

课程介绍

续本达

认识大家

课程概况

概论

开课程机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

课程介绍

续本达

清华大学 工程物理系

2024-09-09 清华

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

认识大家

- 主讲 续本达

2005-2009 在数理基科班学习

2009-2018 在日本神冈地下实验室，学习研究中微子和暗物质

2018-至今 在工程物理系近代物理研究所任教

主要工作：锦屏中微子实验

- ① JUNO 江门中微子实验
- ② 日本 SuperK 超级神冈实验
- ③ 日本 XMASS 暗物质实验

研究兴趣：太阳中微子、中微子质量本源

爱好：大数据分析、高性能计算、系统架构运维

- 助教 刘逸祺

- 未央书院毕业，工物系博士生

学生构成

- 工物班、核班、未央班

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

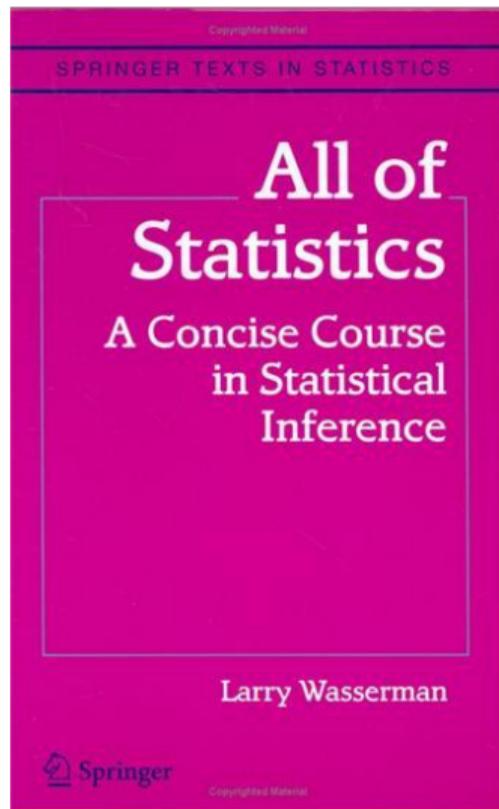
课程概况

- 期中、期末闭卷考试：60%
- 平时作业：32%
 - 每周一收上一周作业
- 课程贡献：8%
 - 课堂抢答、提问、质疑、讨论
 - 指出教学资料中的错误
 - 用 R 语言处理实验课的数据，成为案例
 - 其它贡献
- 课赛结合：2024 Ghost Hunter 中微子数据分析排位赛
 - 一等奖队伍：保送 A+
 - 二等奖队伍：+8%
 - 三等奖队伍：+3%
 - 灵活应用课程中的方法解决竞赛中的问题，可额外计入课程贡献

- 《概率论与数理统计》（第五版）
 - 盛骤，谢式千，潘承毅 编著
 - 高等教育出版社 出版
 - 引入了 R 语言辅助教学



- 给理工科的统计学教材
- 概率论的学习服务于统计应用
- 将机器学习与统计学同等对待
- 注重编程实践
- 紧扣理工科的数据分析主题



- ① 《概率论与数理统计教程》（第二版），茆诗松等编著，高等教育出版社（2011）
 - 在概率论部分讲得非常好。数学严谨性比较强，题目不太容易
- ② 《Information theory, inference and learning algorithm》，David J. Mackay 著，Cambridge Press（2003）
 - 信息论、统计力学、统计推断和机器学习的大统一，适合有物理学基础的同学
- ③ 《测量不确定度导论》，（澳）莱斯 柯卡普等著，曾翔君等译，西安交通大学出版社（2011）
 - 量测技术部分参考
- ④ 《实验物理中的概率和统计》（第二版），朱永生著，科学出版社（2006）
 - 偏向于概率统计在实验中的应用，内容很全。
- ⑤ Statistical Data Analysis, Glen Cowan, Clarendon Press Oxford, 1998
 - 一本非常简要的教材，需要有一定的概率论基础。统计方法以及介绍得非常好。
- ⑥ Data Analysis and Graphics Using R: An Example-Based Approach, 2nd Edition, John Maindonald and John Braun.
 - R 语言参考大全

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

概论

概率统计分析 概率论与数理统计

量测技术 应用到物理实验和测量中

概率论与数理统计

确定现象 一定条件下必然发生的现象

- 例如：太阳从东方升起；上抛物体下落等

随机现象 在一定的条件下并不总是出现相同结果的现象

- 带有随机性、偶然性的现象
- 例如：掷一颗骰子出现的点数；抛硬币出现的正反面等

每次试验前不能预言出现什么结果

不确定性 每次试验后出现的结果不止一个

统计规律性 在相同的条件下进行大量观察或试验时，出现的结果有一定的规律性

概率论

- 研究和揭示随机现象的统计规律性的科学

Probability Theory is Nothing but Common Sense Reduced to Calculations –
Pierre-Simon Laplace

- 统计方法是从大量（实验）数据中总结提取一般规律的必要手段，概率论是统计方法的基础。
 - 桥接数学与应用科学
 - 几乎所有与数据有关的学科的语言
- 在保持一定的数学严谨性的前提下，重点强调统计方法在工程物理系相关学科中的应用。
 - ① 如何根据物理目标完成实验的设计？
 - ② 如何从实验数据中提取物理规律？
 - ③ 如何规范表述实验结果？

- 统计方法是从大量（实验）数据中总结提取一般规律的必要手段，概率论是统计方法的基础。
 - 桥接数学与应用科学
 - 几乎所有与数据有关的学科的语言
- 在保持一定的数学严谨性的前提下，重点强调统计方法在工程物理系相关学科中的应用。
 - ① 如何根据物理目标完成实验的设计？
 - ② 如何从实验数据中提取物理规律？
 - ③ 如何规范表述实验结果？

- 前者来自数学系和统计系，后者来自计算机系
 - 前者觉得后者在重造轮子；后者觉得前者的方法不适用。
- 2000 后事情变了
 - 统计家发现计算家有全新的贡献；计算家承认统计理念的普适性。
- 数据挖掘算法的应用范围超出了统计学家过去的想像
 - 已有的统计学理论比计算家预想得应用广泛

概念映射图

统计学	机器学习	统计学	机器学习
估计	学习	假设 (检验)	-
分类	监督学习	置信区间	-
聚类	非监督学习	有向图	贝叶斯网络
因变量	特征 (工程)	贝叶斯推断	贝叶斯推断
分类函数	假设	频率派统计	-

- 前者来自数学系和统计系，后者来自计算机系
 - 前者觉得后者在重造轮子；后者觉得前者的方法不适用。
- 2000 后事情变了
 - 统计家发现计算家有全新的贡献；计算家承认统计理念的普适性。
- 数据挖掘算法的应用范围超出了统计学家过去的想像
 - 已有的统计学理论比计算家预想得应用广泛

概念映射图

统计学	机器学习	统计学	机器学习
估计	学习	假设 (检验)	-
分类	监督学习	置信区间	-
聚类	非监督学习	有向图	贝叶斯网络
因变量	特征 (工程)	贝叶斯推断	贝叶斯推断
分类函数	假设	频率派统计	-

- 前者来自数学系和统计系，后者来自计算机系
 - 前者觉得后者在重造轮子；后者觉得前者的方法不适用。
- 2000 后事情变了
 - 统计家发现计算家有全新的贡献；计算家承认统计理念的普适性。
- 数据挖掘算法的应用范围超出了统计学家过去的想像
 - 已有的统计学理论比计算家预想得应用广泛

概念映射图

统计学	机器学习	统计学	机器学习
估计	学习	假设（检验）	-
分类	监督学习	置信区间	-
聚类	非监督学习	有向图	贝叶斯网络
因变量	特征（工程）	贝叶斯推断	贝叶斯推断
分类函数	假设	频率派统计	-

- 随机事件与概率
- 随机变量及其分布
- 多维随机变量及其分布
- 随机变量的数字特征
- 蒙特卡罗方法
- 大数定律与中心极限定理
- 样本分布
- 参数估计
- 假设检验
- 方差分析
- 回归分析
- 随机过程
- 测量与不确定度

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

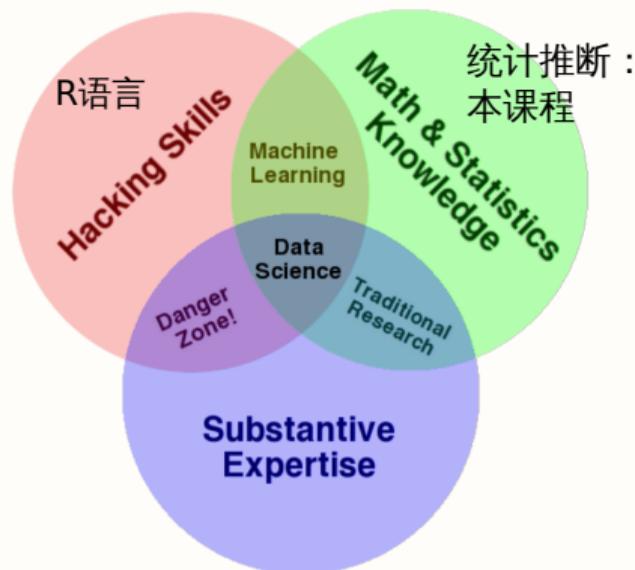
开课动机

为什么要学习概率统计分析？

- ① 优良学风班考察班级学生平均绩点，人数不同的班级适合比较吗？
- ② 买买买剁手时，不同的店家类似的商品，如何比选？
- ③ 拉普拉斯：如何把 Common Sense 用 Calculations 表达出来？

此课在知识体系中的地位：这是一个数据的时代

- 数据时代需要 黑客技术 Hacking Skills、数理统计 Math & Statistics Knowledge、专业知识 Substantive Expertise 等多方面技能。

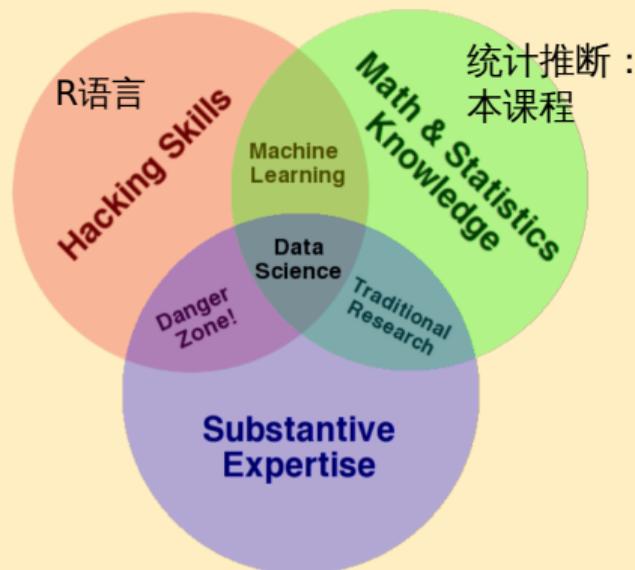


为什么要学习概率统计分析？

- ① 优良学风班考察班级学生平均绩点，人数不同的班级适合比较吗？
- ② 买买买剁手时，不同的店家类似的商品，如何比选？
- ③ 拉普拉斯：如何把 Common Sense 用 Calculations 表达出来？

此课在知识体系中的地位：这是一个数据的时代

- 数据时代需要 **黑客技术** Hacking Skills、**数理统计** Math & Statistics Knowledge、**专业知识** Substantive Expertise 等多方面技能。



课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

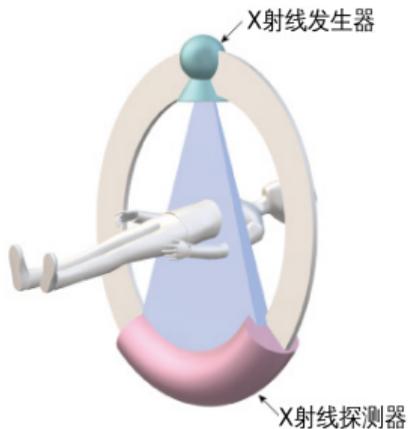
例：精密测量

计算机断层扫描 (Computed Tomography, CT)

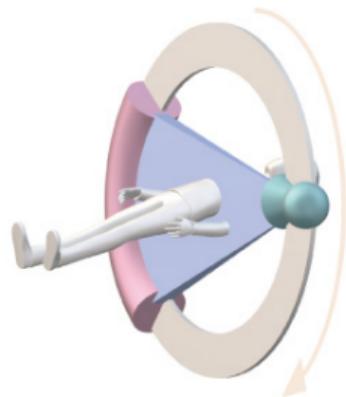
CT 装置



X 射线



转动扫描

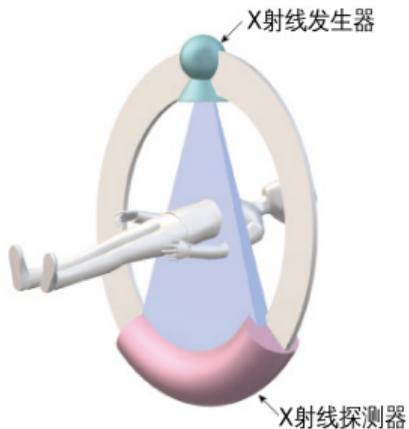


计算机断层扫描 (Computed Tomography, CT)

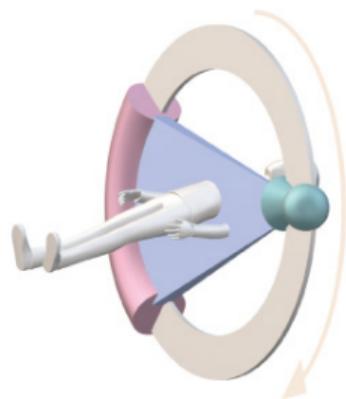
CT 装置



X 射线



转动扫描



* 为了减少 X 射线辐射量，必须使用灵敏的探测器。

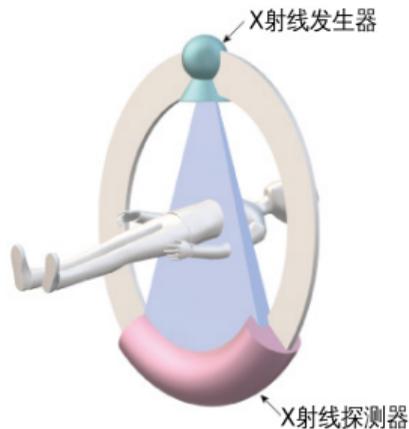
X 射线 $\xrightarrow{\text{闪烁体}}$ 可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

计算机断层扫描 (Computed Tomography, CT)

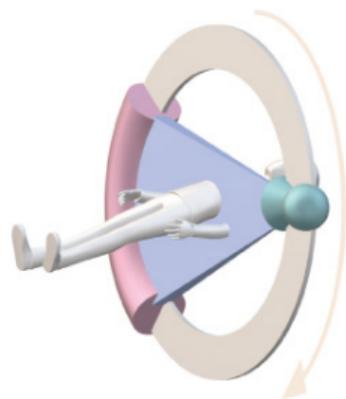
CT 装置



X 射线



转动扫描



- 为了减少 X 射线辐射量，必须使用灵敏的探测器。

X 射线 $\xrightarrow{\text{闪烁体}}$ 可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

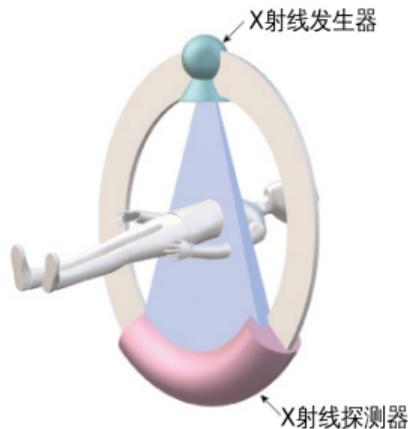
- 光电倍增管是检测极微弱光的器件，在科研与民生各领域应用广泛

计算机断层扫描 (Computed Tomography, CT)

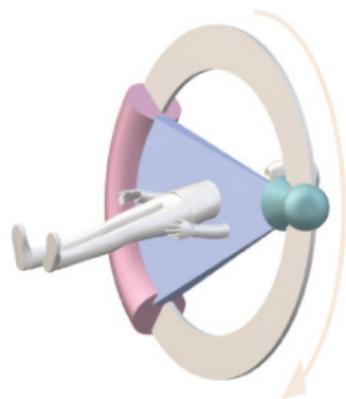
CT 装置



X 射线



转动扫描



- 为了减少 X 射线辐射量，必须使用灵敏的探测器。

X 射线 $\xrightarrow{\text{闪烁体}}$ 可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

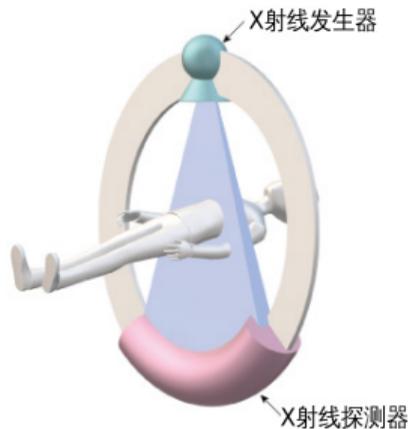
- 光电倍增管是检测极微弱光的器件，在科研与民生各领域应用广泛

计算机断层扫描 (Computed Tomography, CT)

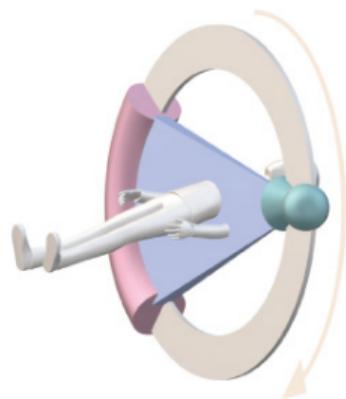
CT 装置



X 射线



转动扫描



- 为了减少 X 射线辐射量，必须使用灵敏的探测器。

X 射线 $\xrightarrow{\text{闪烁体}}$ 可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

- 光电倍增管是检测极微弱光的器件，在科研与民生各领域应用广泛

光电倍增管 (Photomultiplier Tube, PMT) 的光子检测效率

PMT



微弱的可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

光子检测效率 一个光子击中 PMT 后被检测到的概率，表征 PMT 性能

$$\frac{R_1(\text{实验目标})}{R_0(\text{已知参考值})} = \frac{D_1(\text{实验测量计数率})}{D_0}$$

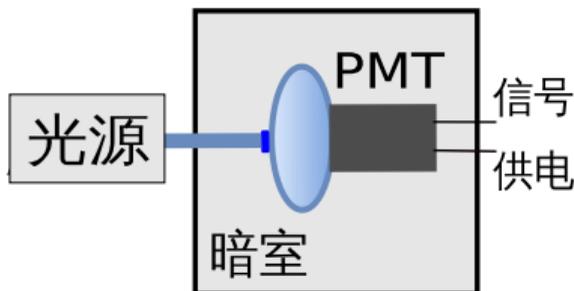
- PMT 工作在极微弱的光环境，测量时间长，光源稳定度不够

光电倍增管 (Photomultiplier Tube, PMT) 的光子检测效率

PMT



PMT 测量标定



微弱的可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

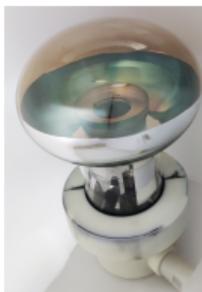
光子检测效率 一个光子击中 PMT 后被检测到的概率，表征 PMT 性能

$$\frac{R_1(\text{实验目标})}{R_0(\text{已知参考值})} = \frac{D_1(\text{实验测量计数率})}{D_0}$$

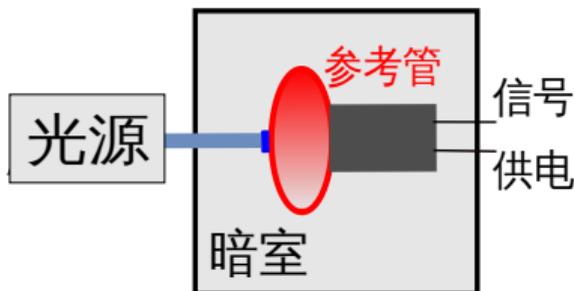
- PMT 工作在极微弱的光环境，测量时间长，光源稳定度不够

光电倍增管 (Photomultiplier Tube, PMT) 的光子检测效率

PMT



光源标定



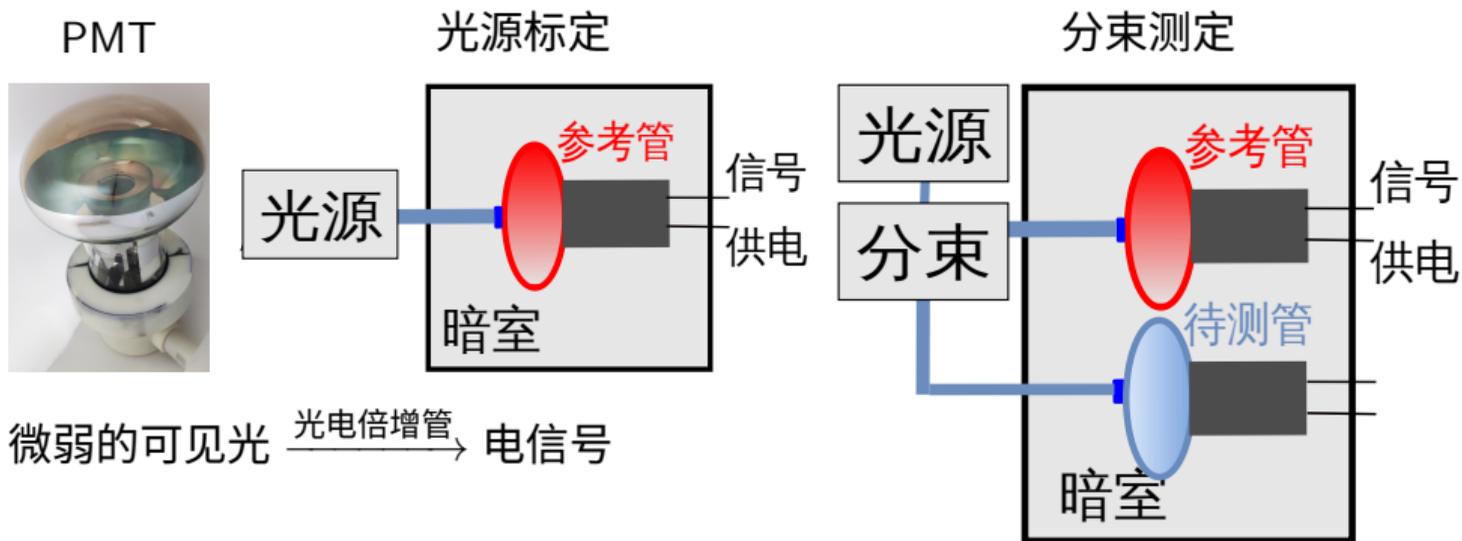
微弱的可见光 $\xrightarrow{\text{光电倍增管}}$ 电信号

光子检测效率 一个光子击中 PMT 后被检测到的概率，表征 PMT 性能

$$\frac{R_1(\text{实验目标})}{R_0(\text{已知参考值})} = \frac{D_1(\text{实验测量计数率})}{D_0}$$

- PMT 工作在极微弱的光环境，测量时间长，光源稳定度不够

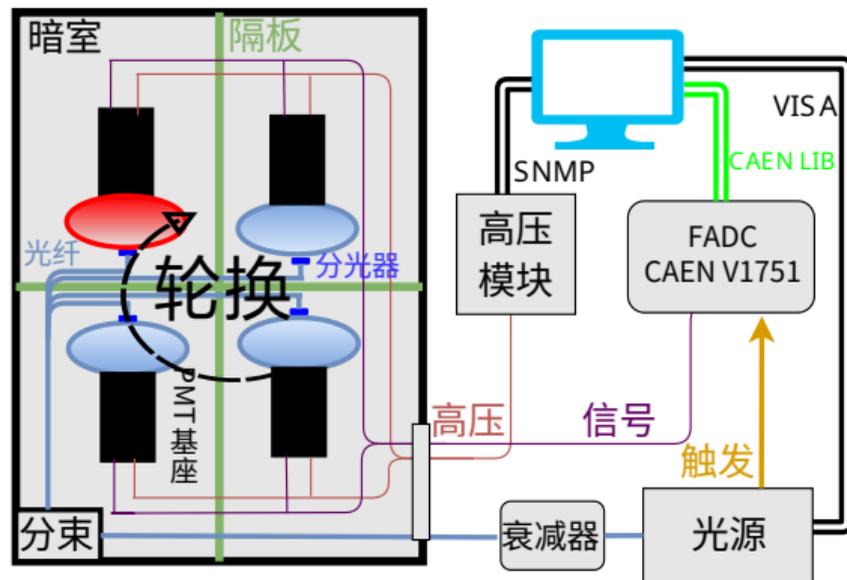
光电倍增管 (Photomultiplier Tube, PMT) 的光子检测效率



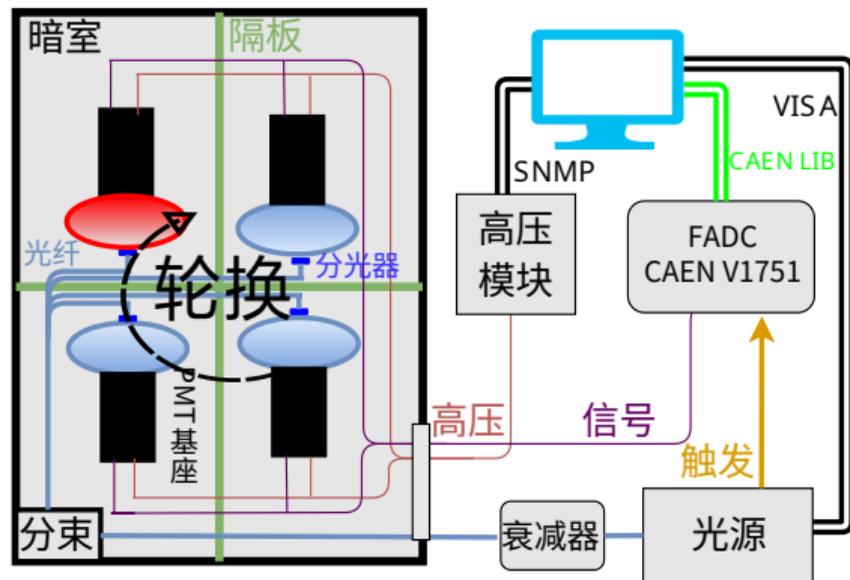
光子检测效率 一个光子击中 PMT 后被检测到的概率，表征 PMT 性能

$$\frac{R_1 (\text{实验目标})}{R_0 (\text{已知参考值})} = \frac{D_1 (\text{实验测量计数率})}{D_0}$$

- PMT 工作在极微弱的光环境，测量时间长，光源稳定度不够



- 将参考管在四个暗室中轮换，同时测量四支 PMT 以降低统计误差
- 疑问：分束器是否均匀分光，使得四个暗室的光强一致？



- 将参考管在四个暗室中轮换，同时测量四支 PMT 以降低统计误差
- 疑问：分束器是否均匀分光，使得四个暗室的光强一致？

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

例：体温计

- 电子体温计使用方便



工作原理：金属氧化物热敏电阻

- 负温度系数型：随着温度升高，电阻值变小

- 电子体温计使用方便



工作原理：金属氧化物热敏电阻

- 负温度系数型：随着温度升高，电阻值变小

別了！水银溫度計！

**我国明确將禁止生产
含汞体温计、血压计产品**

已经取得医疗器械注册证的
含汞体温计和含汞血压计产品，
原注册证在证书有效期内继续有效；
注册证有效期届满可以申请延续注册，
但限定其注册证有效期
不得超过2025年12月31日。

来源：中国市场监管报

获取更多干货，即保护您的合法权益，请联系我们及时删除。

- 电子体温计使用方便



工作原理：金属氧化物热敏电阻

- 负温度系数型：随着温度升高，电阻值变小

別了！水銀溫度計！

**我国明确將禁止生产
含汞体温计、血压计产品**

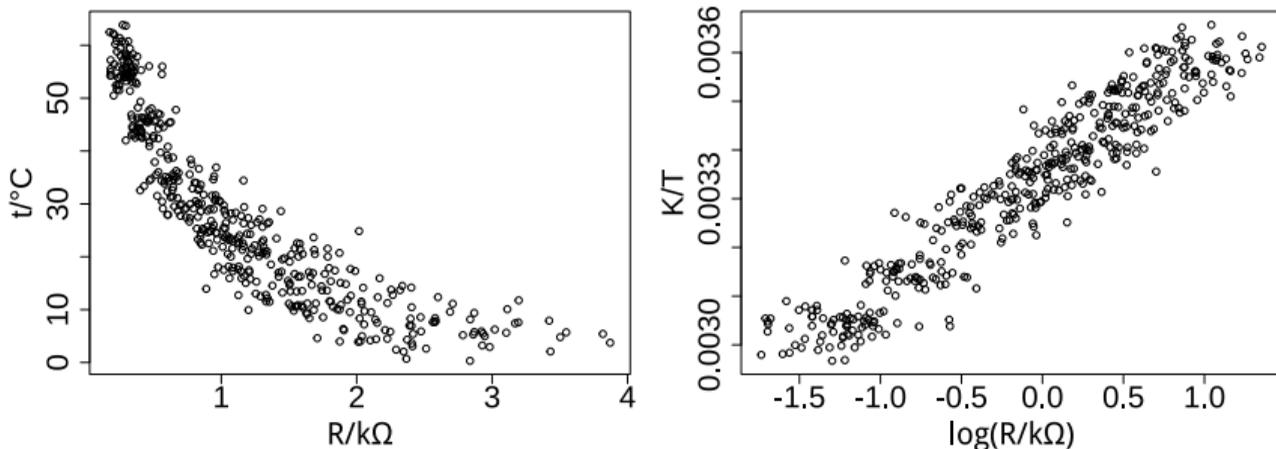


已经取得医疗器械注册证的
含汞体温计和含汞血压计产品，
原注册证在证书有效期内继续有效；
注册证有效期届满可以申请延续注册，
但限定其注册证有效期
不得超过2025年12月31日。

来源：中国市场监管报

资料来源：网络，如侵犯您的合法权益，请联系我们及时删除。 谢谢

```
d = read.table("thermister.csv", col.names=c("温度", "电阻"))
plot(温度 ~ 电阻, data=d, ylab="温度/C", xlab="电阻/Ohm")
```

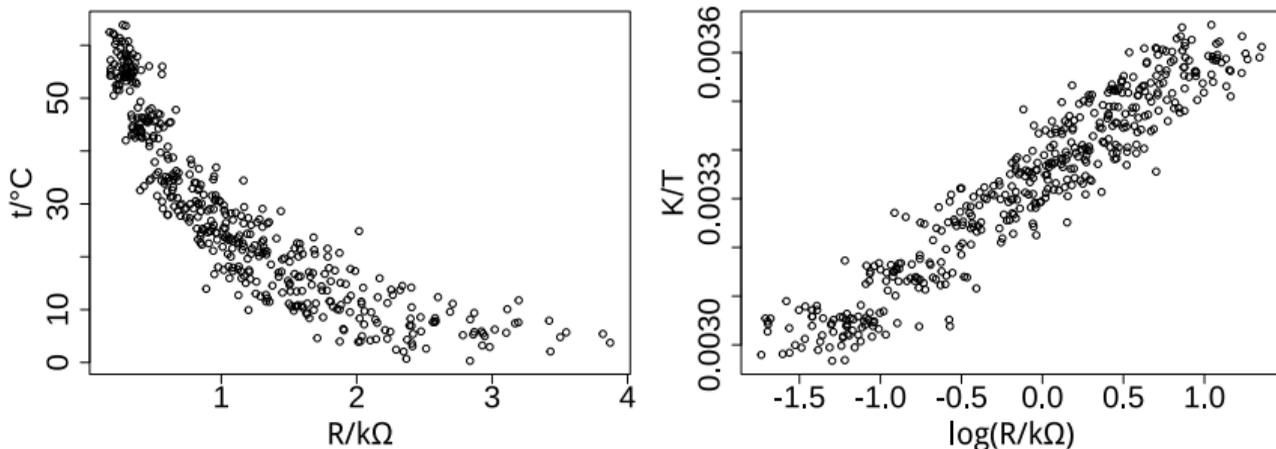


```
plot(1/(温度 + 273.15) ~ log(电阻), data=d, ylab="1/温度 [K]", xlab="log(电阻/[Ohm])")
```

- 温度与电阻值有对应关系

- 变成直线 $T \rightarrow \frac{1}{T + 273.15}$, $R \rightarrow \log \frac{R}{[\Omega]}$

```
d = read.table("thermister.csv", col.names=c("温度", "电阻"))
plot(温度 ~ 电阻, data=d, ylab="温度/C", xlab="电阻/Ohm")
```



```
plot(1/(温度 + 273.15) ~ log(电阻), data=d, ylab="1/温度 [K]", xlab="log(电阻/[Ohm])")
```

- 温度与电阻值有对应关系
- 变成直线 $T \rightarrow \frac{1}{T + 273.15}, R \rightarrow \log \frac{R}{[\Omega]}$

爱因斯坦 1953 给 J.S.Switzer 的信

西方科学的发展是以两个伟大的成就为基础：希腊哲学家发明形式逻辑体系 (在欧几里得几何中)，以及 (在文艺复兴时期) 发现通过系统的实验可能找出因果关系。

- ① 只有逻辑如何？
- ② 只有实验如何？

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

例：辐射计数

计数问题例子：盖革计数器的放射性计量

2023年8月24日至9月11日，日本福岛第一核电站核污染水排海 7788 吨。

早在 2012 年德国海洋研究团队模拟了污水的扩散过程。

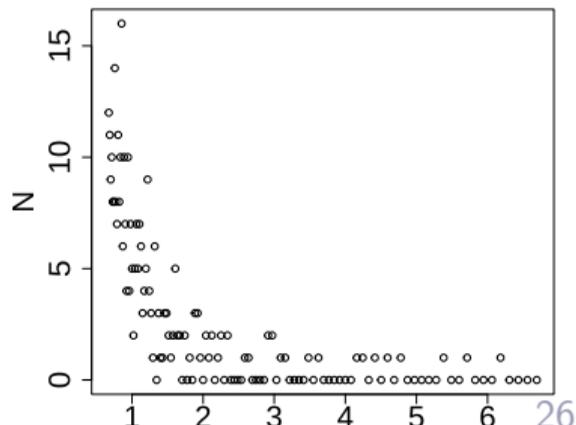
- 环境放射性，由探头计数测量 计数器影片
- 由计数反映辐射强度（为简便，设单位为 1）

计数问题例子：盖革计数器的放射性计量

2023年8月24日至9月11日，日本福岛第一核电站核污染水排海 7788 吨。

早在 2012 年德国海洋研究团队模拟了污水的扩散过程。

- 环境放射性，由探头计数测量 计数器影片
- 由计数反映辐射强度（为简便，设单位为 1）

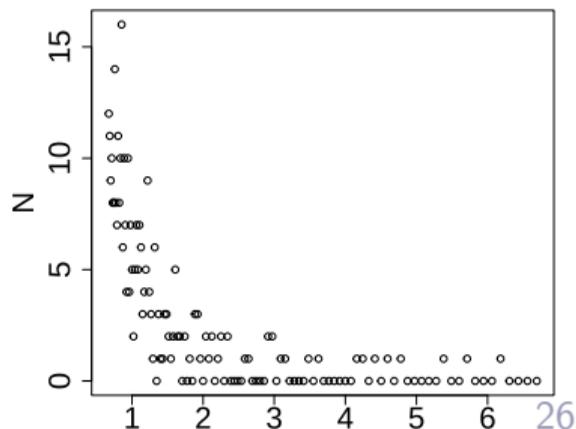


计数问题例子：盖革计数器的放射性计量

2023年8月24日至9月11日，日本福岛第一核电站核污染水排海 7788 吨。

早在 2012 年德国海洋研究团队模拟了污水的扩散过程。

- 环境放射性，由探头计数测量 计数器影片
- 由计数反映辐射强度（为简便，设单位为 1）
- 能否用线性模型描述？



课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

例：雨量



“北京有观测记录的 140 年以来第一大降雨” ——北京市气象局



第 5 号“杜苏芮”台风残余造成北京及华北特大降雨



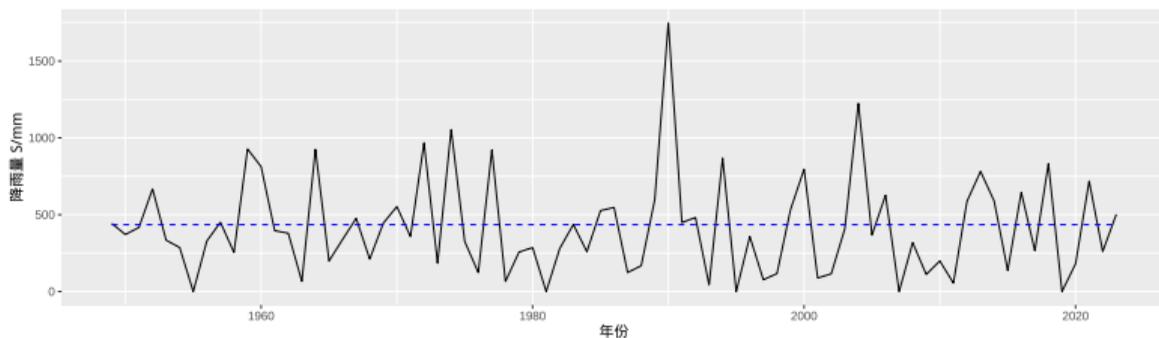
“北京有观测记录的 140 年以来第一大降雨” ——北京市气象局



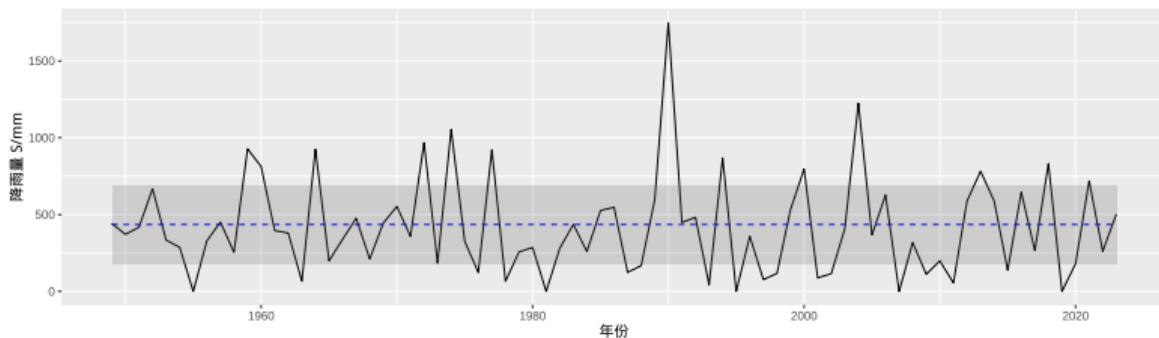
第 5 号“杜苏芮”台风残余造成北京及华北特大降雨

- 设北京 8 月的总降雨量为 S ，是一个随机变量
 - 我们需要刻画 S 的特点，来掌握北京的降雨规律
- 随机变量 S 的期望 $E(S)$ ：直观反映了北京 8 月的平均降雨量。

- 设北京 8 月的总降雨量为 S ，是一个随机变量
 - 我们需要刻画 S 的特点，来掌握北京的降雨规律
- 随机变量 S 的期望 $E(S)$ ：直观反映了北京 8 月的平均降雨量。

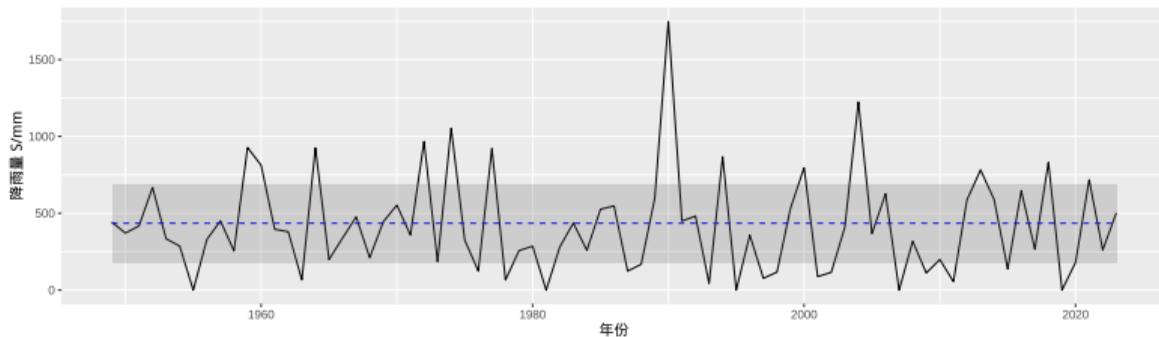


- 设北京 8 月的总降雨量为 S ，是一个随机变量
 - 我们需要刻画 S 的特点，来掌握北京的降雨规律
- 随机变量 S 的期望 $E(S)$ ：直观反映了北京 8 月的平均降雨量。



- 方差 $\text{Var}(S)$ 与标准差：反映了北京 8 月的降雨量的年度变化。

- 设北京 8 月的总降雨量为 S ，是一个随机变量
 - 我们需要刻画 S 的特点，来掌握北京的降雨规律
- 随机变量 S 的期望 $E(S)$ ：直观反映了北京 8 月的平均降雨量。



- 方差 $\text{Var}(S)$ 与标准差：反映了北京 8 月的降雨量的年度变化。

问题

如何从经验的“平均”与“变化”，上升到理论上对雨量期望与方差的理解？

- 北京 8 月降雨量 S 服从什么分布？

- S 受两个因素的影响：

- 北京 8 月的降雨次数 N – 计数：离散型随机变量
- 各 (i) 次下雨的降雨量 R_i ，独立同分布 – 非负的连续型随机变量

$$S = \sum_{i=1}^N R_i$$

- 拆解成两个子问题：

- 降雨次数 N 对总雨量 S 的影响
- 当 N 确定时，单次雨量对总雨量的影响

- 北京 8 月降雨量 S 服从什么分布？
- S 受两个因素的影响：
 - ① 北京 8 月的降雨次数 N – 计数：离散型随机变量
 - ② 各 (i) 次下雨的降雨量 R_i ，独立同分布 – 非负的连续型随机变量

$$S = \sum_{i=1}^N R_i$$

- 拆解成两个子问题：
 - ① 降雨次数 N 对总雨量 S 的影响
 - ② 当 N 确定时，单次雨量对总雨量的影响

- 北京 8 月降雨量 S 服从什么分布？
- S 受两个因素的影响：
 - ① 北京 8 月的降雨次数 N – 计数：离散型随机变量
 - ② 各 (i) 次下雨的降雨量 R_i ，独立同分布 – 非负的连续型随机变量

$$S = \sum_{i=1}^N R_i$$

- 拆解成两个子问题：
 - ① 降雨次数 N 对总雨量 S 的影响
 - ② 当 N 确定时，单次雨量对总雨量的影响

- 北京 8 月降雨量 S 服从什么分布？
- S 受两个因素的影响：
 - ① 北京 8 月的降雨次数 N – 计数：离散型随机变量
 - ② 各 (i) 次下雨的降雨量 R_i ，独立同分布 – 非负的连续型随机变量

$$S = \sum_{i=1}^N R_i$$

- 拆解成两个子问题：
 - ① 降雨次数 N 对总雨量 S 的影响
 - ② 当 N 确定时，单次雨量对总雨量的影响

- 北京 8 月降雨量 S 服从什么分布？
- S 受两个因素的影响：
 - ① 北京 8 月的降雨次数 N – 计数：离散型随机变量
 - ② 各 (i) 次下雨的降雨量 R_i ，独立同分布 – 非负的连续型随机变量

$$S = \sum_{i=1}^N R_i$$

- 拆解成两个子问题：
 - ① 降雨次数 N 对总雨量 S 的影响
 - ② 当 N 确定时，单次雨量对总雨量的影响

课程介绍

续本达

认识大家

课程概况

概论

开课动机

例：精密测量

例：体温计

例：辐射计数

例：雨量

例：核弹

例：核弹

《马兰花开》两弹的研发



《马兰花开》剧照、脚本

钱学森：咱们搞科学的，容不得半点马虎啊。

... ..

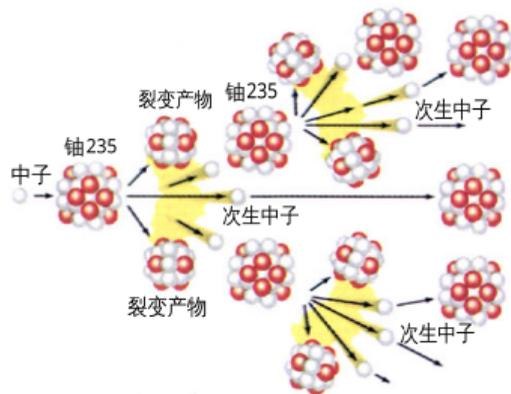
理论组：老邓，第八次计算结果与第九次完全吻合！

邓稼先：中国的第一颗原子弹，就从这里开始吧！

九次计算耗时半年，
最终结果一致



- 核裂变链式反应过程复杂，各步伴随机性
 - 中子在介质里运输被随机散射损耗能量
 - 最终得到确定的“试爆成功”结果
- 要掌握和运用核物理，需要大量的计算



《马兰花开》两弹的研发



《马兰花开》剧照、脚本

钱学森：咱们搞科学的，容不得半点马虎啊。

... ..

理论组：老邓，第八次计算结果与第九次完全吻合！

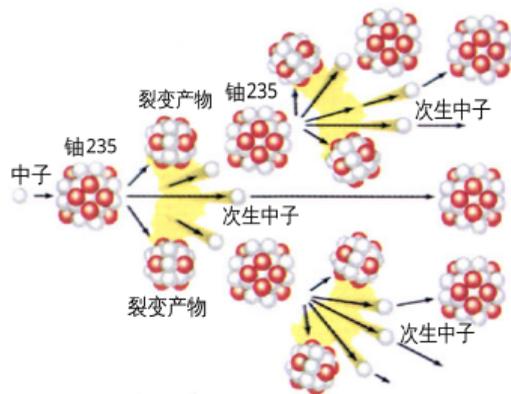
邓稼先：中国的第一颗原子弹，就从这里开始吧！

九次计算耗时半年，
最终结果一致



- 核裂变链式反应过程复杂，各步伴 随机性
- 中子在介质里输运被 随机 散射损耗能量
- 最终得到 确定 的“试爆成功”结果

要掌握和运用核物理，需要大量的 计算



《马兰花开》两弹的研发



《马兰花开》剧照、脚本

钱学森：咱们搞科学的，容不得半点马虎啊。

... ..

理论组：老邓，第八次计算结果与第九次完全吻合！

邓稼先：中国的第一颗原子弹，就从这里开始吧！

九次计算耗时半年，
最终结果一致



- 核裂变链式反应过程复杂，各步伴 随机性
- 中子在介质里输运被 随机 散射损耗能量
- 最终得到 确定 的“试爆成功”结果
要掌握和运用核物理，需要大量的 计算

