**回归分析的基本思想及其初步应用试题及答案-高中数学选修1-2第一章**

1．下列命题中正确的是

(　　)．

①任何两个变量都具有相关关系

②圆的周长与圆的半径具有相关关系

③某商品的需求量与该商品的价格是一种非确定性关系

④根据散点图求得的线性回归方程可能是没有意义的

⑤两个变量的线性相关关系可以通过线性回归方程，把非确定性问题转化为确定性问题进行研究

A．①③④ B．②④⑤

C．③④⑤ D．②③⑤

解析　显然①是错误的，而②中圆的周长与圆的半径的关系为：*C*＝2π*R*，是一种确定性的函数关系，故应选C.

答案　C

2．设两个变量*x*和*y*之间具有线性相关关系，它们的相关系数是*r*，*y*关于*x*的回归直线的斜率是*b*，纵轴上的截距是*a*，那么必有

(　　)．

A．*b*与*r*的符号相同 B．*a*与*r*的符号相同

C．*b*与*r*的符号相反 D．*a*与*r*的符号相反

解析　因为*b*＞0时，两变量正相关，此时*r*＞0；*b*＜0时，两变量负相关，此时*r*＜0.

答案　A

3．在判断两个变量*y*与*x*是否相关时，选择了4个不同的模型，它们的相关指数*R*2分别为：模型1的相关指数*R*2为0.98，模型2的相关指数*R*2为0.80，模型3的相关指数*R*2为0.50，模型4的相关指数*R*2为0.25.其中拟合效果最好的模型是

(　　)．

A．模型1 B．模型2

C．模型3 D．模型4

解析　相关指数*R*2能够刻画用回归模型拟合数据的效果，相关指数*R*2的值越接近于1，说明回归模型拟合数据的效果越好．

答案　A

4．若一组观测值(*x*1，*y*1)，(*x*2，*y*2)，…，(*xn*，*yn*)之间满足*yi*＝*bxi*＋*a*＋*ei*(*i*＝1,2，…，*n*)，且*ei*恒为0，则*R*2为\_\_\_\_\_\_\_\_．

解析　由*ei*恒为0，知*yi*＝*i*，即*yi*－*i*＝0，

故*R*2＝1－＝1－0＝1.

答案　1

5．已知回归直线的斜率的估计值为1.23，样本点的中心为(4,5)，则回归直线方程是\_\_\_\_\_\_\_\_．

解析　由斜率的估计值为1.23，且回归直线一定经过样本点的中心(4,5)，可得－5＝1.23(*x*－4)，

即＝1.23*x*＋0.08.

答案　＝1.23*x*＋0.08

6．某个服装店经营某种服装，在某周内纯获利*y*(元)与该周每天销售这种服装件数*x*之间的一组数据如下表：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| *y* | 66 | 69 | 73 | 81 | 89 | 90 | 91 |

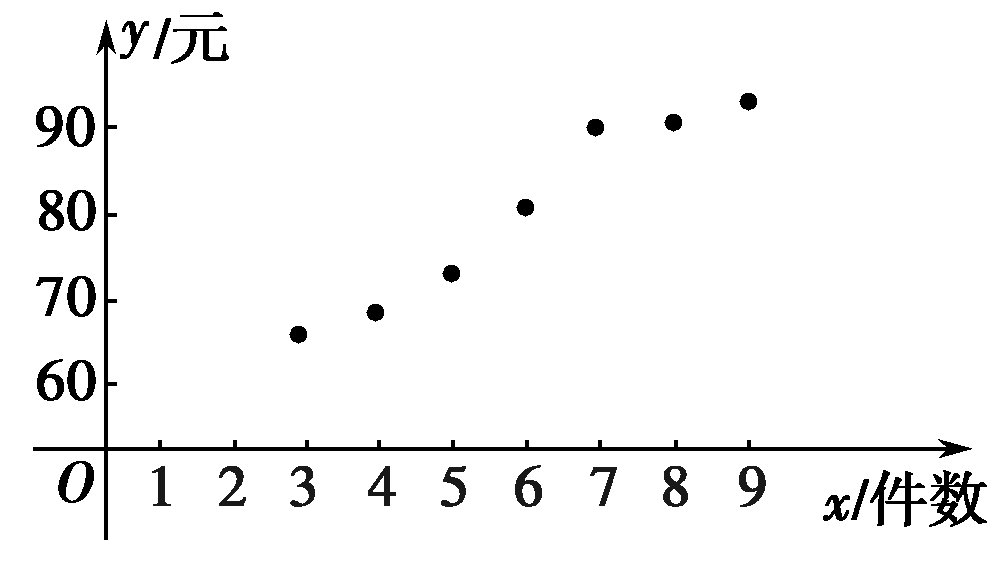
(1)求样本中心点．

(2)画出散点图．

(3)求纯获利*y*与每天销售件数*x* 之间的回归方程．

解　(1)＝6，≈79.86，中心点(6,79.86)．

(2)散点图如下：



(3)因为＝≈4.75，

＝－≈51.36，所以＝4.75*x*＋51.36.

7．为了考察两个变量*x*和*y*之间的线性相关性，甲、乙两位同学各自独立地做了100次和150次试验，并且利用线性回归方法，求得回归直线分别为*l*1和*l*2.已知两个人在试验中发现对变量*x*的观测数据的平均值都是*s*，对变量*y*的观测数据的平均值都是*t*，那么下列说法正确的是(　　)．

A．*l*1和*l*2有交点(*s*，*t*)

B．*l*1与*l*2相交，但交点不一定是(*s*，*t*)

C．*l*1与*l*2必定平行

D．*l*1与*l*2必定重合

解析　都过样本中心点(*s*，*t*)，但斜率不确定．

答案　A

8．某医学科研所对人体脂肪含量与年龄这两个变量研究得到一组随机样本数据，运用Excel软件计算得＝0.577*x*－0.448(*x*为人的年龄，*y*为人体脂肪含量)．对年龄为37岁的人来说，下面说法正确的是(　　)．

A．年龄为37岁的人体内脂肪含量都为20.90%

B．年龄为37岁的人体内脂肪含量为21.01%

C．年龄为37岁的人群中的大部分人的体内脂肪含量为20.90%

D．年龄为37岁的大部分的人体内脂肪含量为31.5%

解析　当*x*＝37时，＝0.577×37－0.448＝20.901≈20.90，由此估计：年龄为37岁的人群中的大部分人的体内脂肪含量为20.90%.

答案　C

9．今年一轮又一轮的寒潮席卷全国．某商场为了了解某品牌羽绒服的月销售量*y*(件)与月平均气温*x*(℃)之间的关系，随机统计了某4个月的月销售量与当月平均气温，数据如下表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 月平均气温*x*(℃) | 17 | 13 | 8 | 2 |
| 月销售量*y*(件) | 24 | 33 | 40 | 55 |

由表中数据算出线性回归方程＝*x*＋中的≈－2.气象部门预测下个月的平均气温约为6 ℃，据此估计，该商场下个月该品牌羽绒服的销售量的件数约为\_\_\_\_\_\_\_\_．

解析　由表格得(，)为(10,38)，又(，)在回归直线

＝*x*＋上，且≈－2，

∴38＝－2×10＋，＝58，所以＝－2*x*＋58，

当*x*＝6时，＝－2×6＋58＝46.

答案　46

10．面对竞争日益激烈的消费市场，众多商家不断扩大自己的销售市场，以降低生产成本．某白酒酿造企业市场部对该企业9月份的产品销量(千箱)与单位成本(元)的资料进行线性回归分析，结果如下：＝，＝71，＝79，

*iyi*＝1 481，＝≈－1.818 2，

＝71－(－1.818 2)×≈77.36，则销量每增加1千箱，单位成本下降\_\_\_\_\_\_\_\_元．

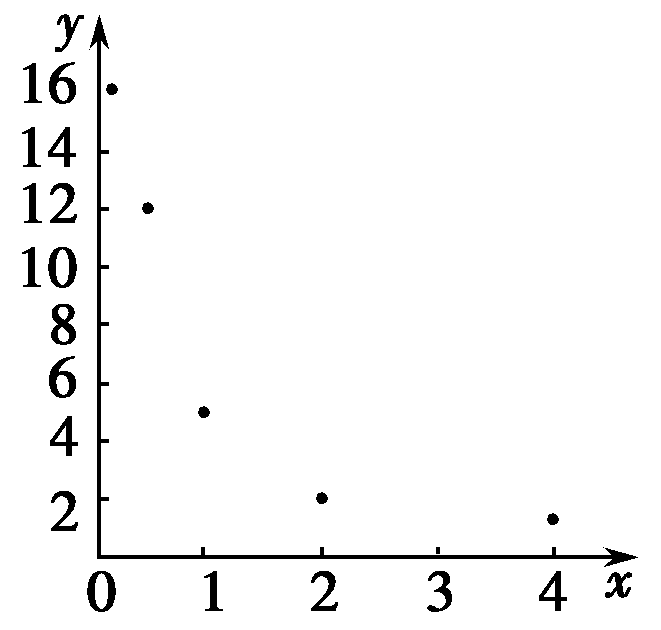
解析　由已知可得，＝－1.818 2*x*＋77.36，销量每增加1千箱，则单位成本下降1.818 2元．

答案　1.818 2

11．在一次抽样调查中测得样本的5个样本点，数值如下表：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | 0.25 | 0.5 | 1 | 2 | 4 |
| *y* | 16 | 12 | 5 | 2 | 1 |

试建立*y*与*x*之间的回归方程．

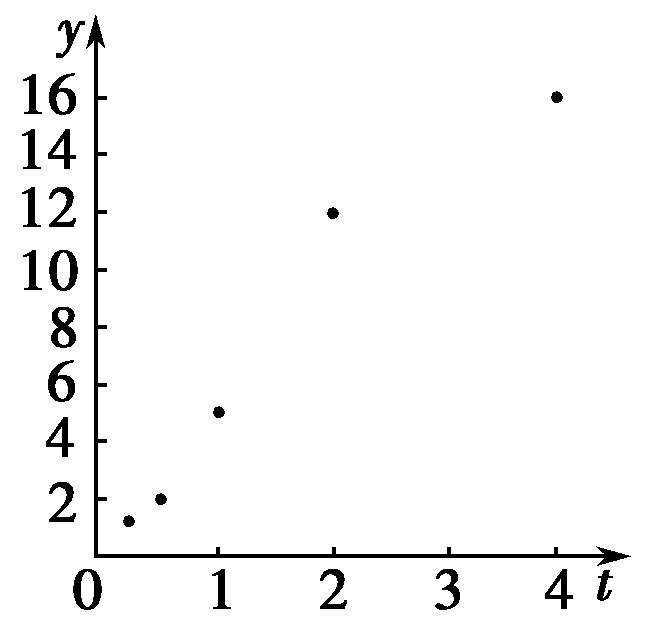
解　由数值表可作散点图如右图．根据散点图可知

*y*与*x*近似地呈反比例函数关系，设*y*＝，令*t*＝，

则*y*＝*kt*，原数据变为：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *t* | 4 | 2 | 1 | 0.5 | 0.25 |
| *y* | 16 | 12 | 5 | 2 | 1 |

由置换后的数值表作散点图如下：



由散点图可以看出*y*与*t*呈近似的线性相关关系．列表如下：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *i* | *ti* | *yi* | *tiyi* | *t* | *y* |
| 1 | 4 | 16 | 64 | 16 | 256 |
| 2 | 2 | 12 | 24 | 4 | 144 |

续表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 3 | 1 | 5 | 5 | 1 | 25 |
| 4 | 0.5 | 2 | 1 | 0.25 | 4 |
| 5 | 0.25 | 1 | 0.25 | 0.625 | 1 |
| ∑ | 7.75 | 36 | 94.25 | 21.312 5 | 430 |

所以＝1.55，＝7.2.

所以＝＝4.134 4，＝－＝0.8.

所以＝4.134 4*t*＋0.8.

所以*y*与*x*的回归方程是＝＋0.8.

12．(创新拓展)某运动员训练次数与成绩之间的数据关系如下：

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 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；

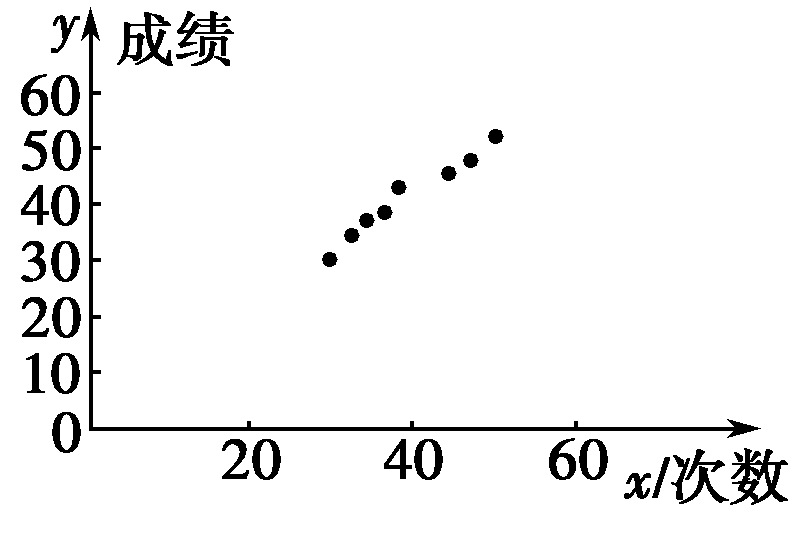
(2)求出回归方程；

(3)作出残差图；

(4)计算相关指数*R*2；

(5)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*)之间的散点图，如图所示，由散点图可知，它们之间具有线性相关关系．



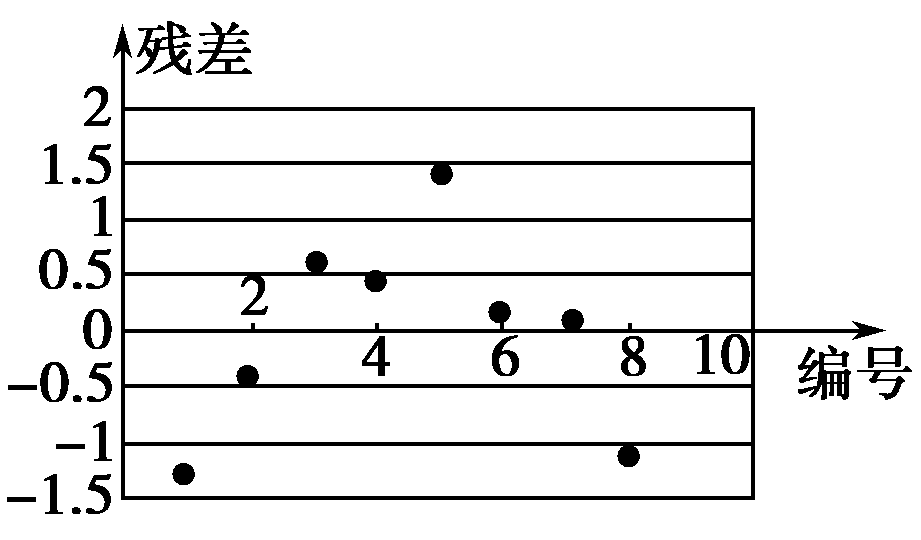
(2)＝39.25，＝40.875，＝12 656，

*iyi*＝13 180，

∴＝＝1.041 5，＝－＝－0.003 88，

∴回归方程为＝1.0415*x*－0.003 88.

(3)作残差图如图所示，



由图可知，残差点比较均匀地分布在水平带状区域中，说明选用的模型比较合适．

(4)计算得相关指数*R*2＝0.985 5，说明了该运动员的成绩的差异有98.55%是由训练次数引起的．

(5)由上述分析可知，我们可用回归方程＝1.041 5*x*－0.003 88作为该运动员成绩的预报值．

将*x*＝47和*x*＝55分别代入该方程可得*y*≈49和*y*≈57.

故预测该运动员训练47次和55次的成绩分别为49和57.