

# “大雾实验工具”的开发

## 《程序设计进阶与实践》大作业报告

姓名（组长）	孙旭磊	学号	PB21000270
姓名（组员1）	秦沁*	学号	PB21111630
姓名（组员2）	赵弈	学号	PB21000033
姓名（组员3）	鲍政廷	学号	PB21111741
姓名（组员4）	张学涵	学号	PB21000079
项目名称	大雾实验工具——绘制图像&计算不确定度&生成计算公式		

## 1 项目需求分析

## 2 项目功能设计

### 2.1 总体功能说明

大雾实验工具是本组成员 2022 春季学期程序设计进阶与实践的大作业项目。本工具搭建于网页平台，支持任何设备自由访问。传入实验数据后，本工具立刻完成绘制图像、计算不确定度、生成计算公式一系列操作，并将最终结果整理成一份 Word 文档，下载后即可直接使用。本工具支持一级大物的所有实验（如表 2），大大提升了学生们撰写实验报告的效率。由于本工具只是将传入的实验数据进行自动分析，故不会造成抄袭、造假等学术不端问题。

### 2.2 具体功能点说明

使用本工具时，用户只需输入他们做实验时测量到的原始数据，而无需任何额外的计算处理，用户所要做的只有按照规定的格式上传 Excel 文档。本工具支持 `xlsx`, `csv` 等各种格式的数据表格。具体而言，每个实验都会有一张示例数据表供用户参考，如图 1 的界面所示。用户也可以直接下载示例数据，并直接在其基础上进行修改。因此，本工具没有任何学习成本，是一款即点即用、免安装的简单轻应用。

#### 绘制图像

本工具根据输入的数据以及实验原理，自动生成美观的实验图像。本工具支持平滑去噪、数



图 1: “拉伸法测钢丝杨氏模量”的工具界面

据拟合、双 y 图等多种图像生成需求，如图 2 所示。

#### 计算不确定度

本工具在生成的 Word 文档中渲染了各种公式，如图 3 所示。用户可以直观看到不确定度每一步的计算过程，并在自己的报告中直接使用这些算式与结果。

#### 生成计算公式

在 Word 文档中除了有已经渲染好的公式外，我们还提供了它们的 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ 源码，如图 4 所示。这极大方便了用 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ , Markdown 等排版实验报告的用户，他们再也不需要手动敲入每一个算式了。

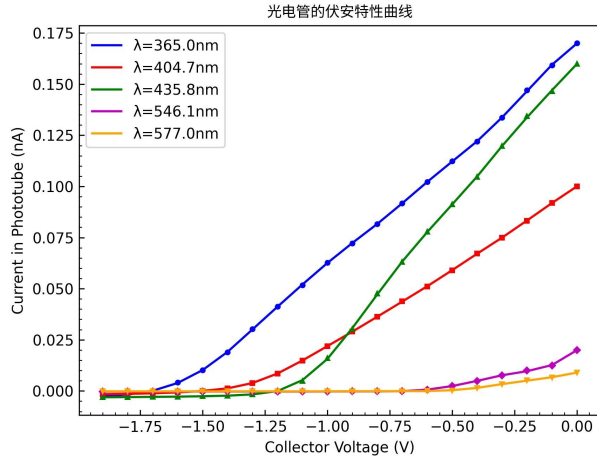


图 2: 平滑连接的光电效应伏安特性曲线

$$\begin{aligned} &0.293)^2 + (0.295 - 0.293)^2 \}^{5-1} \} \backslash \mathrm{mm} \backslash \\ &\&=0.0015811 \backslash \mathrm{mm} \\ &\backslash \mathrm{end{aligned}} \\ &\$ \$ \end{aligned}$$

钢丝直径 d 的 B 类不确定度:

$$\begin{aligned} &\$ \$ \\ &\Delta_{B,d} = \sqrt{\Delta_{\text{仪}}^2 + \Delta_{\text{估}}^2} \\ &= \sqrt{0.004^2 + 0.005^2} \backslash \mathrm{mm} = 0.0064031 \backslash \mathrm{mm} \\ &\$ \$ \end{aligned}$$

钢丝直径 d 的展伸不确定度:

$$\begin{aligned} &\$ \$ \\ &\backslash \mathrm{begin{aligned}} \\ &U_{d,P} = \sqrt{\left( t_P \frac{\sigma_d}{\sqrt{n}} \right)^2 + \left( k_P \frac{\Delta_{B,d}}{C} \right)^2} \\ &= \sqrt{\left( 2.78 \times \frac{0.0015811}{\sqrt{5}} \right)^2 + \left( 1.96 \times \frac{0.0064031}{3} \right)^2} \backslash \mathrm{mm} \\ &\&= 4.6222 \times 10^{-3} \backslash \mathrm{mm}, P = 0.95 \end{aligned}$$

图 4: 不确定度算式的L<sup>A</sup>T<sub>E</sub>X源码

钢丝直径 d 的平均值:

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = \frac{0.291 + 0.292 + 0.293 + 0.294 + 0.295}{5} \text{ mm} = 0.293 \text{ mm}$$

钢丝直径 d 的标准差:

$$\begin{aligned} \sigma_d &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2} \\ &= \sqrt{\frac{(0.291 - 0.293)^2 + (0.292 - 0.293)^2 + (0.293 - 0.293)^2 + (0.294 - 0.293)^2 + (0.295 - 0.293)^2}{5-1}} \text{ mm} \\ &= 0.0015811 \text{ mm} \end{aligned}$$

钢丝直径 d 的 B 类不确定度:

$$\Delta_{B,d} = \sqrt{\Delta_{\text{仪}}^2 + \Delta_{\text{估}}^2} = \sqrt{0.004^2 + 0.005^2} \text{ mm} = 0.0064031 \text{ mm}$$

钢丝直径 d 的展伸不确定度:

$$\begin{aligned} U_{d,P} &= \sqrt{\left( t_P \frac{\sigma_d}{\sqrt{n}} \right)^2 + \left( k_P \frac{\Delta_{B,d}}{C} \right)^2} = \sqrt{\left( 2.78 \times \frac{0.0015811}{\sqrt{5}} \right)^2 + \left( 1.96 \times \frac{0.0064031}{3} \right)^2} \text{ mm} \\ &= 4.6222 \times 10^{-3} \text{ mm}, P = 0.95 \end{aligned}$$

图 3: 不确定度计算的详细过程

## 2.3 功能点设计细节

本工具后端使用 Python 编写, 使用的包与模块如表 1 所示。前端由 HTML 编写, 并使用了 Flask Web 应用框架。

### 2.3.1 图像绘制

图像由 matplotlib 绘制。我们的规范如下:

- 面向绘图对象作图: `fig, ax = matplotlib.pyplot.subplots()`
- 设置副刻度为主刻度的一半, 主刻度为默认: `ax.xaxis.set_minor_locator(matplotlib.ticker.AutoMinorLocator(2))`
- 刻度朝内: `matplotlib.rcParams["xtick.direction"] = matplotlib.rcParams["ytick.direction"] = "in"`

表 1: 本工具使用的全部 Python 包与模块

Python 包或模块	用途
chardet	检测用户上传的数据表格的编码
collections	通过 <code>namedtuple</code> 使代码更清晰
Flask	Web 应用框架
latex2mathml	L <sup>A</sup> T <sub>E</sub> X 代码转换为 MathML 代码
lxml	渲染 MathML 代码以用于 Word
math, numpy	不确定度数字运算
matplotlib	绘制物理图像
os, random, shutil	后台文件操作与管理
pandas	数据表格处理
python-docx	生成 Word 文档
scipy	数据拟合
sympy	不确定度符号运算
time, threading	定时删除生成的 Word 文档
traceback	打印运行错误以便调试

- 若一张图有且只有一组点线，则点使用红色 (`color="r"`)，线使用蓝色 (`color="b"`)，且线覆盖在点的上面；若一张图有多组点线，则同一组点线的颜色应当相同，并依次使用蓝(b)、红(r)、绿(g)、紫(m)、橙(orange)、青(c)。
- 点的类型使用实心圆 (`"o"`)，若一张图有多组点线，则依次使用实心圆(o)、正方形(s)、上三角(^)、菱形(D)、下三角(v)、星号(\*)。
- 线条粗细使用 `linewidth=1.5`，点的大小使用 `markersize=3`，可视数据量、数据组数适当调整，但应保持统一性。
- 绘制双y轴图使用 `twinx()` 方法。
- 只有一组点线的图，一般不显示图例。
- 图像字体: `SourceHanSansSC-Regular.otf`

- 轴标签和标题中的物理量名称与单位应使用L<sup>A</sup>T<sub>E</sub>X。

### 2.3.2 数据处理

无论使绘制图像时的线性拟合，还是计算不确定度的大小，都绕不开数据处理。我们利用 `pandas`, `scipy`, `sympy` 等包自主编写了 `calc.py` 应用程序接口，它提供以下函数：

**科学计数法输出** `numlatex`: (`num`: float, `prec`: int = 5) -> str

返回一个数的科学计数法形式的L<sup>A</sup>T<sub>E</sub>X代码

`num`: 要转成科学计数法的数字

`prec`: 有效数字位数（默认值：5）

**不确定度计算** `analyse`: (`data`: `pandas.DataFrame`, `delta_b1`: float = 0., `delta_b2`: float = 0., `symbol`: str = "x", `unit`: str = "", `confidence_C`: float = 3., `confidence_P`: float = 0.95) -> `collections.namedtuple` ("AnalyseData", ["average", "averagex", "averagex2", "sigma", "sigmax", "sigmax2", "delta\_b", "delta\_bx", "delta\_bx2", "unc", "uncx", "uncx2"])

返回一组数据的平均值、标准差、不确定度

`data`: 要处理的一组实验数据

`delta_b1`: 仪器最大允差  $\Delta_{\text{仪}}$ （默认值：0）

`delta_b2`: 估读最大允差  $\Delta_{\text{估}}$ （默认值：0）

`symbol`: 数据的物理符号（默认值："x"）

`unit`: 数据的单位（默认值："")

`confidence_C`: 置信系数  $C$ （默认值：3）

`confidence_P`: 置信概率  $P$ （缺省值：0.95）

`AnalyseData.average/averagex/averagex2`: 数据的平均值/平均值计算过程的L<sup>A</sup>T<sub>E</sub>X代

码/平均值计算过程的 MathML 代码

AnalyseData.sigma/sigma/sigma2: 数据的标准差/标准差计算过程的 $\text{\LaTeX}$ 代码/标准差计算过程的 MathML 代码

AnalyseData.delta\_b/delta\_bx/delta\_bx2: 数据的B类不确定度/B类不确定度计算过程的 $\text{\LaTeX}$ 代码/B类不确定度计算过程的 MathML 代码

AnalyseData.unc/uncx/uncx2: 数据的展伸不确定度/展伸不确定度计算过程的 $\text{\LaTeX}$ 代码/展伸不确定度计算过程的 MathML 代码

AnalyseData.r/rx/rx2: 线性拟合的相关系数/相关系数计算过程的 $\text{\LaTeX}$ 代码/相关系数计算过程的 MathML 代码

AnalyseData.s\_m/s\_mx/s\_mx2: 拟合直线的斜率标准差/斜率标准差计算过程的 $\text{\LaTeX}$ 代码/斜率标准差计算过程的 MathML 代码

AnalyseData.s\_b/s\_bx/s\_bx2: 拟合直线的截距标准差/截距标准差计算过程的 $\text{\LaTeX}$ 代码/截距标准差计算过程的 MathML 代码

最小二乘法线性回归 analyse\_lsm: (data\_X: pandas.DataFrame, data\_Y: pandas.DataFrame, symbol\_X: str = "X", symbol\_Y: str = "Y", unit\_m: str = "", unit\_b: str = "") -> collections.namedtuple("AnalyseLsmData", ["m", "mx", "mx2", "b", "bx", "bx2", "r", "rx", "rx2", "s\_m", "s\_mx", "s\_mx2", "s\_b", "s\_bx", "s\_bx2"])

返回最小二乘法拟合数据得到的直线参数

data\_X: x轴数据 (自变量数据)

data\_Y: y轴数据 (因变量数据)

symbol\_X: 自变量物理符号 (默认值: "X")

symbol\_Y: 因变量物理符号 (默认值: "Y")

unit\_m: 斜率的单位 (默认值: "")

unit\_b: 截距的单位 (默认值: "")

AnalyseData.m/mx/mx2: 拟合直线的斜率/斜率计算过程的 $\text{\LaTeX}$ 代码/斜率计算过程的 MathML 代码

AnalyseData.b/bx/bx2: 拟合直线的截距/截距计算过程的 $\text{\LaTeX}$ 代码/截距计算过程的 MathML 代码

### 3 测试、运行情况

### 4 设计、开发过程中的难点

### 5 小组分工

### 6 总结与收获

### 7 参考资料

表 2: 全部大雾实验工具

实验分类	实验 ID	实验名称	开发者	子实验
通用	0	通用工具	孙旭磊	

## A Appendix

未完成：斜率计算过程的 $\text{\LaTeX}$ 与  $\text{MathML}$  代码，截距与 Pearson's  $r$  的计算公式错误