# 江苏省仪征中学2021-2022学年度第二学期高二数学学科导学案

## 9.1.2线性回归方程

研制人：陆烽琴 审核人：鲁媛媛

班级： 姓名： 学号： 授课日期：

**一、学习目标**

**1.**了解随机误差、残差、残差图的概念

**2.**会通过分析残差判断线性回归模型的拟合效果

**3.**掌握建立线性回归模型的步骤

**4.**掌握非线性回归转化为线性回归的方法，会求非线性回归方程，并作出预测

重点、难点：求线性回归方程.

二、课前自学

**1.线性回归方程**

(1)线性回归模型：

其中*a*和*b*是模型的未知参数， 称为 ，

(2) 对于一组具有线性相关关系的数据(*x*1，*y*1)，(*x*2，*y*2)，…，(*xn*，*yn*)，回归直线的斜率和截距的最小二乘估计公式分别为

回归系数＝＝

回归截距＝－， 其中(，)称为样本点的中心．

**2. 残差的概念**

对于样本点(*x*1，*y*1)，(*x*2，*y*2)，…，(*xn*，*yn*)而言，它们的随机误差为，*i*＝1，2，…，*n*，称为相应于点(*xi*，*yi*)的**残差**．

**3. 刻画回归效果的方式**

(1)残差图法

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，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．

(2)残差平方和法

残差平方和 (*yi*－*i*)2，残差平方和越小，模型拟合效果越好．

(3)利用*R*2刻画回归效果

*R*2＝1－；*R*2表示解释变量对于预报变量变化的贡献率．*R*2越接近于1，表示回归的效果越好.

三、问题探究

例1．根据下表中的数据，得到的回归方程为＝*x*＋9，则＝(　　)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | 4 | 5 | 6 | 7 | 8 |
| *y* | 5 | 4 | 3 | 2 | 1 |

A.2 B.1 C.0 D.－1

例2. 教材P152 例5

例3. 已知某种商品的价格*x*(元)与需求量*y*(件)之间的关系有如下一组数据：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | 14 | 16 | 18 | 20 | 22 |
| *y* | 12 | 10 | 7 | 5 | 3 |

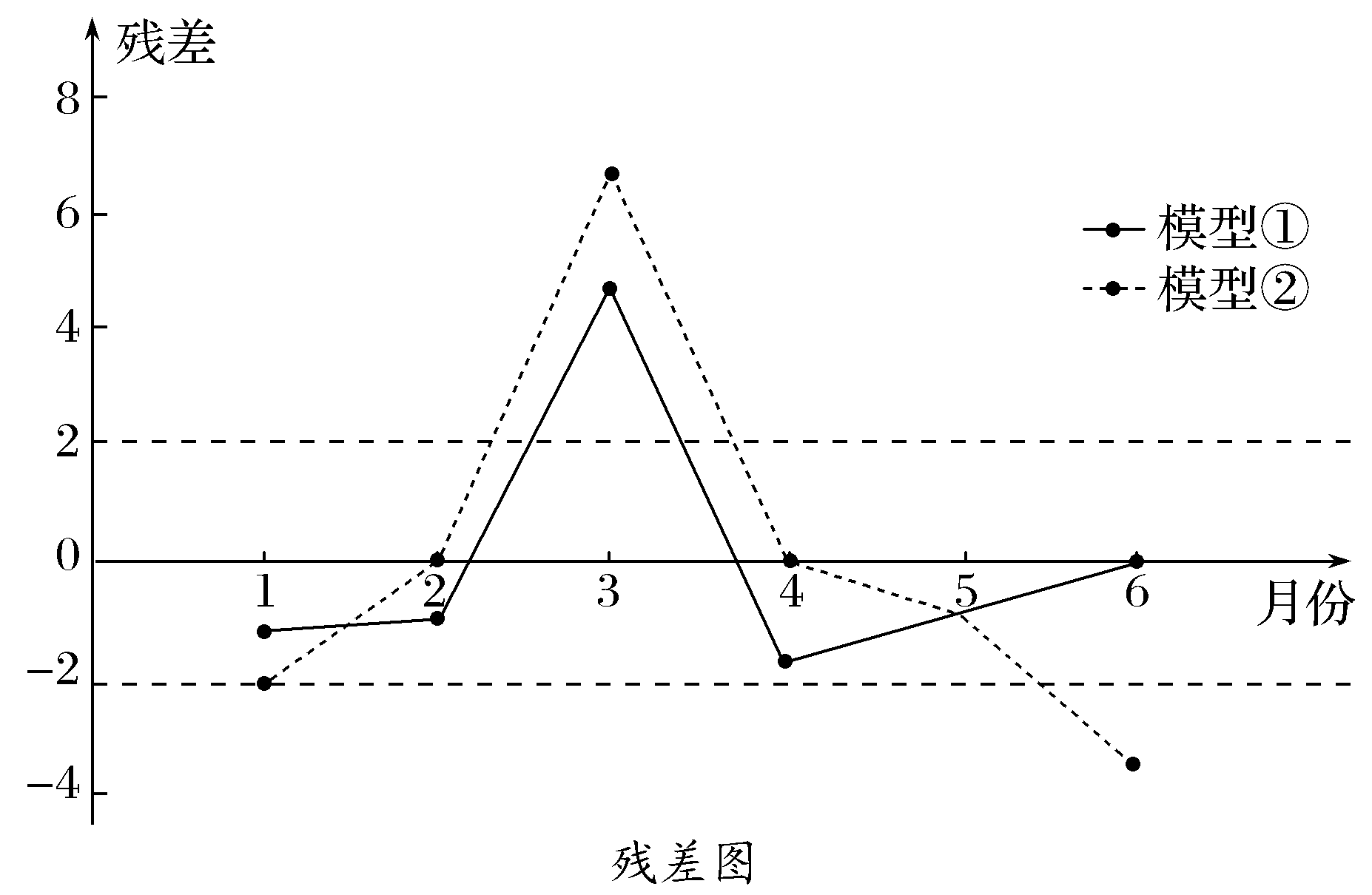
求*y*对*x*的回归直线方程，并说明回归模型拟合效果的好坏．

例4.某互联网公司为了确定下一季度的前期广告投入计划，收集了近6个月广告投入量x(单位：万元)和收益y(单位：万元)的数据如下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 月份 | 1 | 2 | 3 | 4 | 5 | 6 |
| 广告投入量/万元 | 2 | 4 | 6 | 8 | 10 | 12 |
| 收益/万元 | 14.21 | 20.31 | 31.8 | 31.18 | 37.83 | 44.67 |

他们用两种模型①＝x＋，②*y*＝*a*e*bx*分别进行拟合，得到相应的回归方程并进行残差分析，得到如图所示的残差图及一些统计量的值：

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | *xiyi* | *x* |
| 7 | 30 | 1 464.24 | 364 |



(1)根据残差图，比较模型①，②的拟合效果，应选择哪个模型？并说明理由.

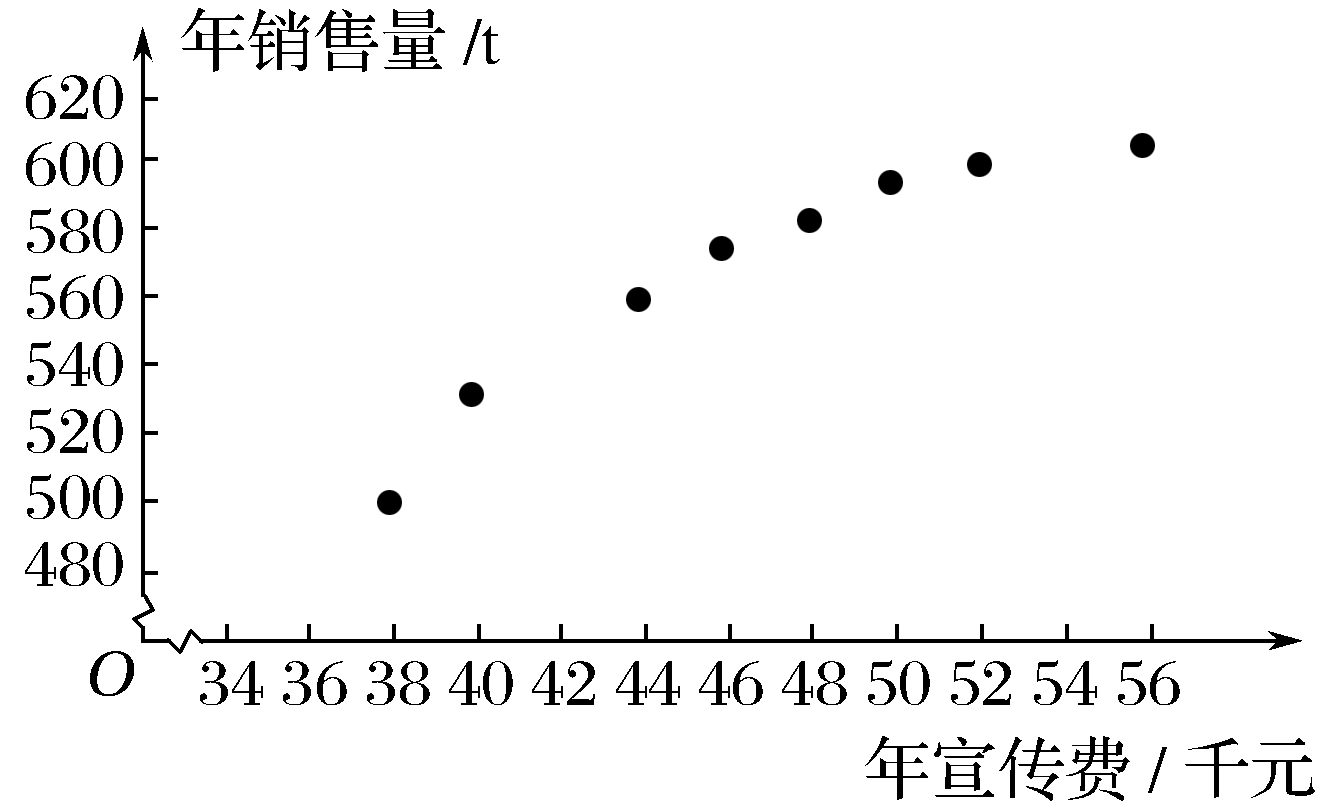
(2)残差绝对值大于2的数据被认为是异常数据，需要剔除：

(ⅰ)剔除异常数据后，求出(1)中所选模型的回归方程；

(ⅱ)广告投入量x＝18时，(1)中所选模型收益的预报值是多少？

附：对于一组数据(x1，y1)，(x2，y2)，…，(xn，yn)，其回归直线＝x＋的斜率和截距的最小二乘估计分别为：＝＝，＝－.

例5. 某公司为确定下一年度投入某种产品的宣传费，需了解年宣传费*x*(单位：千元)对年销售量*y*(单位：t)和年利润*z*(单位：千元)的影响，对近8年的年宣传费*xi*和年销售量*yi*(*i*＝1，2，…，8)数据作了初步处理，得到下面的散点图及一些统计量的值.



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | (*xi*－)2 | (*wi*－)2 | (*xi*－)·(*yi*－) | (*wi*－)·(*yi*－) |
| 46.6 | 563 | 6.8 | 289.8 | 1.6 | 1 469 | 108.8 |

表中*wi*＝，＝*wi*.

(1)根据散点图判断，*y*＝*a*＋*bx*与*y*＝*c*＋*d*哪一个适宜作为年销售量*y*关于年宣传费*x*的回归方程类型(给出判断即可，不必说明理由)?

(2)根据(1)的判断结果及表中数据，建立*y*关于*x*的回归方程；

(3)已知这种产品的年利润*z*与*x*，*y*的关系为*z*＝0.2*y*－*x*.根据(2)的结果回答下列问题：

①年宣传费*x*＝49时，年销售量及年利润的预报值是多少？

②年宣传费*x*为何值时，年利润的预报值最大？

四、反馈：课本P154练习1.2

五、小结：