

## 8 データの検証

### 仮説検定の考え方



A社で新しい飲料を開発し、最終候補としてPとQが残った。消費者30人を偏りなく選んでどちらを好むか調査したところ、Pを好む人が24人、Qを好む人が6人であった。

一般に、Pの方が好まれそうだと判断してもよいか考えてみよう。

PとQを好む人は半々であると仮定する。このとき、30人中24人以上がPを好むと答えることがどのくらい起こり得るかを調べてみる。

そのために、表裏が同じ割合で出る30枚の硬貨を同時に投げて表が出た枚数を記録することを1000回行った。右の表はその結果である。この表から、24回以上表が出る割合を求めると、

表が出た枚数	回		
0~6	0	16	134
7	1	17	131
8	2	18	99
9	14	19	46
10	22	20	35
11	50	21	14
12	63	22	5
13	115	23	3
14	118	24	2
15	145	25	1
		26~30	0

$$\frac{2}{1000} + \frac{1}{1000} = \frac{3}{1000} = 0.003 \quad 0.3\%$$

となり、ほとんど起こり得ない。

よって、Pの方が好まれると判断することが妥当である。

**注** ある場合が起こりにくいかどうかは、その場合を含めてより外れた場合全体が起こりにくいかどうかで判断する。

このように、あるデータが与えられたとき、仮説を立て、それが妥当かを判定する統計的手法を**仮説検定**という。仮説検定では、最初に立てた仮説Aを否定する仮説Bを考える。仮説Bを前提とするとき、与えられたデータが得られる割合が極めて小さければ、仮説Bは否定され、最初の仮説Aは妥当といえる。そうでなければ、結論を保留する。



では、仮説Aは「Pの方が好まれる」、仮説Bは「PとQを好む人の割合は同程度である」となる。