拒绝域

假设检验的基础原理，是构造“拒绝域”。

首先，我们有一个原假设 H_0，然后，根据 H_0，在状态空间 Ω 中取一个子集 $\mathcal{R}$，如果采样 $X \in \mathcal{R}$，则拒绝假设 H_0。

分析这个过程：H_0 是固定的，根据 H_0 确定的 $\mathcal{R}$ 也是固定的。随机性是在哪里引入的呢？是在“采样”那一步。

这里就是第一个混淆点：从直观的角度出发，所谓假设检验，最好能够求出“假设成立的概率”，即 $P(H_0 \mid X)$。但是这个概率是不存在的，因为 H_0 是确定的，其是否成立也是确定的，对它求概率是没有意义的。

有意义的概率一定来自于 X。根据假设实际上是否成立，以及 X 是否落入拒绝域（是否拒绝），有如下四种概率：

1. $P(X \in \mathcal{R} \mid \overline{H_0})$：$H_0$ 不成立，正确地拒绝了 H_0。
2. $P(X \in \mathcal{R} \mid H_0)$：$H_0$ 成立，错误地拒绝了 H_0。
3. $P(X \notin \mathcal{R} \mid \overline{H_0})$：$H_0$ 不成立，但是未能拒绝 H_0。
4. $P(X \notin \mathcal{R} \mid H_0)$：$H_0$ 成立，且未能拒绝 H_0。

不难看出，如果情况都是 1 和 4，那假设检验就非常成功。第二个概率对应的错误被称作：**第一类错误**，即错误地拒绝了成立的假设；第三个概率对应的错误被称作 **第二类错误**，即未能拒绝错误的假设。

> 在贝叶斯学派中，我们不认为 H_0 是确定的，此时可以使用贝叶斯公式计算 $P(H_0 \mid X)$：
>
> $$P(H_0 \mid X) = \frac{P(X \mid H_0)P(H_0)}{P(X)}$$
>
> 此时遇到一个困难：$P(H_0)$ 是什么？这是“不考虑采样时 H_0 成立的概率”，被称作“先验概率”，**是可以人为选择的**，反映我们对 H_0 最初的判断。而运用公式计算 $P(H_0 \mid X)$，就是在得到采样后根据采样修正对 H_0 的判断，得到的概率被称为“后验概率”。

接下来，一个自然的问题是：怎么设计这个拒绝域？

设计的原则是**控制第一类错误发生的概率**。正确的拒绝域设计，需要保证 $P(X \in \mathcal{R} \mid H_0) \leq \alpha$。此时，$\alpha$ 被称作**显著性水平**，它代表了第一类错误的概率上界。

自然的疑问是：第二类错误呢？这就是假设检验的偏心之处：它对第二类错误完全没有限制。这带来一个结果：

只有得到拒绝结果的假设检验是有说服力的。

如果得到的结果是 $X \notin \mathcal{R}$，因为我们对 $P(X \notin \mathcal{R} \mid \overline{H_0})$ 完全未知，所以此时，对 H_0 做出任何判断都是不负责任的。

所以上文始终没有提到“接受”假设。准确来说，“接受”应该是“未能拒绝”。

但是一般来说，假设检验的最终目标应该是支持某个假设，而不是反对某个假设。如果我们能够拒绝原假设，实际就是支持了原假设的对立面，也即 $\overline{H_0}$，记作 H_1。这被称作**备择假设**。在假设检验过程中，**备择假设才是我们想支持的假设**。

它们两个的关系非常有趣：拒绝 H_0 能推出认同 H_1，但未能拒绝 H_0 却不能拒绝 H_1。

正态参数的假设检验

点值假设

对于等式类型的假设，一般来说，只能把 $\theta = \theta_0$ 作为原假设。为什么呢?

因为备择假设 $\theta \neq \theta_0$ 对应的情况是无界的，但是 $\mathcal{R}$ 要求 $P(X \in \mathcal{R} \mid H_0)$ 有上界，难以构造合理的 $\mathcal{R}$。

这个时候，我们能够得到这样的结果:

- 拒绝假设，“说明 θ 和 θ_0 有差异”
- 接受假设，“不能说明 θ 和 θ_0 有差异”，但不是说能够说明无差异。

单侧假设

对于单侧型的假设，以 $\theta \leq \theta_0$ 为例，构造拒绝域的方法是：根据临界值 θ_0 构造拒绝域，使得 $P(X \in \mathcal{R} \mid \theta = \theta_0) = \alpha$。构造时要保证概率对 θ 的单调性，使得 $P(X \in \mathcal{R} \mid \theta \leq \theta_0) \leq \alpha$。

但是这个时候就出现问题：我应该选择 $H_0 : \theta \leq \theta_0$ 为原假设还是 $H_1 : \theta \geq \theta_0$ 为原假设？会不会出现矛盾的情况?

不妨都试一下:

拒绝 H_0，未能拒绝 H_1

这很简单，既然拒绝了 H_0，我们可以支持 H_1。

未能拒绝 H_0，拒绝 H_1

同样地，我们支持 H_0。

未能拒绝 H_0, H_1

我认为这是不矛盾的。这说明我们的采样还不够多，不足以对假设做出判断。

在答题的时候，我们回答“不能认为 H_0 成立”，并不隐含“能认为 H_1 成立”。

同时拒绝 H_0, H_1

这可能发生吗?

在我们使用的假设检验方法中，只要 $\alpha < 0.5$，就不可能出现这种情况。证明:

我们构造的拒绝域 $\mathcal{R}$ 总是满足如下形式:

$$\mathcal{R}_0 = \{x \mid f(x) \geq F_\alpha, f \sim F(\theta_0)\}$$
$$\mathcal{R}_1 = \{x \mid f(x) \leq F_{1-\alpha}, f \sim F(\theta_0)\}$$

其中 $\mathcal{R}_i$ 是假设 H_i 的拒绝域。由于分界点 θ_0 相同，F_α 和 $F_{1-\alpha}$ 是同一分布的分位数，在 $\alpha < 0.5$ 时，必然有:

$$F_{1-\alpha} < F_\alpha$$

那么:

$$\mathcal{R}_0 \cap \mathcal{R}_1 = \{x \mid F_\alpha \le f(x) \le F_{1-\alpha}, f \sim F(\theta_0)\} = \varnothing$$

所以同时拒绝两个假设在单侧假设的情况下是不可能发生的。

但是，如果假设变得复杂，比如多个分段区间；或者采用了奇怪的拒绝域设计，还能够保证“不会同时拒绝两个假设”吗?

我不太清楚了。我个人偏向认为存在反例。但是出现这种情况意味着什么？也许说明检验方法设计的比较失败?

总结

对于目前能够遇到的问题，可以通过这样的流程得出结论:

1. 尝试所有可能的原假设
2. 如果得到拒绝结果，则支持对应的备择假设
3. 如果所有假设都无法拒绝，则认为“无法支撑题设”。但是，这不代表“支撑题设的反面”。