

源于社区 服务社区

 中国DevOps社区峰会 2023 · 广州



量化管理在软件开发中的应用

张玉洁-禅道-产品经理





张玉洁

禅道软件（青岛）有限公司·产品经理

- 禅道项目管理软件
- ZenDAS统计分析工具



什么是量化管理

经验管理

依赖于管理者的经验判断来选择和实施各种措施以管理目标的达成

管理者的经验有丰富与匮乏的区别，经验也有其适用的范围，有时正确，有时又可能错误

例如：医生看病

优点：简单，有效

缺点：要有经验丰富管理者

适用场景：需要快速响应



并非要对所有的过程都进行量化管理，而是有选择的，有的过程是量化管理，有的过程也可能是经验管理

量化管理

现象用数字进行刻画，从这些数字之间的关系推导出背后的道理，从而识别出管理的措施，用以管理目标的达成

这种推理是一种具有更高确定性的推理方式

优点：科学准确

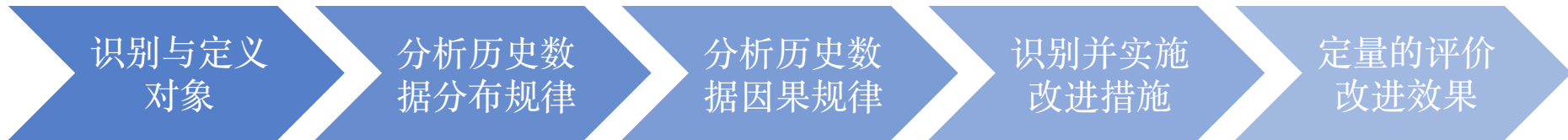
缺点：成本高

适用场景：高水平高质量需求的场景





什么是量化管理





目录

1 小刘的困扰

2 数据收集难

3 数据无规律

4 数据分析难

5 Q&A



小刘的困扰

- 迭代难题
- 照葫芦画瓢
- 困难重重

迭代难题



基本信息

需求 49

任务 65

Bug 60

工时统计

进度 100% 

计划开始 2023-03-27

实际开始 2023-03-24

计划完成 2023-03-31

实际完成 2023-04-03

预计 140.5工时

可用工日 5天

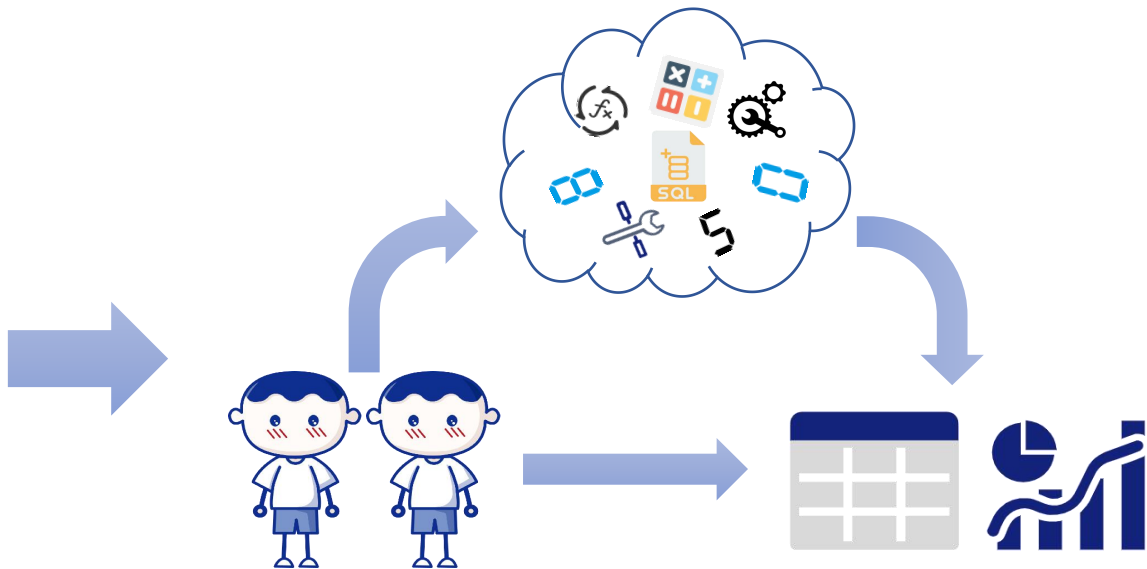
消耗 148工时

可用工时 212.5工时

剩余 0工时



照葫芦画瓢





困难重重

- 数据收集难
- 数据无规律
- 数据分析难





数据收集难

- 度量角色划分
- 度量指标定义

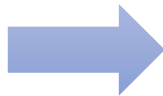




度量角色划分

- 度量的采集和定义是由两个角色共同协作完成的。
- 度量定义者是上游，数据收集者是下游。
- 度量定义非常关键，直接影响数据收集的准确性。

定义度量



数据收集



度量指标定义

名称	完成研发需求数
范围	迭代
对象	研发需求
目的	规模估算
描述	迭代中已完成的研发需求数
单位	个
采集时间	迭代完成后



小刘的定义有什么问题？





度量指标定义

名称	完成研发需求数
范围	迭代
对象	研发需求
目的	规模估算
描述	迭代中已完成的研发需求数
单位	个
采集时间	迭代完成后



- 定义描述不清晰。
- 用excel的方式储存。



度量指标定义

The screenshot displays a BI system interface with a sidebar on the left containing navigation items: 地盘, 项目集, 产品, 项目, 执行, 测试, DevOps, 看板, 资产库, 文档, BI (highlighted), 组织, 办公, and 更多. The main content area shows a list of metrics under the '执行' (Execution) tab. The selected metric is '按执行统计的已完成研发需求数'.

详情

基本信息

范围	执行
对象	需求
目的	规模估算
名称	按执行统计的已完成研发需求数
代号	count_of_finished_story_in_execution
单位	个
描述	按执行统计的已完成研发需求数是指状态为已关闭且关闭原因为已完成的研发需求的数量。这个度量项可以反映执行团队在开发过程中的进展和交付能力。已完成研发需求数越多, 说明执行团队在该时间段内取得了更多的开发成果。
计算规则	执行中研发需求的个数和 状态为已关闭 关闭原因为已完成 过滤已删除的研发需求 过滤已删除的执行 过滤已删除的项目

创建编辑信息

创建者	system 于 2023-10-30 21:24:36
-----	------------------------------

编辑信息

2023-10-30 21:24:36





数据无规律

- 小刘的数据
- 数据的清洗



小刘的数据

对数据实施正态性检验。

正态性检验的作用是确定数据是否具有正态分布的特征。如果数据满足正态分布假设，那么一些统计方法和假设检验的结果将更加可靠和准确。



ZenDAS 迭代数据分析

项目管理

数据管理

用户管理

迭代数据

分析报告

+

数据 66*25

生成数据 + 添加数据

	C1 名称	C2 工期偏差率	C3 交付效率	C4 消耗工时	C5 完成需求数	C6 需求清晰度	C7 产品经理能力	C8 开发人员能力	C9 测试人员能力
1	迭代-01	0.07	0.74	42	31	1.39	0.80	0.79	0.74
2	迭代-02	0.26	0.77	58	45	1.04	0.82	0.82	0.49
3	迭代-03	0.20	1.20	30	36	1.14	0.77	0.77	0.76
4	迭代-04	0.32	0.78	47	37	1.42	0.92	0.91	0.73
5	迭代-05	0.06	0.04	46	2	0.93	0.78	0.61	0.75
6	迭代-06	0.21	0.76	37	28	0.77	0.68	0.68	0.66
7	迭代-07	0.31	0.73	57	42	0.68	0.86	0.85	0.73
8	迭代-08	0.34	0.95	36	34	0.66	0.89	0.89	0.91
9	迭代-09	0.17	0.71	48	34	1.07	0.64	0.63	0.90
10	迭代-10	0.10	0.87	38	33	0.91	0.48	0.48	0.89
11	迭代-13	0.26	0.74	54	40	1.01	0.69	0.68	1.00
12	迭代-14	0.33	0.76	47	36	1.17	1.00	1.00	0.91
13	迭代-15	0.16	0.49	49	24	1.08	0.70	0.69	0.71
14	迭代-16	0.30	1.10	34	37	1.13	0.72	0.73	0.38
15	迭代-17	0.59	0.60	54	32	0.79	1.00	1.00	0.60
16	迭代-18	0.80	1.03	27	28	0.50	1.00	1.00	0.80
17	迭代-19	0.26	0.85	45	38	0.93	0.93	0.94	0.63
18	迭代-20	0.58	0.94	39	37	0.60	0.96	0.95	0.57
19	迭代-21	0.40	0.80	49	39	0.43	0.78	0.79	0.72
20	迭代-22	0.05	0.80	41	33	0.52	0.63	0.64	0.50
21	迭代-23	0.06	0.60	56	34	1.43	0.67	0.67	1.00
22	迭代-24	0.31	0.96	39	38	0.66	0.60	0.60	1.00
23	迭代-25	0.62	0.88	44	39	0.36	1.00	1.00	1.00



小刘的数据——正态性检验

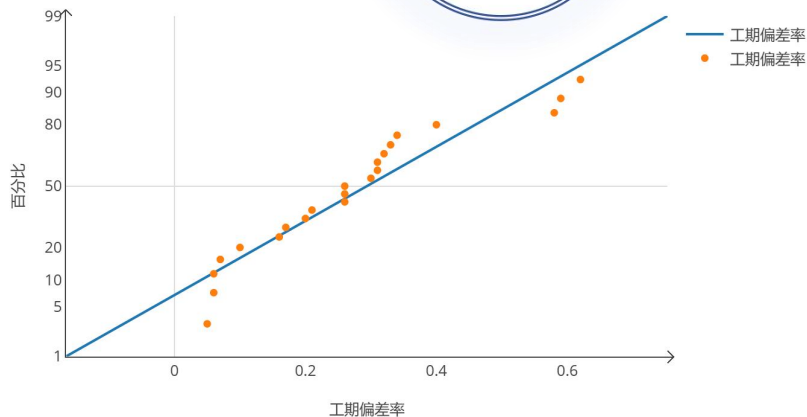
正态性检验的p值是用来评估数据是否服从正态分布的概率。

如果正态性检验的p值较大（通常 >0.05 ），则表示数据在统计意义上可以被认为是服从正态分布的。

工期偏差率的 正态性检验

	均值	标准差	计数	D	p 值
工期偏差率	0.2939	0.1974	23	0.35	0.0466

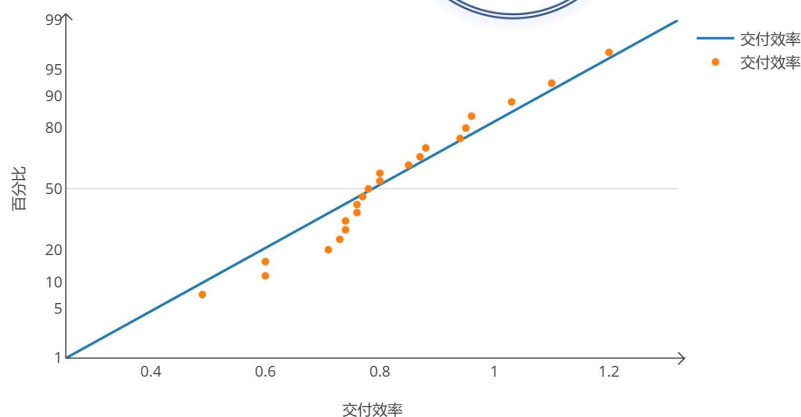
工期偏差率的 正态概率图



交付效率的 正态性检验

	均值	标准差	计数	D	p 值
交付效率	0.7870	0.2295	23	0.327	0.0359

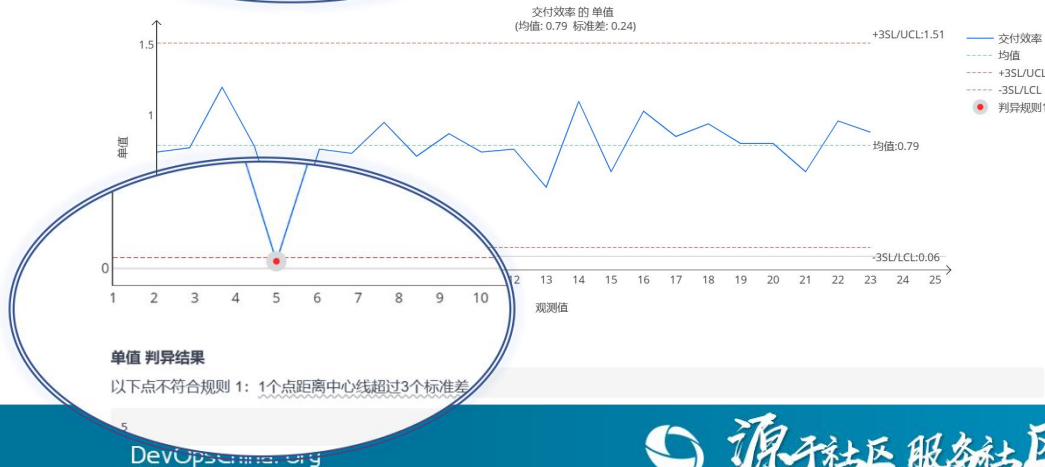
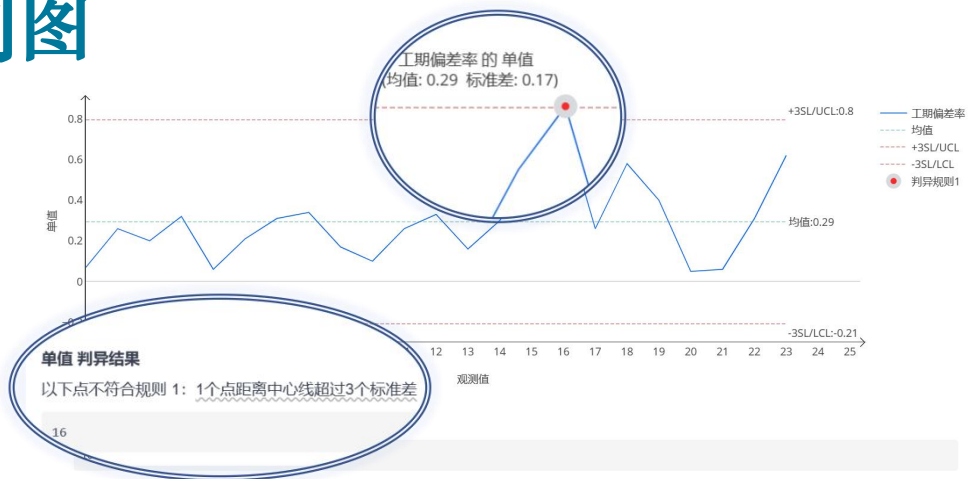
交付效率的 正态概率图



小刘的数据——控制图

控制图分析可以提供关于过程稳定性、能力和异常情况的重要信息，帮助组织监控和改进其过程，以实现质量控制和持续改进的目标。

- 中心线：中心线代表数据的中心或平均值。
- 控制限：控制图上通常会绘制上下两条控制限，用于标识过程的稳定性范围。





数据的清洗

1.寻找迭代交付效率低的原因。

存在异常迭代，迭代临近发版，时间主要用来解bug。

解决方案：移除异常迭代。

2.寻找工期偏差率高的原因。

开发小组长在迭代全部任务完成之后没有及时关闭迭代。

解决方案：修正数据。



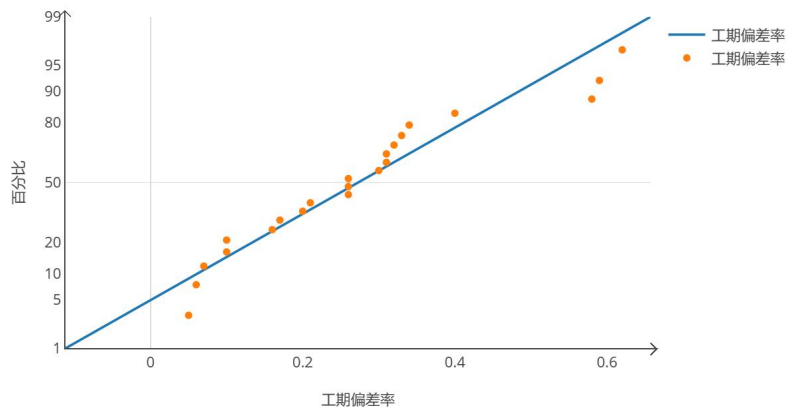
数据的清洗

清洗之后，P值>0.05，符合正态分布。

工期偏差率的 正态性检验

	均值	标准差	计数	AD	p 值
工期偏差率	0.2727	0.1653	22	0.5941	0.1084

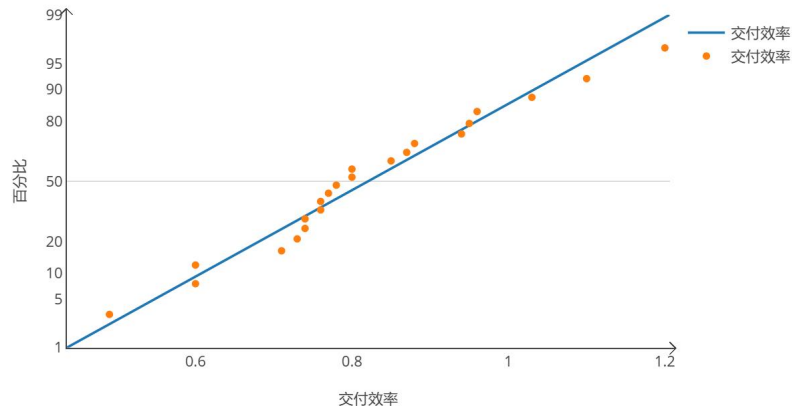
工期偏差率的 正态概率图



交付效率的 正态性检验

	均值	标准差	计数	AD	p 值
交付效率	0.8209	0.1656	22	0.3488	0.4432

交付效率的 正态概率图





数据分析难

- 相关性分析
- 数据建模
- 数据预测
- 数据监控



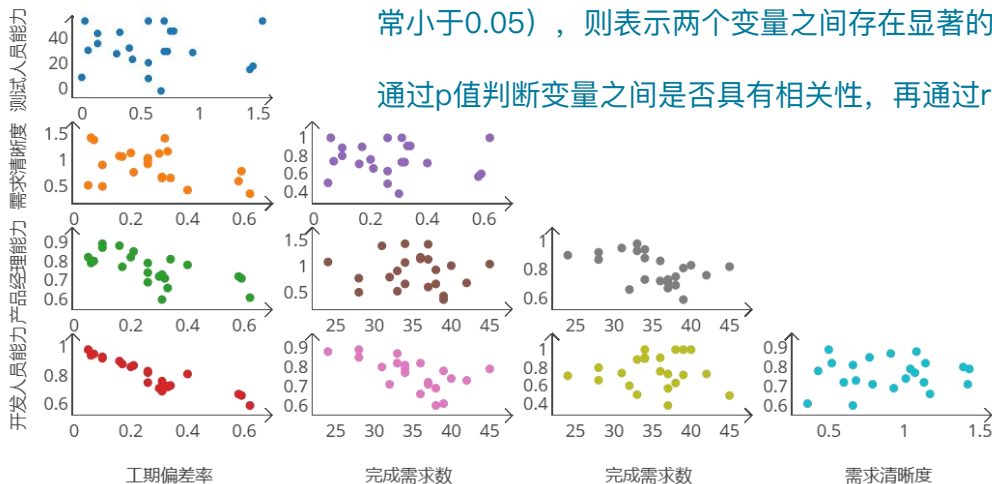
相关性分析

相关性分析是一种统计方法，用于衡量和描述两个或多个变量之间的相关关系和相关程度。

样本	N	相关性r	p值
工期偏差率 - 完成需求数	22	0.3914	0.0717
工期偏差率 - 需求清晰度	22	-0.4232	0.0497
工期偏差率 - 产品经理能力	22	-0.6538	0.0010
工期偏差率 - 开发人员能力	22	-0.9146	0.0000
工期偏差率 - 测试人员能力	22	-0.0556	0.8059
完成需求数 - 需求清晰度	22	-0.0749	0.7405
完成需求数 - 产品经理能力	22	-0.5871	0.0041
完成需求数 - 开发人员能力	22	-0.4675	0.0282
完成需求数 - 测试人员能力	22	-0.0189	0.9334
需求清晰度 - 产品经理能力	22	0.0875	0.6985
需求清晰度 - 开发人员能力	22	0.2851	0.1985
需求清晰度 - 测试人员能力	22	0.0227	0.9200
产品经理能力 - 开发人员能力	22	0.8111	0.0000
产品经理能力 - 测试人员能力	22	-0.1828	0.4155
开发人员能力 - 测试人员能力	22	-0.0040	0.9858

r 是相关性系数，取值范围在 -1 到 1 之间。 r 的绝对值越接近1，表明两个变量之间的相关性越强，反之越弱。 p 值用于评估两个变量之间的相关关系的显著性。如果相关性检验的 p 值较小（通常小于0.05），则表示两个变量之间存在显著的相关性。

通过 p 值判断变量之间是否具有相关性，再通过 r 值看相关程度。





数据的建模——回归分析

回归方程

工期偏差率 = 1.4020 - 1.3024 开发人员能力 - 0.0910 需求清晰度

作用：

- **预测与优化：**基于给定的因子值，回归方程可以用于预测目标变量的值。这可以帮助确定在给定因子条件下，目标变量的表现。然后通过对因子进行调整和优化，可以寻找最佳的因子组合，以达到期望的目标结果。
- **因子优先级排序：**回归方程中的系数还可以用于确定因子的优先级排序。系数绝对值较大的因子对目标变量的影响更大。
- **过程控制和调整：**通过实时监测因子值和目标变量的变化，并结合回归方程的预测能力，可以进行过程控制和调整。如果因子值偏离了预期的范围，导致目标变量与期望的预测值之间存在较大差异，可以及时采取措施来调整和控制因子，以实现目标的达成。



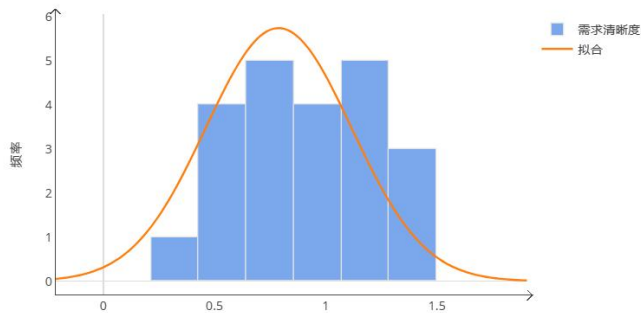
数据的预测——数据分布

根据影响历史数据，我们无法直接推断出未来目标变量与影响因子的具体数值，但可以拟合出数据的分布。根据数据分布，预测目标值的概率密度分布，通过目标值的概率调整影响因子取值，获取改进方案。

数据汇总

Anderson 值	0.579	均值	0.895	标准差	0.3213
方差	0.1032	偏度	0.0809	峰度	-0.9704
计数	22	最小值	0.36	10%分位数	0.451
下四分位数	0.645	中位数	0.92	上四分位数	1.1325
90%分位数	1.411	最大值	1.43	四分位差	0.4875

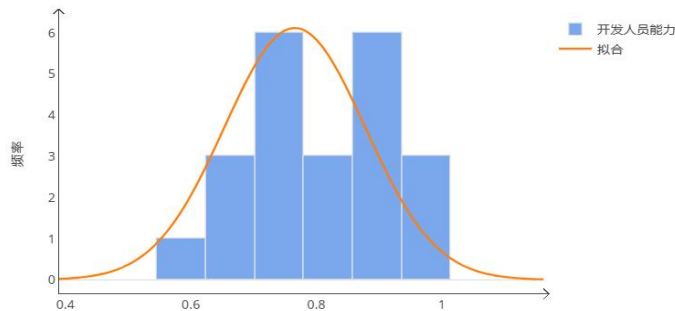
需求清晰度的 图形化汇总



数据汇总

Anderson 值	0.4797	均值	0.8045	标准差	0.1097
方差	0.012	偏度	-0.1236	峰度	-1.047
计数	22	最小值	0.59	10%分位数	0.663
下四分位数	0.7175	中位数	0.815	上四分位数	0.905
90%分位数	0.947	最大值	0.98	四分位差	0.1875

开发人员能力的 图形化汇总



数据的预测——蒙特卡罗模拟

设定分析方法：蒙特卡罗模拟



假设变量

服从分布或值 *

a

需求清晰度

分布

正态分布,0.895,0.3213,0.36,1.43



b

开发人员能力

分布

正态分布,0.8045,0.1097,0.59,0.98



+ 添加变量

使用决策变量

☒ 否

☐ 是

输出

预测名称

公式 ① *

☒

y1

工期偏差率

$1.4020-1.3024*b-0.091*a$

条件公式



蒙特卡罗模拟又称为统计模拟法、随机抽样技术。依赖计算机从输入变量的概率分布中大量、随机抽样，以得到输出变量的分布。

例：

目标：到达公司的时间概率分布

影响因子：起床耗时,洗漱耗时,路程耗时...

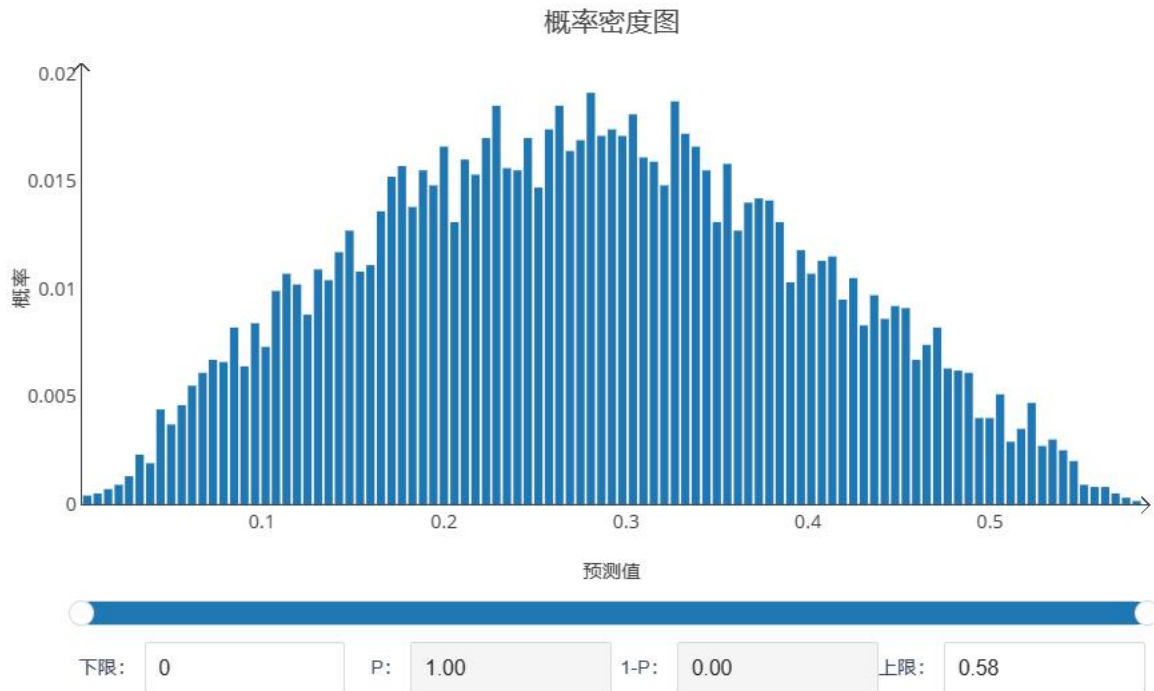




数据的预测——蒙特卡罗模拟

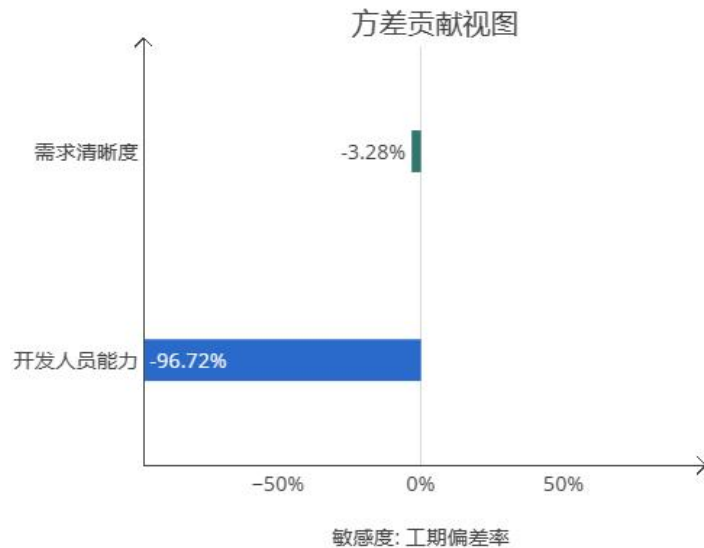
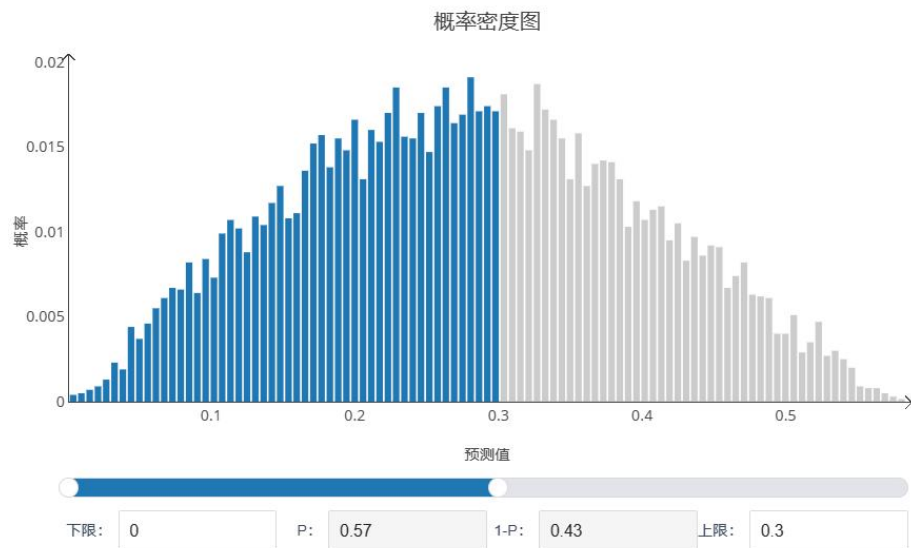
工期偏差率的基本描述

统计值	预测值
试验次数	10000
均值	0.2818
中位数	0.2806
方差	0.0142
标准差	0.1193
偏度	0.0807
峰度	2.3027
最大值	0.5805
最小值	0.0039



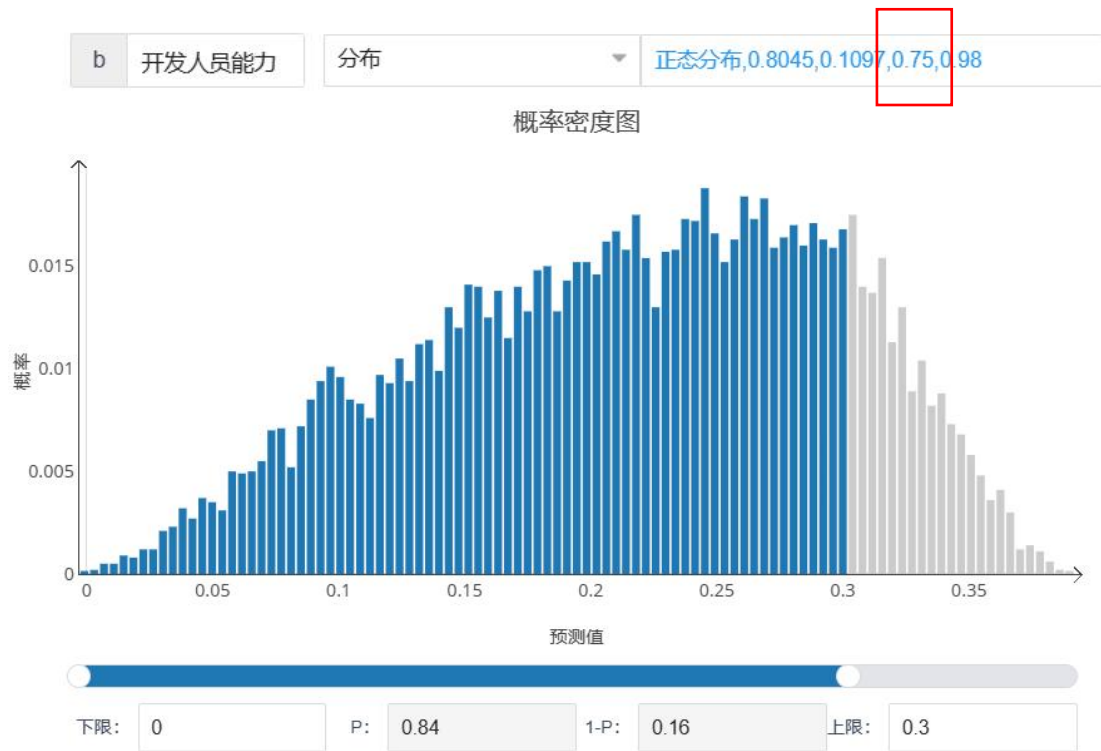


数据的预测——蒙特卡罗模拟





数据的预测——蒙特卡罗模拟



- 预测值达到目标期望后，可以针对性的实施改进措施。
- 改进后通过实时监测因子和目标变量的变化。



数据的监控



- 实时监控过实时监测因子和目标变量。评价改进措施有无效果。
- 如果因子值偏离了预期的范围，可以及时采取措施来调整和控制因子，以实现目标的达成。



数据的监控





总结

通过运用量化管理知识和统计分析工具，以更科学的方式做管理。



源于社区 服务社区

THANKS!

