



假设检验-1

我们这里不涉及任何具体的统计分布

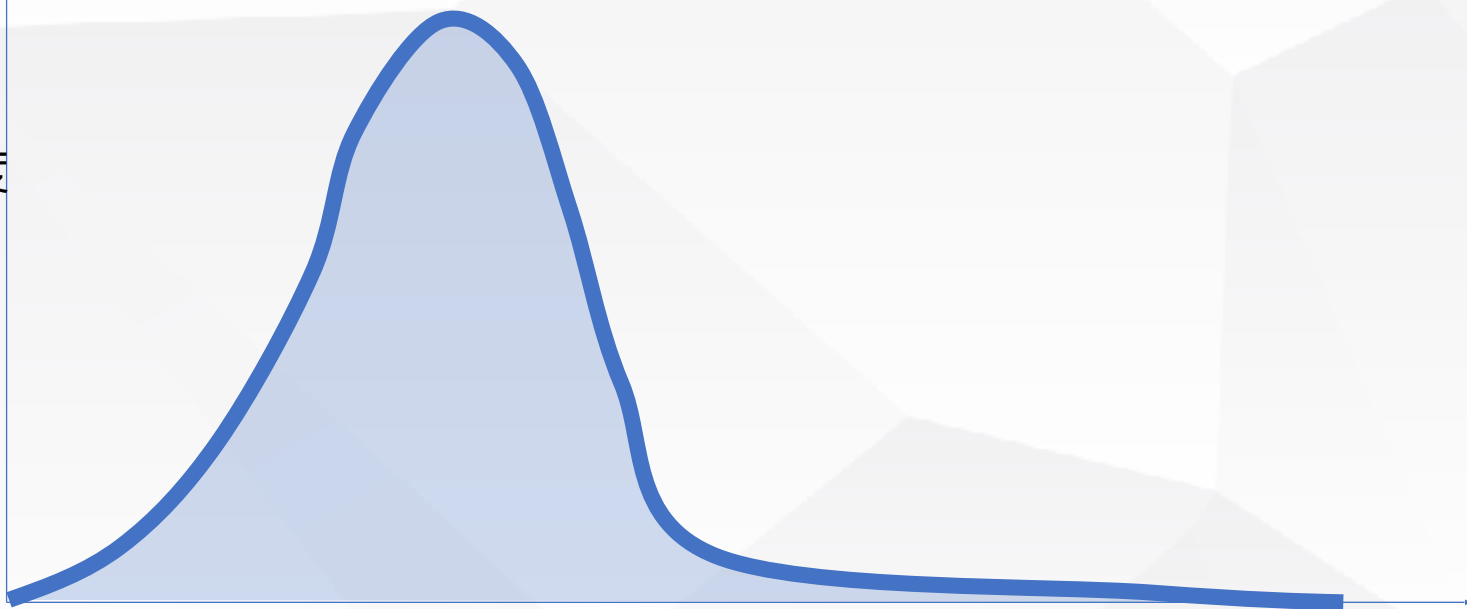


假设检验-2.原假设

原假设：没喝酒

正常人酒精浓度检测结果分布

概率密度



检测仪测得酒精浓度度数

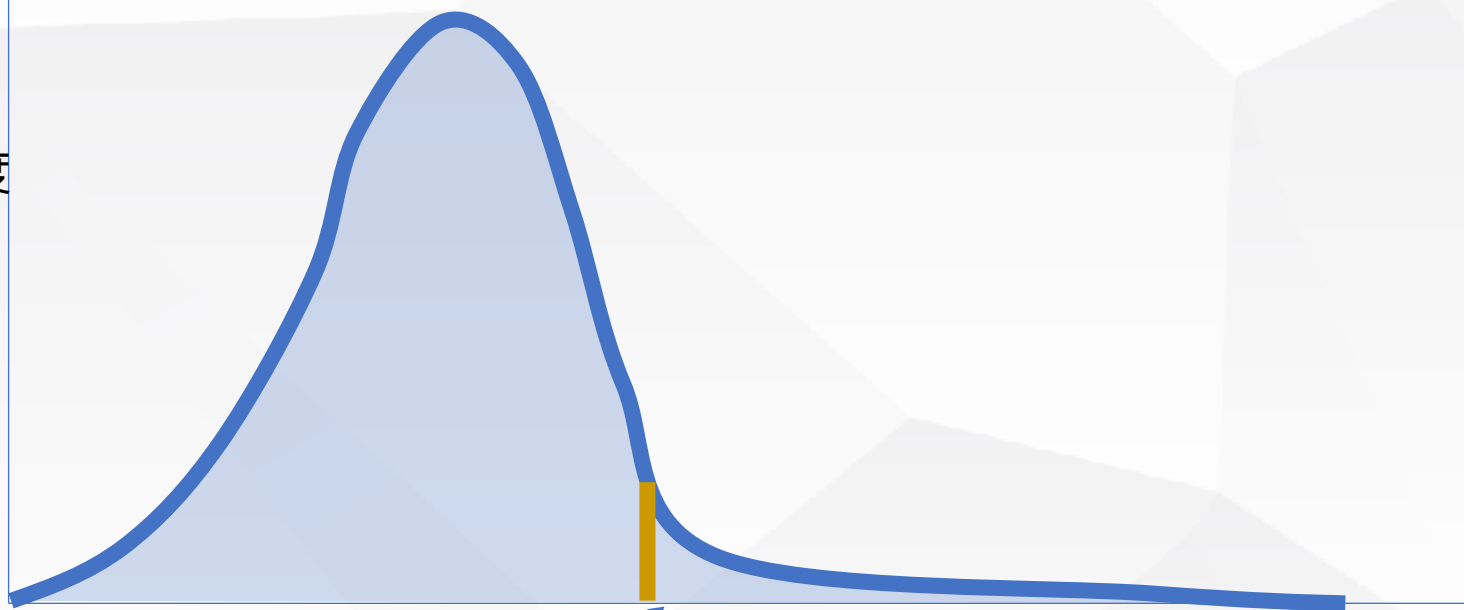


假设检验-3.P-value

原假设：没喝酒

正常人酒精浓度检测结果分布

概率密度



某司机经过检测仪测得酒精浓度度数

P-value: 在原假设成立时, 出现观测值以及更极端情况的**概率**

注意事项: 现在还没有定义备择假设, 跟不存在判断是否拒绝原假设



假设检验-4.备择假设

原假设：没喝酒

备择假设：喝了酒

正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

概率密度



检测仪酒精浓度度数 mg/100ml



假设检验-5.Threshold分类阈值

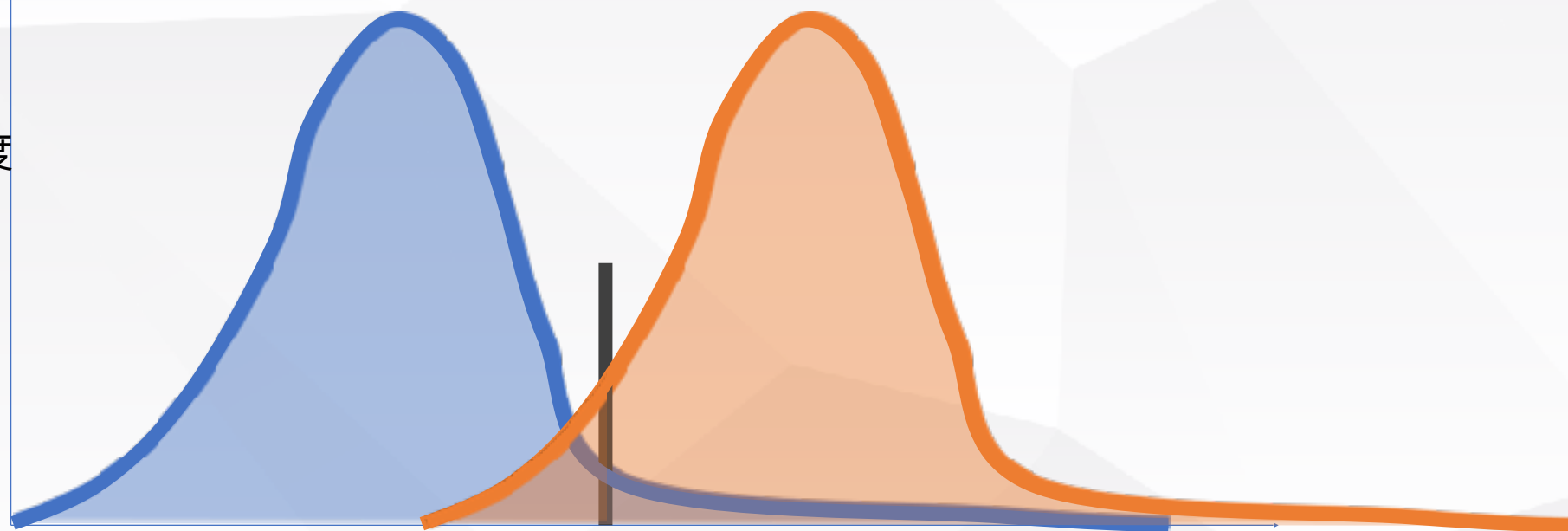
原假设：没喝酒

备择假设：喝了酒

正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

概率密度



检测仪酒精浓度度数

Threshold酒驾检测标准：
分类的阈值，阈值可人为改变



假设检验-6.设立分类阈值的结果

原假设：没喝酒

备择假设：喝了酒

正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

True Positive

False Positive

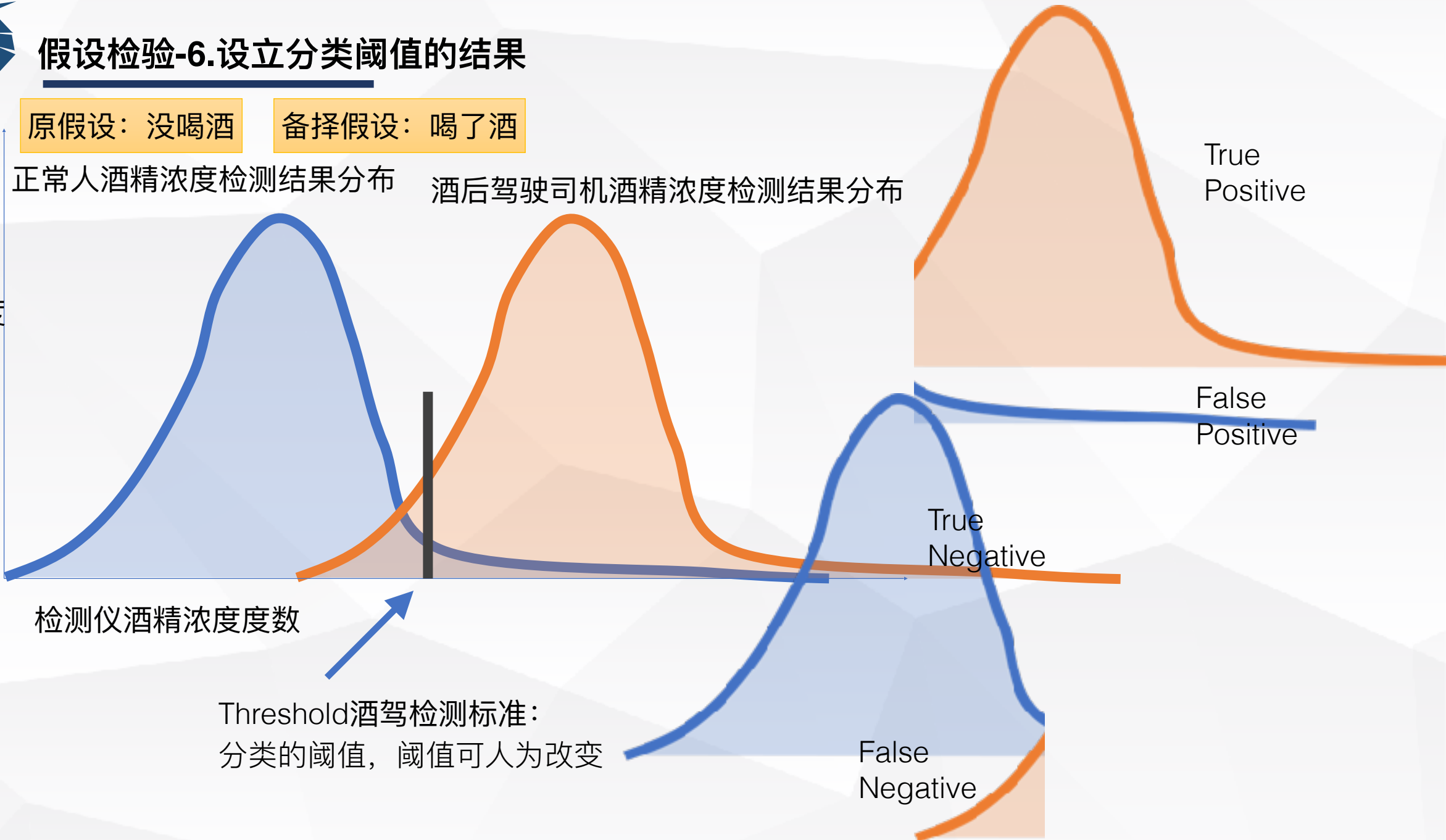
True Negative

False Negative

检测仪酒精浓度度数

Threshold酒驾检测标准：
分类的阈值，阈值可人为改变

概率密度





假设检验-7.定义阈值的分类效果:显著性水平

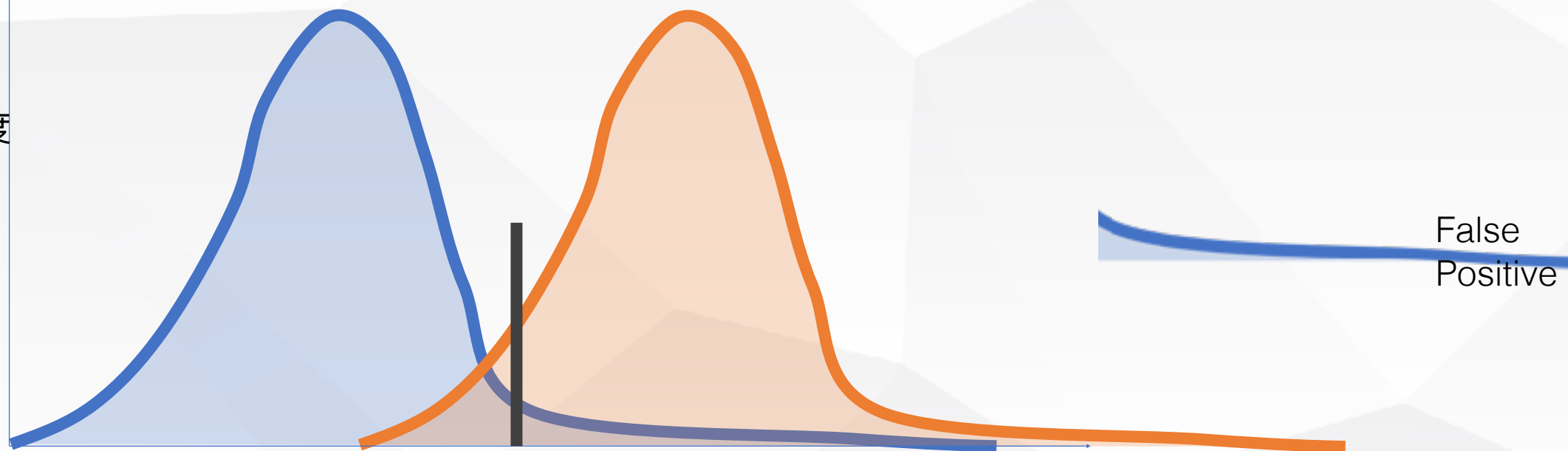
原假设：没喝酒

备择假设：喝了酒

正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

概率密度



检测仪酒精浓度度数

Threshold酒驾检测标准：
分类的阈值，阈值可人为改变

定义阈值的分类效果：
用犯错(Type I error)的概率来衡量

白话：能不能显著地区分？减少不确定性？



什么是好的分类-8. alpha, beta, and power

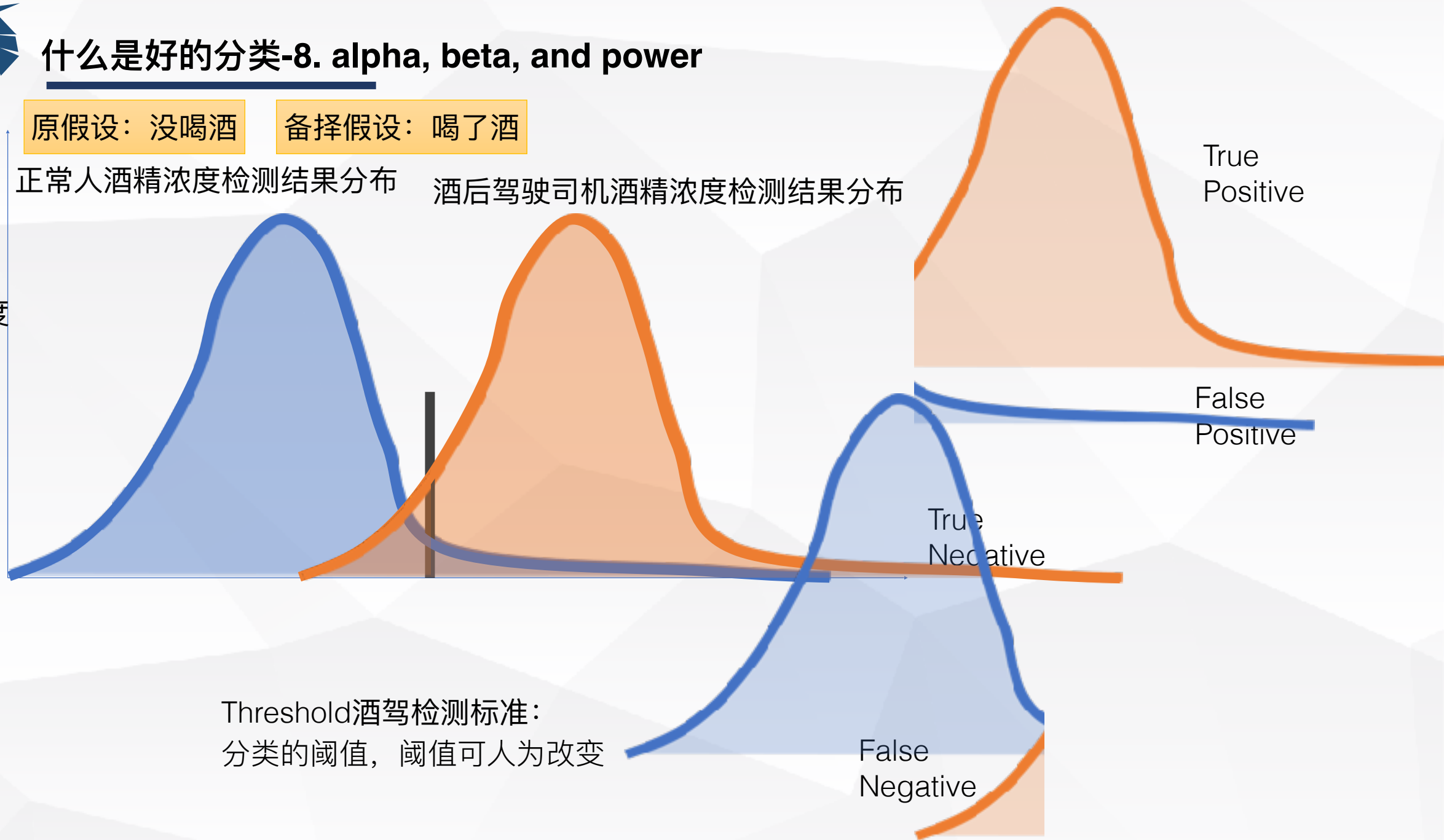
原假设：没喝酒

备择假设：喝了酒

正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

概率密度





什么是好的分类-9.ROC曲线

原假设：没喝酒

备择假设：喝了酒

正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

概率密度





什么是好的分类-10.什么样的ROC曲线更好

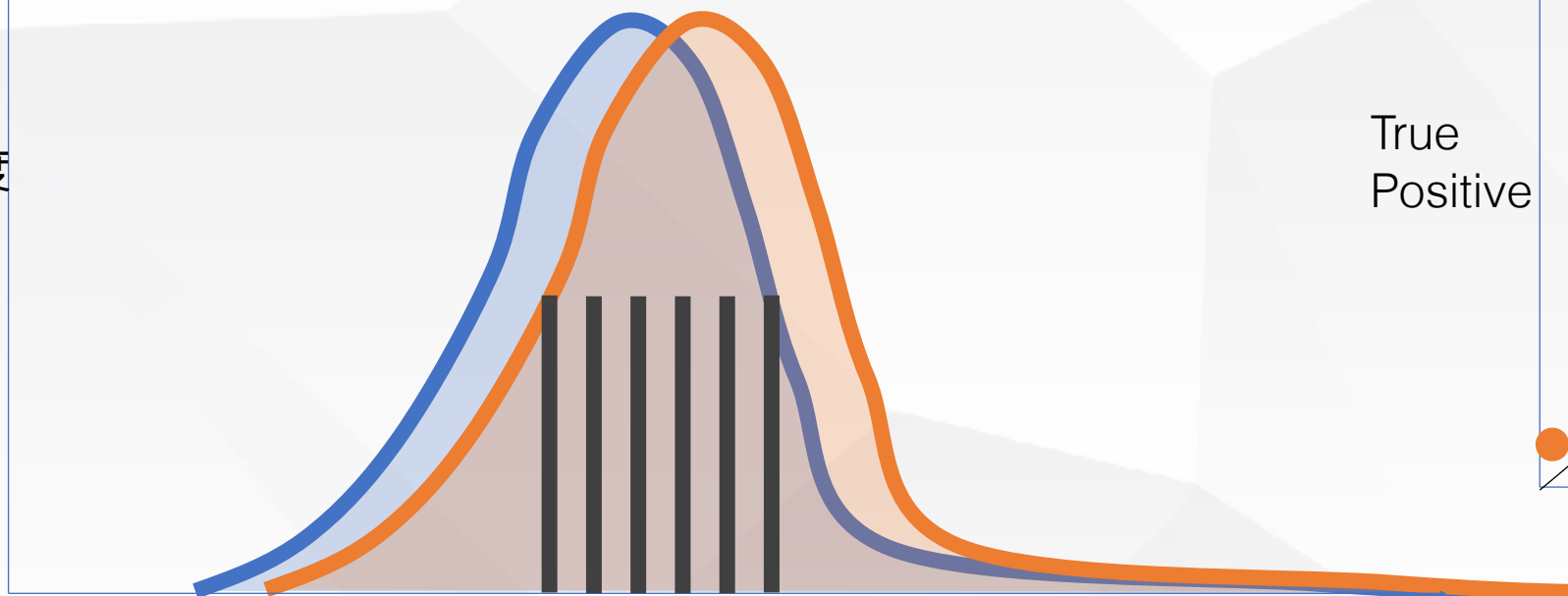
原假设：没喝酒

备择假设：喝了酒

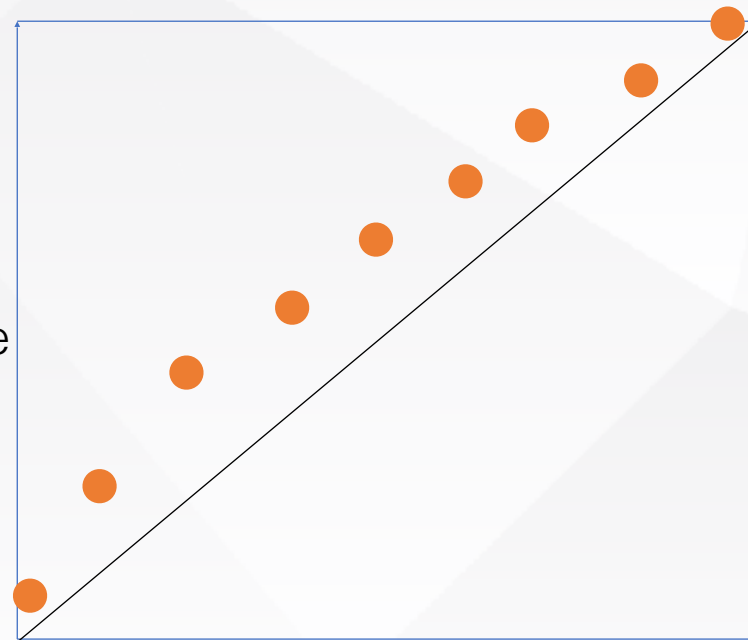
正常人酒精浓度检测结果分布

酒后驾驶司机酒精浓度检测结果分布

概率密度



True
Positive



False
Positive

1. 曲线单调增
2. 每个点都在45度线上方
3. 这条曲线离45度线越远，表示分类效果越好
4. 曲线下方的面积，就是 AUC (area under curve)



什么是好的分类-11.ROC曲线历史

- 最早用于二战时期衡量雷达识别效果(Receiver Operating Curve)
- ROC衡量的，是将分类事件转化为可衡量的量化标准的有效性
- 显著性水平衡量的是取某个分类阈值对分类的结果的不确定性



什么是好的分类-12 confusion matrix

	识别结果：喝酒了	识别结果：没喝酒
实际：喝酒了		
实际：没喝酒		

目标：抓酒驾

测试仪若显示浓度超标：Positive, 阳性

测试仪若显示浓度正常：Negative, 阴性



什么是好的分类-13.弃真存伪

Confusion Matrix混淆矩阵

	识别结果：喝酒了 (Positive)	识别结果：没喝酒 (Negative)
实际：喝酒了 (Positive)	True Positive	False Negative
实际：没喝酒 (Negative)	False Positive	True Negative

目标：抓酒驾

测试仪若显示浓度超标：Positive, 阳性

测试仪若显示浓度正常：Negative, 阴性



什么是好的分类

- 查酒驾结果，以测试仪显示结果是否 $\geq 20\text{mg}/100\text{ml}$ 为标准。
- 一共查了200人，其中，170人显示超过 $20\text{mg}/100\text{ml}$ ，其中163人证实喝酒，7人确实没喝酒。剩余30人显示低于 $20\text{mg}/100\text{ml}$ ，但交警时候发现，其中有3人也喝过酒，只是采取了一些特殊方式蒙骗了测试仪，其余27人没喝过酒。
- 怎样填写上页的矩阵？



什么是好的分类

	识别结果：喝酒了 (Positive)	识别结果：没喝酒 (Negative)
实际：喝酒了 (Positive)	True Positive 163	False Negative 3
实际：没喝酒 (Negative)	False Positive 7	True Negative 27

目标：抓酒驾

测试仪若显示浓度超标：Positive, 阳性

测试仪若显示浓度正常：Negative, 阴性



什么是好的分类

召回率，覆盖率，Recall rate：实际有多少比例被抓到

	识别结果：喝酒了 (Positive)	识别结果：没喝酒 (Negative)
实际：喝酒了 (Positive)	True Positive 163	False Negative 3
实际：没喝酒 (Negative)	False Positive 7	True Negative 27

目标：抓酒驾

测试仪若显示浓度超标：Positive，阳性

测试仪若显示浓度正常：Negative，阴性



什么是好的分类

召回率，覆盖率， Recall rate：实际有多少比例被抓到

准确度， Precision：预测喝酒的有多少是真的喝了酒

	识别结果：喝酒了 (Positive)	识别结果：没喝酒 (Negative)
实际：喝酒了 (Positive)	True Positive 163	False Negative 3
实际：没喝酒 (Negative)	False Positive 7	True Negative 27

目标：抓酒驾

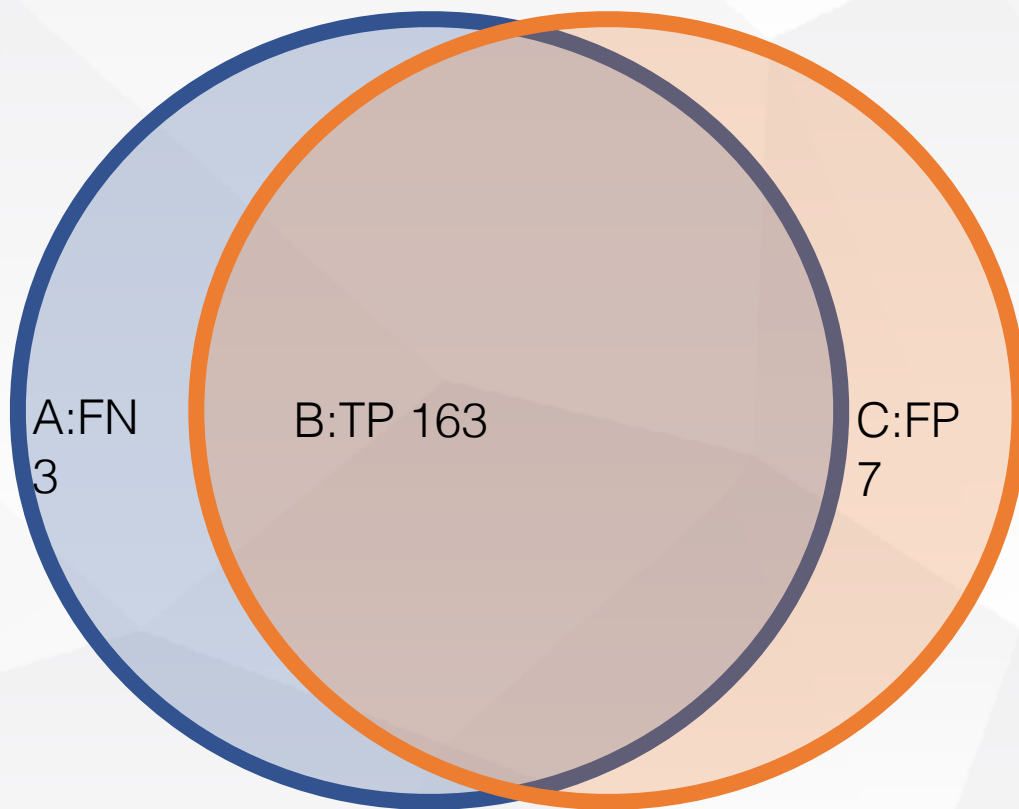
测试仪若显示浓度超标：Positive，阳性

测试仪若显示浓度正常：Negative，阴性



什么是好的分类

D:TN
27



实际喝酒

识别喝酒

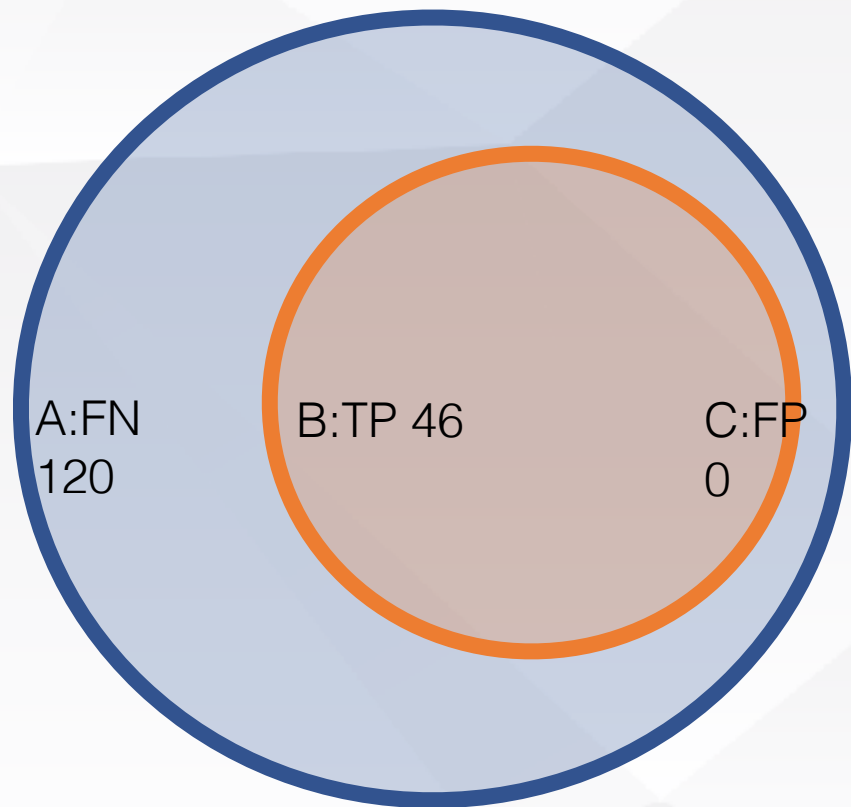
$$\text{召回率: } R = \frac{B}{A+B}$$

$$\text{准确率: } P = \frac{B}{C+B}$$



什么是好的分类

D:TN
80



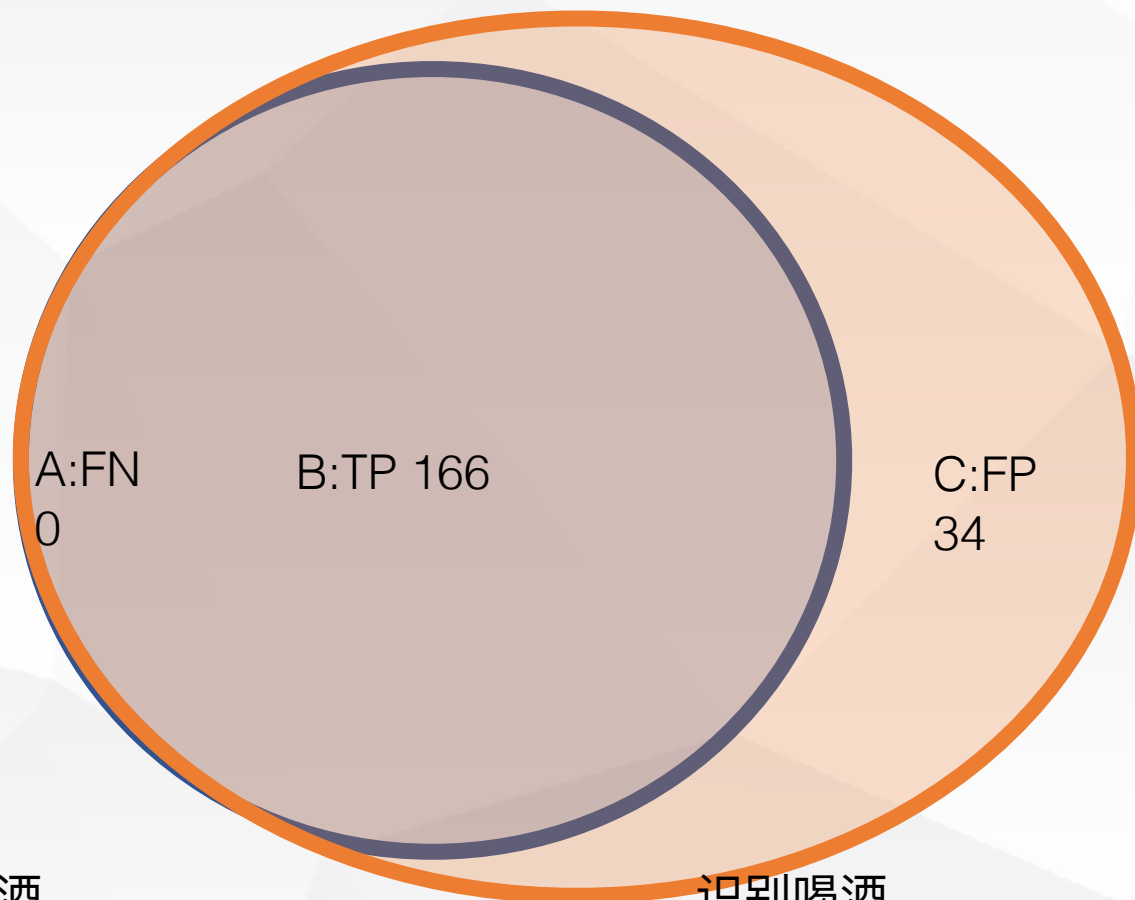
实际喝酒

检测仪1结果

识别喝酒

召回率: $R = \frac{B}{A+B} = 46/166$ 准确率: $P = \frac{B}{C+B} = 46/46$

D:TN
0



实际喝酒

检测仪2结果

识别喝酒

召回率: $R = \frac{B}{A+B} = 166/166$ 准确率: $P = \frac{B}{C+B} = 166/200$

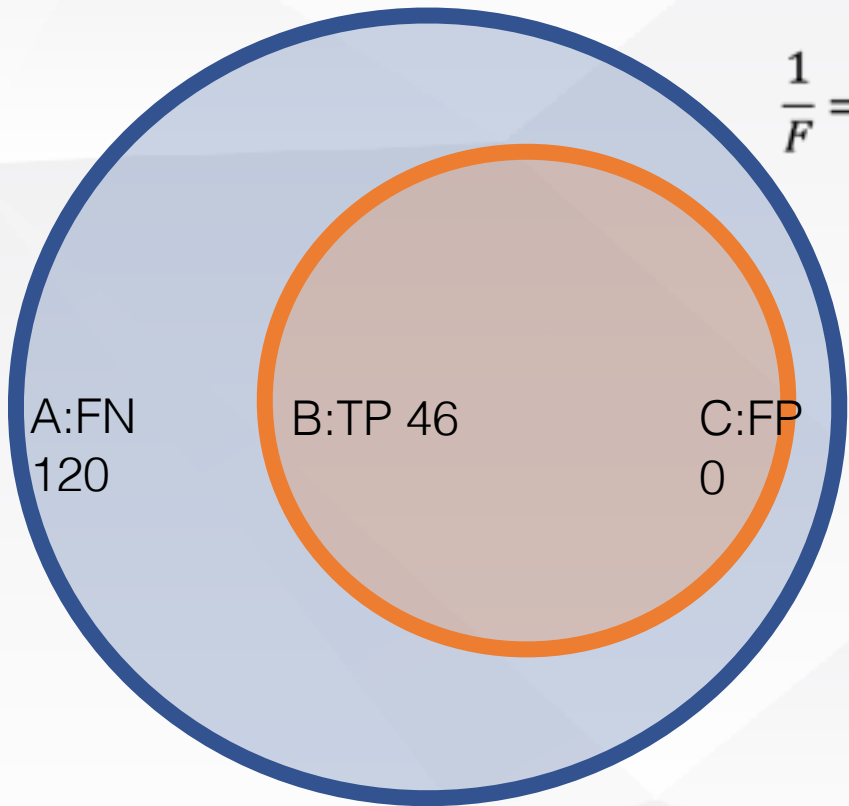


什么是好的分类

哪台检测仪好?

$$F = \frac{2 R * P}{R + P}$$
$$\frac{1}{F} = \frac{\frac{1}{R} + \frac{1}{P}}{2}$$

D:TN
80



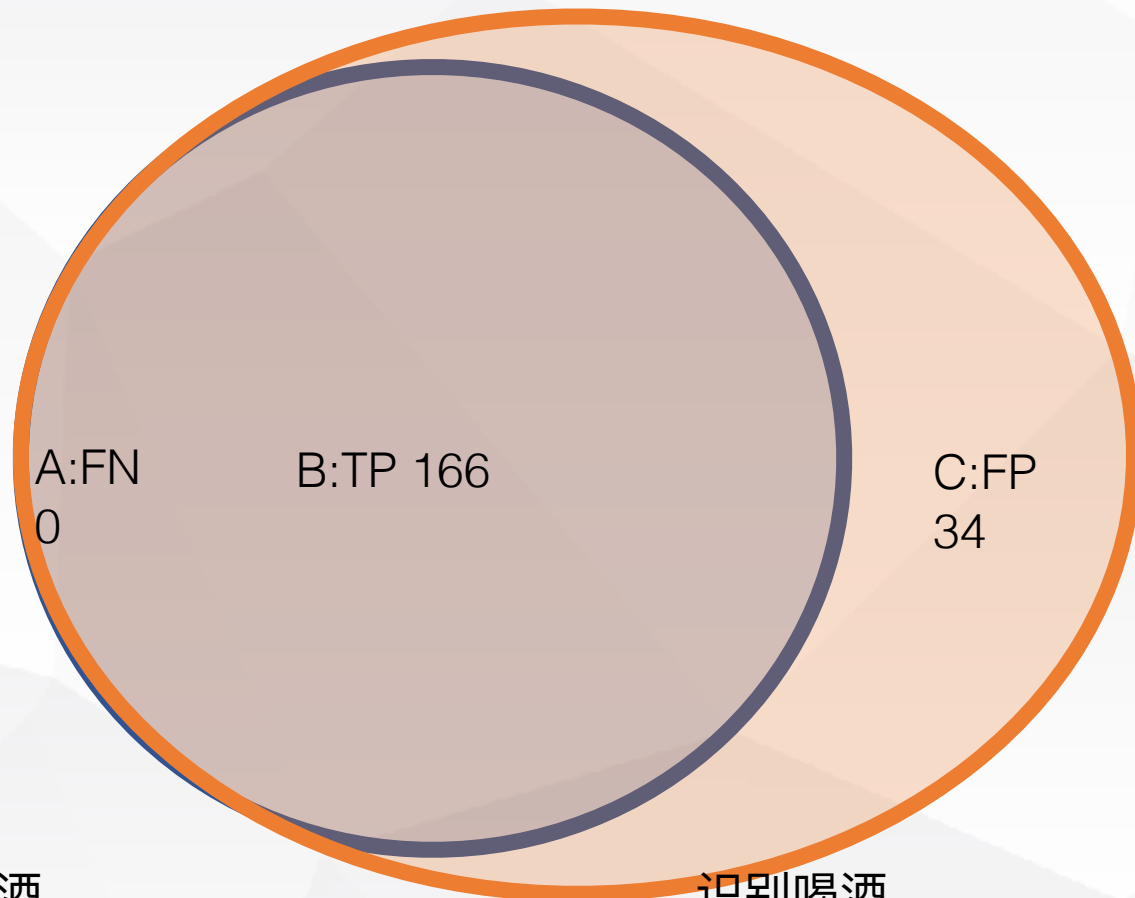
检测仪1结果

识别喝酒

实际喝酒

召回率: $R = \frac{B}{A+B} = 46/166$ 准确率: $P = \frac{B}{C+B} = 46/46$

D:TN
0



检测仪2结果

识别喝酒

召回率: $R = \frac{B}{A+B} = 166/166$ 准确率: $P = \frac{B}{C+B} = 166/200$



什么是好的分类

- 动手