# **p值是什么以及如何计算**

1. 杨程瑜 20202241107

## **1 p值是什么**

### **1.1 p值的历史**

1925年，英国遗传学家兼统计学家罗纳德·爱尔默·费希尔出版了《研究者的统计方法》（Statistical Methods for Research Workers）一书。这本书的书名在当时看起来并不会“畅销”，但实际上这本书却取得了巨大的成功，而且还使费雪成为现代统计学之父。在这本书中，他着眼于研究人员如何将统计检验理论应用于实际数据，以便基于数据得出他们所发现的结论。当使用某个统计假设来做检验时，该检验能够概述数据与其假设的模型之间的兼容性，并生成一个p值。 费雪建议，作为一个方便的指南，研究人员可以考虑将p值设为0.05。对于这一点，他专门论述道：“在判断某个偏差是否应该被认为是显著的时候，将这一阈值作为判断标准是很方便的。”他还建议，p值低于该阈值的结论是可靠的，因此不要把时间花在大于该阈值的统计结论上。因此，费雪的这一建议诞生了p小于0.05等价于所谓的统计显著性，这成了 “显著”的数学定义。

### **1.2 p值的定义**

p值是假设检验中**假设零假设为真时观测到至少与实际观测样本相同极端的样本的概率**。很小的p值说明在零假设下观测极端结果的很小。

### **1.3 p值的作用**

p值的作用：p值就是用来**判断H0假设是否成立的依据**。因为期望值是基于H0假设得出的，如果观测值与期望值越一致，则说明检验现象与零假设越接近，则越没有理由拒绝零假设。如果观测值与期望值越偏离，说明零假设越站不住脚，则越有理由拒绝零假设，从而推出对立假设的成立。

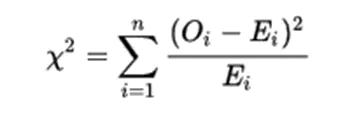
## **2 p值如何计算**

p值又叫概率值，是让科学家验证猜想是否成功的一个数值。这个值和显著性水平有直接关系，后者是验证科学研究的数据是否有统计上的显著性的重要指标。可以画个表，在计算其他统计值以后来找出P值。如果要计算其值，卡方是其中要先找出的统计值之一。

### **2.1 理论过程**

#### **2.1.1 计算chi-square（卡方），比较两组数据**

chi-square（卡方）的计算公式如下



其中o是观察数据或实际数据，e是期望值。

#### **2.1.2 决定数据自由度**

自由度是研究中涉及的变量个数，变量个数受到总类别数的限制。

#### **2.1.3 选择显著度**

显著度由计算者自己决定，一般定为0.05，或5%。这表示5%的总时间内实验结果会因为完全随机而偏离你的假设，而剩余95%的绝大部分时间内都符合你的假设。

#### **2.1.4 用一个卡方分布来找到p值**

卡方分布表会基于自由度和卡方值给出p值，但是通常来说不是所有卡方值都在分布表上有，所以找出最接近的卡方值即可。

### **2.2 实际例子**

如果将上述概念性的过程套用到实际例子中，比如抛硬币实验，我们就更好理解p值是如何计算出来的

比如我们想知道我们抛的硬币是否是正常的硬币（正反概率应各为50%），那么在掷硬币这个例子中，我们可以设定：

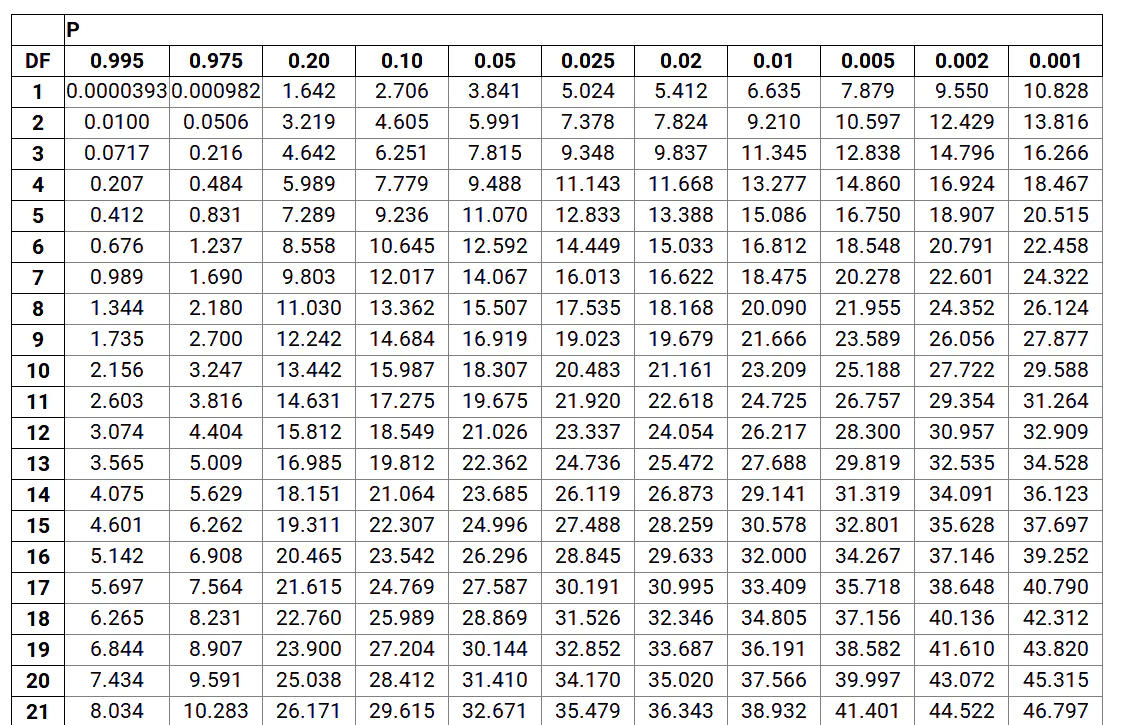
H0: 手中的硬币是常规硬币

H1: 手中的硬币做过手脚

如果手中硬币是常规硬币，我们知道正面和反面出现的概率各为50%，所以如果我投掷10次硬币，则正面和反面出现的次数各位5次。正面5次，反面5次，就是我们对于投掷10次硬币的期望值（expected value）。

现在我们开始投掷硬币，出现的是正面3次，反面7次。这个结果就是我们对于投掷10次硬币的观测值（observed valued），即实际的结果。

通过分析期望值和观测值的差距，我们就可以判断出硬币是否正常。而这个期望值和观测值差距的判断方法就是chi-square。

那么在此例子中其卡方为(3-5)^2/5+(7-5)^2/5=1.6

上图即为卡方分布表，左上角的α表示错误拒绝H0假设的概率（即虚无假设事实上成立，但我们计算出的结果却错误判断虚无假设不成立的概率）。n代表自由度（degree of freedom），即独立变量数减1，在这个例子中，独立变量数为2（正面和反面），所以自由度为1（2-1=1）。

假设置信度为95%，即错误拒绝H0的概率为0.05。展开解释就是，我们有95%的概率确信检验结果正确，有5%的概率会错误拒绝虚无假设。（我们总说的p值与0.05比较就是这个啦，其实不一定时0.05，根据具体情况可以设置不一样的值，只是大部分时候都用0.05）

对照着卡方分布表，找到1所在的行（我们计算出的chi-square自由度是1），发现1.6是介于1.323和2.706之间，查表得出其p值为0.25到0.1之间，大于0.05，所以我们不能拒绝H0。换句话说，H0成立，即硬币是常规硬币，没有做手脚。

总的思路是，做出H0,H1这对互斥的假设，计算出H0为真时的期望值，统计出实际的观测值，通过期望值和观测值求得chi-square（卡方），再通过卡方查表，得到p值。根据p值与α（1-置信度）的比较，如果p-value<α，则拒绝H0，推出H1成立；如果p-value>α，则接受H0，推出H1不成立。