

假设给人们吃三种不同的食物，然后做测试，分数如下：

食物1	食物2	食物3
3	5	5
2	3	6
1	4	7

我们想知道的是：食物类型是否会影响分数。是否有总体均值1=总体均值2=总体均值3。如果不相等就意味着不同事物对人们的测试结果有影响。

先建立假设

$H_0$ ：均值相同，事物不会产生差别

$H_1$ ：均值不等，事物会产生差别

$\alpha = 10\%$

在零假设成立的前提下，求出F统计量到达如此极端的概率是多少

F统计量是组间平方和同除以其自由度，有时也称为组间均分（mean squares between, MSB），然后除以组内平方和除以其自由度：


$$F\text{-statistic} = \frac{\frac{SSB}{m-1}}{\frac{SSW}{m(n-1)}}$$

如果分子比分母大得多，那就说明波动大多数来自于各组之间。此时我们应该相信，总体均值之间存在差异。如果这个值很大，那就意味着零假设成立的概率较低。如果值很小，意味着组内波动比组间波动在总波动中占比更多，这就意味着，差异可能只是随机产生的。

以上面的表为例：

$$F = \frac{\frac{24}{2}}{\frac{6}{6}} = 12$$

F分布实际上是两个卡方分布之比。分子的自由度是2，分母的自由度是6，查表：



\	df <sub>1</sub> =1	2	3	4	5	6	7	8	9	10	12
df <sub>2</sub> =1	39.86346	49.50000	53.59324	55.83296	57.24008	58.20442	58.90595	59.43898	59.85759	60.19498	60.701
2	8.52632	9.00000	9.16179	9.24342	9.29263	9.32553	9.34908	9.36677	9.38004	9.39157	9.401
3	5.53832	5.46238	5.39077	5.34264	5.30916	5.28473	5.26619	5.25167	5.24000	5.23041	5.219
4	4.54477	4.32456	4.19086	4.10725	4.05038	4.00975	3.97897	3.95494	3.93567	3.91988	3.899
5	4.06042	3.77972	3.61948	3.52020	3.45298	3.40451	3.36790	3.33928	3.31628	3.29740	3.268
6	3.77395	3.46330	3.28876	3.18076	3.10731	3.05455	3.01446	2.98304	2.95774	2.93693	2.904
7	3.58943	3.25744	3.07407	2.96053	2.88334	2.82739	2.78493	2.75158	2.72468	2.70251	2.668
8	3.45792	3.11312	2.92380	2.80643	2.72645	2.66833	2.62413	2.58935	2.56124	2.53804	2.501

这个临界值是3.46。而我们得到的值是12。零假设前提下，得到如此极端的值的概率非常低。我们拒绝零假设，给人们不同事物，得到的分数可能会有很大差异。