第9章

统　计

统计是研究如何收集、处理、分析数据并作出结论的科学，它是分析并认识客观现象的有力工具．在本章中，学生将在必修课程已经学习的统计知识的基础上，通过对典型案例的讨论，了解和使用一些常用的统计方法，例如回归分析和独立性检验：利用线性回归模型研究两个变量之间的随机关系，进行预测；利用2×2列联表检验两个随机变量的独立性．在本章的教学中，应通过具体案例引导学生参与数据分析的全过程．

本章重点提升学生的逻辑推理、数学建模、数学运算、数据分析素养.

第1课时　变量的相关性(1)



知识技能

1. 通过实例，了解变量间的相关关系，会区别变量之间的函数关系与相关关系．

2. 通过收集实际问题中两个有关联变量的数据，作出散点图，通过散点图能够判断两个变量的相关性．

思想方法

1. 通过实例感受运用统计方法处理相关关系的过程：随机抽取样本，运用数学理论和方法处理样本数据，建立模型，再运用数学模型进行预测和估计．

2. 利用散点图发现线性相关的特点：数据落在一条直线附近．

核心素养

1. 通过实例，体会相关关系不是函数关系，但相关关系可以用函数关系近似地表达，从而进行辩证思维的训练，发展数据分析素养．

2. 在运用散点图解决问题的过程中，发展数学运算与直观想象素养．



重点：相关关系的概念．

难点：利用散点图直观认识两个变量之间的线性相关关系．



问题导引

预习教材P139～141，思考下面的问题：

1. 两个变量间的关系分类有哪些？

提示　两个变量间的关系分为三类：一类是确定性的函数关系；另一类是变量间确实存在关系，但又不具备函数关系所要求的确定性，它们的关系是带有随机性的，这种关系就是相关关系；第三类是两变量没有任何关系．

2. 怎样确定两个变量*x, y*是否具有相关关系？

提示　(1) 散点图法：通过散点图，观察它们的分布是否存在一定规律，直观地判断；如果发现点的分布从整体上看大致在一条直线附近，那么这两个变量就是线性相关的，注意不要受个别点的位置的影响．

(2) 表格、关系式法：结合表格或关系式进行判断．

(3) 经验法：借助积累的经验进行分析判断．

3. 什么是线性相关关系？怎样理解正相关和负相关？

即时体验

1. 判断(正确的打“√”，错误的打“×”)：

(1) 粮食产量和当年的降雨量是一种函数关系；(×)

(2) 两个变量之间产生相关关系的原因受许多不确定的随机因素的影响；(√)

(3) 需要通过样本来判断变量之间是否存在不同关系；(√)

(4) 相关关系是一种因果关系，具有确定性．(×)

2. 下列两个变量之间具有相关关系但不具有函数关系的是(C)

A. 圆的半径与面积

B. 匀速行驶的车辆的行驶路程与时间

C. 人的身高与体重

D. 人的身高与视力

3. 若“名师出高徒”成立，则名师与高徒之间存在的关系是(A)

A. 相关关系 B. 函数关系

C. 无任何关系 D. 不能确定

4. 已知两个变量*x, y*，下列散点图中，*x, y*呈正相关趋势的是(A)





一、 问题情境

公安人员在勘察案发现场时，总是非常仔细地搜寻犯罪嫌疑人的脚印，一个重要的原因就是可以根据脚的大小来推测其身高．从直觉上看，成年人的身高与脚长(脚趾到脚跟的长度)存在着某种关系．

相关研究发现，我国成年人的身高(单位：cm)与脚长(单位：cm)有着下面的关系：

人的身高＝6.876×脚长＋误差．

(1) 若案犯留下的脚印长度为25cm，你能推断出什么结论？

答：6.876×25＝171.9(cm)，可以推断出脚印长度为25cm的人的身高在171.9cm上下波动．

(2) 人的脚长与身高之间具有某种关联，它们之间是确定的函数关系吗？

答：两个脚长一样的人，他们的身高并不一定相同，也就是说，人的脚长与身高之间并不是确定的函数关系．

生活中还存在着大量的这类问题：人的体重与身高之间的关系，商品的销售量与广告宣传费之间的关系，农作物的产量与施肥量之间的关系，家庭的支出与收入之间的关系……

这类问题中的两个量之间是一种怎样的关系呢？[1]

二、 数学建构

(一) 生成概念

人的体重与身高有关，一般来说，身高越高，体重越重，但不能用一个函数来严格地表示身高与体重之间的关系[2]；商品的销售量与广告宣传费有关，一般来说，广告宣传费越高，销售量就越大，但不能用一个函数来确切地表示商品的销售量与广告宣传费之间的关系；农作物的产量与施肥量有关，一般来说，在一定范围内，施肥量越多，农作物的产量就越高，但不能用一个函数来准确地表示产量与施肥量之间的关系；家庭的支出与收入有关，一般来说，收入高的家庭，往往支出也较多，但不能用一个函数来精确地表示家庭支出与收入之间的关系．

通常把上述问题中的身高、广告宣传费、施肥量、家庭收入称为自变量，与之对应的体重、商品的销售量、农作物的产量、家庭支出称为相应的因变量．

问题1　根据问题情境中的几个问题，你能说出自变量与因变量之间是一种怎样的关系吗？

对于上述几个问题，在多次重复观测中，自变量取一定值，因变量不一定取一个确定的值与之对应，而是有或多或少的差异．这是因为作为因变量的事物，除受问题中的自变量的影响外，还受到其他许多因素的影响．这些因素中有些是可知的，有些可能难以明确．

像这样，两个变量之间具有一定的联系，但又没有确定性函数关系，这种关系称为相关关系．

问题2　相关关系和函数关系有什么区别？你还能举出具有这种关系的其他例子吗？

(函数关系中，对于确定的自变量，因变量具有确定性；而相关关系中，对于确定的自变量，因变量的取值具有随机性．在生活还有很多具有相关关系的两个变量：例如未成年人的年龄与身高、网速与下载文件所需的时间、吸烟与患呼吸道疾病之间的关系……)

问题3　在分析上面几个问题时，大多靠“直觉”与“经验”进行判断，如何进行更为科学、严密的推断呢？怎样利用统计理论和方法处理这些问题呢？

(随机抽取数据，获得样本数据，处理样本数据，建立数学模型，再利用模型进行预测与估计．)

问题4　在一次对人体脂肪含量和年龄关系的研究中，研究人员获得了一组样本数据：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 年龄/岁 | 23 | 27 | 39 | 41 | 45 | 49 | 50 |
| 脂肪 | 9.5 | 17.8 | 21.2 | 25.9 | 27.5 | 26.3 | 28.2 |
| 年龄/岁 | 53 | 54 | 56 | 57 | 58 | 60 | 61 |
| 脂肪 | 29.6 | 30.2 | 31.4 | 30.8 | 33.5 | 35.2 | 34.6 |

根据上述数据，人体的脂肪含量与年龄之间有怎样的关系？有没有直观的方式来进行表示？

用*x*轴表示年龄，*y*轴表示脂肪，建立平面直角坐标系，将表中数据构成的点在坐标系中标出(一组样本数据就对应着一个点)．今后我们称这类图为散点图．



通过散点图可以看出，随着年龄的增长，脂肪含量也在上升，这些点散布在一条直线附近，说明脂肪含量*y*与年龄*x*具有相关关系．我们将具有这种特性的相关关系称为线性相关关系．

问题5　观察下列4幅图，图象有什么特点？



图1



图2



图3



图4

图1呈从左下向右上方向发展的趋势(上升趋势)，称这两个变量之间正相关．这就像函数中的增函数，即一个变量从小到大，另一个变量也从小到大．

图2呈从左上逐渐向右下方向发展的趋势(下降趋势)，称这两个变量之间负相关．这就像函数中的减函数，即一个变量从小到大，另一个变量从大到小．

前面两幅图中点的分布呈条状，后面两幅图很乱．从数学的角度解释：图1, 2中的点的分布从整体上看大致在一条直线附近，具有线性相关关系；图3, 4中的两个变量是非线性相关关系．[3]

(二) 理解概念

1. 变量之间的相关关系

(1) 变量之间的关系，常见的有两类：一类是函数关系，另一类是相关关系．

相关关系与函数关系的异同点：

相同点：均是指两个变量的关系．

不同点：函数关系是一种确定的关系，相关关系是一种非确定关系．

(2) 两个变量之间产生相关关系的原因是受许多不确定的随机因素的影响．

(3) 需要通过样本来判断变量之间是否存在相关关系．

2. 判断相关性的常用统计图是散点图．

三、 数学运用

例1[4]　(1) 下列两个变量之间的关系是相关关系的是(C)

A. 正方体的棱长与体积

B. 单位面积的产量为常数时，土地面积与总产量

C. 日照时间与水稻的亩产量

D. 电压一定时，电流与电阻

(2) 下列说法中正确的有①②⑤.(填序号)

①函数关系是一种确定性关系；②相关关系是一种非确定性关系；③任何两个变量都具有相关关系；④圆的周长与该圆的半径具有相关关系；⑤某商品的需求量与该商品的价格是一种非确定性关系．

(见学生用书课堂本P86)

[处理建议]　根据函数关系、相关关系定义的差异性辨析．相关关系是一种非确定的关系，函数关系是一种确定的关系．

[规范板书]　解析　(1) A, B, D中的两个变量之间的关系是确定的，不是相关关系；对于C，日照时间会影响水稻的亩产量，但不是唯一的影响因素，它们之间是相关关系．

(2) 任何两个变量不一定都具有相关关系，故③错；圆的周长与该圆的半径是函数关系，而不是相关关系，故④错．

[题后反思]　相关关系与函数关系均是指两个变量之间的关系，但函数关系是一种确定的关系，相关关系是一种非确定的关系，具有随机性的特征．判断变量之间是否具有相关关系，在没给出具体数值的前提下，可根据感觉直观判断，此时要用到已有的知识或生活中的经验．

　(1) 下列两个变量之间的关系不是函数关系的是(B)

A. 角度和它的正切值

B. 人的右手一拃长和身高

C. 正方体的棱长和表面积

D. 真空中做自由落体运动的物体的下落距离和下落时间

(2) 下列两个变量之间的关系是相关关系的是(C)

A. 匀速直线运动的物体的运动时间与位移

B. 学生的成绩和体重

C. 路上酒后驾驶的人数和交通事故发生的多少

D. 水的体积和质量

[规范板书]　解析　(1) 角度和它的正切值有确定的关系*y*＝tan*x*；人的右手一拃长和身高不具有确定的关系，故不是函数关系；正方体的棱长和表面积有关系*S*＝6*a*2；真空中做自由落体运动的物体的下落距离和下落时间有确定的关系*s*＝*gt*2.

(2) 匀速直线运动的物体的运动时间与位移的关系是函数关系；学生成绩与体重之间不具有相关性；路上酒后驾驶的人数和交通事故发生的多少的关系是相关关系；水的体积与质量的关系是函数关系．

[题后反思]

|  |  |  |
| --- | --- | --- |
|  | 确定性函数关系 | 相关关系 |
| 相同点 | 都是指变量之间的关系 |
| 不同点 | 是一种确定的关系 | 是一种非确定的关系 |
| 是两个可控变量之间的关系 | 一个是可控变量，另一个是随机变量；或两个都是随机变量 |
| 是一种因果关系 | 不一定是因果关系，可能是伴随关系 |
| 是一种理想关系模型 | 是更为一般的情况 |

例2　(多选)下列散点图中，两个变量*x, y*具有相关关系的有[5](CD)



 (见学生用书课堂本P86)

[处理建议]　根据散点图的图形特征解答．

[规范板书]　解析　C中的点大致分布在一条直线附近，D中的点大致分布在一条曲线附近，所以C, D中的两个变量具有相关关系．

[题后反思]　在散点图中，从整体上，如果散点大部分分布在一条曲线附近，那么就说两变量具有相关关系；从整体上，如果散点大致在一条直线附近，那么就说两变量具有线性相关关系．

　下列散点图中，两个变量之间是正相关的是(A)



[处理建议]　根据散点图的图形特征以及正相关、负相关的概念解答．

[规范板书]　解析　对于A，散点图中的点从左向右是上升的，且在一条直线附近，是正相关；对于B，散点图中的点从左向右是下降的，且在一条直线附近，是负相关；对于C, D，散点图中的点不成带状分布，没有明显的相关关系．

　下列反映两个变量的相关关系中，不同于其他三个的是(C)

A. 名师出高徒 B. 水涨船高

C. 月明星稀 D. 登高望远

[规范板书]　解析　名师出高徒是正相关，水涨船高是正相关，月明星稀是负相关，登高望远是正相关．

　下列关于散点图的说法中正确的是(C)

A. 一定可以看出变量之间的变化规律

B. 一定不可以看出变量之间的变化规律

C. 可以看出正相关与负相关有明显区别

D. 看不出正相关与负相关有什么区别

[规范板书]　解析　给出一组样本数据，总可以作出相应的散点图，但不一定能分析出两个变量的关系，但是通过散点图可以看出正相关与负相关有明显区别．

[题后反思]　并非所有散点图都呈现出规律性的特征，因此由散点图不一定能看出变量之间的规律．另外，若变量具有线性相关关系，在所呈现的散点图中，可根据散点呈左下向右上方向发展的趋势，还是呈左上向右下方向发展的趋势，区分正相关与负相关．

例3　某个男孩的年龄(单位：岁)与身高(单位：cm)的统计数据如下表所示：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 年龄/岁 | 1 | 2 | 3 | 4 | 5 | 6 |
| 身高/cm | 78 | 87 | 98 | 108 | 115 | 120 |

(1) 作出散点图；

(2) 判断*y*与*x*是否具有线性相关关系．[6]

(见学生用书课堂本P86)

[处理建议]　在平面直角坐标系中作出对应的散点，根据散点图，结合*y*随*x*的变化情况进行判断：*y*随*x*的增大而增大，且散点图中点的分布从整体上看大致在一条直线附近，那么这两个变量具有线性相关关系．

[规范板书]　解　(1) 用*x*轴表示年龄，*y*轴表示身高，作出散点图(如图)．



(例3)

(2) 由图象可知，所有数据点大致在一条直线附近，所以*y*与*x*具有线性相关关系．

[题后反思]　由两个变量相应值的对应关系，作出散点图．观察散点图，根据各点是否分布在某条直线周围，判断变量之间是否具有线性相关关系．再结合图象走向或结合*y*随*x*的变化情况，判断是正相关还是负相关．

　有人收集了春节期间平均气温(单位：℃)与某取暖商品销售额(单位：万元)的有关数据，具体数据如下表所示：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 平均气温/℃ | －2 | －3 | －5 | －6 |
| 销售额/万元 | 20 | 23 | 27 | 30 |

该商品销售额与平均气温有(C)

A. 确定性关系 B. 正相关关系

C. 负相关关系 D. 函数关系

[规范板书]　解析　方法1：根据实际生活经验可知，气温与取暖商品的销售额线性相关，且负相关．

方法2：由表中数据可知，*y*随*x*的减小而增大，是负相关关系．

方法3：可作出散点图，根据图象走向进行判断．

[题后反思]　根据变量的取值判断时，观察一个变量是否随另一个变量的增大而增大(或增大而减小)，从而判断两个变量之间是否具有相关关系．

　两对变量*A*和*B, C*和*D*的取值分别对应如表1和表2，画出散点图，判断它们是否具有相关关系．若具有相关关系，说出它们相关关系的区别．

表1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *A* | 26 | 18 | 13 | 10 | 4 | －1 |
| *B* | 20 | 24 | 34 | 38 | 50 | 64 |

表2

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *C* | 0 | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| *D* | 541.67 | 602.66 | 672.09 | 704.99 | 806.71 | 908.59 | 975.42 | 1034.75 |

[处理建议]　将提供的数据整合成散点图，结合正相关、负相关、线性相关的概念解决问题．

[规范板书]　解　散点图分别如图所示：

 

图1 图2

(变式2)

从图中可以看出两图中的点都分布在一条直线附近，因此两图中的变量都具有相关关系．

图1中*A*的值由大变小时，*B*的值却是由小变大的，故*A*和*B*负相关．

图2中*C*的值由小变大时，*D*的值也是由小变大的，故*C*和*D*正相关．

[题后反思]　运用相关性的概念解决生活中的问题时，要善于将生活中的问题转化为数学问题，例如对数据的处理，将图表数据转化为散点图，从图形中直观感受出正相关与负相关．

　有时候，一些东西吃起来口味越好，对我们的身体越有害，下表给出了某种食品不同类型的数据．第二行数据表示此种食品所含热量的百分比，第三行数据表示由一些美食家以百分制给出的对此种食品口味的评价．

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 品　牌 | *A* | *B* | *C* | *D* | *E* | *F* | *G* | *H* | *I* | *J* |
| 所含热量的百分比 | 25 | 34 | 20 | 19 | 26 | 20 | 19 | 24 | 19 | 13 |
| 口味记录 | 89 | 89 | 80 | 78 | 75 | 71 | 65 | 62 | 60 | 52 |

(1) 作出散点图．

(2) 你能从散点图中发现两者之间的近似关系吗？

(3) 如果近似成线性关系，请画出一条直线来近似地表示这种线性关系．

(4) 为什么人们更喜欢吃位于直线上方的食品而不是下方的？

[规范板书]　解　(1) 散点图如图所示：



(变式3)

(2) 从上图看两者近似成线性相关关系．

(3) 直线如上图所示．

(4) 因为当直线上方的食品和下方的食品所含热量相同时，直线上方的食品口味更好，所以人们更喜欢吃位于直线上方的食品而不是下方的．

四、 课堂练习

1. 对于给定的两个变量的统计数据，下列说法中正确的是(C)

A. 都可以分析出两个变量的关系

B. 都可以用一条直线近似地表示两者的关系

C. 都可以作出散点图

D. 都可以用确定的表达式表示两者的关系

2. (多选)下列两个量之间的关系是相关关系的有(BC)

A. 考试号与考生考试成绩

B. 勤能补拙

C. 水稻产量与气候

D. 正方形的边长与正方形的面积

3. 某商场五天内某种T恤衫的销售情况如下表所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 第*x*天 | 1 | 2 | 3 | 4 | 5 |
| 销售量*y*/件 | 19 | 39 | 59 | 79 | 104 |

下列说法中正确的是(B)

A. *y*与*x*负相关 B. *y*与*x*正相关

C. *y*与*x*不相关 D. *y*与*x*成正比例关系

4. 某公司近年来科研费用*x*(单位：万元)与公司所获的利润*y*(单位：万元)之间有如下的统计数据：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *x*/万元 | 2 | 3 | 4 | 5 |
| *y*/万元 | 18 | 27 | 32 | 35 |

(1) 作出散点图；

(2) 判断*y*与*x*是否具有线性相关关系．

解　(1) 散点图如图所示：



(第4题)

(2) 由图象可知，所有数据点大致在一条直线附近，所以*y*与*x*具有线性相关关系．

五、 课堂小结

1. 相关关系与函数关系均是指两个变量之间的关系，但函数关系是一种确定的关系，相关关系是一种不确定的关系．

2. 除了根据直观感觉判断变量之间的关系，还可以借助散点图来判断．若各点大致分布在一条直线附近，则两个变量线性相关．在散点图中，各点呈上升趋势(*y*随*x*的增大而增大)，两个变量正相关；各点呈下降趋势(*y*随*x*的增大而减小)，两个变量负相关.



[1] 通过学生熟悉的函数关系，引导学生关注生活中两个变量之间还存在的相关关系．体会研究变量之间相关关系的重要性，感受数学来源于生活．

[2] 身高并不是决定体重的唯一因素．例如生活中的饮食习惯、睡眠时间、体育锻炼时间等也是影响体重的重要因素．

[3] 数形结合，扫清了思维障碍，体现数学的简约美，发展直观想象素养．

[4] 对函数关系与相关关系概念的理解．

[5] 对散点图及变量间的相关关系概念的理解．

[6] 通过作散点图，理解线性相关的概念．