

# 高性能

## 前言：

个人感觉这个还是挺有必要学一下的，无论是继续发展生信还是想走计算机的路，命令行都是绕不开的技能，都必须掌握。

## 课程体验：

徐老师上课还可以，但是对于对电脑很生疏或者对命令行抗拒的人可能会跟不上他上课节奏，可能需要课后多实操玩玩命令行，重点在 linux 操作和并行的概念理解，那两节 C++ 课可能会难理解，问题不大。

## 考试：

选择题不是一般的简单，如 pwd, hpc 目的, chmod, >

名词解释：全局变量，并行效率，高性能计算，数据竞争，

论述：Linux 发行版，PATH 目的怎么临时怎么永久，语言比较，module 用处，openmp 干嘛的，简单的累加平方，要小心什么

## 复习：

如果是为了过考试，考前两三天甚至都够了，考的太简单了，不过有一小道 C++ 的编程题，写一个简单的平方求和，我触发被动写成 python 的代码了，这题错了，这个课复习也是只能看 ppt，不过我个人对于纯 PPT 复习很抗拒，自己总结了一个涵盖 ppt 内容的类似课本的文档，在附录里

成绩分析单								
类型	名称	分数<60	60≤分数<70	70≤分数<80	80≤分数<90	90≤分数	总人数	排名
课程	高性能计算导论	0	0	12	19	5	36	1
教学班	2023生物信息学	0	0	12	19	5	36	1
生物医学工程学院		95.3	4.5	48	3	专业基础课	必修	正常考试 白分制 1168329

英语 C

复习：  
原题，不过爱课原题一共有 80 道左右，看你们自己有什么方法来记住了，两篇作文记得要按照考研形式写（作文格式）

成绩分析单								
类型	名称	分数<60	60≤分数<70	70≤分数<80	80≤分数<90	90≤分数	总人数	排名
课程	综合高级英语	5	2	13	30	8	2644	128
教学班	模块一C1	5	2	13	30	8	58	1

## 导论

### 前言：

又杂又广，而且会有一种复习了等于白看的感觉（即使你看了好几遍）

课程体验：前期很多理论课参杂实验课，后期都是实验课，而且几乎都是讲几句就动手实操。

### 考试：

考试出的选择题很多是没见过，是你上课听课，课后用他上课 ppt 好好学也见不到的知识点。且 25 年考试改了，往年综合题是从“点阵法，比对，UPGMA 建树”出一个，25 年老师用 AI 出了一个全程考 BLAST 的情景题（25 分）：大概是科学家在哪里发现了什么序列，想找找同源改用什么工具。考了六大基本工具和打分矩阵，最后考了一个很离谱的 BLAST 高级工具，这个工具上课不会讲到，实验课不会实操到，就像是 ppt 上一个知识扩展，但是值 8 分，正常复习的几乎不可能看到这个东西。最后一题是：“生信学啥，你对什么领域感兴趣，这个领域一般怎么用生信技能”。往年是有附加题的，但是 25 年是实打实的 100 分，非常不幸了。不过简答题是原题，后文说明

### 复习：

这门课我用的贺师兄总结文档结合上课 ppt 进行复习，文档几乎把所有 ppt 的重点要点涵盖了，但是就像前面说的，还是有很多是见不到的。

要点：有个公众号里面收录了基础医学的生信考试经验，里面的一些题目是这次考试的原题，可以去看看。我们考的“蛋白质预测方法，一般生信流程，甲基化，GO 数据库构成原理”

### 另话：

不理解为什么这一课考的如此糟糕，虽然实际上也就差第一名四五分，但是写卷子的时候我感觉自我良好，简答题都是公众号的原题我都背到了，不知道为啥考的不太好，不过导论我复习的时间确实很少，那段时间很忙

成绩分析单



类型	名称	分数<60	60≤分数<70	70≤分数<80	80≤分数<90	90≤分数	总人数	排名
课程	生物信息学导论	1	0	5	27	3	36	12
教学班	2023生物信息学	1	0	5	27	3	36	12

## 生信语言

### 提一下：

平时的作业可以多敲敲代码，不要一上来就 AI，特别难的情况可以配合 AI 学习（询问 AI 一些函数的作用，省自己查阅帮助文档时间），而不是直接让他给答案然后 copy，这样十次作业下来，对 python 和 R 的理解基本上比较牢固。

### 考试：

同样，25 年有改动，从 20 道题变为了 60 道题。往年选择题都考的很简单，但是应教务，选择题增加为 50 道，所有难免会出现一两道不会的情况，不过分值不大。简答题很简单，问问输出，问问该用什么函数等。有一个删除字典怎么删怎么合并，R 基础绘图上标题的函数等。编程题 R 是读文件，然后处理数据框，最后画图，有点难度，Python 是读 fa 文件，输出用户想要的 A, B 区间，如果平时作业一套 AI 完成，可能会 fa 文件是什么都不知道。简答题是全英，不过就是问问你用了什么库或包，解决了什么问题，课程建议

### 复习：

对于已经经常敲代码的，不用费很大劲，主要是复习一些平时敲代码也没机会注意的小细节，这些选择题会考，比如删除有缺失值的一行 `na.omit`，字典的值可以是任意数据类型，键不能是可变的等等，R 的选择题很简单，python 多多少少会有一两道不清楚的。这两个我都没用 ppt 复习，因为 ppt 是清朝 ppt，而且涵盖了很多考试我认为不会考的复杂函数和参数和用法（事实上确实没考）。R 的话多敲敲代码，选择题简答题编程题基本没什么问题，python 体量大一点，可以扣细节的地方也多，可以看看贺师兄的总结。我主要是用 AI 和贺师兄的复习文档进行复习。

### 提示：

这个课，python 部分是郭老师和贺老师一起上，但是郭老师出的卷子，贺老师上课的内容他一点都没出。

成绩分析

类型	名称	分数<60	60≤分数<70	70≤分数<80	80≤分数<90	90≤分数	总人数	排名
课程	生物信息学应用语言	1	0	7	15	13	36	2
教学班	2023生物信息学	1	0	7	15	13	36	2

# 统计

## 前言（重要）:

统计很重要，现在是数据横行的时代，做什么都躲不开处理数据，处理数据需要统计知识。无论是生信还是其他领域，都有大量数据需要处理，可以说统计学是以后做科研的基本功

## 课程:

自学很重要吧只能说，老师非常想努力将AI融入课堂，但是我觉得目前还不够成熟，我建议自己去学习一些统计知识，可以通过实操项目等方式，让理论得到实践，这样印象更深刻。总之统计这门课很重要

## 考试:

唉，跟上一届学长总结的一样，考的东西实际上和课上ppt的东西关联不是很大，但是ppt全学懂了也不可能考差说是，后面老师会给一个复习大纲ppt，就相当于给每一章都划重点，然后让我评价一下就是考试内容和此复习大纲所要求内容重合度为63.41%，总之不可能考差，但是仅凭这个也别想的很高。简答题有一个腹黑的题：生物性状或生物界中举例哪些是正态分布，泊松分布，二项分布。其他都算是课上学到的：统计量和参数，随机误差，主成分分析，交互作用四个名词解释。简答题是显著性检验判定方法，方差分析假定条件，卡方分布是什么与适合和拟合。然后抽象的两道计算题都是出错了，第一道Z检验表给错了，第二道回归分析最小二乘法给的数据也是错的，会算出r大于1，而且计算量很大，九位数除于十一位数，都是散乱数据。

## 复习:

怎么说，我感觉我复习的内容大部分都没有考，感觉超额复习了，复习了特别多补考的内容（ppt上）然后考试的时候又比较大意，考的不是很好，所以不留复习思路了

成绩分析单

类型	名称	分数<60	60≤分数<70	70≤分数<80	80≤分数<90	90≤分数	总人数	排名
课程	生物信息学统计方法	0	0	2	21	13	36	5
教学班	2023生物信息学	0	0	2	21	13	36	5

## 高通量

### 前言：

要继续发展生信这条路，高通量这门课很重要，学懂了才算入门生信，是作为在生信学习的基础，比如后续的空转和 single cell 等。贺莎老师很好，特别负责，改卷子很认真，上课很认真。

### 课程：

初期接触会有难度，虽然贺老师会努力的讲，但是对于刚接触专业课的来说，还是会难理解，觉得抽象。除非有些人提前做过一些单细胞数据啥的可能好懂。最主要的是把每节课内容串起来，十章内容之间关联性非常强，一定要串起来学

### 考试：

我觉得我们这一届试卷出得比较难，考了二代测序三代测序特点我手快写错了，简答题：外显子章节的贝叶斯公式（默公式，写后验先验和得分公式），Kmer 长度影响，BS-seq 要做的处理，BSseq 和 RNAseq 数据处理异同，宏基因的稀疏分析，BQSR 原理，GATK 流程

综合题：BS-seq 统计方法（fisher 和逻辑）统计的零假设，GATK 流程，通路富集，VQSR 原理（属于回归和分类），单细胞通路富集分析，DAVID（fisher）方法，轨迹分析

### 复习：

老 7 套复习资料挺有用，没错就是你手上那一堆资料。但是最主要的是理解，理解很重要，我感觉贺老师这一届出卷难在侧重理解了，根据往年口径来说，简答题几乎背流程就全写完，但是这一次几乎都是要理解才答的出来，没有大量直接问你某个技术的流程这样的题了，不知道以后会怎么样，不过其实有一份资料把全部上课 ppt 整合起来了的。

贺莎老师每一节课，每一节课都会边上课边暗示考试点，所以上课认真听有好处，可以在他说要考的地方 ppt 做点备注日后用

成绩分析单

类型	名称	分数<60	60≤分数<70	70≤分数<80	80≤分数<90	90≤分数	总人数	排名
课程	高通量测序数据分析	1	0	13	11	11	36	1
教学班	2023生物信息学	1	0	13	11	11	36	1

附件：高性能课本 part1，高性能课本 part2

Contact: welon0725@outlook.com

Last update: January 25, 2025