**编号**



**学士学位论文**

**一元线性模型三种检验的等价性分析**

学院名称： 数学与统计学院

专业班级： 应用统计学

学生姓名： 吴明强

学 号： 20140106006

指导教师： 孙瑞德

完成日期： 2018 年 4 月 16 日

# 摘 要

两个变量建立一元线性回归模型，要对两个变量的线性相关性进行检验，检验的方法有三种，F-检验、t-检验、r-检验，论文给出了三种检验方法的具体做法，并证明了三种检验方法的等价性．

关键词**：**线性相关；相关性检验；等价性

# **Abstract**

**The Equivalence Analysis Of Three Kinds Of Test Of Unary Linear Model**

The establishment of one element regression model for research variables，to test the linear correlation of two variables，there are three kinds of methods for testing，F-test、t-test、r-test，the paper gives the concrete methods of three methods，the equivalence of the three methods is proved．

**Key words**：Linear；Correlation test；Equivalence

目 录

[摘 要 I](#_Toc511162969)

[Abstract II](#_Toc511162970)

[引言 1](#_Toc511162971)

[1一元线性模型及其假设 1](#_Toc511162972)

[1.1一元线性模型 1](#_Toc511162973)

[1.2一元线性模型的假设 1](#_Toc511162974)

[1.3参数的普通最小二乘估计 2](#_Toc511162975)

[1.4参数的普通最小二乘估计的性质 3](#_Toc511162976)

[2.三种检验方法 4](#_Toc511162977)

[2.1三种检验方法的预备 4](#_Toc511162978)

[2.1.1参数估计值的分布 4](#_Toc511162979)

[2.1.2三个离差及其关系 4](#_Toc511162980)

[2.1.3检验的预备定理 4](#_Toc511162981)

[2.2 f-检验法 7](#_Toc511162982)

[2.3 t-检验法 8](#_Toc511162983)

[2.4 r-检验法 9](#_Toc511162984)

[3.三种检验方法的等价性分析 9](#_Toc511162985)

[3.1 f-检验与t-检验的等价性 9](#_Toc511162986)

[3.2 r-检验与f-检验的等价性 9](#_Toc511162987)

[3.3 实例分析 10](#_Toc511162988)

[参考文献 13](#_Toc511162989)

[致 谢 14](#_Toc511162990)

# 引言

在实际问题的研究中，经常需要研究某一变量与影响它的某一最主要因素的关系，如影响粮食产量的因素非常多，但在众多因素中，施肥量是一个重要的因素，往往需要研究施肥量这一因素与粮食产量之间的关系；在消费研究中，影响消费的因素很多，但可以只研究国民收入与消费额之间的关系，因为国民收入是影响消费的最主要因素；保险公司在研究火灾损失的规律时，把火灾发生地与最近的消防站的距离作为一个最主要的因素，研究火灾损失与火灾发生地距离最近消防站的距离之间的关系[1]．如果两个变量大致呈线性关系，可以建立一元线性回归模型．两个变量是否大致呈线性关系，可以看散点图，散点图的优点是直观明了，缺点是没有进行定量分析，缺乏数据支撑．本文用数据分析方法检验两个变量是否大致呈线性关系，大致呈线性关系时，线性相关的程度也可以从有关数据体现出来．本文给出了三种检验方法，F检验、t检验、r检验，并证明了三种检验的等价性．

# 1一元线性模型及其参数的估计

## 1.1一元线性模型

若自变量，因变量之间大体上呈直线关系，它们的关系可用下面的式子来描述：

，

这就是一元线性回归模型，其中、为未知参数，误差项为随机变量，并设

[2]

假设样本观测值为，，…，，则线性回归模型可以表示为：

．

## 1.2一元线性模型的假设

（1）变量是非随机变量，观测值是常数．

（2）高斯-马尔科夫（Gauss-Markov）条件[3]：

在（1）、（2）的基础上，如果再假设误差项服从正态分布，模型就是正态模型，这时，，相互独立．

## 1.3参数的普通最小二乘估计

由经验回归方程，如果参数的值为，，处的离差为：

，

n个观测值，离差的平方和为：

、的普通最小二乘估计、，就是使达到最小值时的、．

由为、的多项式，对、求偏导得：

整理得：

称上式为正规方程组．其中、分别为观测值，，…，与，，…，的平均值．

解正规方程组得、的普通最小二乘估计为[2]：

若记：，

，

则（1.3.2）式可以表示为：

．

## 1.4参数的普通最小二乘估计的性质

（1）线性性：*、*是观测值的线性函数．

（2）无偏性：*E()=，E(=*．

（3）最优性：在、的所有线性无偏估计中，*、*的方差最小[5]．

证明：

的最小二乘估计记为：

*=*，

则．

假设是的任一线性无偏估计，记：

，

，

因为是的线性无偏估计，故有，

从而： ．

对求方差得：

由于：

，

，

所以：，

从而：，

故：，当且仅当时，取最小值，所以在的所有线性无偏估计中的方差最小，即是的最优线性无偏估计．

同理可证是的最优线性无偏估计．

# 2.三种检验方法

## 2.1三种检验方法的预备

### 2.1.1参数估计值的分布

对正态模型[6]：

### 2.1.2三个离差及其关系

总离差：，总离差平方和记为．

回归离差：，回归离差平方和记为．

残差离差：，残差平方和记为．

可以证明[7]：．

### 2.1.3检验的预备定理

定理一[8]：对正态模型，时，．

证明：

由知，当时，有，

从而，

故，即当时，；

由得，

所以当时，，证毕．

定理二：正态假定下，，且、、相互独立．

证明：

考虑如下形式的正交矩阵：

正交，应当满足：

由的前n-2行与的第n列正交得到：

，，…，，共有n-2个式子，

由的第一行与的第列正交得：，

即，由此可得．

由的前行与的第列正交得：

，，…，，共个式子；

由的前行与的第列内积为1得：

，，…，，共个式子；

由的前行与的前列中不同列正交得：

，其中1，共有个式子；

以上共有个式子．

由于待定参数有个，约束等式个数有个，所以只要，未知参数就不少于约束等式个数，故满足要求的正交矩阵一定存在．

记：，，令，

即：，

*……*

，

，

由于、、均服从正态分布，而可由、、线性表出，所以也服从正态分布，向量服从多元正态分布，其期望和方差分别为：

，

……

，

，

*)* ，

，

由以上得，，，…，相互独立，而且：

*，……， ~* ，

……

*~* ，

*~*

由于：

可得：

因为只与有关，只与有关，只与有关，而、、相互独立，所以、、相互独立．又因为，所以、、也相互独立．

## 2.2

在回归模型中，当与之间具有较强的线性关系时，用经验回归方程应能更

准确的表示的值，此时用经验回归方程预测的值，应与观测值相差很小，而与的差的平方和可以用 表示，考虑到：

，

即，可以看出，当线性关系强时偏小，此时，所以可以用与的比值来度量线性关系的强弱，即越大线性关系越强，反之，越小线性关系越弱．

再由线性回归模型，可以看出检验是否具有线性关系等同于检验是否为0，故可以建立假设：

，

若为真，没有线性关系，此时应偏小．

若为假，*具*有线性关系，此时应偏大．

从定理一、定理二知，在正态假定下相互独立，且

，时，．

从而当成立时有：

，

当成立时，应偏小，此时是小概率事件，故对显著性水平有：

时，拒绝，认为，线性关系显著．

时，接受，认为，线性关系不显著．

时，再取样本，重新检验．

## 2.3

正态假定下，由 ，且、相互独立，可知：

，

故对假设：

，

如果为真，，此时

可以看出：

当，应偏小，应偏小，所以偏小时应接受，偏大时应拒绝．

对显著性水平：

时，拒绝，认为，线性关系显著，

时，接受，认为，线性关系不显著，

时，再取样本，重新检验．

## 2.4

一元线性回归模型，样本的相关系数为：

再由得：

对假设：，

当为真时，应较小，此时较小，若较大应拒绝，反之，接受．

将样本与表中的数值相比较，当时，接受，认为的线性关系不显著，当时，拒绝，认为的线性关系显著，当时，再取样本，重新检验．

# 3.三种检验方法的等价性分析

## 3.1

在中，统计量为：

，界点为，

在中，统计量为：

，界点为，

由于、 ，所以和等价[9]．

## 3.2

在中，统计量为：

可以证明统计量是的增函数，在中，越大越认为线性关系越强，而中也是越大线性关系越强，中界点的值是利用，由界点的值解出界点的值，故与r-检验等价．又由和等价，所以也与等价．

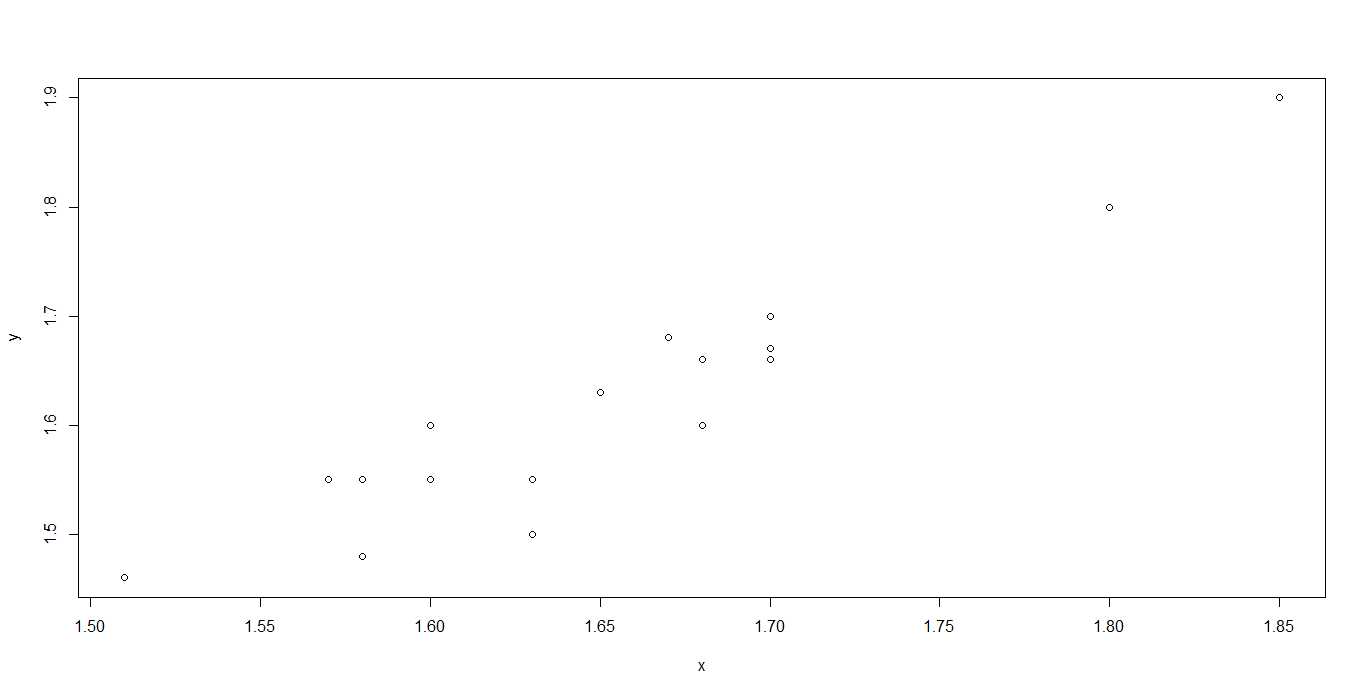
## 3.3 实例分析

我校某班学生身高和小臂长测量值见下表，分析身高和小臂长之间的线性关系．

表1 身高与臂展长数据

|  |  |  |
| --- | --- | --- |
| 编号 | 身高(米)(自变量) | 臂展长（米）（因变量y） |
|  |  |  |
| 1 | 1.60 | 1.55 |
| 2 | 1.58 | 1.55 |
| 3 | 1.57 | 1.55 |
| 4 | 1.65 | 1.63 |
| 5 | 1.68 | 1.60 |
| 6 | 1.68 | 1.66 |
| 7 | 1.51 | 1.46 |
| 8 | 1.70 | 1.67 |
| 9 | 1.60 | 1.60 |
| 10 | 1.63 | 1.50 |
| 11 | 1.70 | 1.70 |
| 12 | 1.80 | 1.80 |
| 13 | 1.67 | 1.68 |
| 14 | 1.70 | 1.66 |
| 15 | 1.63 | 1.55 |
| 16 | 1.63 | 1.55 |
| 17 | 1.85 | 1.90 |
| 18 | 1.58 | 1.48 |
| 19 | 1.65 | 1.63 |
| 20 | 1.60 | 1.55 |

做表中数据的散点图如下[10]：



从散点图可以看到大体上呈线性关系．

接下来对进行一元回归建模，其中：

，，，，

SSR= 0.1909254，SSE= 0.02472959

由以上数据得：

,

样本回归方程为：

．

(1)：

查表知，在显著水平时，，而样本，故在显著水平时，拒绝，认为线性关系显著．

(2)：

查表知，在显著水平数据时，，而，故在显著水平时，拒绝，认为线性关系显著．

(3)：

查表知，在显著水平数据时，，而，故在显著水平时，拒绝，认为线性关系显著．

(4)由（1）（2）（3）看到三种检验分别得到了相同的检验结果，可以证明三种检验具有等价性，即在对一元线性回归模型检验时，不必固守成规，可以选用这三种检验方法的任意一种，得到相同的结果．

（5）另取九组数据数据用模型预测结果见下表：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 编号 | 身高(米)(自变量) | 实际臂展长（米）（因变量y） | 预测臂展长（米）（因变量y） | 绝对误差 | 相对误差 |
| 1 | 1.500082 | 1.50 | 1.4211 | 0.0789 | 0.0526 |
| 2 | 1.700066 | 1.70 | 1.6769 | 0.0231 | 0.0136 |
| 3 | 1.550039 | 1.55 | 1.4850 | 0.0650 | 0.0419 |
| 4 | 1.660037 | 1.66 | 1.6257 | 0.0343 | 0.0207 |
| 5 | 1.570053 | 1.57 | 1.5106 | 0.0594 | 0.0378 |
| 6 | 1.650030 | 1.65 | 1.6129 | 0.0371 | 0.0225 |
| 7 | 1.599995 | 1.60 | 1.5489 | 0.0511 | 0.0319 |
| 8 | 1.670044 | 1.67 | 1.6385 | 0.0315 | 0.0189 |
| 9 | 1.550039 | 1.55 | 1.4850 | 0.0650 | 0.0419 |

由上表中预测值和真实值的绝对误差和相对误差可以看出，建立的模型能较好的反应身高和臂展长之间的关系．

# 参考文献

[1] 何晓群.实用回归分析[M].北京:高等教育出版社,2016:35.

[2] 徐国祥.计预测和决策-五版[M].上海:上海财经大学出版社,2016.1:29～31.

[3] 王黎明,陈颖,杨楠.应用回归分析[M].上海:复旦大学出版社,2008.6:4～6

[4] 何书元.数理统计[M].北京:高等教育出版社,2012.1:235～237.

[5] 宫春子.统计学原理[M].北京:机械工业出版社,2014.3:167～168.

[6] 周纪芗.实用回归分析方法[M].上海:上海科学技术出版社,1990:32.

[7] 周纪芗.回归分析[M].上海:华东师范大学出版社,1993:8.

[8] 茆诗松.概率论与数理统计(第二版)[M].中国统计出版社.200.7:362.

[9] 王天营.一元线性回归分析中三种检验的等价性研究[J].

[10]张良均.R语言数据分析与挖掘实战[M].北京:机械工业出版社,2015.9:37～40

# 

# 致 谢

在这里我要感谢学院给了我这个宝贵的机会，让我学有所用，感谢我的论文指导老师孙瑞德教授，在论文的写作过程中给了我极大的帮助和支持，并在忙碌的教学工作中，挤出宝贵的时间审阅和修改我的论文；感谢我的父母培育我长大成人，供我完成学业，也感谢学校的老师和领导在生活和学习中给与我的教诲和帮助．

毕业在即，在今后的生活中，我会继续不懈的努力，争取取得更大的进步．