一、教学目标

1. 核心素养

通过学习独立性检验的基本思想及其初步应用，初步形成基本的数据分析能力，培养数学运算能力．

2. 学习目标

（1）1.1.1.1 了解分类变量的概念

（2）1.1.1.2 了解等高条形图、列联表概念，学会用列联表、等高条形图直观判断分类变量的关系

（3）1.1.1.3 了解独立性检验基本思想，初步学会用独立性检验的基本思想判断分类变量的关系

3. 学习重点

了解独立性检验基本思想，初步学会用独立性检验的基本思想判断分类变量的关系

4. 学习难点

了解独立性检验基本思想，初步学会用独立性检验的基本思想判断分类变量的关系

二、教学设计

（一）课前设计

1. 预习任务

任务 1

阅读教材 P 10 — P 12，思考什么是分类变量，列联表如何画？

任务 2

有哪些方法可以直观判断两个分类变量是否有关系？

2. 预习自测

1. 下列不是分类变量的是 （ ）

A. 近视 B. 身高 C. 血压 D. 药物反应

解： B. 判断一个量是否是分类变量，只需看变量的不同值是否表示个体的不同类别，A,C,D 选项的不同值都可以表示个体的不同类别，只有 B 选项的不同值不表示个体的不同类别．

2. 下面是一个 2×2 列联表

	不健康	健　康	总计
不优秀	a	21	73
优　秀	2	25	27
总　计	b	46	100

则表中 a,b 的之分别是（　　）

A. 94 ， 96 B. 52 ， 50 C. 52 ， 54 D. 54 ， 52

解： C

（二）课堂设计

1. 知识回顾

（ 1 ）线性回归方程： $\hat{y}=\hat{b}x+\hat{a}$,

其中 ： $\hat{b}=\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}=\dfrac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}$ ， $\hat{a}=\bar{y}-\hat{b}\bar{x}$

（ 2 ）回归分析：是对具有相关关系的两个变量进行的统计分析的一种常用方法 .

（ 3 ）线性回归模型： $y=bx+a+e$ 其中 a 和　为模型的未知参数， e 称为随机误差 .

2. 问题探究

问题探究一　什么是分类变量？

● 活动一　理论研究，概念学习 — 分类变量

在现实生活中，会遇到各种各样的变量，如果要研究它们之间的关系，观察下面两组变量，分析在取不同的值时表示的个体有何差异？

变量的不同 “ 值 ” 表示个体所属的不同类别，像这样的变量成为分类变量 .

（1） 分类变量也称为属性变量或定性变量，它的不同值表示个体所属的不同类别 .

（2） 分类变量的取值一定是离散的，如性别只取男、女两个值 .

（3） 可以把分类变量的不同取值用数字表示，如用 0 表示男， 1 表示女，这是性别变量就成了取值为 0 和 1 的随机变量，但这些数字的大小没有意义 .

分类变量是大量存在的，例如是否吸烟，宗教信仰，国籍等

问题探究二 如何研究两个分类变量之间是否有关系？ 重点、难点知识 ★▲

在日常生活中，我们常常关心两个分类变量之间是否有关系 . 例如，吸烟与患肺癌是否有关系？性别是否对喜欢数学课程有影响？

● 活动一 实例探究，引出问题

例 1 为调查吸烟是否对肺癌有影响，某肿瘤研究所随机地调查了 9965 人，得到如下结果：

表格 1

那么吸烟是否对患肺癌有影响？估计吸烟者与不吸烟者患病的可能性差异？

● 活动二 实例探究，引出概念

1. 列联表

类似于上面的表格这样列出两个分类变量的频数表，称为列联表 . 即

列联表是两个或者两个以上分类变量的频数表，书中仅限于研究两个分类变量的列联表，并且每个分类变量只取两个值，这样的列联表成为 2×2 列联表 . 一般的，假设有两个分类变量 和 Y，它们的取值分别为 和 ，其样本频数列联表为：

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

其中 $a+b+c+d$ 是样本容量 .

● 活动三　利用旧知，研究问题

利用频率分布表判断 ；

由患肺癌在吸烟者与不吸烟者中的频率差异可粗略估计吸烟对患肺癌有影响 ；

● 活动四　学习新知，对比研究

与表格相比，图形更能直观的反映出两个分类变量间是否相互影响，我们常用等高条形图展示列联表数据的频率特征 .

2. 等高条形图

利用等高条形图来分析两个分类变量之间是否具有相关关系，可以形象、直观地反映两个分类变量之间的总体状态和差异大小，进而判断它们之间是否具有相关关系 .

（ 1 ）绘制等高条形图时，列联表的行对应的是高度，两行的数据不相等，但对应的条形图的高度是相同的；两行的数据对应不同的颜色 .

（ 2 ）等高条形图中由两个高度相同的矩形，每一个矩形中都有两种颜色，观察下方颜色区域的高度，如果两个高度相差比较明显，就判断两个分类变量之间有关系 .

下图是吸烟与是否患肺癌的等高条形图

深圳小学家长群:254317299　深圳初中家长群 :90482695　深圳高中家长群 :175743089
更多资料详见: http://sz.jiajiaoban.com/　咨询电话：4000-121-121

由条形图可以发现，在吸烟样本中，患肺癌的频率要高些，因此直观上可以认为吸烟更容易引发肺癌 .

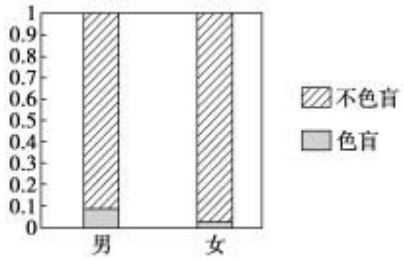
例 2　在调查的 480 名男人中有 38 人患色盲，520 名女人中有 6 名患色盲，试利用图形来判断色盲与性别是否有关？

【知识点：分类变量，等高条形图】

详解　根据题目给出的数据作出如下的列联表：

	色盲	不色盲	总计
男	38	442	480
女	6	514	520
总计	44	956	1 000

根据列联表作出相应的等高条形图：



从等高条形图来看在男人中患色盲的比例要比在女人中患色盲的比例大得多，因而，我们认为性别与患色盲是有关系的 .

点拨 ：利用数形结合的思想，借助等高条形图来判断两个分类变量是否相关是判断变量相关的常见方法之一 . 一般地，在等高条形图中，　与 $\frac{c}{c+d}$ 相差越大，两个分类变量有关系的可能性就越大 .

问题探究三　如何从统计学方面研究两个分类变量之间是否有关系？　重点、难点知识 ★▲

通过数据和图形分析，我们得到的直观判断是 “ 吸烟和患肺癌有关 ” 那么这种判断是否可靠？我们通过统计分析回答这个问题 .

为研究的一般性 ， 在列联表中用字母代替数字

	不患肺癌	患肺癌	总计
不吸烟	a	b	a+b
吸烟	c	d	c+d
总计	a+c	b+d	a+b+c+d

为了回答上述问题，我们先假设 H_0：吸烟与患肺癌没有关系 ， 那么吸烟样本中不患肺癌的比例应该与不吸烟样本中相应的比例差不多，即：

，即 .

因此， 越小，说明吸烟与患肺癌之间关系越弱； 越大，说明吸烟与患肺癌之间关系越强 .

为了使不同样本容量的数据有统一的评判标准，构造一个随机变量

$$K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \quad (1)$$

（其中 $n=a+b+c+d$ 为样本容量 . ）

若 H_0 成立，即 “ 吸烟与患肺癌没有关系 ” ，则 应该很小 .

根据 表 1 中的数据，利用 公式（ 1 ） 计算得到 的观测值为

$$k=\frac{9965(7775\times 49-42\times 2099)^2}{7817\times 2148\times 9874\times 91}\approx 56.632$$

这个值到底能告诉我们什么呢？

统计学家经过研究后发现，在 H_0 成立的情况下，
（ 2 ）

在 成立的情况下， K^2 的观测值大于 6.635 的概率非常小，近似为 0.010 ，是个小概率事件 .

现在 K^2 的观测值 $k \approx 56.632$，远远大于 6.635，所以有理由断定 H_0 不成立，即认为 “ 吸烟与患肺癌有关系 ”. 但这种判断会犯错误，犯错误的概率不会超过 .

上面这种利用随机变量 K^2 来确定是否能以一定把握认为 “ 两个分类变量有关系 ” 的方法，称为两个分类变量的 独立性检验 .

3. 课堂总结

【知识梳理】

(1) 变量的不用 “ 值 ” 表示个体所属的不同类别，像这样的变量成为分类变量 .

(2) 列出两个分类变量的频数表，称为列联表 .

(3) 设 ：吸烟与患肺癌没有关系 ，那么吸烟样本中不患肺癌的比例应该与不吸 烟样本中相应的比例差不多，即：

$\frac{a}{a+b} \approx \frac{c}{c+d}$，即 $ad \approx bc$.

因此， 越小，说明吸烟与患肺癌之间关系越弱； 越大，说明吸烟与患肺癌之间关系越强 .

为了使不同样本容量的数据有统一的评判标准，构造一个随机变量

$$K^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)} \quad (1)$$

（其中 $n = a+b+c+d$ 为样本容量 . ）

若 H_0 成立，即 “ 吸烟与患肺癌没有关系 ” ，则 K^2 应该很小 .

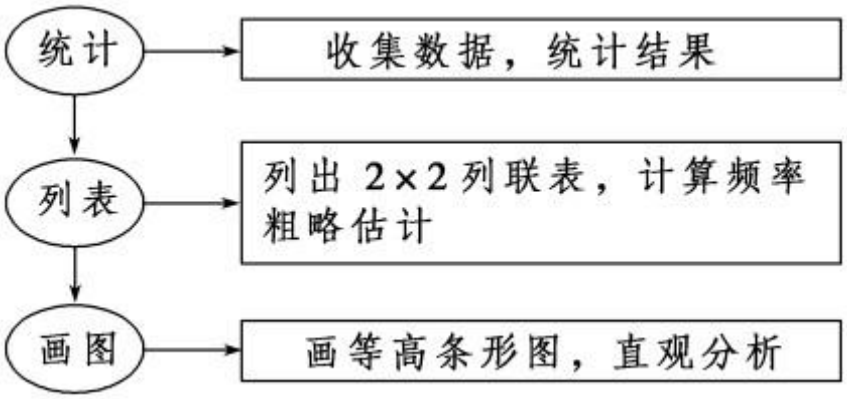
【重难点突破】

(1) 列联表与等高条形图

列联表由两个分类变量之间频率大小差异说明这两个变量之间是否有关联关系，而利用等高条形图能形象直观地反映它们之间的差异，进而推断它们之间是否具

有关联关系．一般地，在等高条形图中，　　与　　相差越大，两个分类变量有关系的可能性就越大．

(2) 利用等高条形图判断两个分类变量是否有关的步骤：



4. 随堂检测

1. 独立性检验中，可以粗略地判断两个分类变量是否有关的是（　　）

A. 残差　B. 等高条形图　C. 假设检验的思想　D. 以上都不对

【知识点：独立性检验】

解：B.

2. 分类变量 X 和 Y 的列联表如下，则（　　）

	Y_1	Y_2	合计
X_1	a	b	a + b
	c	d	c + d
合计	a + c	b + d	a + b + c + d

A.　　越小，说明　　与　　的关系越弱

B.　　越大，说明　　与 Y 的关系越强

C. $(ad-bc)^2$ 越大，说明 X 与 Y 的关系越强

D. $(ad-bc)^2$越接近于　，说明　　与 Y 关系越强

【知识点：独立性检验】

解：C　解：因为 $K^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，　越大，

越大，　犯错误的概率的越小，说明　与 Y 的关系越强．

3.. 在一次独立性检验中，得出 2×2 列联表如下：

	y 1	y 2	合计
x 1	200	800	1000
x 2	180	m	180+ m
合计	380	800+ m	1180+ m

最后发现，两个分类变量 x 和 y 没有任何关系，则 m 的可能值是（　）

A.200 B.720 C.100 D.180

【知识点：独立性检验】

解：B　分类变量 x 和　没有任何的关系，所以

，得到 $m=720$，故选 B.

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

A.99.9% B.95% C.90% D.0

附表：

	0.050	0.010	0.001
k	3.841	6.635	10.828

【知识点：独立性检验】

解：A　因为所求的 $k^2=13.097>10.828$，故可能性为 99.9%，所以选 A.

5. 某校为了研究学生的性别和对待某一活动的态度（支持和不支持两种态度）的关系，运用 2×2 列联表进行独立性检验，经计算 K 2 ＝ 7.069 ，则至少有_ 的把握认为 “ 学生性别与是否支持该活动有关系 ”.

附：

P（K 2 ≥ k 0）	0.100	0.050	0.025	0.010	0.001
k 0	2.706	3.841	5.024	6.635	10.828

【知识点：独立性检验】

解： 99 %

（三）课后作业

基础型　自主突破

1. 下面说法正确的是（　　）

A. 统计方法的特点是统计推断准确、有效

B. 独立性检验的基本思想类似于数学上的反证法

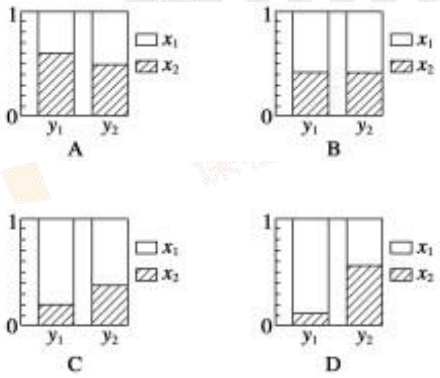
C. 任何两个分类变量有关系的可信度都可以通过查表得到

D. 不能从等高条形图中看出两个分类变量是否相关

【知识点：独立性检验】

解： B

2. 观察下列各图，其中两个分类变量 x ， y 之间关系最强的是（　　）



【知识点：独立性检验】

解： D

3. 确定结论 “ X 与 Y 有关系 ” 的可信度为 % 时，则随机变量 的观测值 k 必须（ ）

A. 大于 B. 大于 C. 小于 D. 大于 2.706

【知识点：独立性检验】

解： B 通过表中的数据可知可信度为 95% 时 $k^2>3.841$

P（K 2 ≥ k 0）	0.05	0.025	0.010	0.005	0.001
k 0	3.841	5.024	6.635	7.879	10.828

4. 想要检验是否喜欢参加体育活动是不是与性别有关，应该检验（ ）

A. H 0 ：男性喜欢参加体育活动 B. H 0 ：女性不喜欢参加体育活动

C. H 0 ：喜欢参加体育活动与性别有关 D. H 0 ：喜欢参加体育活动与性别无关

【知识点：独立性检验】

解： D

5. 对分类变量 X 与 Y 的随机变量 K^2 的观测值 K ，说法正确的是（ ）

A . K 越大，″ X 与 Y 有关系 ” 可信程度越小 ；

B . K 越小，″ X 与 Y 有关系 ” 可信程度越小 ；

C . K 越接近于 0 ，″ X 与 Y 无关 ” 程度越小

D . K 越大，″ X 与 Y 无关 ” 程度越大

【知识点：独立性检验】

解： B

能力型 师生共研

6. 若有 99.9% 的把握说事件 与事件 B 有关，那么具体算出的 的观测值 k 一定满足（ ）

A.　　B. $k<10.828$　C. $k>6.635$　D. $k<6.635$

【知识点：独立性检验】

解：　A

7. 假设有两个分类变量 X 和 Y ，它们的值域分别为 {x1 ， x2} 和 {y1 ，y 2 } ，其 2×2 列联表为：

Y	y 1	y 2	总计
X			
x 1	a	b	a ＋ b
x 2	c	d	c ＋ d
总计	a ＋ c	b ＋ d	a ＋ b ＋ c ＋ d

对同一样本，以下数据能说明 X 与 Y 有关的可能性最大的一组为（ D ）

A. a ＝ 5 ， b ＝ 4 ， c ＝ 3 ， d ＝ 2

B. a ＝ 5 ， b ＝ 3 ， c ＝ 4 ， d ＝ 2

C. a ＝ 2 ， b ＝ 3 ， c ＝ 4 ， d ＝ 5

D. a ＝ 3 ， b ＝ 2 ， c ＝ 4 ， d ＝ 5

【知识点：独立性检验】

解：　D

8. 某校为了研究学生的性别和对待某一活动的态度（ 支持与不支持 ）的关系，运用 2 ×2 列联表进行独立性检验，经计算 K 2 =7.069 ，则所得到的统计学结论为：有　把握认为 “ 学生性别与支持该活动有关系 ”

P（ $K^2 \geq k$ 0 ）	0.100	0.050	0.025	0.010	0.001
k .	2.706	3.841	5.024	6.635	10.828

【知识点：独立性检验】

解：　99 %　【解析】根据 $6.635<7.069<10.828$，所以犯错误率低于　，

所以应 该有 99% 的把握，认为 “ 学生性别与支持该活动有关系 ” ，

探究型　多维突破

9. 某班主任对全班 50 名学生学习积极性和对待班级工作的态度进行了调查，统计数据如下表所示：

	积极参加班级工作	不太主动参加班级工作	合计
学习积极性高	18	7	25
学习积极性一般	6	19	25
合计	24	26	50

(1) 如果随机抽查这个班的一名学生，那么抽到积极参加班级工作的学生的概率是多少？抽到不太主动参加班级工作且学习积极性一般的学生的概率是多少？

(2) 试运用独立性检验的思想方法点拨：学生的学习积极性与对待班级工作的态度是否有关系？并说明理由 .（ 参考下表 ）

P（K2≥k）	0.50	0.40	0.25	0.15	0.10	0.05	0.025	0.010	0.005	0.001
k	0.455	0.708	1.323	2.072	2.706	3.841	5.024	6.635	7.879	10.828

【知识点：独立性检验，古典概型】

解：(1) 积极参加班级工作的学生有 24 人，总人数为 50 人，概率为

；

不太主动参加班级工作且学习积极性一般的学生有 19 人，概率为 $\frac{19}{50}$.

(2) $K^2=\frac{50\times(18\times19-6\times7)^2}{24\times25\times24\times26}=\frac{150}{13}\approx11.5$，

∵ ，

∴ 有 的把握说学习积极性与对待班级工作的态度有关系 .

10. 2016 年夏季奥运会将在巴西里约热内卢举行，体育频道为了解某地区关于奥运会直播的收视情况，随机抽取了　　名观众进行调查，其中　　岁以上的观众有　　名，下面是根据调查结果绘制的观众准备平均每天收看奥运会直播时间的频率分布表（ 时间：分钟 ）：

分组			$[40,60)$		$[100,120)$
频率	0.1	0.18	0.22	0.2	0.05

将每天准备收看奥运会直播的时间不低于 80 分钟的观众称为 “ 奥运迷 ” ，已知 “ 奥运迷 ” 中有 10 名　　岁以上的观众 .

（ 1 ）根据已知条件完成下面的　　列联表，并据此资料你是否有 95% 以上的把握认为 “ 奥运迷 ” 与年龄有关？

	非 “ 奥运迷 ”	“ 奥运迷 ”	合计
40 岁以下			
岁以上			
合计			

（ 2 ）将每天准备收看奥运会直播不低于　　分钟的观众称为 “ 超级奥运迷 ” ，已知 “ 超级奥运迷 ” 中有 2 名　　岁以上的观众，若从 “ 超级奥运迷 ” 中任意选取　人，求至少有 1 名 40 岁以上的观众的概率 .

附：$K^2=\dfrac{n(ad-bc)^2}{(a+b)(a+d)(a+c)(b+d)}$

k	3.841	6.635

【知识点：独立性检验 ， 概率统计】

深圳小学家长群:254317299　　深圳初中家长群 :90482695　　深圳高中家长群 :175743089
更多资料详见: http://sz.jiajiaoban.com/　　咨询电话 : 4000-121-121

解：（ 1 ）由频率分布表可知，在轴取的 100 人中，“ 奥运迷 ” 有 25 人，从完成 2×2 列联表如下：

	非 “ 奥运迷 ”	“ 奥运迷 ”	合计
岁以下	30		
40 岁以上			55
合计	75	25	

$$K^2=\frac{100\times(30\times10-45\times15)^2}{75\times25\times45\times55}=\frac{100}{33}\approx3.030$$.

因为 ，所以没有 95% 以上的把握认为 “ 奥运迷 ” 与年龄有关 .

（ 2 ）由频率分布表可知， “ 超级奥运迷 ” 有 5 人，从而所有可能结果所组成的基本事件空间为：

其中 a_i 表示男性， $i=1,2,3,b_i$ 表示女性， $i=1,2$. Ω 由 个基本事件组成，且是等可能的，用 A 表示事件 “ 任意选 人，至少有 1 名 岁以上观众 ” ，则 ，即事件 A 包含 个基本事件，所以 $P(A)=\frac{7}{10}$.

（四）自助餐

1. 在等高条形图中，下列哪两个比值相差越大，要推断的论述成立的可能性就越大（ ）

A. $\frac{a}{a+b}$ 与 $\frac{d}{c+d}$ B. 与 $\frac{a}{c+d}$ C. 与 D. $\frac{a}{a+b}$ 与 $\frac{c}{b+c}$

【知识点：独立性检验】

深圳小学家长群:254317299 深圳初中家长群 :90482695 深圳高中家长群 :175743089
更多资料详见: http://sz.jiajiaoban.com/ 咨询电话：4000-121-121

解：　　C

2. 为了调查中学生近视情况，某校　　名男生中有　　名近视，140名女生中有70名近视．在检验这些中学生眼睛近视是否与性别有关时用什么方法最有说服力（　　）

A. 平均数　B. 方差　C. 回归分析　D. 独立性检验

【知识点：独立性检验】

解：　D　本例考查学生眼睛的“近视”与“性别”两件事情之间是否存在相关性，从给出的数据可以列出2×2列联表，所以适合用独立性检验．

3. 在一个 2×2 列联表中，由其数据计算得 K 2 的观测值 k ＝ 7.097 ，则这两个变量间有关系的可能性为　（　　）

A.99% B.99.5% C.99.9% D. 无关系

【知识点：独立性检验】

解：　A　由表格数据可知 k ＝ 7.097>6.635 ，所以这两个变量间有关系的可能性为 99%

4. 某疾病研究所想知道吸烟与患肺病是否有关，于是随机抽取　　名成年人调查是否抽烟及是否患有肺病得到2×2列联表，经计算得$K^2=5.231$，已知在假设吸烟与患肺病无关的前提条件下，

．则该研究所可以（　　）

A. 有95%以上的把握认为“吸烟与患肺病有关”

B. 有95%以上的把握认为“吸烟与患肺病无关”

C. 有99%以上的把握认为“吸烟与患肺病有关”

D. 有　　以上的把握认为“吸烟与患肺病无关”

【知识点：独立性检验】

解： A 因为 $K^2=5.231>3.841$，而 ，故有有 以上的把握认为 “ 吸烟与患肺病有关 ”

5.2016 年 3 月 9 日至 15 日，谷歌人工智能系统 “ 阿尔法 ” 迎战围棋冠军李世石，最终结果 “ 阿尔法 ” 以总比分 4 比 1 战胜李世石 . 许多人认为这场比赛是人类的胜利，也有许多人持反对意见，有网友为此进行了调查，在参加调查的 2548 名男性中有 1560 名持反对意见， 2452 名女性中有 1200 名持反对意见，在运用这些数据说明 “ 性别 ” 对判断 “ 人机大战是人类的胜利 ” 是否有关系时，应采用的统计方法是（ ）

A. 茎叶图 B. 分层抽样 C. 独立性检验 D. 回归直线方程

【知识点：独立性检验】

解 ： C 这是独立性检验，因为这里有两个分类变量，一个是性别分为男女，一个是意见分为支持和反对，这样就构成一个 联表，用独立性检验来验证 “ 人机大战是人类的胜利 ” 是否有关系 .

6. 某人研究中学生的性别与成绩、视力、智商、阅读量这 4 个变量之间的关系，随机抽查 52 名中学生，得到统计数据如表 1 至表 4 ，这与性别有关联的可能性最大的 变量是（ ）

表1

成绩 性别	不及格	及格	总计
男	6	14	20
女	10	22	32
总计	16	36	52

表2

视力 性别	好	差	总计
男	4	16	20
女	12	20	32
总计	16	36	52

A. 成绩　B. 视力　C. 智商　D. 阅读量

【知识点：独立性检验】

解：　D　　由表中数据可得：

表 1：$K^2=\frac{52\times(6\times22-10\times14)}{16\times36\times20\times32}\approx0.009$；表 2：$K^2=\frac{52\times(4\times20-12\times16)}{16\times36\times20\times32}\approx1.769$；

表 3：$K^2=\frac{52\times(8\times24-12\times8)}{16\times36\times20\times32}\approx1.3$；表 4：　　.

其中　　最大，所以阅读量与性别有关联的可能性最大．

7. 如下表是对于喜欢足球与否的统计列联表依据表中的数据，得到 K^2　.

【知识点：独立性检验】

解：　　.

8. 若由一个 2×2 列联表中的数据计算得 K 2 的观测值 k ＝ 4.013 ，那么在犯错误的概率不超过 _____ 的前提下认为两个变量之间有关系 .

【知识点：独立性检验】

解：　0.05　因随机变量 K 2 的观测值 k ＝ 4.013 ＞ 3.841 ，因此，在犯错误的概率不超过 0.05 的前提下，认为两个变量之间有关系 .

9. 如果 K 2 的观测值为 6.645 ，可以认为 “ x 与 y 无关 ” 的可信度是 _____.

【知识点：独立性检验】

解：　　1%

10. 某学校对该校学生作了一项调查发现 ： 在平时的模拟考试中 ， 性格内向的学生 426 人中 332 人在考前心情紧张 ， 性格外向的学生 594 人中有 213 人在考前心情紧张 ， 作出等高条形图 ， 利用图形判断考前心情紧张与性格类别是否有关系 .

【知识点：独立性检验】

解：作列联表如下 ：

	性格内向	性格外向	总计
考前心情紧张	332	213	545
考前心情不紧张	94	381	475
总计	426	594	1020

相应的等高条形图如图所示 ：

图中阴影部分表示考前心情紧张与考前心情不紧张中性格内向的比例 ， 从图中可以看出考前心情紧张的样本中性格内向占的比例比考前心情不紧张样本中性格内向占的比例大 ， 可以认为考前紧张与性格类型有关 .

11. 甲、乙两所学校高三年级分别有 1200 人， 1000 人，为了了解两所学校全体高三年级学生在该地区六校联考的数学成绩情况，采用分层抽样方法从两所学校一共抽取了 110 名学生的数学成绩，并作出了频数分布统计表如下：

甲校	分组	[70 ， 80 ）	[80 ， 90 ）	[90 ， 100 ）	[100 ， 110 ）
	频数	3	4	8	15
	分组	[110 ， 120 ）	[120 ， 130 ）	[130 ， 140 ）	[140 ， 150]
	频数	15	x	3	2

乙校	分组	[70，80）	[80，90）	[90，100）	[100，110）
	频数	1	2	8	9
	分组	[110，120）	[120，130）	[130，140）	[140，150]
	频数	10	10	y	3

	甲校	乙校	总计
优秀			
非优秀			
总计			

（1）计算 x，　　的值；

（2）若规定考试成绩在［120，150］内为优秀，请分别估计两所学校数学成绩的优秀率；

（3）根据以上统计数据完成 2×2 列联表，并判断是否有 90% 的把握认为两所学校的数学成绩有差异．

【知识点：独立性检验，分层抽样，概率统计】

解：（1）x＝10，y＝7；（2）甲乙分别为；25%，40%（3）见解析．

试题分析：（1）由题为分层抽样，可确定出甲乙两个学校分别抽取的人数，然后结合频数表，可求出 x，y 的值；

（2）由题给出了优秀的标准，结合给出的表格，可分别求甲乙学校的数学成绩的优秀率，（即由每个学校优秀的人数除以它们的人数）；

（3）由题为独立性检验；可先做出二列联表，再代入独立性检验的公式，求出　　，对应参考值可下结论．

试题解析：（1）甲校抽取 $110\times\frac{1200}{2200}=$ 60 人，

乙校抽取 $110\times\frac{1000}{2200}=50$ 人，

故 x ＝ 10 ， y ＝ 7 ，

（ 2 ） 估计甲校优秀率为 $\frac{15}{60}=25\%$ ，

乙校优秀率为 $\frac{20}{50}$ ＝ 40%.

（ 3 ）表格填写如图，

	甲校	乙校	总计
优秀	15	20	35
非优秀	45	30	75
总计	60	50	110

k 2 ＝ $\frac{110(15\times30-20\times45)^2}{60\times50\times35\times75}$ ≈2.83 ＞ 2.706

又因为 1 － 0.10 ＝ 0.9 ，故有 90% 的把握认为两个学校的数学成绩有差异 .

学而思1对1 课程咨询：4000-121-121