课后习题

题9.1 合金的强度与其中的碳含量有比较密切的关系，今从生产中收集了一批数据如表1。

表1

|  |  |
| --- | --- |
|  | 0.10 0.11 0.12 0.13 0.14 0.15 0.16 0.17 0.18 |
|  | 42.3 41.5 45.2 45.5 44.9 47.6 48.9 55.1 50.2 |

试先拟合一个函数 ，再用回归分析对它进行检验。

题9.2 体重约70kg的某人在短时间内喝下2瓶啤酒后，隔一定时间测量他的血液中酒精含量（毫克/百毫升），得到数据如下：

表2 血液中酒精含量

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 时间(小时) | 0.25 | 0.5 | 0.75 | 1 | 1.5 | 2 | 2.5 | 3 | 3.5 | 4 | 4.5 | 5 |
| 酒精含量 | 30 | 68 | 75 | 82 | 82 | 77 | 68 | 68 | 58 | 51 | 50 | 41 |
| 时间(小时) | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |  |
| 酒精含量 | 38 | 35 | 28 | 25 | 18 | 15 | 12 | 10 | 7 | 7 | 4 |  |

（1）依据数据作出人体血液中酒精含量与酒后时间的散点图，从图形上看能否选择多项式函数进行回归分析？为什么？

（2）建立人体血液中酒精含量与酒后时间的函数关系；

（3）对照《车辆驾驶人员血液、呼气酒精含量阈值与检验》国家标准，车辆驾驶人员血液中的酒精含量大于或等于20毫克／百毫升，小于80毫克／百毫升为饮酒驾车；血液中的酒精含量大于或等于80毫克／百毫升为醉酒驾车，此人在短时间内喝下1瓶啤酒后，隔多长时间开车是安全的？

题9.3根据表3中给出的1971~2010人口统计数据，选择恰当的函数进行回归分析，并利用所得到的公式估计2011~2019年我国人口数，最后与新近统计数据进行比较和分析。

表3 1971~2010人口统计数据

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 年份 | 人口（亿） | 年份 | 人口（亿） | 年份 | 人口（亿） | 年份 | 人口（亿） |
| 1971 | 8.5229 | 1981 | 10.0072 | 1991 | 11.5823 | 2001 | 12.7627 |
| 1972 | 8.7177 | 1982 | 10.1654 | 1992 | 11.7171 | 2002 | 12.8453 |
| 1973 | 8.9221 | 1983 | 10.3008 | 1993 | 11.8517 | 2003 | 12.9227 |
| 1974 | 9.0859 | 1984 | 10.4357 | 1994 | 11.985 | 2004 | 12.9988 |
| 1975 | 9.2420 | 1985 | 10.5851 | 1995 | 12.1121 | 2005 | 13.0756 |
| 1976 | 9.3717 | 1986 | 10.7507 | 1996 | 12.2389 | 2006 | 13.1448 |
| 1977 | 9.4974 | 1987 | 10.9300 | 1997 | 12.3626 | 2007 | 13.2129 |
| 1978 | 9.6259 | 1988 | 11.1026 | 1998 | 12.4761 | 2008 | 13.2802 |
| 1979 | 9.7542 | 1989 | 11.2704 | 1999 | 12.5786 | 2009 | 13.3474 |
| 1980 | 9.8705 | 1990 | 11.4333 | 2000 | 12.6743 | 2010 | 13.41 |

题9.4 海水温度随着深度的变化而变化，海面温度较高，随着深度的增加温度越来越低，这样也就影响了海水的对流和混合，使得深层海水中的氧气越来越少，这是潜水员必须考虑的问题，同时根据这一规律也可对海水鱼层作一个划分.现在通过实验测得一组海水深度与温度的数据如下：

表4 海水温度随着深度的变化数值

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| t | 23.5 | 22.9 | 20.1 | 19.1 | 15.4 | 11.5 | 9.5 | 8.2 |
| h | 0 | 1.5 | 2.5 | 4.6 | 8.2 | 12.5 | 16.5 | 26.5 |

要求：（1）找出温度与海水深度之间的一个近似函数关系；

（2）找出温度变化最快的深度位置（实际上该位置就是潜水员在潜水时，随着下潜深度的不同，需要更换吸入气体种类的位置，也是不同种类鱼层的分界位置）。

题9.5 一矿脉有13个相邻样本点，人为地设定一原点，现测得各样本点对原点的距离 ，与该样本点处某种金属含的一组数据如表5，画出散点图观测二者的关系，试建立合适的回归模型。

表5

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2 | 3 | 4 | 5 | 7 | 8 | 10 |
|  | 106.42 | 109.20 | 109.58 | 109.50 | 110.00 | 109.93 | 110.49 |
|  | 11 | 14 | 15 | 16 | 18 | 19 |  |
|  | 110.59 | 110.60 | 110.90 | 110.76 | 111.00 | 111.20 |  |

题9.6 某商场若干月的库存占用资金（）、广告投入的费用（）、员工薪酬（）以及销售额（）等数据如下，根据这些数据建立销售额与其它三个变量之间的线性回归模型。

= [75.0 77.6 80.7 76 79.5 81.8 98.3 67.7 74 151 90.8 102.3 115.6 125 137.8 175.6 155.1]；

= [30.1 31.3 33.9 29.6 32.5 27.9 24.8 23.6 33.9 27.7 45.5 42.6 40 45.8 51.7 67.2 65.1]；

= [21.0 21.4 22.9 21.4 21.5 21.7 21.5 21 22.4 24.7 23.2 24.3 23.1 29.1 24.6 27.5 26.4 ]；

= [1090.3 1133 1242.1 1003.2 1283.2 1012.2 1098.8 826.3 1003.3 1554.6 1199 1483.1 1407.1 1551.3 1601.2 2311.7 2126.6]；

题9.7开放式作业1：工业蒸汽量预测

样本数据和预测要求见

https://tianchi.aliyun.com/competition/entrance/231693/introduction?spm=5176.12281949.1003.11.493e7a74PW6sOX

题9.8开放式作业3：北京PM2.5浓度预测

样本数据和预测要求见

https://www.pkbigdata.com/common/cmpt/%E5%8C%97%E4%BA%ACPM2.5%E6%B5%93%E5%BA%A6%E5%9B%9E%E5%BD%92%E5%88%86%E6%9E%90%E8%AE%AD%E7%BB%83%E8%B5%9B\_%E7%AB%9E%E8%B5%9B%E4%BF%A1%E6%81%AF.html

或者数据源https://archive.ics.uci.edu/ml/datasets/Beijing+PM2.5+Data

题9.9开放式作业5：地震预测

样本数据和预测要求见https://www.kaggle.com/c/LANL-Earthquake-Prediction/overview

当前数据科学挑战赛/竞赛（Data Science Challenge / Competition）主要有：

Kaggle/ DrivenData/ Codalab/ Challenge Data/ crowdAI/ EvalAI/ Numerai/ SIGNATE/ Unearthed/ GoogleAI/ 和鲸/ Analytics Vidhya/ techGig/ 阿里云天池/ HeroX/ DataCastle/ AI Challenger/ Biendata