$
x_2	c	d	$c+d$
总计	$a+c$	$b+d$	$a+b+c+d$

则下列说法正确的是 ( )

A. $ad-bc$ 越小，说明 X 与 Y 关系越弱

B. $ad-bc$ 越大，说明 X 与 Y 关系越强

C. $(ad-bc)^2$ 越大，说明 X 与 Y 关系越强

D. $(ad-bc)^2$ 越接近于 0，说明 X 与 Y 关系越强

3. 独立性检验是研究( )的方法. ( )

A. 解释变量与预报变量

B. 随机误差与预报变量

C. 分类变量

D. 解释变量与随机误差

4. 在一个 2×2 列联表中，由其数据计算得 $K^2=13.097$，则其两个变量间有关系的可能性为 ( )

A. 99%　B. 95%　C. 90%　D. 无关系

5. 列联表

	y_1	y_2	合计
x_1	214	175	389
x_2	451	597	1048
合计	665	772	1437

则随机变量 $K^2=$________.

6. 在独立性检验中，一般地，随机变量 $K^2\leqslant$________时，就认为没有充分的证据显示两个分类变量有关系.

7. 在 500 个人身上试验某种血清预防感冒的作用，把一年中的记录与另外 500 个未用血清的人作比较，结果如下：

	未感冒	感冒	合计
处理	252	248	500
未处理	224	276	500
合计	476	524	1000

问：该种血清能否起到预防感冒的作用？

8. 冶炼某种金属可以用旧设备和改造后的新设备，为了检验用这两种设备生产的产品中所含杂质的关系，调查结果如下表所示：

	杂质高	杂质低
旧设备	37	121
新设备	22	202

试作统计分析推断.

能力训练

9. 调查者通过询问 72 名男女大学生在购买食品时，是否看营养说明得到的数据如下表所示：

	看营养说明	不看营养说明	合计
男大学生	28	8	36
女大学生	16	20	36
合计	44	28	72

问：大学生的性别和是否看营养说明之间有没有关系？

10. 在研究某种新措施对猪白痢的防治效果问题时，得到以下数据：

	存活数	死亡数	合计
新措施	132	18	150
对照	114	36	150
合计	246	54	300

试问：新措施对防治猪白痢是否有效？

反思提高

单元测试

一、选择题

1. 对具有相关关系的两个变量作统计分析的一种常用方法是　(　　)

A. 回归分析　B. 相关系数分析

C. 残差分析　D. 相关指数分析

2. 两个变量相关性越强,相关系数 r　(　　)

A. 越接近 0　B. 越接近 1

C. 越接近 -1　D. 绝对值越接近 1

3. 回归直线必过　(　　)

A. $(0,0)$　B. $(\overline{x},0)$　C. $(0,\overline{y})$　D. $(\overline{x},\overline{y})$

4. 若散点图中所有样本点都在一条直线上,解释变量与预报变量的相关系数为　(　　)

A. 0　B. 1　C. -1　D. -1 或 1

5. 两个变量有线性相关关系且正相关,则回归直线方程中,$\hat{y}=bx+a$ 的系数 b　(　　)

A. $b>0$　B. $b<0$　C. $b=0$　D. $b=1$

6. 两个变量有线性相关关系且残差的平方和等于 0,则　(　　)

A. 样本点都在回归直线上

B. 样本点都集中在回归直线附近

C. 样本点比较分散

D. 不存在规律

7. 把两个分类变量的频数列出,称为　(　　)

A. 三维柱形图　B. 二维条形图

C. 列联表　D. 独立性检验

8. X 与 Y 两个分类变量有关系的可能性越大,随机变量 K^2 的值　(　　)

A. 越大　B. 越小

C. 不变　D. 可能越大也可能越小

9. 由列联表

	y_1	y_2	合计
x_1	43	162	205
x_2	13	121	134
合计	56	283	339

则随机变量 K^2 的值为　(　　)

A. 5.039　B. 2.021　C. 7.469　D. 12.452

10. 下面是一个 2×2 列联表

	y_1	y_2	合计
x_1	a	21	73
x_2	2	25	27
合计	b	46	

则表中 a,b 处的值分别为　(　　)

A. 94、96　B. 52、50

C. 52、54　D. 54、52

二、填空题

11. 样本的中心点是________.

12. 相关系数 $r=$________.

13. 相关指数 $R^2=$________.

14. 随机变量 K^2 的值为________.

三、解答题

15. 某种产品的广告费支出 x 与销售额 y(单位:百万元)之间有如下对应数据

x	2	4	5	6	8
y	30	40	60	50	70

(1)画出散点图,判断是否具有线性相关关系;

(2)求回归方程;

(3)估计广告费支出 7 百万元时的销售额.

16. 为考察某种药物预防疾病的效果，进行动物试验，得到如下的列联表：

	患病	未患病	总计
服用药	214	175	389
没服药	451	597	1048
总计	665	772	1437

请问能有多大把握认为药物有效？

17. 对 196 个接受心脏搭桥手术的病人和 196 个接受血管清障手术的病人进行 3 年的跟踪研究，调查他们是否又发作过心脏病，调查结果如下表所示：

	又发作心脏病	未发作	总计
心脏搭桥手术	39	157	196
血管清障手术	29	167	196
合计	68	324	392

试根据上述数据比较这两种手术对病人又发作心脏病的影响有没有差别.