



SAS FORUM | 中国用户大会暨
CHINA 2017 | 商业分析领袖峰会

泰康人寿两核大数据项目经验分享

泰康集团数据信息中心大数据部总经理 周雄志



SAS FORUM | 中国用户大会暨
CHINA 2017 | 商业分析领袖峰会



尊重生命 全心为您



让保险更安心、更便捷、更实惠，
让人们更健康、更长寿、更富足，
让泰康成为人们幸福生活的一部分。

版权所有 © 泰康保险集团股份有限公司

保险 资管 医养

让保险更安心、更便捷、更实惠
让人们更健康、更长寿、更富足
让泰康成为人们幸福生活的一部分

背景

核保

核赔

安心

实惠

便捷

偿付能力

风险选择

时效

目录

1 项目概述

- 1.1 项目目标
- 1.2 项目团队

ONE

2 模型开发

- 2.1 模型开发步骤
- 2.2 数据提取
- 2.3 建模数据准备
- 2.4 变量选择
- 2.5 模型结果
- 2.6 模型效果

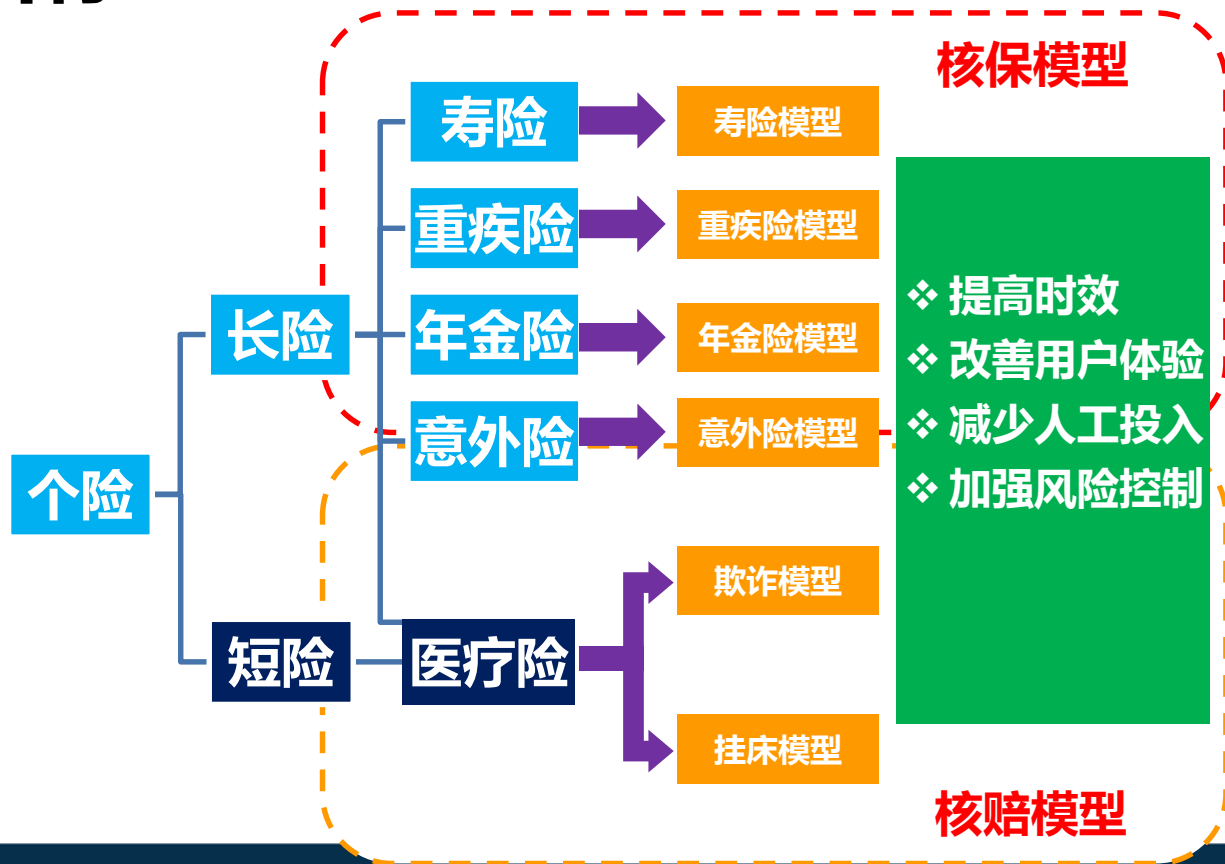
TWO

3 模型应用

- 3.1 核保应用
- 3.2 核赔应用
- 3.3 应用展望

THREE

1.1 项目目标



1.2 项目团队



目录

1 项目概述

- 1.1 项目目标
- 1.2 项目团队

ONE

2 模型开发

- 2.1 模型开发步骤
- 2.2 数据提取
- 2.3 建模数据准备
- 2.4 变量选择
- 2.5 模型结果
- 2.6 模型效果

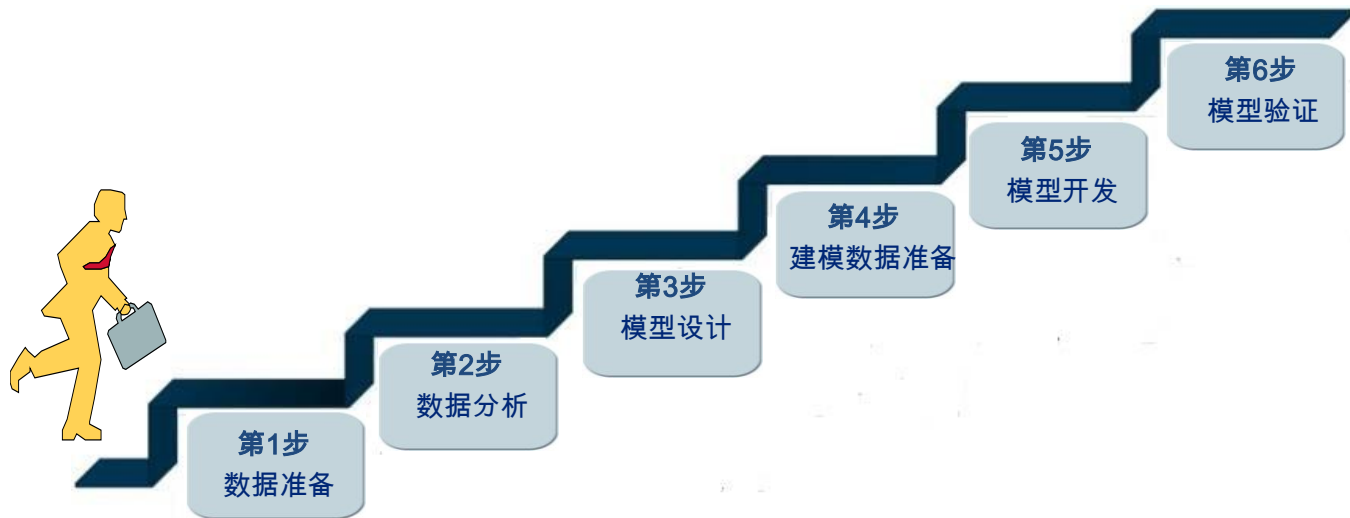
TWO

3 模型应用

- 3.1 核保应用
- 3.2 核赔应用
- 3.3 应用展望

THREE

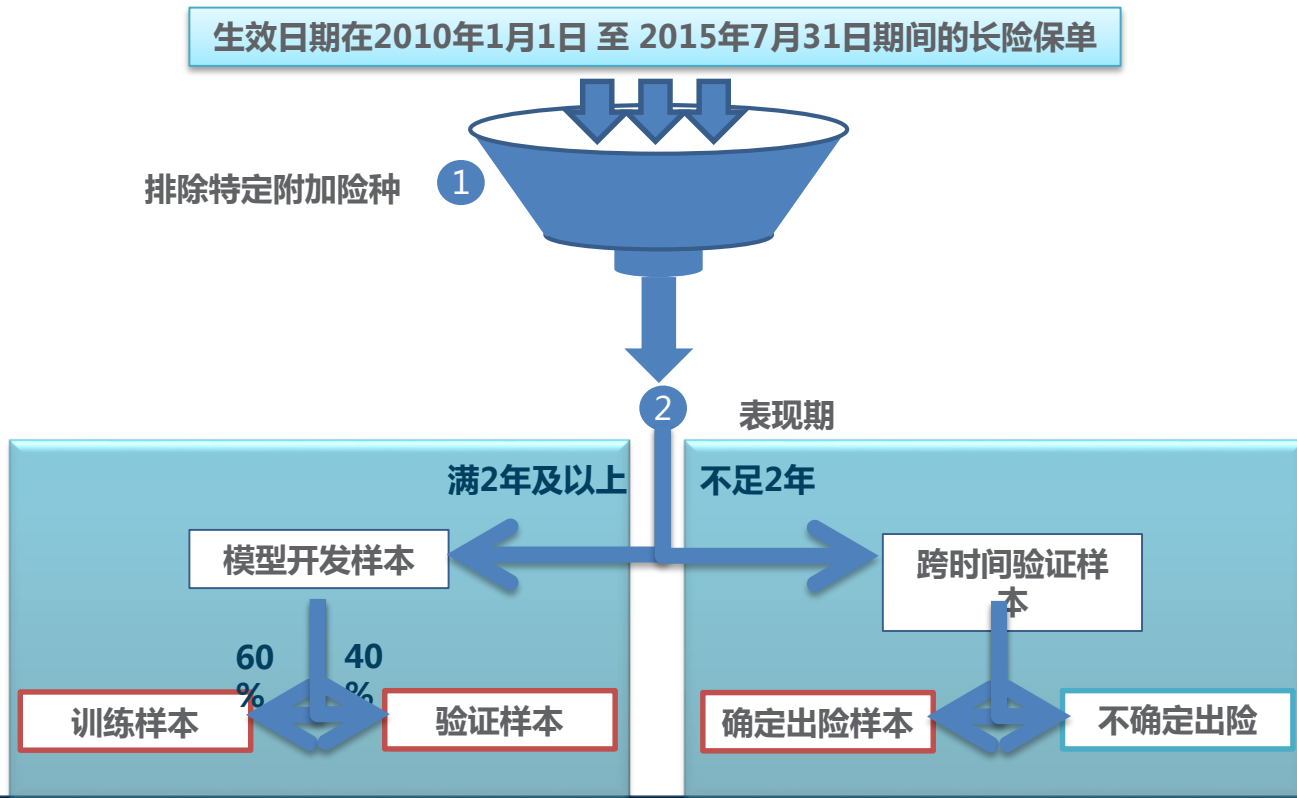
2.1 模型开发步骤



2.2.2 数据准备：核赔模型为例

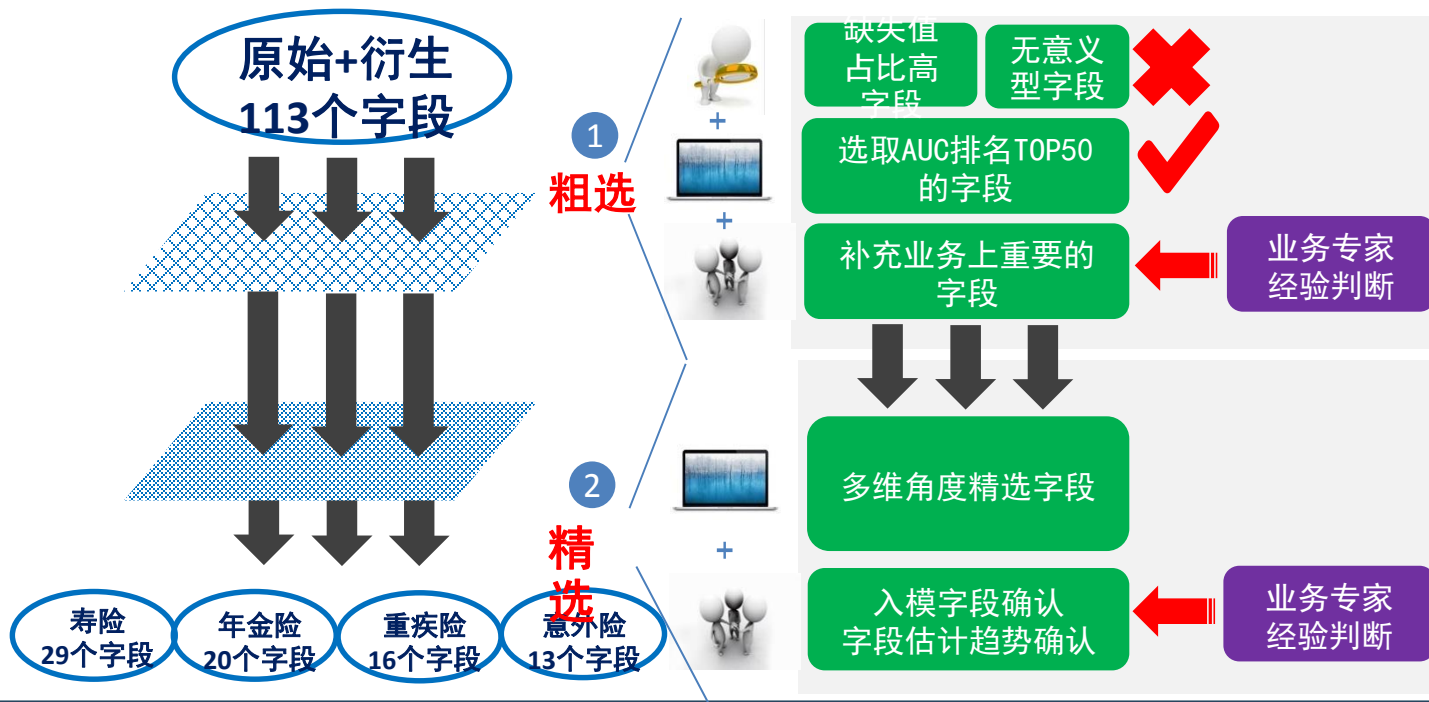


2.3.1 建模数据准备：核保寿险模型



2.4.1 变量选择

变量筛选流程：人工 + 自动 + 合作

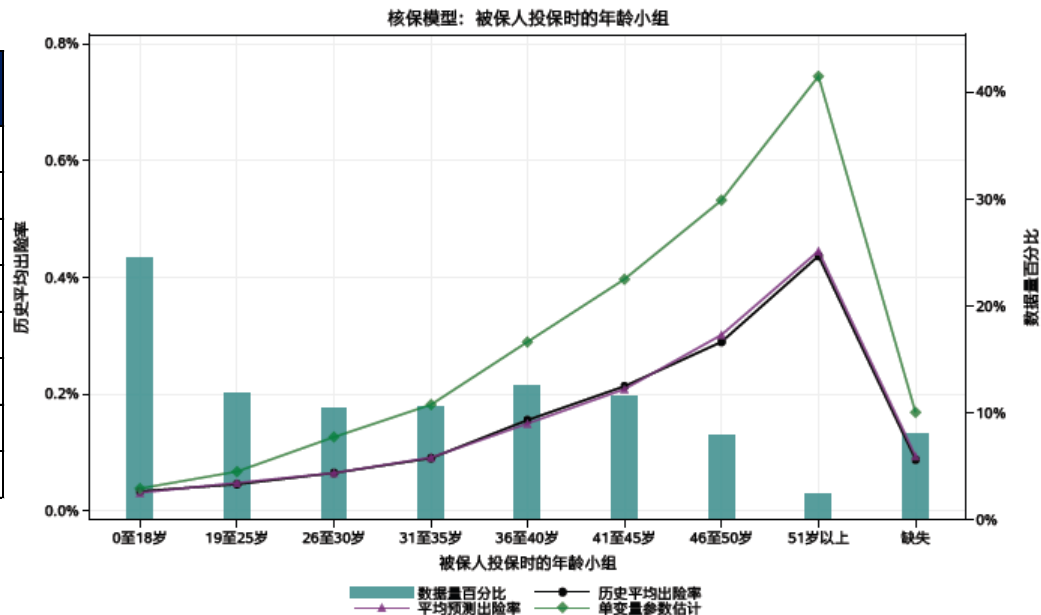


2.4.2 变量筛选

以核保寿险模型为例

❖ 被保险人投保时的年龄越大，首两年出险的风险越高。

被保险人投保时的年龄小组	参数估计值
(0,18]	0.000000000
(18,25]	0.562089191
(25,30]	1.200106942
(30,35]	1.565610615
(35,40]	2.031595813
(40,45]	2.350160753
(45,50]	2.643519028
(50,Inf)	2.981883176

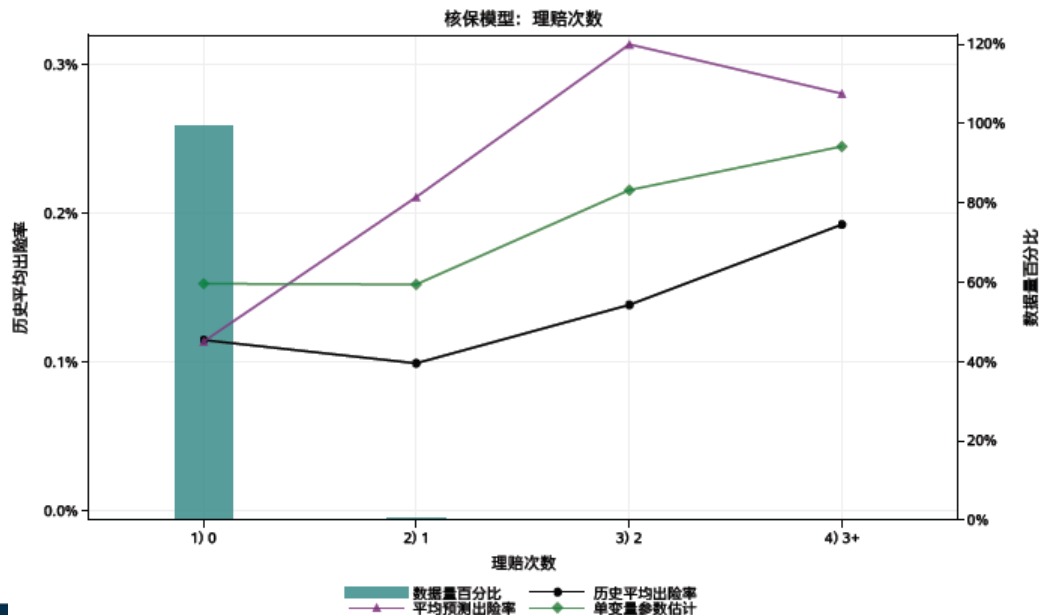


2.4.2 变量筛选（续）

以核保寿险模型为例

❖ 医疗险既往理赔次数越多，首两年出险的风险越高。

理赔次数	参数估计值
=0	0.000000000
=1	-0.004354233
=2	0.345121331
=3	0.473443322
>3	0.473443322



2.5.1 最终入模字段

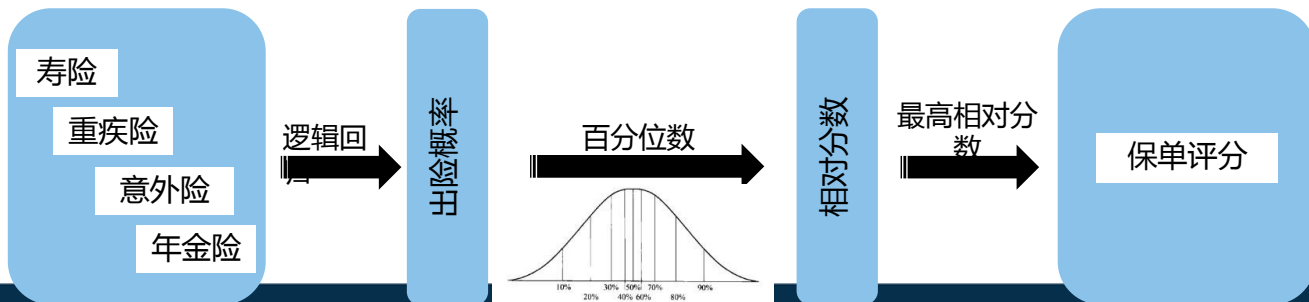
	寿险核保
客户信息	<ul style="list-style-type: none">投保人学历被保险人年龄理赔次数近两年自己购买重疾总数 (投).....
保单信息	<ul style="list-style-type: none">保障年期基本保额附加险种比例 (本保单附加险/主险)累计投保额度 (医疗) (被).....
代理人信息	<ul style="list-style-type: none">精英代理人标识是否营销员自保单代理人工作年限营销员年龄.....
销售渠道地区	<ul style="list-style-type: none">本部/郊县标识渠道分类分公司.....

	理赔欺诈
客户信息	<ul style="list-style-type: none">被保险人出险年龄.....
保单信息	<ul style="list-style-type: none">基本保额累计投保额度 (津贴)产品类型_x_等待期.....
代理人信息	<ul style="list-style-type: none">代理人名下理赔调查阳性保单件数 (医疗险).....
销售渠道地区	<ul style="list-style-type: none">渠道分类.....
理赔信息	<ul style="list-style-type: none">ICD疾病分组疾病分类 (意外/疾病)最近一次出险时间间隔被保人提交的住院总天数.....
医院及费用信息	<ul style="list-style-type: none">药费费用.....

2.5.3 绝对分数和相对分数

评分策略：**绝对分数 + 相对分数**

- 绝对分数 — 出险概率
 - 结合保额等其它信息可用于预测赔付金额，为来年计划准备金提供理论依据
- 相对分数 — 出险概率分布的百分位数
 - 相对分数将不同险种的出险概率统一到同样的尺度（1分-100分）
 - 便于比较分析同一张保单上不同险种责任的风险，准确描绘出整张保单的风险等级



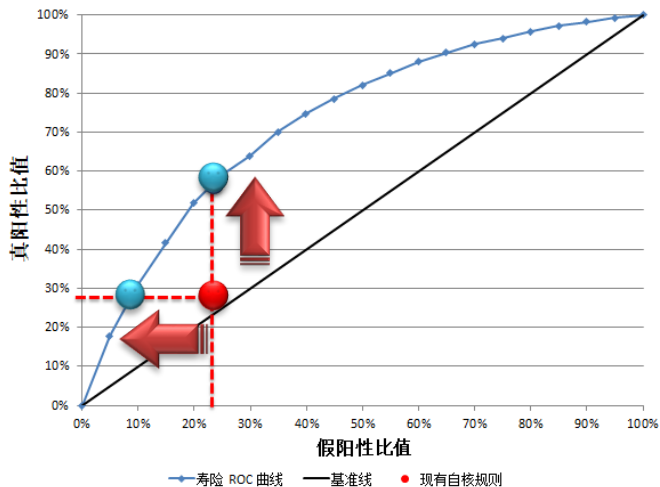
2.6.1 模型效果：以核保寿险模型为例



区分能力

模型验证：ROC 曲线

AUC=0.73

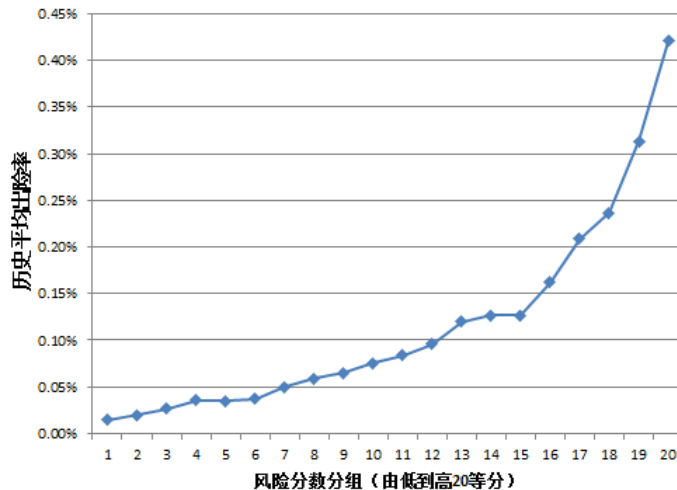


❖ 模型的区分能力优于现有自核规则。



排序能力

模型验证：风险评分与实际出险率分布对比图



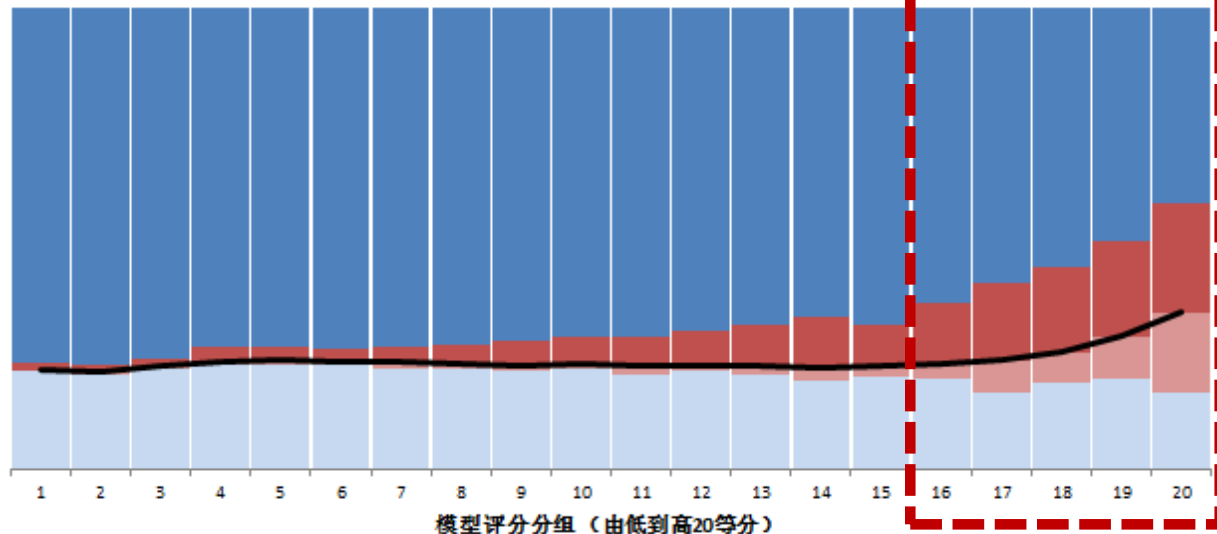
❖ 按照风险评分模型的风险评分，将模型验证样本由低到高均分成20组。随着风险分数的提高，实际出险率越来越高。

2.6.1 模型效果：优化自核规则



- 如果严格利用模型评分来核保，可以在保证自核率不变的前提下，将自核通过保单的出险率降低约一倍，从而大幅度降低公司的潜在的赔付金额。

现有自核规则 vs 模型评分



目录

1 项目概述

- 1.1 项目目标
- 1.2 项目时间
- 1.3 项目成果
- 1.4 项目团队

ONE

2 模型开发

- 2.1 模型开发步骤
- 2.2 数据提取
- 2.3 建模数据准备
- 2.4 变量选择
- 2.5 模型结果
- 2.6 模型效果

TWO

3 模型应用

- 3.1 核保应用
- 3.2 核赔应用
- 3.3 应用展望

THREE

3.1 核保模型应用

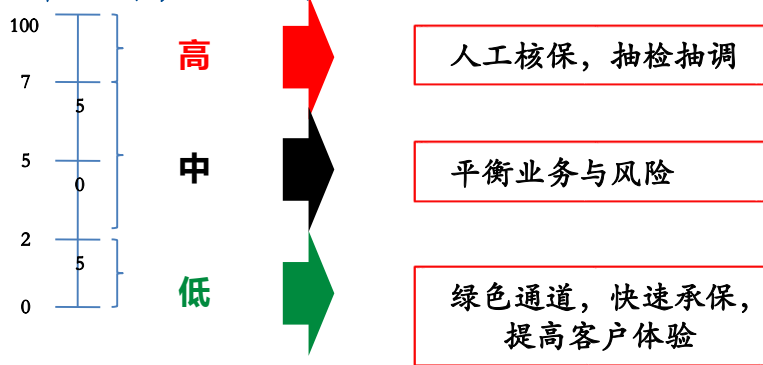
1、初级阶段---规则优先，评分监测

模型评分 自核结果	低	中	高
Y	√ (1)	√ (2)	√ (3)
N	× (4)	× (5)	× (6)

承保后出险前调查

进行后期统计分析

2、高级阶段---评分为准，优化规则



3.2 核赔模型应用

精准风险识别，提升自核效率

模型评分 自核结果	低	高
Y	通过	通过
N	二次自动审核	人工审核
二次规则审核		
Y	通过	
N	人工审核	

二次审核，提升自核通过率，提升效率；
人工审核，大数据提示风险因素，提升核赔质量。

3.3 风险量化应用展望

个性化费率、产品



优质客户筛选



专属产品

优惠费率



提升业务价值，
客户满意



SAS FORUM | 中国用户大会暨
CHINA 2017 | 商业分析领袖峰会

Thank You



SAS FORUM | 中国用户大会暨
CHINA 2017 | 商业分析领袖峰会

