



Agile Development

A++

敏捷开发小手册



宋世傑 (Edmond S.K. SUNG) 彭国明



# A++ 敏捷开发小手册

— 从个人提升到以数据说话







# Contents

|           |  |           |
|-----------|--|-----------|
| <b>1</b>  | <b>如何改善敏捷开发质量</b>                      | <b>17</b> |
|           | 某软件开发公司敏捷开发过程改进案例 . . . . .            | 17        |
|           | 如何改善 . . . . .                         | 19        |
|           | 附件 . . . . .                           | 26        |
|           | XP . . . . .                           | 26        |
|           | 乔布斯谈 Dr.Juran . . . . .                | 28        |
| <br>      |  |           |
| <b>I</b>  | <b>个人提升</b>                            | <b>29</b> |
| <br>      |  |           |
| <b>2</b>  | <b>如何创新</b>                            | <b>33</b> |
|           | 创新 Creativity . . . . .                | 34        |
|           | 创造的精神 Spirit of Creating . . . . .     | 34        |
|           | 目标怎样设? 什么样的目标才算理想? . . . . .           | 37        |
|           | FRITZ 把这改变过程简单分为三步 . . . . .           | 39        |
|           | References . . . . .                   | 40        |
| <br>      |  |           |
| <b>3</b>  | <b>克服拖延症</b>                           | <b>41</b> |
|           | References . . . . .                   | 45        |
| <br>      |  |           |
| <b>II</b> | <b>自我管理开发过程</b>                        | <b>47</b> |
| <br>      |  |           |
| <b>4</b>  | <b>三点估算</b>                            | <b>51</b> |
|           | 估算是一个范围, 不是一个数 . . . . .               | 51        |
|           | 从单点到三点估算 . . . . .                     | 52        |
|           | 总结 + 解读分析结果 . . . . .                  | 53        |
|           | 利用蒙特卡洛模拟 10 个步骤 (三角形分布) 的总分布 . . . . . | 53        |
|           | 分析 10 步骤模拟结果 . . . . .                 | 55        |
|           | 附件 . . . . .                           | 55        |
|           | 蒙特卡洛 (Monte Carlo) 模拟 . . . . .        | 55        |
| <br>      |  |           |
| <b>5</b>  | <b>个人量化管理</b>                          | <b>57</b> |
|           | 如何降低项目进度偏差 . . . . .                   | 57        |

|  |            |
|--|------------|
| 从量化培训开始 . . . . .  | 58         |
| 互动培训 . . . . .   | 58         |
| 从估算、策划到监控 . . . . .  | 59         |
| 总结 . . . . .   | 62         |
| 应该如何改善 . . . . .   | 63         |
| 附件 . . . . .   | 65         |
| 挣值分析法 (Earned Value) . . . . .                                     | 65         |
| Tips: 如何简单记录工时 . . . . .   | 66         |
| PSP - 个体软件过程 Personal Software Process 简介 . . . . .                | 66         |
| References . . . . .   | 68         |
| <b>6 估算软件规模 . . . . .</b>  | <b>69</b>  |
| 为什么要估规模 . . . . .  | 69         |
| 为什么不应用代码行数 (LOC) . . . . .   | 69         |
| 怎样估 SiFP . . . . .   | 71         |
| 从故事点转简化功能点 . . . . .   | 72         |
| 附件 . . . . .   | 74         |
| 简化功能点 (SiFP) 简介 . . . . .  | 74         |
| 逻辑文件 Logical File . . . . .  | 76         |
| 基本过程 Elementary Process . . . . .                                  | 76         |
| 实例: 识别基本过程 (EP elementary process) . . . . .                       | 77         |
| 1. 潜水学校: 开发项目 . . . . .  | 78         |
| 例子一: 答案与解读 . . . . .   | 80         |
| 2. 潜水学校:FEM 项目 . . . . .   | 83         |
| 例子二: 答案与解读 . . . . .   | 84         |
| 与国际功能点 (IFPUG) 的偏差 . . . . .                                       | 87         |
| <b>7 估算工作量 . . . . .</b>   | <b>89</b>  |
| 估算: 靠模型, 还是靠专家? (Model-based Vs Expert-based estimation) . . . . . | 89         |
| 为什么估算开发工作量/工期这么困难? . . . . .                                       | 91         |
| 估算软件开发的困难 . . . . .  | 91         |
| 附件 . . . . .   | 94         |
| 估算砌乐高积木时长 - 2015 年学生实验数据 . . . . .                                 | 94         |
| References . . . . .   | 99         |
| <b>III 利用评估开始过程改进 . . . . .</b>                                    | <b>101</b> |
| <b>8 获取高层支持 . . . . .</b>  | <b>105</b> |
| 案例 . . . . .   | 105        |
| 第一版方案书 . . . . .   | 106        |
| 第二版与中层确认, 然后向高层汇报 . . . . .  | 106        |
| <b>9 寻找改进机会 . . . . .</b>  | <b>111</b> |
| 评估中访谈公司培训专员 . . . . .  | 112        |



|           |     |
|-----------|-----|
| 总结访谈的问题发现 | 112 |
| 管理层的理解与支持 | 114 |
| 总结        | 114 |
| 附件        | 115 |
| CMMI 评估案例 | 115 |

## IV 从做好迭代回顾开始 117

|                                  |            |
|----------------------------------|------------|
| <b>10 做好准备</b>                   | <b>121</b> |
| 缺陷排除率 (DRE)                      | 122        |
| 自动化统计分析                          | 124        |
| 附件                               | 127        |
| 质量成本 COQ (Cost of Quality)       | 127        |
| Brehm 1956 决策影响实验                | 127        |
| 粮食分配实验                           | 127        |
| 开发项目工作量 (成本) 分布                  | 128        |
| Reference                        | 128        |
| <b>11 根因分析</b>                   | <b>131</b> |
| Esquire QC 圈                     | 131        |
| 用二八原则 (Pareto 图) 识别主要源头          | 131        |
| 建立目标                             | 134        |
| 根因分析 Root Cause analysis (鱼骨图分析) | 135        |
| 改进措施                             | 137        |
| 改进效果                             | 138        |
| A3 报告                            | 139        |
| 技术总监经验之谈                         | 140        |
| 总结                               | 140        |
| 附件                               | 141        |
| 5 Why 例子                         | 141        |
| FMEA 实例                          | 141        |
| <b>12 如何做好迭代回顾</b>               | <b>145</b> |
| 从项目冲刺的回顾复盘开始                     | 145        |
| 回顾流程                             | 147        |
| 怎样开始                             | 148        |
| 常见问题                             | 149        |
| 客户反馈                             | 150        |
| 总结                               | 150        |
| 附件                               | 150        |
| 游戏 1: 应与否                        | 150        |
| 游戏 2: 时间表                        | 152        |
| 游戏 3: 颜色点                        | 153        |
| Asch 1951 群众压力实验                 | 154        |
| KJ 分析方法步骤                        | 156        |

|  |                |
|--|----------------|
| References . . . . .                       | 156            |
| <b>13 从定性到定量</b>                           | <b>157</b>     |
| 使用缺陷排除率，配合蒙特卡洛预测模型做定量分析 . . . . .          | 157            |
| 怎样开始 . . . . .                             | 159            |
| 某公司经过培训后第一轮回顾例子 . . . . .                  | 161            |
| 总结 . . . . .                               | 165            |
| 附件 . . . . .                               | 166            |
| 水晶球蒙特卡洛预测模型 . . . . .                      | 166            |
| <br><b>V 基础：代码质量</b>                       | <br><b>171</b> |
| <b>14 测试驱动开发</b>                           | <b>175</b>     |
| 为什么我们要测试驱动开发 . . . . .                     | 175            |
| 每小步验证 . . . . .                            | 176            |
| 用 TDD 开发程序 . . . . .                       | 177            |
| 单元测试与 TDD 的好处 . . . . .                    | 179            |
| Reference: . . . . .                       | 179            |
| <b>15 代码重构</b>                             | <b>181</b>     |
| 如何学好 SOLID，打好基础做重构 . . . . .               | 181            |
| 如何提升学员的兴趣与动力 . . . . .                     | 181            |
| 什么时候做重构 . . . . .                          | 182            |
| 代码规范 . . . . .                             | 183            |
| 持续集成 . . . . .                             | 183            |
| 附件 . . . . .                               | 184            |
| 小孩利用乌龟玩几何游戏 (Turtle Geometry) 例子 . . . . . | 184            |
| 温伯格的教学实验 . . . . .                         | 187            |
| References . . . . .                       | 187            |
| <b>16 持续集成</b>                             | <b>189</b>     |
| 验收测试 . . . . .                             | 189            |
| 测试自动化 . . . . .                            | 190            |
| 持续集成 Continuous Integration . . . . .      | 191            |
| 团队协作 . . . . .                             | 192            |
| 持续交付 Continuous Delivery . . . . .         | 193            |
| 附件 . . . . .                               | 193            |
| 持续集成步骤 . . . . .                           | 193            |
| 技术债务例子 . . . . .                           | 194            |
| References . . . . .                       | 195            |
| <b>17 结对编程与同行评审</b>                        | <b>197</b>     |
| 培训 . . . . .                               | 198            |
| 总结 . . . . .                               | 200            |
| 附件 . . . . .                               | 201            |

|   |            |
|---|------------|
| CONTENTS  | 11         |
| 代码可读性例子 . . . . .                                   | 201        |
| <b>18 软件需求</b>                                      | <b>203</b> |
| 附件 . . . . .  | 205        |
| 英国煤矿生产问题的研究 . . . . .                               | 205        |
| <b>19 客户参与</b>                                      | <b>211</b> |
| 常见问题 . . . . .                                      | 211        |
| 识别利益相关者 . . . . .                                   | 212        |
| 用户故事卡 . . . . .                                     | 212        |
| 用例与场景 . . . . .                                     | 213        |
| 举例 . . . . .  | 215        |
| 附件 . . . . .  | 216        |
| 利益相关者计划检查单 (Stakeholders Plan Checklist): . . . . . | 216        |
| 客户声音: Kano Diagram . . . . .                        | 219        |
| References . . . . .                                | 219        |
| <b>20 团队协作</b>                                      | <b>221</b> |
| 各自负责自己开发的模块 . . . . .                               | 222        |
| 团队合作 . . . . .                                      | 223        |
| 按时上下班 . . . . .                                     | 223        |
| 什么导致系统崩溃 . . . . .                                  | 224        |
| 技术团队的持续性 . . . . .                                  | 224        |
| 驱动团队改进 . . . . .                                    | 225        |
| 保持团队稳定 . . . . .                                    | 226        |
| 附件 . . . . .  | 227        |
| 平衡心态 Cognitive Dissonance . . . . .                 | 227        |
| References . . . . .                                | 227        |
| <b>VI 度量与分析</b>                                     | <b>229</b> |
| <b>21 做问卷调查</b>                                     | <b>233</b> |
| 案例: 为某全国快餐连锁做问卷调查 . . . . .                         | 233        |
| 总结 . . . . .  | 235        |
| <b>22 教小孩学统计分析</b>                                  | <b>237</b> |
| 从考试成绩到贸易逆差 . . . . .                                | 237        |
| 分析例子: 连续与分组数据 . . . . .                             | 242        |
| 总结 . . . . .  | 244        |
| 附件 . . . . .  | 244        |
| 如何画茎叶图 (Stem and leaf plot) . . . . .               | 244        |
| References . . . . .                                | 245        |
| <b>23 假设检验</b>                                      | <b>247</b> |
| 假设检验的步骤 . . . . .                                   | 248        |
| 检验单个正态总体的均值 . . . . .                               | 248        |

|                                     |            |
|-------------------------------------|------------|
| 比较两个正态总体的均值 . . . . .               | 251        |
| 检验单个正态总体的均值, 方差未知 . . . . .         | 252        |
| 方差分析 (ANOVA test) . . . . .         | 253        |
| 分组 (分类) 数据 . . . . .                | 256        |
| 附件 . . . . .                        | 257        |
| 检验单个正态整体均值相关方程式 . . . . .           | 257        |
| 比较两个正态总体的均值 . . . . .               | 258        |
| 检验单个正态总体的均值, 方差未知 . . . . .         | 258        |
| Tukey-Kramer 相关方程式 . . . . .        | 258        |
| <b>24 预测模型 . . . . .</b>            | <b>261</b> |
| 芬兰某商业银行的软件维护数据统计分析 . . . . .        | 261        |
| 总结报告 . . . . .                      | 261        |
| 数据分析步骤 . . . . .                    | 263        |
| 回顾与总结 . . . . .                     | 270        |
| 附件 . . . . .                        | 271        |
| 案例部分变量列表 . . . . .                  | 271        |
| X 与 Y 相关性图 . . . . .                | 272        |
| 利用 Wrapper 降低模型的变量 . . . . .        | 274        |
| 如何做回归分析 . . . . .                   | 276        |
| References . . . . .                | 281        |
| <b>25 控制图 . . . . .</b>             | <b>283</b> |
| 建立基线 . . . . .                      | 284        |
| 噪音还是信号 . . . . .                    | 284        |
| 控制图的应用 . . . . .                    | 287        |
| 总结 . . . . .                        | 288        |
| 附件 . . . . .                        | 289        |
| 如何画 ImR 控制图 . . . . .               | 289        |
| 各种控制图 . . . . .                     | 290        |
| 过程能力 (Process Capability) . . . . . | 290        |
| 细分例子 . . . . .                      | 291        |
| References . . . . .                | 292        |
| <b>VII 总结 . . . . .</b>             | <b>293</b> |
| <b>26 领导的工作 . . . . .</b>           | <b>297</b> |
| 我们很注重质量与客户满意度 . . . . .             | 297        |
| 拉吉夫终生受用的一件事 . . . . .               | 299        |
| 团队根因分析 . . . . .                    | 302        |
| 从定性提升到定量管理 . . . . .                | 302        |
| <b>27 东北虎 . . . . .</b>             | <b>305</b> |
| 结语 (Epilogue) . . . . .             | 305        |

## Part I

# 个人提升



自我改善效率基本功





## Part II

# 自我管理开发过程



软件开发工程师自我改进过程



## Part III

# 利用评估开始过程改进



团队必须先获得管理层的关注与支持

## Part IV

# 从做好迭代回顾开始





做好迭代回顾让团队向量化管理迈进一大步



## Part V

### 基础：代码质量



这部分参照 Jeffery 先生总结极限编程 (XP) 的三层图 (在第一章里), 利用实例, 先从内层的测试驱动开发 (TDD), 重构 (Refactoring), 与结对编程 (PairProgramming) 等核心最佳实践开始, 然后到中层的持续集成 (Continuous-Integration), 最后是外层的团队各角色协作/互补 (Whole Team) 和如何做好需求分析等。



## Part VI

# 度量与分析





当团队开始养成每天收集数据的习惯，在迭代回顾时分析数据，公司便可以开始准备度量与分析。

(注: 要做好度量与分析，跟过程改进一样，要针对目标，制定度量计划。) 本部分开头会以问卷调查为例，介绍度量计划的主要元素，

然后会介绍统计分析，控制图，假设检验等数据分析的基本功。

最后利用某银行软件维护数据分析案例，简单介绍数据收集后，做分析的主要步骤。

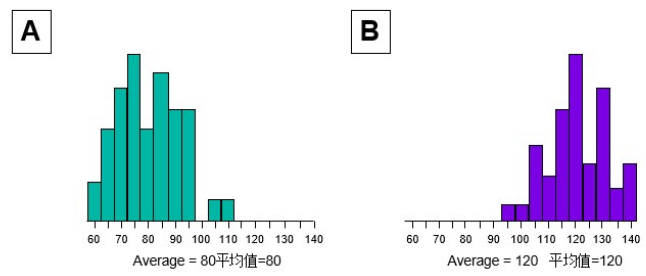
注：请不要误以为度量与分析必须由管理层驱动。若要团队能从定性升到定量管理，并能持续，而且有效果，必须从团队开始，而不是靠管理者总体策划



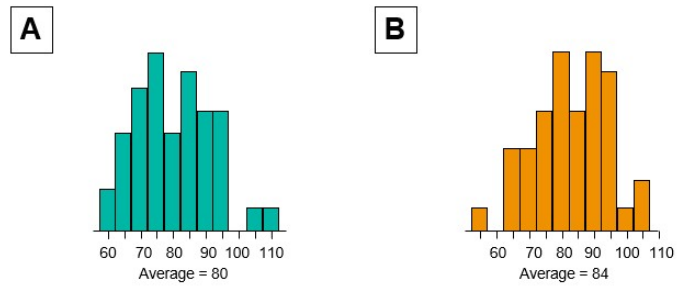
# Chapter 23

## 假设检验

上章介绍了双样本 T 检验，是假设检验的一种；下面介绍假设检验，及它帮我们解决什么问题。



想比较 A B 两组实验结果是否有显著差异。如果像上图应该看到 B 很明显比 A 高。  
但如果像下图，我就很难看得出来：



假设检验利用统计分析方法帮我们区分 - 两组数据是否有显著的差异。

### 假设检验的步骤

零假设: A 与 B 没有显著差异

备选假设: A 与 B 有显著差异

挑选对应的检验方法后，

可以利用工具计算出的 P 值来判断：

如果 P 值大于 0.05，就不能拒绝零假设

如果 P 值小于 0.05，就可以拒绝

因 A 与 B 都包含非常多样本数据，我们只可以利用 A 的随机抽样，和 B 的随机抽样数值，来判断 A 与 B 是否有显著差异，这便是双样本 T 检验。

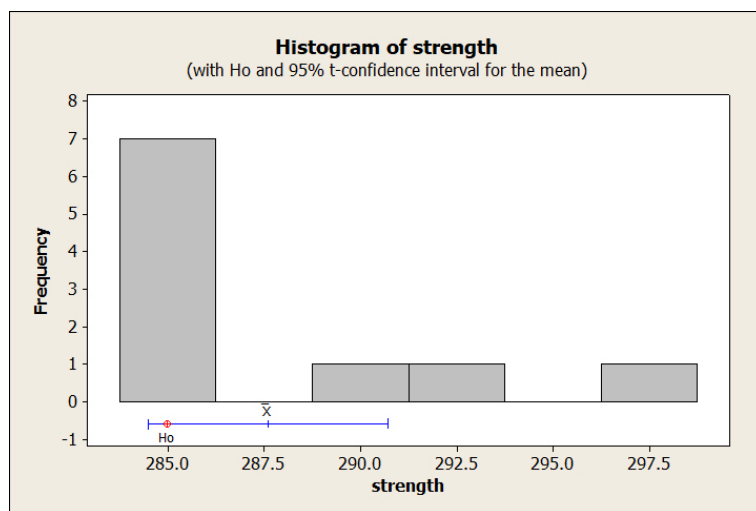
要深入了解便要与数据，先举个比较简单的假设检验实例：

### 检验单个正态总体的均值

按历史数据，某厂生产的钢丝折断力是正态分布（平均值 = 285，标准差 = 4），更换了新的钢材供货商，随机抽取以下 10 样本，想判断更换供货商后折断力和原先是否有显著提升？

289, 286, 285, 284, 286, 285, 285, 286, 298, 292

从下面样本钢丝折断力柱状图看到 10 个样本中有七个接近 285，有 3 个是明显大于 285，应如何判断是否有显著差异呢？



按假设检验步骤：

1. 零假设 \*：新钢材的折断力小于或等于 285 ( $H_0: \mu \leq 285$ )



要了解 P 值是什么？为什么低于 0.05 可以拒绝？首先要理解的会发生两类错误

检验可能犯错误，所谓犯错误就是检验的结论与实际情况不符，有两种情况：

1. 实际情况是  $H_0$  成立，而检验的结果表明  $H_0$  不成立，即拒绝了  $H_0$ ，这时称该检验犯了第一类错误 (Type I error) 或“弃真”错误
2. 实际情况是  $H_0$  不成立， $H_1$  成立，而检验的结果表明  $H_0$  成立，即接受了  $H_0$ ，这时称该检验犯了第二类错误 (Type II error)，或称“取伪”错误

If the decision is:  
估计是:

|       |  |  |
|-------|--|--|
|       | $H_0$                                  | $H_a$                                  |
| $H_0$ | Right Decision<br>✓                    | $\alpha$ Risk<br>Type I Error<br>第一类错误 |
| $H_a$ | $\beta$ Risk<br>Type II Error<br>第二类错误 | Right Decision<br>✓                    |

If the correct answer is:  
实际是:

这两类错误  $\alpha$  与  $\beta$  之间是有关系：

如果  $\alpha$  风险（概率从本来 0.05 降低）下降，便会引起  $\beta$  风险提高，所以不能同时降低  $\alpha$  与  $\beta$ 。 $\alpha$  不能无限降低，一般选  $\alpha = 0.05$ 。

|       |               |                |
|-------|---------------|----------------|
|       | $H_0$         | $H_a$          |
| $H_0$ | ✓             | 第一类错误 $\alpha$ |
| $H_a$ | 第二类错误 $\beta$ | ✓              |

我们可以简单理解 P 值为发生第一类错误的概率，如果低于 5% 我们就有信心拒绝零假设。

假设检验用于比较两组数据有没有显著差异的例子；

### 比较两个正态总体的均值

一名研究人员假设，大学为男生提供的运动项目的平均数量大于大学为女生提供的运动项目的平均数量。对大学为男生和女生提供的体育项目的随机抽样显示，设定临界概率为 5%（Alpha  $\alpha=0.05$ ），是否有足够的证据支持这种说法？（假设方差都一样， $\sigma = 3.3$ ）。

| 男生 Males |    |    |    |    | 女生 Females |    |    |    |    |
|----------|----|----|----|----|------------|----|----|----|----|
| 6        | 11 | 11 | 8  | 15 | 6          | 8  | 11 | 13 | 8  |
| 6        | 14 | 8  | 12 | 18 | 7          | 5  | 13 | 14 | 6  |
| 6        | 9  | 5  | 6  | 9  | 6          | 5  | 5  | 7  | 6  |
| 6        | 9  | 18 | 7  | 6  | 16         | 10 | 7  | 8  | 5  |
| 15       | 6  | 11 | 5  | 5  | 16         | 10 | 7  | 8  | 5  |
| 9        | 9  | 5  | 5  | 8  | 7          | 5  | 5  | 6  | 5  |
| 8        | 9  | 6  | 11 | 6  | 9          | 18 | 13 | 7  | 10 |
| 9        | 5  | 11 | 5  | 8  | 7          | 8  | 5  | 7  | 6  |
| 7        | 7  | 5  | 10 | 7  | 11         | 4  | 6  | 8  | 7  |
| 10       | 7  | 10 | 8  | 11 | 14         | 12 | 5  | 8  | 5  |

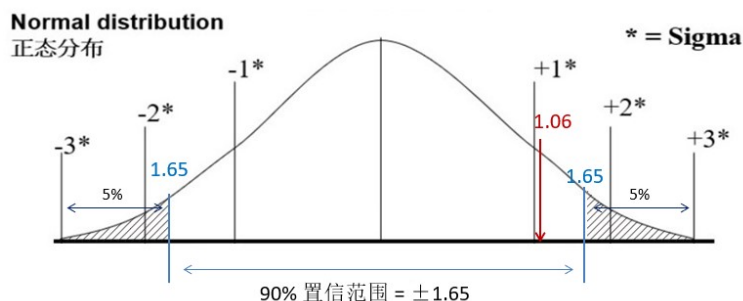
1. 零假设：男女运动项目平均数量没有区别（ $H_0: \mu_1 \leq \mu_2$ ）
2. 备选假设：男生运动项目平均数量比女生多（ $H_1: \mu_1 > \mu_2$ ）

男生的均值是 8.6；女生是 7.9；相差是 0.7

假定使用 90% 置信区间：

正态分布  $\pm 1.65\sigma$  (标准差) 内的面积是总面积的 90%，从下图看到 1.06 落在 90% 置信区间之内，所以不能拒绝零假设。





(有关计算 1.06 的方程式，详见附件)

使用工具跑出来的 P 值是 0.146，大于 0.05，所以不能拒绝零假设

以上两个例子都是假定：分布的方差已知

但很多时，我们不一定知道分布的方差或标准差是多少？我们便需要从样本数估计分布的方差（或标准差）。下面是一个结对 T 检验 (Paired T test) 例子

### 检验单个正态总体的均值, 方差未知

营养师正在观察如果每天饮食都添加某类矿物质，是否会改变人的胆固醇水平。对随机选取 6 位对象在实验之前进行了水平测试。然后对他们进行为期 6 周的测试，在它们大食物里添加矿物质补充，实验之后再测试，结果显示在下表 (胆固醇水平以每分升毫克为单位)。能否得出结论，在  $\alpha = 0.1$  时，胆固醇水平是否发生了变化。

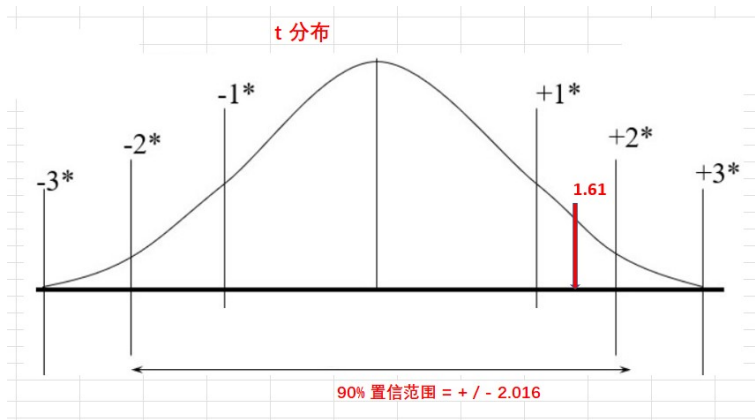
| 对象 Subject | 1   | 2   | 3   | 4   | 5   | 6   |
|------------|-----|-----|-----|-----|-----|-----|
| 前( $X_1$ ) | 210 | 235 | 208 | 190 | 172 | 244 |
| 后( $X_2$ ) | 190 | 170 | 210 | 188 | 173 | 228 |

1. 零假设：没有变化 ( $H_0: \mu_D = 0$ )
2. 备选假设：有显著变化 ( $H_1: \mu_D \neq 0$ )

注意：与上面大学男女生的各 50 个数据成一组，两组 50 个数之间没有对应关系

这例子不一样，因前后是一一对应，我们需要用结对 T 检验 (Paired T test)。因标准差不是已知，是从样本估算，所以要用 T 检验，非 Z 检验。

可以使用相关方程式，计算出 T 值  $= 1.61$ ，



从上图看到 1.61 落在 90% 置信区间 (2.016) 之内，所以不能拒绝零假设。

也可以利用工具，计算双样本 T 检验 (Paired T test) 的 P 值，因 P 值 = 0.165，大于 0.05，判断不能拒绝零假设

如果想比较超过两组数据便要用方差分析 (ANOVA)

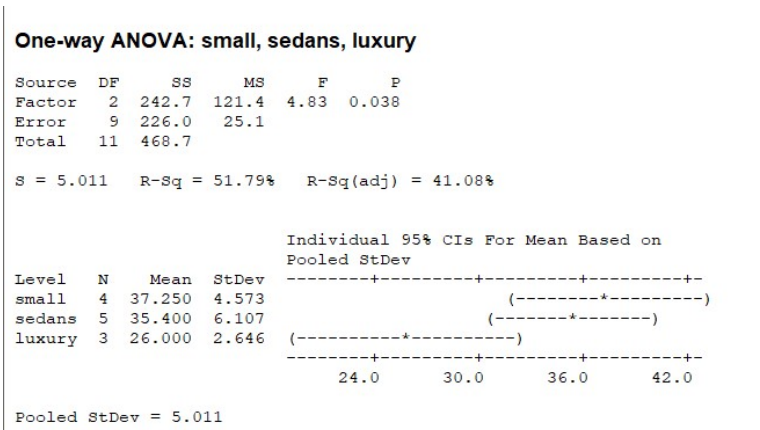
## 方差分析 (ANOVA test)

例子一

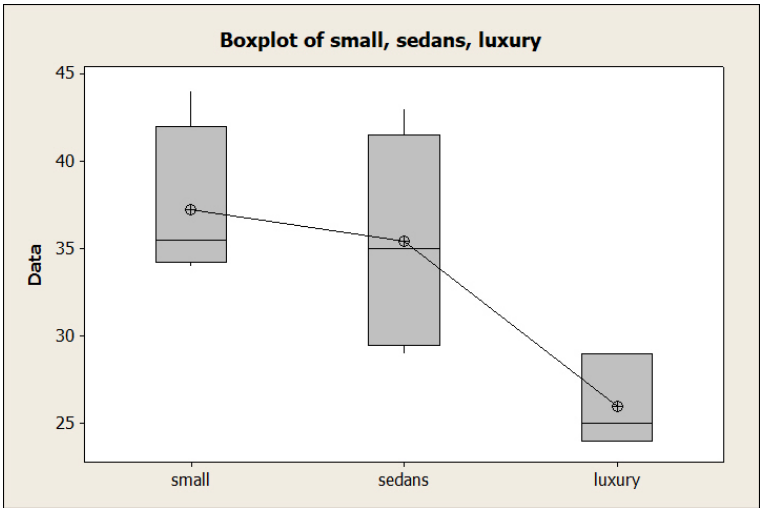
比较三种车型（小轿车，舒适型，豪华车）的省油系数有没有显著差异？

| Worksheet 5 *** |       |        |        |  |
|-----------------|-------|--------|--------|--|
| ↓               | C1    | C2     | C3     |  |
|                 | small | sedans | luxury |  |
| 1               | 36    | 43     | 29     |  |
| 2               | 44    | 35     | 25     |  |
| 3               | 34    | 30     | 24     |  |
| 4               | 35    | 29     |        |  |
| 5               |       | 40     |        |  |
| 6               |       |        |        |  |

用统计工具跑单因子方差分析得出 P 值是 0.0385，低于 0.05，所以我们便可以拒绝零假设



从下面工具的简单画图的也看到豪华型的系数范围比其他两种有显著差异：



方差分析只能识别有没有显著差异，但若若要明确哪些地方有差异，便要用 Tukey-Kramer 做多组之间比较

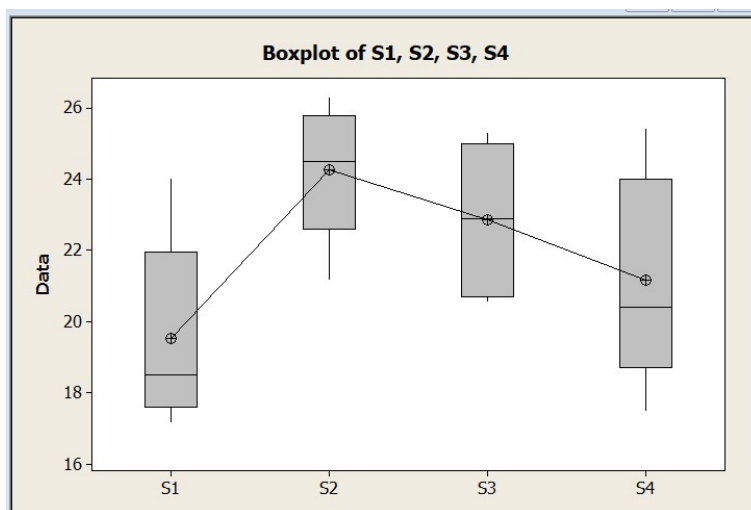
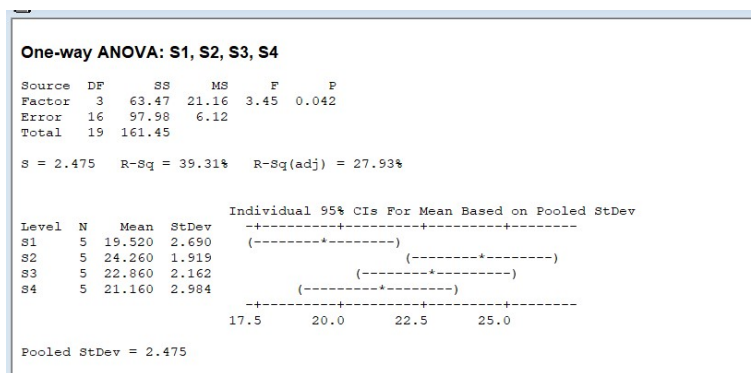
方差分析 2 (ANOVA test)

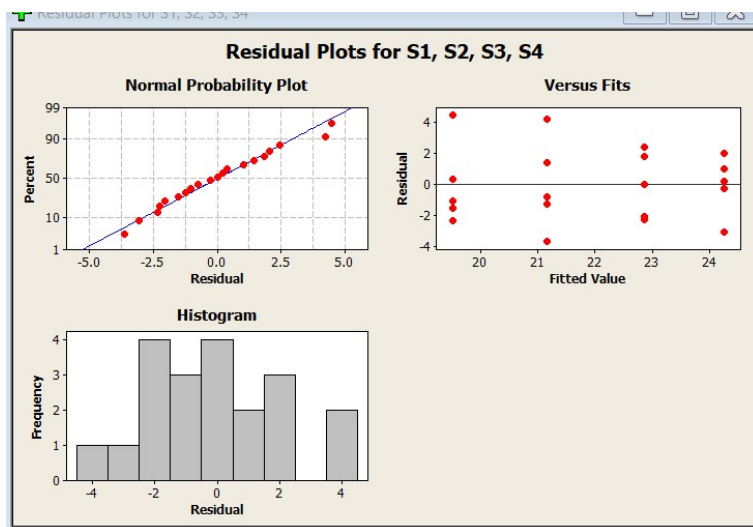
某降落伞工厂想比较 4 家布料供货商的布料拉力（因拉力不足会影响使用者的性命安全）  
下面是它们随机抽样的结果：

| 布料供货商(Supplier)                 | 1     | 2     | 3     | 4     |
|---------------------------------|-------|-------|-------|-------|
| 随机样本拉力                          | 18.5  | 26.3  | 20.6  | 25.4  |
|                                 | 24.0  | 25.3  | 25.2  | 19.9  |
|                                 | 17.2  | 24.0  | 20.8  | 22.6  |
|                                 | 19.9  | 21.2  | 24.7  | 17.5  |
|                                 | 18.0  | 24.5  | 22.9  | 20.4  |
| 样本均值 Sample Mean                | 19.52 | 24.26 | 22.84 | 21.16 |
| 样本标准差 Sample Standard Deviation | 2.69  | 1.92  | 2.13  | 2.98  |

用工具跑方差分析 P 值是 0.042，少于 0.05，判断拒绝零假设

- Minitab 出来方差分析结果表，与上面的总结表能一一对应；





- 从 4 组的分布，你可能看出第一组与第二组的差异很明显，其他差不多。但当差异没有这么明显，便需要用统计分析判断。

### Tukey-Kramer 多组比较

实例（比较 4 供货商布料的拉力）：

$$\#|\bar{X}_1 - \bar{X}_2| = |19.52 - 24.26| = 4.74$$

1.  $|\bar{X}_1 - \bar{X}_3| = |19.52 - 22.84| = 3.32$
2.  $|\bar{X}_1 - \bar{X}_4| = |19.52 - 21.16| = 1.64$
3.  $|\bar{X}_2 - \bar{X}_3| = |24.26 - 22.84| = 1.42$
4.  $|\bar{X}_2 - \bar{X}_4| = |24.26 - 21.16| = 3.10$
5.  $|\bar{X}_3 - \bar{X}_4| = |22.84 - 21.16| = 1.68$

用 Tukey-Kramer 相关方程式（详见附件），得出临界范围为 4.47，所以判断唯一供货商 1 与 2 的布料拉力是有显著差异。

### 分组（分类）数据

上面都是用于连续数据的例子

假设检验也可以用于分组（分类）数据

### 卡方检验 (Chi Square test)

想研究三家医院针在三类病人感染的主因:（外科手术 Surgical Site，肺炎 Pneumonia，血液 Bloodstream）最终导致病人死亡

是否有显著差异？

|   | C1-T     | C2           | C3        | C4          |
|---|----------|--------------|-----------|-------------|
|   | hospital | SurgicalSite | Pneumonia | Bloodstream |
| 1 | A        | 41           | 27        | 51          |
| 2 | B        | 35           | 3         | 40          |
| 3 | C        | 169          | 106       | 109         |

1. 零假设：那家医院与感染主因没有关系
2. 备选假设：有显著关系

卡方检验得出 P 值等于零，所以可以拒绝零假设，医院与感染主因有关系

Chi-Square Test: SurgicalSite, Pneumonia, Bloodstream

Expected counts are printed below observed counts  
Chi-square contributions are printed below expected counts

|       | SurgicalSite | Pneumonia | Bloodstream | Total |
|-------|--------------|-----------|-------------|-------|
| 1     | 41           | 27        | 51          | 119   |
|       | 50.45        | 27.06     | 40.56       |       |
|       | 1.680        | 0.026     | 2.459       |       |
| 2     | 35           | 3         | 40          | 78    |
|       | 32.09        | 18.24     | 26.65       |       |
|       | 0.125        | 12.751    | 6.440       |       |
| 3     | 169          | 106       | 109         | 384   |
|       | 161.55       | 89.89     | 132.19      |       |
|       | 0.309        | 2.889     | 4.067       |       |
| Total | 245          | 136       | 200         | 581   |

Chi-Sq = 30.755, DF = 4, P-Value = 0.000

- 1: 零假设: 医院 (Hospitals) 与感染种类 (Infections) 之间没有相关
- 2: 为了计算卡方值 (Chi-square), 首先使用以下公式计算每个列联表的预计值 E(Expected Value),

$$E = \frac{(\text{row sum}) * (\text{column sum})}{\text{grand total}}$$

再用下面公式从预计值与实际值 O(Observed value), 计算卡方值:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

- 3: 计算得出卡方值 = 30.696 (注), 比关键值 9.5 高

(对应 DF=4 (= (3-1)\*(3-1)),  $\alpha = 0.05$ , 从  $\chi^2$  统计表得出 9.5),

所以拒绝零假设, 医院 (Hospitals) 与感染种类 (Infections) 之间相关。

(注: 对应 Minitab 工具自动算出的 30.755)

## 附件

### 检验单个正态整体均值相关方程式

平均值分布的方 = 样本方差 / n, 并可以用以下方程式估算 (如果样本数 n 越大, 样本均值的方差就越少):

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

$$\bar{X} = 287.6$$

$$\frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} = \frac{\sqrt{10}(287.6 - 285)}{4} = 2.06$$

比较两个正态总体的均值

男生:  $\bar{X}_1 = 8.6$  ;  $\bar{\sigma}_1 = 3.3$

女生:  $\bar{X}_2 = 7.9$  ;  $\bar{\sigma}_2 = 3.3$

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(8.6 - 7.9) - 0}{\sqrt{\frac{(3.3)^2}{50} + \frac{(3.3)^2}{50}}} = 1.06$$

检验单个正态总体的均值, 方差未知

$$\bar{D} = \sum D/n = 100/6 = 16.7$$

( $D = X_1 - X_2$ )

$\mathbf{H}_0$  的拒绝域为:

$$\frac{\sqrt{n}(\bar{D} - \mu_0)}{S_D} > t_{\alpha}(n-1)$$

- 查 student-t 分布表,  $t_{0.05}(5) = 2.016$
- 从样本估计方差

$$s_D^2 = \frac{\sum (D_i - \bar{D})^2}{(n-1)} = \frac{\sum D_i^2 - n \times \bar{D}^2}{(n-1)} = \frac{4890 - 6 \times (16.7)^2}{(6-1)} = 645$$

$$s_D = 25.4$$

- 计算得:

$$t = \frac{\sqrt{n}(\bar{D} - \mu_0)}{S_D} = \frac{\sqrt{6}(16.7 - 0)}{25.4}$$

$$= 1.61$$

**Tukey-Kramer** 相关方程式

- 计算显著差异的临界值:

$$\text{Critical range} = Q_u \sqrt{\frac{MSW}{2} \left( \frac{1}{n_j} + \frac{1}{n_j} \right)}$$

对应布料拉力例子:

$$\alpha = 0.05, c = 4, n - c = 20 - 4 = 16$$

从统计表得出对应  $Q_u = 4.05$  ;

MSW = 6.1 (如何计算 MSW, 详见下面” 如何计算 MSW”)

代入上面方程式：

$$\begin{aligned}\text{Critical range} &= 4.05 \sqrt{\left(\frac{6.1}{2}\right)\left(\frac{1}{5} + \frac{1}{5}\right)} \\ &= 4.47\end{aligned}$$

如何计算 **MSW**，参考 **ANOVA** 方程式

- 总离差平方和 Total Variation = Sum of Squares for Total (SST):
- SST 由组内离差平方和 (SSW)，和组间离差平方和 (SSA)，双加出来：
- 如果组间离差远远大于组内离差，表示组间差异大，应拒绝零假设；反之，如差异不大，便难以拒绝零假设，没有显著差异。
- 单因素方差分析计算  $F = \text{MSA} / \text{MSW}$ ，如果 F 值大于临界值，拒绝零假设

统计工具，如 Minitab，能直接算出下面总结报告：

| 原因 | 自由度   | 离差平方和 (Sum of Squares) | (Mean Square)                   | F                       |
|----|-------|------------------------|---------------------------------|-------------------------|
| 组间 | c - 1 | SSA                    | $\text{MSA} = \text{SSA}/(c-1)$ | $\text{MSA}/\text{MSW}$ |
| 组内 | n - c | SSW                    | $\text{MSW} = \text{SSW}/(n-c)$ |                         |
| 总  | n - 1 | SST                    |                                 |                         |

相关方程式

\* 总离差平方和

$$\text{SST} = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{\bar{X}})^2$$

$$\text{SST} = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{\bar{X}})^2$$

$$\text{注：总体平均数 } \bar{\bar{X}} = \frac{\sum_{j=1}^c \sum_{i=1}^{n_j} X_{ij}}{n}$$

- 组间离差平方和

$$\text{SSA} = \sum_{j=1}^c n_j (\bar{X}_j - \bar{\bar{X}})^2$$

$$\text{注：组 } j \text{ 的平均值 } \bar{X}_j$$

- 组内离差平方和

$$\text{SSW} = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

$$\text{注：组 } j \text{ 的平均值 } \bar{X}_j$$

- 单因素方差分析均方和

$$\text{MSA} = \frac{\text{SSA}}{c-1}$$

$$\text{MSW} = \frac{\text{SSW}}{n-c}$$

$$\text{MST} = \frac{\text{SST}}{n-1}$$

实例（比较 4 供货商布料的拉力）：



- 组间自由度 =  $c - 1 = 4 - 1 = 3$
- 组内自由度 =  $n - c = 20 - 4 = 16$
- 总自由度 =  $n - 1 = 20 - 1 = 19$

$$\text{总体平均数 } \bar{\bar{X}} = \frac{\sum_{j=1}^c \sum_{i=1}^{n_j} X_{ij}}{n} = \frac{438.9}{20} = 21.945$$

$$SSA = \sum_{j=1}^c n_j (\bar{X}_j - \bar{\bar{X}})^2$$

$$= (5)(19.52 - 21.945)^2 + (5)(24.26 - 21.945)^2 + (5)(22.84 - 21.945)^2 + (50)(21.16 - 21.945)^2$$

$$= 63.285$$

$$SSW = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

$$= (18.5 - 19.52)^2 + \dots + (18.0 - 19.52)^2 + (26.3 - 24.26)^2 + \dots + (24.5 - 24.26)^2$$

$$+ (25.4 - 21.16)^2 + \dots + (20.4 - 21.16)^2$$

$$= 97.5$$

$$SST = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{\bar{X}})^2$$

$$= (18.5 - 21.945)^2 + (24 - 21.945)^2 + \dots + (20.4 - 21.945)^2$$

$$= 160.7895$$

$$MSA = \frac{SSA}{c-1} = \frac{63.285}{4-1} = 21.1$$

$$MSW = \frac{SSW}{n-c} = \frac{97.5}{16} = 6.1$$

$$F = MSA / MSW = 21.1 / 6.1 = 3.46$$

能对应文中 Minitab 出来的结果：

$$MSA = 21.16, MSW = 6.12, F = 3.45$$

## Part VII

### 总结



管理者应如何支持团队做好敏捷开发，为公司增值