

# 在大学物理实验教学中应用R软件进行数据统计分析

## ——以“迈克尔逊干涉仪”实验为例

武红 李峰

南京邮电大学 理学院

**摘要:** 大学物理实验经常需要利用统计技术进行数据分析和解释。本文以迈克尔逊干涉实验为案例,介绍了一种创新的R软件应用方法,以有效解决该实验中遇到的大量实验数据处理困难的问题。该方案的统计指标可作为评估实验结果准确性和精确性的量化指标,从而为改进大学物理实验教学方法提供客观的评价框架。

**关键词:** 迈克尔逊干涉仪; R软件; 相关系数; 线性回归

**【DOI】** 10.12252/j.issn.2096-6288.2024.01.087

### 引言

作图法是大学物理实验中常见的数据处理方法之一。然而,对于数据量较大且精度要求较高的实验,作图法难以直观准确地反映变量之间的统计学关系,亟须更加专业的统计学方案来分析实验数据。R语言软件是一款专业的统计学软件,可以更好的对实验数据进行统计分析。R语言可将数据拟合到各种回归模型中,从而确定最适合、最能代表实验现象的模型。此外,R语言还能计算斜率、截距和相关系数等统计参数,有助于深入理解基本物理原理。这些统计指标可作为评估实验结果准确性和精确性的量化指标<sup>[1-4]</sup>。

迈克尔逊干涉仪是一种用于测量可见光波长的高精度仪器<sup>[5-7]</sup>。在利用迈克尔逊干涉仪测量可见光波长的实验中,学生需要完成以下几个操作步骤:1、观察屏的调节;2、动反射镜位置的初始调节;3、最亮光点重合的调节;4、干涉同心圆环的调节;5、消除回程差;6、读数。由于实验流程较长,对学生的实验操作技巧提出了很高的要求,这使得每一个操作步骤都有可能带来人为误差或系统误差。因此,从统计学意义上分析实验结果的可行度就很有必要。通过利用统计R软件,我们可以对收集到的迈克尔逊干涉仪实验数据全面的统计分析。首先对数据进行预处理,确保数据集的完整性和质量。随后,采用线性回归模型建立干涉条纹变化量和反射镜移动距离两者之间的关系,并评估测量的精确度。通过编程,我们能够一键式地分析全班学生的实验结果,从而高效地处理大量数据,并对实验误差进行准确的分类。这一创新方法不仅提升了实验数据分析的实时性,也为后续指导学生实验和制定客观的评分标准提供了有力支持。这是利用电子化信息手段提升实验教

学效果的一种新颖且有效的尝试。

### 一、实验原理

当两列光波重叠时,若产生稳定的、明暗相间的条纹,则称为干涉现象。当两列光波的光程差为波长整数倍时,产生亮纹,当光程差为半波长奇数倍时,产生暗纹。实验中,利用分光镜将入射光束分成两路。两束光沿着各自的路径传播,在光屏处重新相遇,产生的干涉图案取决于两束光的光程差。若移动干涉仪的其中一个反射镜,即改变两束光的光程差,干涉条纹便会发生移动,并且干涉条纹级数的变化与反射镜移动的距离满足下面公式:

$$\Delta d = \Delta k \times \frac{\lambda}{2} \quad (1)$$

其中, $\Delta d$ 是反射镜的移动距离, $\Delta k$ 是干涉明或暗环级数的变化, $\lambda$ 是入射光的波长。因此,通过测量 $\Delta d$ 与 $\Delta k$ ,便可测量得到入射光波长。

为了对数据进行线性回归处理,干涉条纹的级数与反射镜移动的距离呈线性关系,公式(1)可以表示成线性方程的形式 $Y = \beta_0 + \beta_1 X$ 。由于实验误差 $\varepsilon$ 无法消除,因此第*i*次实验的测量变量 $Y_i$ 由下面的一元线性回归方程来表示 $Y_i = \beta_0 + \beta_1 X_i + \varepsilon$ 。其中, $\beta_0$ , $\beta_1$ 是回归系数, $\varepsilon$ 是在测量过程中产生的误差。因此,可得可见光波长 $\lambda$ 的估计值为 $\lambda = 2 \times \frac{\beta_1}{N}$ 。由最小二乘法可得,参数 $\beta_1$ 的估计量为:

$$\beta_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (2)$$

此外，已知测量值的相关系数 $r$ 为

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \cdot \sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \quad (3)$$

因此，测量的不确定度定义为

$$u_r(R) = \frac{u(R)}{R} = \sqrt{\frac{1}{r^2} - 1} \quad (4)$$

最终，我们将对实验数据的显著性 $p$ 值，线性回归系数 $\beta_0$ 、 $\beta_1$ ，样本决定系数 $R^2$ ，相关系数 $r$ ，和不确定

度 $u_r$ 等方面对实验数据进行综合分析，为学生的实验结果提供客观的统计学评分标准。

## 二、实例分析

一个小班共10位同学完成了利用迈克尔逊干涉原理测量可见光波长实验。为了方便讨论，我们将对前两位同学的实验数据进行比较，见表1。首先，我们对实验数据的关系的显著性 $p$ 值进行检验。一般认为 $p < 0.05$ 表示两个变量之间的关系具有显著性，相反 $p > 0.05$ 则认为两个变量之间的关系不显著。检验结果表明，第一组实验数据的 $p = 2.51 \times 10^{-11}$ ，第二组实验数据的 $p = 5.09 \times 10^{-6}$ 。P值分析结果显示两组数据都具有较好的显著性，其中第一组的实验数据的显著性更高。

表1 迈克尔逊干涉测量数据

冒出或消失条纹数		0	50	100	150	200	250
同学A	动镜 $M_1$ 的位置/mm	$d_0 = 50.07068$	$d_1 = 50.08669$	$d_2 = 50.10280$	$d_3 = 50.11896$	$d_4 = 50.13513$	$d_5 = 50.15144$
	150个条纹，动镜 $M_1$ 移动的距离/mm	$\Delta d_1 = d_3 - d_0 = 0.04828$		$\Delta d_2 = d_4 - d_1 = 0.04844$		$\Delta d_3 = d_5 - d_2 = 0.04864$	
	$\Delta d$ /mm	0.04845					
同学B	动镜 $M_1$ 的位置/mm	$d_0 = 49.89255$	$d_1 = 49.90963$	$d_2 = 49.92986$	$d_3 = 49.94284$	$d_4 = 49.95932$	$d_5 = 49.97090$
	150个条纹，动镜 $M_1$ 移动的距离/mm	$\Delta d_1 = d_3 - d_0 = 0.05029$		$\Delta d_2 = d_4 - d_1 = 0.04969$		$\Delta d_3 = d_5 - d_2 = 0.04104$	
	$\Delta d$ /mm	0.04701					

其次，基于R语言对两组数据分别进行线性回归分析。图1(a, b)分别给出了两组实验数据的散点图和估计方程的曲线。回归分析结果显示，两组数据的估计方程分别为 $y_i = 50.050 + 0.0165x_i$ 和 $y_i = 49.880 + 0.0173x_i$ 。分别利用公式(1)计算，可得两组实验测量的可见光波长分别为646.0 nm和626.8 nm。后者比前者小了约3%。然而，考虑到镜片个体的差异性，我们无法从散点图和回归方程中判断哪位同学的实验结果更优。

我们借助统计分析R软件对数据的决定系数 $R^2$ 值进行拟合程度分析。决定系数 $R^2$ 值的大小决定了自变量和因变量两者间相关的密切程度。决定系数越大，则自变量对因变量的解释程度越高，观察点在回归直线附近越密集。结果表明两组数据的样本决定系数 $R^2$ 值分别为1.0000和0.9954。因此从统计角度来看，第一组实验数据的拟合效果更好。

接着，我们对两组实验进行相关系数 $r$ 分析。一般认为 $r < 0.3$ 为弱相关， $0.3 \leq r \leq 0.7$ 为中相关， $r > 0.7$ 为强相关。利用公式5分析数据，结果表明第一组实

验的相关系数 $r_1 = 0.9999$ ，测量的不确定度 $u_{r1}(R) = 0.0014$ ，第二组实验的相关系数 $r_2 = 0.9981$ ，测量的不确定度 $u_{r2}(R) = 0.0304$ 。从测量的不确定度来看第二组实验的不确定度更大些，这与从统计上的样本决定系数的角度得到的结论是一致的。

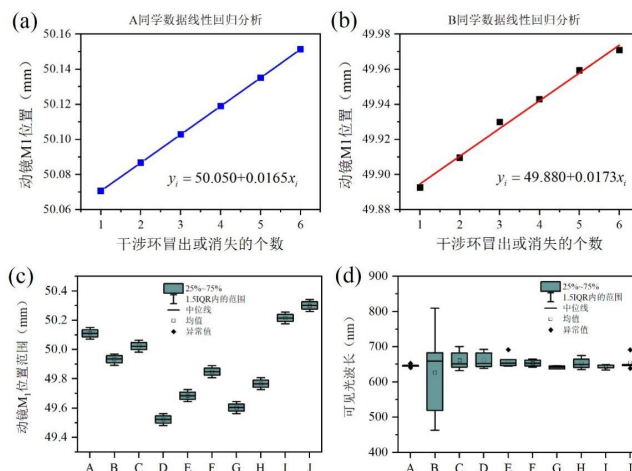


图1 (a, b) 采用线性回归分析第A和B号学生的实验数据。(c) 全班所有同学的原始实验数据分布图。(d) 全班同学测量可见光波长结果分布图。

采用同样的统计方法，我们将小班中所有同学的实验数据进行统计学分析。如图1(c)所示，整体来看10位同学的数据分布较为合理，其动镜的位置范围在49.4~50.4 mm范围内浮动。其中，D同学的动镜位置距离分束镜最近，J同学的动镜位置距离分束镜最远。最后，对本次实验所有同学测量的可见光波长结果进行统计学分析，如图1(d)所示。统计发现，同学们计算得到的可见光波长分布在650nm上下。其中，B同学计算的结果经过平均后同样处在650nm附近，但是仔细分析数据会发现其原始数值的偏差明显。说明B同学未能很好掌握本次实验，需教师帮助该生找到数据偏大的原因。导致这一差异的原因可能来自观察干涉环冒出或消失的时候计数错误，同时出现多数和少数的情况。因此，虽然平均后的可见光波长计算结果与其他同学基本一致，但不确定度较大，即实验数据的可行度较低，教师需要重点关注B同学，帮助其掌握操作迈克尔逊仪器的实验技能。

```
###本实验的R语言代码的主体部分：
# 迈克耳孙干涉仪，测量可将光波长
setwd ("C:/Users/LI/Desktop")
library (readxl)
MichelsonCircle = read_excel ("Michelson.
xlsx") ###Michelson为实验数据
####第1组数据 dist1
XX=MichelsonCircle$circle #数据的x轴
YY=MichelsonCircle$dist1 #数据的y轴
rr=cor (XX, YY) #相关系数r
scatter.smooth (YY~XX, main= "A同学数据线性
回归分析", xlab= "干涉条纹环增加的次m", ylab=
"D^2")
fit1 = lm (YY~XX, data=MichelsonCircle) #
线性拟合
summary (fit1) #线性
拟合的结果，决定系数R^2
Ur1= ((1/rr^2-1) / (length (YY) -2)) ^0.5
```

#不确定度ur

## 结语

本文以迈克尔逊干涉为例，详细阐述了R软件在实验中的运用。R软件在处理全班同学的实验数据方面具有显著优势，能够快速获得大量数据的统计学特征，为实验结果的优劣提供客观评判标准。将此方法融入课程，有助于培养学生的批判性思维、数据驱动决策以及对实验技术的深入理解，进而提升大学物理实验的教学质量。

## 参考文献

- [1] 宋豪豪, 贾俊梅, 范洪雁, 洪志敏. R软件在高校成绩分析中的应用[J]. 无线互联科技 2023, 16, 99-102.
  - [2] 李宇春, 戴忠, 武莹, 缪楚宇, 邢睿哲. 基于R软件的课程评价可视化分析[J]. 计算机工程 2023, 52(11), 58-60.
  - [3] 薛毅, 陈立萍. 统计建模与R软件[M]. 清华大学出版社, 2015.
  - [4] 何晓群, 刘文卿. 应用回归分析[M]. 北京: 中国人民大学出版社, 2019.
  - [5] 郭佳璐, 方波, 李剑敏, 王震, 高艳姣, 蔡晋辉. 基于迈克尔逊干涉法的太赫兹波长精密测量. [J]. 计量学报 2022 43(9), 1147-1153
  - [6] 穆晓东. 迈克尔逊干涉仪自动测量系统设计与实现[J]. 现代信息科技 2023 7(11) 172-178.
  - [7] 姜盛, 武进敏, 鲁溟峰, 魏洪涛, 范军芳, 张峰, 陶然. 基于分数傅里叶域信号处理的迈克尔逊干涉法测量波长[J]. 计量学报 2023 44(11) 1652-1658
- 作者简介: 武红(1986.05—), 女, 汉族, 山西忻州人, 博士, 副教授, 研究生导师, 主要从事超材料、超表面、光子晶体器件的理论及实验研究。
- 基金项目: 南京邮电大学教学改革研究项目(重点), 项目号JG00721JX14; 大学物理情境教学法的研究与应用, 项目号JG00719JX31。