

8 - 假设检验

在科学中，假设是一个可以通过实验证明或反驳的陈述。

例子 加利略·伽利莱 (1564-1642) 验证了这样的假设：任何一对从同一高度落下的物体到达地面的时间相同。

为了验证这一点，伽利略从比萨斜塔从不同大小的球体上方将它们释放。每次，伽利略都看到这些球同时落地。

在统计学中，假设是：

例如，一包Old Dutch番茄酱味薯片的平均质量为255克。

我们用来检验统计假设的证据是：

- 要检验关于总体均值 μ 的主张，我们用来自样本数据的_____。
- 要检验关于总体比例 p 的主张，我们用来自样本数据的_____。
- 要检验关于总体方差的一个主张 σ^2 ，我们从样本数据中使用_____。

在统计学中，假设检验是一种正式程序，使用样本统计量来检验关于相应总体参数的断言。

假设检验基于以下原理：

只有当样本数据提供的支持不太可能是偶然产生时，才足以支持一个假设。

Example 假设我掷硬币6次，出现了5次“正面”。我能因此断定这枚硬币偏向正面吗？

Answer: No. 抛出 5 个或更多正面的概率对于公平硬币是： $\frac{6+1}{2^6} = 0.1094$ 由于 $0.1094 > 5\%$ ，我们说在 5% 的显著性水平下证据不具有统计显著性。据我们所知，这枚硬币并未偏向正面。

1094 > 5%，我们说在 5% 的显著性水平下证据不具有统计显著性。据我们所知，这枚硬币并未偏向正面。



对于 μ 的拒绝域法

Example (Two-Tailed) 网络工程师声称通过网络传输一个 1 Gb 文件的平均时间为 12.44 秒。你通过在 10 个随机时刻发送一个 1 Gb 文件并记录以下传输时间来检验这一断言。在显著性水平为 5% 下检验工程师的说法。

12.0	13.5	11.8	12.4	13.8
11.9	12.2	12.9	13.0	12.3

第 1 步：（陈述主张）— 哪个总体参数？主张是什么？

第 2 步：（假设）— 记录原假设和备择假设

步骤 3：（检验统计量）——计算合适的检验统计量

步骤 4: (拒绝域) —— 哪些检验统计量的取值会让我们拒绝 H_0 ?

第 5 步: (判定) – 拒绝 H_0 还是 未能拒绝 H_0 ?

Step 6: (结论) —— 使用原始主张的措辞陈述你的结论。

Example (Left-Tailed) 一周后，网络工程师声称将 1 Gb 文件通过网络传输的平均时间已被降低到低于10 秒。你通过在 5 个随机时间点发送一个 1 Gb 文件并记录以下传输时间来检验该主张。以 5% 的显著性水平检验工程师的主张。

9.0	8.5	10.1	8.7	9.5
-----	-----	------	-----	-----

p -检验方法用于 μ

定义 给定样本统计量的 p 值是指在原假设为真的前提下，随机获得与该值一样或更极端的样本统计量的概率。（此处“极端”是指 朝向备择假设所指定方向。）

如果 p 值低于 α ，我们会拒绝 H_0 。

示例（右尾检验） BCIT 的一项研究旨在确定变量的平均值：

X = amount of time a BCIT student works (at a job) each week

该研究对学生的随机样本进行了数据收集。样本统计如下：

$$n = 100; \bar{X} = 11.2; s = 4.5$$

在5%显著性水平下检验这一主张：平均 μ 大于每周10小时。

关于总体 p 的推断

关于 p 的推断基于以下原则：

样本比例 p 在满足 $np \geq 5$ 和 $nq \geq 5$ 时，近似服从正态分布。

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

示例 有46位历任美国总统中有7位是左撇子。一般人群中，10%的人是左撇子。检验这一主张：美国总统是否来自一个比一般人群更可能为左撇子的总体。

Example (Wald Conf. Interval) 假设对 $n = 1000$ 名随机抽取的加拿大选民进行的一项民意调查显示, $X = 512$ 将投票给蓝色政党。

a. 确定加拿大人中将投票给蓝党的人口比例的95%置信区间。

b. 使用传统方法检验假设：一半选民将投票给蓝党。

关于步骤1和2（主张与假设）的笔记

- 虚无假设始终涉及相等。它对该主张/问题给出一个默认答案。
- 备择假设 总是涉及不等式。
- 样本统计量（即 p 、 \bar{X} 、 s^2 ）*never* 不出现在假设中。

Example 将下列每一项陈述翻译为适当的假设。

- a) BCIT 学生的平均年龄为 24.2 岁。
- b) 这枚硬币至少有 50% 的概率正面朝上。
- c) 加拿大人的平均收入低于 \$100 000。
- d) 普通加拿大人的受教育年限最多为15.5年。

关于第6步（陈述结论）的说明

If the original claim was H_0 and we 拒绝 H_0 :

“证据足够充分，可以拒绝……的假设”

If the original claim was H_0 and we 未能拒绝 H_0 :

“证据不足以驳回该主张……”

如果原始主张是 H_1 而我们 拒绝 H_0 :

“证据足以接受该主张……”

如果原始主张是 H_1 而我们 未能驳回 H_0 :

“证据不足以接受……这一主张。”

假设检验中的错误

第一类错误——在原假设实际上为真时却拒绝它。

第二类错误——在原假设实际上为假时未能拒绝原假设。

示例 在加拿大的法院中，除非被证明有罪，否则你被推定为无罪。

H_0 :

H_1 :

第一类错误是

第二类错误是

		真实情况	
		H_0 为真	H_0 为假 (H_1 为真)
决定	未能拒绝 H_0		
	拒绝 H_0		

我们如何降低犯第一类错误的概率？

我们如何降低犯第二类错误的概率？