

效用与决策

闫巩固 博士

本章内容

- ▣ 效度的实际意义与用途
- ▣ 一些基本概念
 - 基础率，选择率，命中率，分数线，效度
- ▣ 泰勒-罗素期望表
- ▣ 布罗格登表

效用问题

- ▣ 有些测验是实际表现的有效预测源（效标效度的有用水平），有些则不是。
- ▣ 问题是：它们的实际有用性（价值）如何？
- ▣ 测验的有用性取决于：
 - 预测的绩效或表现的价值
 - 用测验做出的每种选择决策的价值

使用一个测验可以确定

- ▣ 效度
 - 效标效度越高越好，是这样吗（假设效标度量了我们想度量的构想）？
- ▣ 基础率
- ▣ 选择率

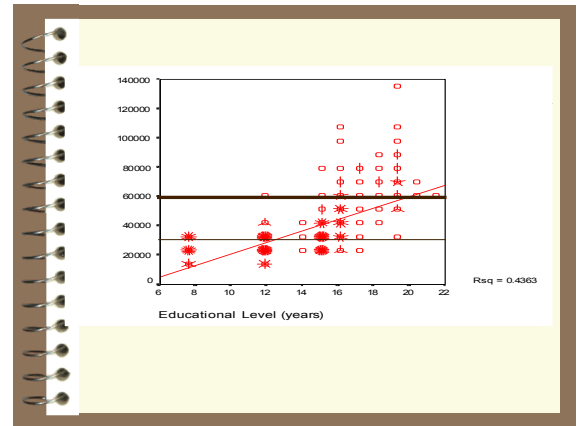
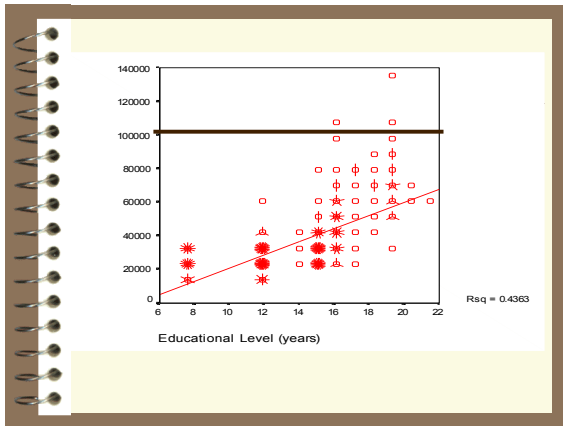


基础率(Base Rate, BR)

- ▣ 定义：在未经选择的团体中某一现象的出现率。或自然发生率。
 - 50岁的妇女患乳腺癌的机会（22/万）
 - 大学生的自杀概率
 - 北京的交通事故率
 - 月收入在20000元的机会
 - 胜任某工作的可能性
- ▣ 基础率=工作成功人数/总人数

基础率与测验的有用性

- ▣ 定义成功或某一现象
- ▣ 在效标分数上确定切分线(Cut score)
- ▣ 原则：测验要有用，必须比单纯依据基础率做出更多的正确决定。
- ▣ 当要预测的现象具有很高或很低的基础率时，基础率的重要性尤为重要（常识）
- ▣ 一般来说，基础率越接近50%，测验的贡献越大。



基础率与决策的正确性

■ 分数线(cut score or cutoff score)

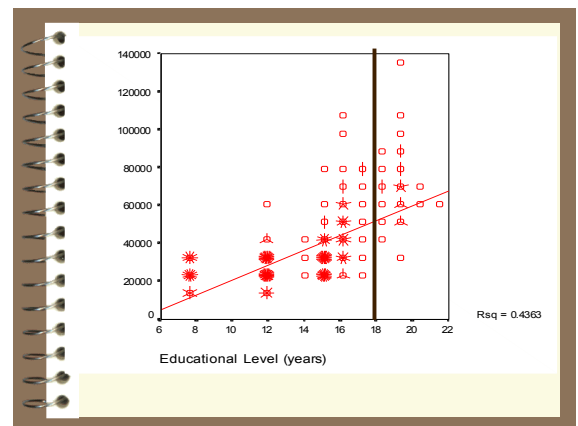
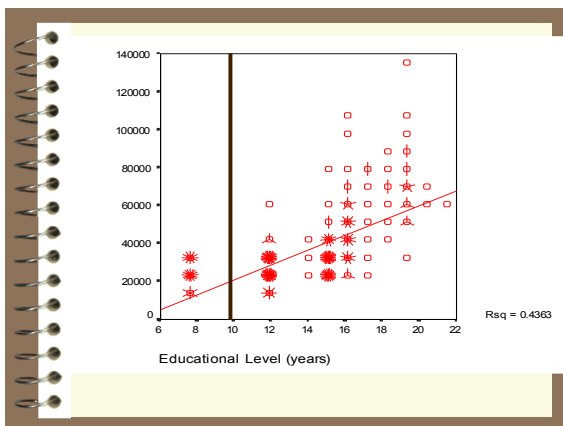
■ 改善空间

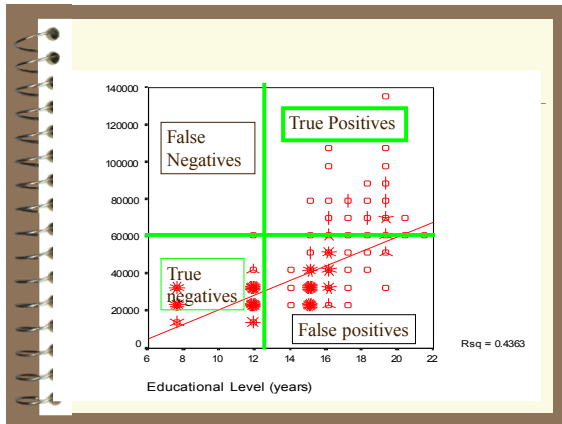
录取率(Selection ratio, SR)

■ 从申请者中选择的人数比例

■ 录取率=录取人数/总人数

■ 一般来讲，在测验效度不变的情况下，录取率越高，决策的正确性越低。反之亦然。





决策理论术语定义

- 基础率: $(A+B)/\text{总体}$
- 选择率: $(A+C)/\text{总体}$
- 敏感性: $A/(A+B)$
- 特异性: $D/(C+D)$
- 准确性: $(A+D)/\text{总体}$
- 录取效率: $A/(A+C)$
- 拒绝效率: $D/(B+D)$

		测验结果		
		录取	不录取	小计
工作实 际表现	成功	A, 38	B, 22	60
	失败	C, 7	D, 33	40
小计		45	55	100

例子:

		测验结果		
		录取	不录取	小计
工作实 际表现	成功	A, 38	B, 22	60
	失败	C, 7	D, 33	40
小计		45	55	100

A=命中 (击中)
B=漏报
C=虚报
D=命中 (正确拒绝)

基础率= $A+B=60\%$
检测率= $A/(A+B)=63\%$, 又称敏感性(sensitivity)
特异性= $D/(C+D)=82.5\%$, 特异性(specificity)
准确性=总命中率= $(A+D)/(A+B+C+D)=71\%$
录取效率=正命中率= $A/(A+C)=84\%$

员工录用问题

- 历史上, 雇主只考虑击中的人数; 即你根据测验选择的员工多少是成功的。
- 有时, 在录取率很高的情况下 (需要大量人手), 也会考虑虚报情况。
 - 如筛选飞行员, 医生, 核电站操作员
- 至少在美国, 组织越来越关注漏报问题, 为什么?

当漏报太大时,

- 在市场劳动力紧缺的情况下, 你可能失去有能力的员工;
- 人们的工作权利是相等的, 但选拔性测验往往会忽略那些有能力的妇女、少数民族以及有缺陷的人
- 你会被起诉。歧视?

测验的效用评价

- 一个测验之所以是有用的, 是因为用它比根本不用它更有可能选择出在效标上“成功”的一组人
- 有效性增加=正命中率-基础率= $84\%-60\%=24\%$
- 有两个现成的表可以帮助我们评价一个测验的效用

BR, SR与效度

- $P(TP) = BR * SR + r * ((BR * (1 - BR) * SR * (1 - SR))^{1/2})$
- $P(TP)$ = 正命中率
- BR = 基础率
- SR = 选择率或录取率
- R = 效度系数

Taylor-Russell期望表 (BR=.40, 1939)

	Selection-Rate										
	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95
0.00	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
0.05	0.44	0.43	0.43	0.42	0.42	0.42	0.41	0.41	0.41	0.40	0.40
0.10	0.48	0.47	0.46	0.45	0.44	0.43	0.42	0.42	0.41	0.41	0.40
0.20	0.57	0.54	0.51	0.49	0.48	0.46	0.45	0.44	0.43	0.41	0.41
0.30	0.65	0.61	0.57	0.54	0.51	0.49	0.47	0.46	0.44	0.42	0.41
0.40	0.73	0.69	0.63	0.59	0.56	0.53	0.50	0.48	0.45	0.43	0.41
0.50	0.81	0.76	0.69	0.64	0.60	0.56	0.53	0.49	0.46	0.43	0.42
0.60	0.89	0.83	0.75	0.69	0.64	0.60	0.55	0.51	0.48	0.44	0.42
0.70	0.95	0.90	0.82	0.76	0.69	0.64	0.58	0.53	0.49	0.44	0.42
0.80	0.99	0.96	0.89	0.82	0.75	0.68	0.61	0.55	0.49	0.44	0.42
0.90	1.00	1.00	0.97	0.91	0.82	0.74	0.65	0.57	0.50	0.44	0.42
0.95	1.00	1.00	0.99	0.96	0.87	0.77	0.66	0.57	0.50	0.44	0.42
1.00	1.00	1.00	1.00	1.00	1.00	0.80	0.67	0.57	0.50	0.44	0.42

成功比例

Taylor-Russell期望表 (BR=.60, 1939)

效度	.1	.3	.5	.7	.9	选择率
.00	.60	.60	.60	.60	.60	
.10	.67	.64	.63	.62	.61	
.30	.79	.73	.69	.66	.62	
.50	.90	.82	.76	.70	.64	
.70	.97	.91	.84	.75	.66	
.90	1.00	.99	.94	.82	.67	
成功比例						

泰勒-罗素期望表

- 低录取率加上高效度：导致测验选择出成功个体的比例更高
- 低（或无）效度或高录取率：导致测验选择出很少成功的个体

例子：

- 假设你是一家公司的人事部经理。有100个人申请某一职位。根据以往公司招聘员工的资料，你了解到申请该职位的平均成功率为60%。现在你手头上有个针对该职位的测验，效度为.7，而因额度限制，你只能招收30人。
- 你将预期招到的人中有多少是成功的？
- 如果胜任该职位但未被录取的人提出诉讼，你将如何给出合理的解释？

		决 策		
		选择	拒绝	小计
绩	成功	27	33	60
	失败	3	37	40
效 小计		30	70	100

根据泰勒-罗素表，正命中率=.91。91*30=27.3
在所有被拒绝的人中漏报的比率=33/70=.47

布罗格登 (Brogden) 表

不使用基础率

估计被选择组的平均效标Z分数

$$r_{xy} = \text{实际节省} / \text{最大节省} = (S - U) / (P - U)$$

- S=根据测验分数选择出来的人的平均绩效水平

- U=随机选取的人的平均绩效水平，即效度为0。

- P=完美选择的人的平均绩效水平，即效度为1。

当 $r=0$ 时，等于没有用任何测验，被选者的成功率等于基础率，被选者的平均效标分数等于全体申请者的平均效标分数（ $Z=0$ ）；当 $r_{xy}=1$ 时，相当于受测者的实际效标成绩选人，其被选者的平均效标分数最高。

例子：

一公司须选择50名销售员，有200人报名。

对所有人进行测试。

从200人中随机选50名，定义为U

从剩下的150人中选出测验得分最高的50名，构成S。

- 雇佣所有200人。计算U和S组的每月平均销售额。假设U=30，S=40。

- 最后从所有200名中选出每月平均销售额最高的50名，形成P组。假设P=60。

$$r_{xy} = (S - U) / (P - U) = (40 - 30) / (60 - 30) = 1/3$$

绩效增加或收益

Brogden表

(1946,1949)

效度	.1	.3	.5	.7	.9	选择率
.00	.00	.00	.00	.00	.00	
.25	.44	.29	.20	.12	.05	
.50	.88	.58	.40	.25	.10	
.75	1.32	.87	.60	.37	.15	
1.00	1.76	1.16	.80	.50	.20	

效标的平均Z分数

低录取率加上高效度：导致测验选择出具有更高效标分数的个体（平均）

低（或无）效度或高录取率：导致测验选择出具有更低效标分数的个体（平均）

一个测验之所以是有用的，是因为用它比根本不用它更有可能选择出在效标上“成功”的一组人

效用理论

一种在选拔中运用测验估计生产力(收益/年)的方法。以金钱数量表示。

效用最大化:

- 预测源的效度：高

- 选择率：低

- 基础率：接近50%

- 选拔项目的成本：低

- 差的选拔（招聘，培训、低生产力）的成本：高

$$\text{生产率收益} = K * r * SD_y * Z_s$$

K=被选择者数量

R=效度系数

SDy=校标的标准差

Zs = 被选择者的平均测验分数

选拔效用的例子

■ 通用电器（GE）
公司流水线工人
研究

■ 使用两个测验：
一般能力和机械
能力

	低于平 均	平均水 平	高于平 均
两个测 验均高 于平均	18%	38%	65%
一个测 验均高 于平均	21%	35%	29%
两个测 验均低 于平均	61%	21%	6%