

统计课程表

<p>第一天 上午</p>	<p>数据分析前的准备工作——数据清洗与探索</p>	<ul style="list-style-type: none"> ➤ 如何快速查找缺失数据？ ➤ 数据缺失比较多怎么办，只能删掉吗？——介绍几种常见的处理缺失值的方法 ➤ 我想对缺失的数据进行填补，可以用什么办法？——缺失值填补技术及其实现 ➤ 如何快速发现数据中的异常值？ ➤ 利用JMP软件快速定位、隐藏、排除异常值 ➤ 发现异常值后，我可以直接把它们删除掉吗？——介绍处理异常值的几种方法 ➤ 数据的可视化探索——JMP软件教你不一样的数据展示方式 ➤ 利用JMP软件的绘图功能，真正实现人与图形的交互，想到即可见（原来我以前的做图都弱爆了） ➤ 利用JMP软件的绘图功能，绘制会动的动态图，让你的报告更精彩 ➤ 利用JMP的制表功能，让你快速生成你想要的任意表格形式 ➤ 利用JMP软件交互、动态地探索、展示数据，让你快速发现数据之间的关联，迅速找到研究方向
<p>第一天 下午</p>	<p>常用基础统计分析方法</p>	<ul style="list-style-type: none"> ➤ 为什么我的文章中采用t检验，结果频频被退回修改（t检验真的可以“踢”遍全世界？） ➤ 卡方检验可以用于所有“率”的比较？ ➤ 等级资料到底是用卡方检验还是秩和检验？ ➤ 一张图教你看懂基础统计分析方法的选择 ➤ 发现一种新的治疗方法，可以用t检验说明它跟传统治疗方法疗效一样吗？（什么情况下用t检验，什么情况下用等效性检验） ➤ 方差分析后的两两比较方法这么多，到底该选择哪个？ ➤ 如何从成千上万的基因中筛选出有意义的基因（利用FDR法对P值进行校正） ➤ 发现了一个新基因，如何评价它对疾病诊断有多大的价值（诊断试验的ROC曲线分析） ➤ 如何寻找诊断指标的cut-off值？ ➤ 两种影像诊断技术，如何评价它们的诊断结果一致性如何（Kappa一致性系数的计算及比较） ➤ 两个指标，到底要看相关性，还是一致性？（相关与一致性的区别）

<p>第二天 上午</p>	<p>线性回归与 logistic回归</p>	<p>线性回归分析思路</p> <ul style="list-style-type: none"> ➤ 为什么有的因素我在从专业认为应该有影响，但是回归分析却总是检验不出来（线性回归的多因素分析思路） ➤ 为什么有的因素在单因素分析中有统计学意义，多因素分析却无意义？ ➤ 逐步回归到底可不可靠？为什么统计学家不建议采用逐步回 ➤ 危险因素筛选中，可以同时有分类资料和连续资料吗？（一般线性模型的分析） ➤ 血压值是一直随着年龄增加而升高吗？还是中年升高、老年又降低？（如何探索两个变量之间的真实变化关系） <p>Logistic回归分析思路</p> <ul style="list-style-type: none"> ➤ 为什么有的指标在单因素分析中有意义，而多因素分析没有意义，我到底该选择哪个？（解释单因素分析与多因素分析的差 ➤ 我知道了一个初产年龄和流产次数，可以预测一个女性发生乳腺增生的概率吗（利用logistic回归预测疾病发生情况） ➤ 我想在报告中把预测结果展示的炫酷一些，怎么办呢？（绘制分类树，可视化展示预测分类）
<p>第二天 下午</p>	<p>生存资料分析 与重复测量分析</p>	<p>生存资料分析</p> <ul style="list-style-type: none"> ➤ 为什么我觉得两条生存曲线差别挺大，但是log-rank检验却无统计学意义？（生存资料组间比较的方法选择） ➤ 我可以预测一个人的死亡时间吗？可以预测一个人1年后死亡概率有多大吗？（利用Weibull模型预测死亡发生率） ➤ 我有一堆疾病发生时间的数据，只能列个频数表？（利用参数模型将手中已有数据变废为宝） ➤ 如何分析癌症死亡的危险因素到底有哪些？（肿瘤分析中的Cox回归多因素分析） <p>重复测量数据</p> <ul style="list-style-type: none"> ➤ 对患者在治疗后观察了4个周，我可以在每个周分别用t检验比较吗？（目前对重复测量数据的常见误用方法） ➤ 我想了解患者服药后4个周的疗效变化趋势是什么样的，可以做到吗？（采用重复测量方差分析探索随时间变化的趋势） ➤ 不同治疗的两组，观察了4次，发现A组变化慢，B组变化快。怎么发现两组变化趋势是不是一致呢？（采用重复测量方差分析探索两组变化趋势是否相同）