



清华大学
Tsinghua University



粒子物理与核物理实验中的 数据分析

第零章：课程介绍

杨振伟
清华大学



课程信息

➤ 上课时间地点

周一第3大节 (13: 30-15: 05) : 统计数据分析理论部分

周三第1大节 (08: 00-09: 35) : 数据分析实践部分

授课地点: 旧水304 (+ 瞩目会议+网络学堂+微信群)

➤ 任课教师

王喆、杨振伟、张黎明

➤ 课程考核

理论部分: 期末考试(闭卷): 40%

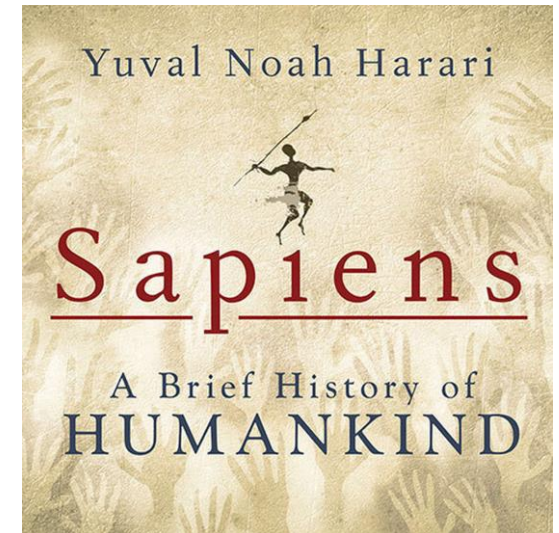
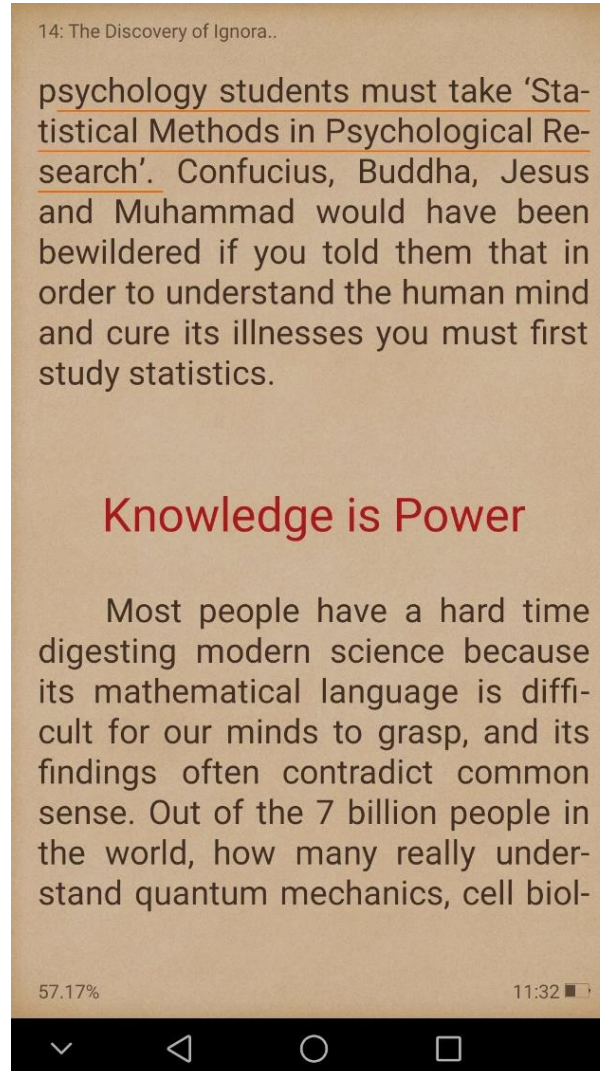
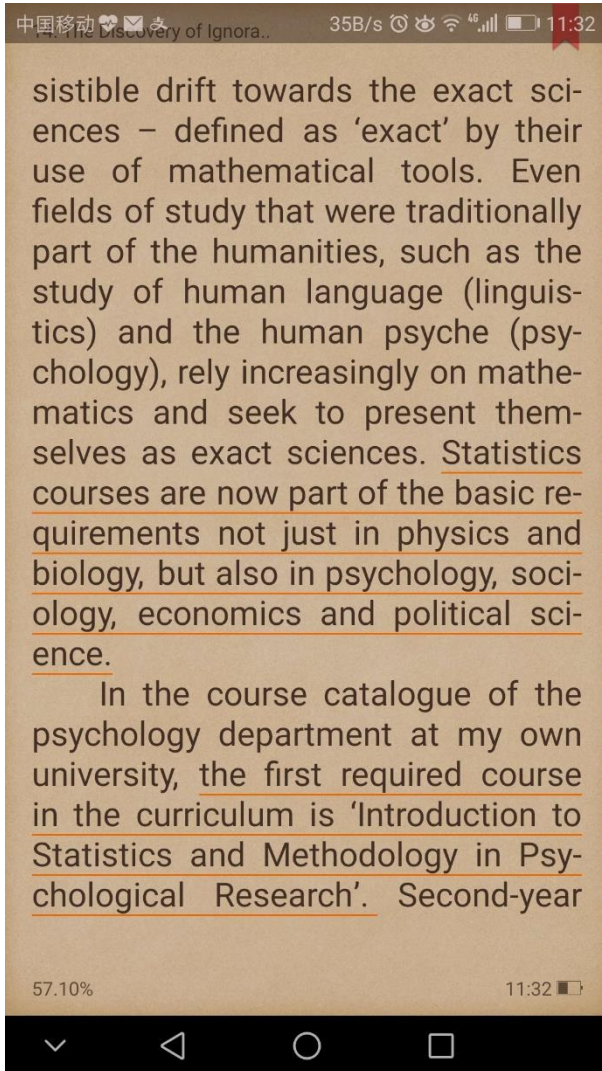
实践部分 (期中+大作业+期末) : 40% (待调整, 约第3周确定最终形式)

平时作业与课堂表现: 20%

参考书目

- G. Cowan, *Statistical Data Analysis*, Clarendon, Oxford, 1998
- F. James, *Statistical Methods in Experimental Physics*, 2nd ed., World Scientific, 2006
- Luca Lista, *Statistical Methods for Data Analysis in Particle Physics*, Springer, 2017
- 朱永生, 《实验物理中的概率和统计》(第二版), 科学出版社, 2006

统计数据分析的重要性



粒子与核物理实验的特点

高！大！快！

➤ 高事例率

LHCb (40 MHz) , STAR (~10 MHz), 超级神冈 (~10 kHz)...

➤ 高统计量

E949 实验 (10^{12} 事例), 北京谱仪实验 (10^7 事例)...

➤ 高信息量

每个事例包含诸如漂移室每根丝的电荷与时间信息, 量能器每块闪烁体的电荷与时间信息...

➤ 大数据量

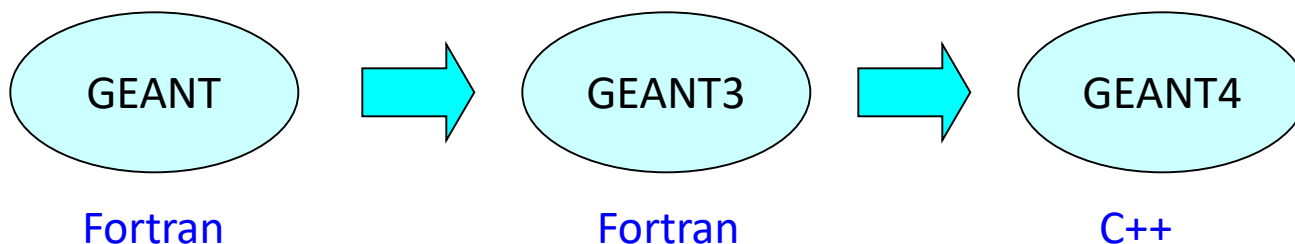
LHCb实验 (20 TB/秒→20 MB/秒→1.7 TB/天)

➤ 快速模拟与数据重建与分析

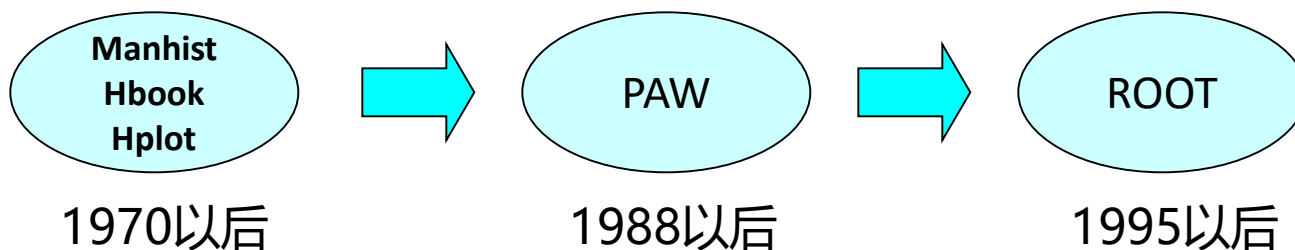
要求在实验过程实时重建事例的物理图象, 分析结果

数据分析手段的发展

➤ 探测器模拟



➤ 数据分析软件包



填图与绘图功能分立，分析功能简单

填图与绘图功能已整合，并具有较强的分析功能

除具有PAW拥有的功能，还结合了C++语言的优势

课程纲要

➤ 统计数据分析理论部分

- 概率, 随机变量
- 蒙特卡罗方法
- 统计检验, 多变量分析 (神经网络)
- 参数估计
- 区间估计
- 不确定度的传递与估计
- 系统不确定度的估计
- 解谱方法 (选讲)

➤ 统计数据分析实践部分

- 简要介绍C++等基本要素
- 介绍Linux操作系统以及必要的shell 脚本写作
- 介绍使用ROOT, GEANT4

课程目的

实验数据 → 统计分析 → 物理结论

➤ 介绍实验数据分析的基本数学工具

- 概率与统计理论

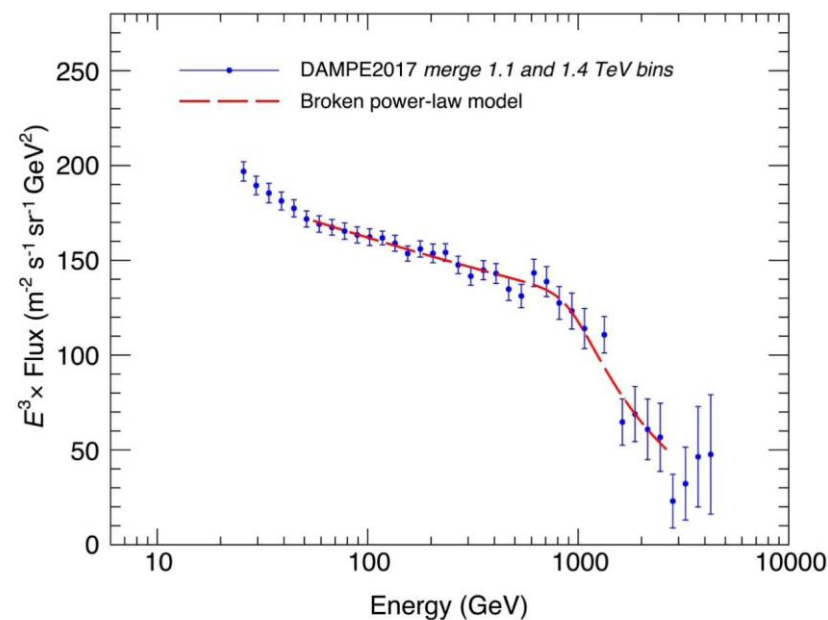
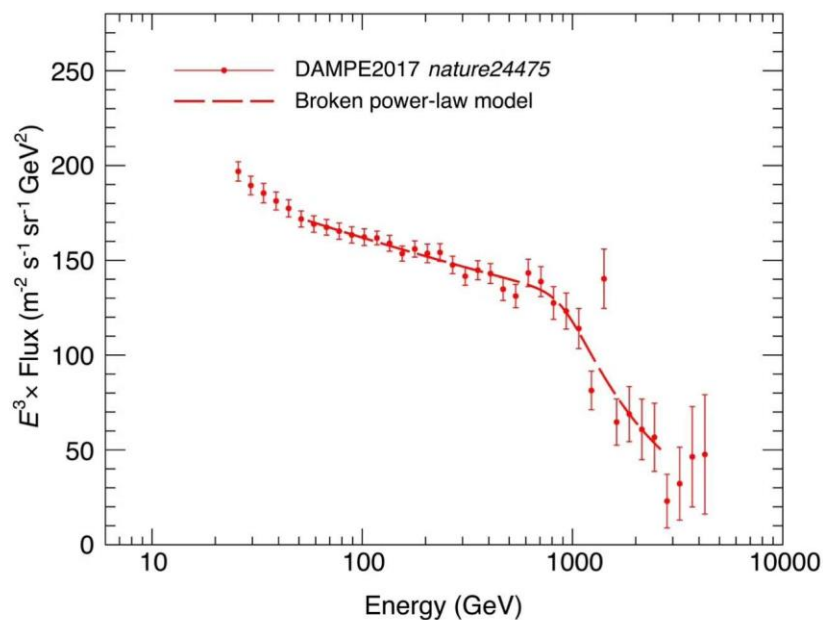
- ✓ 假设检验
- ✓ 参数估计, 区间估计
- ✓ 系统不确定度的分析

➤ 通过实践熟悉实验数据的具体应用

- 熟悉常用的数据分析软件包

- ✓ 基本计算机语言, Linux环境与工具
- ✓ 直方图、散点图、等高线图等绘图技能
- ✓ 参数拟合、假设检验
- ✓ 蒙特卡罗模拟

一个例子



得到实验数据后，如何给出信息？如何给出结论？