

## 2. 統計的な検定の意味

e-stat（政府統計ポータルサイト）にある、国民健康・栄養調査の2018年のデータを見ると、20歳男性の平均身長は169.1センチメートル、20歳女性の平均身長は157.0センチメートルである。両者は統計的に差がある（より厳密に言えば、差がないとは言えない）と言えるだろうか。平均値だけを見れば、169.1と157.0には12.1の差がある。ここで問題なのは、「統計的に」という言葉である。

データを見ると、20歳男性の平均は15人の身長データで、20歳女性の平均は17人の身長データでそれぞれ計算されている。また、19歳のデータを見ると男性20人の平均は174.0センチメートル、女性15人の平均は156.6センチメートルであり、20歳男女の平均と異なっている。男性に関しては約5センチメートルも差がある。2018年当時の19歳と20歳との間で成長期に栄養状態が異なっていたと考えるのは難しく、20人程度の平均値であれば5センチメートル程度の平均値の差はおそらく偶然発生したと考えるのが自然だろう。一方、20歳男女の約12センチメートルの差については、男女で身長差があるのは当たり前だと考える人がほとんどだと思われる。

統計的に意味があるかどうかを考える上で重要なのは、分析の結果（ここでは平均値を計算した結果）が、偶然発生したものか、あるいは偶然それが生じる確率は極めて低いため偶然とは言い難いものか、という違いである。後者の場合を統計的に意味のある差（有意な差）と表現する。相関係数も同様である。世の中には偶然相関が発生する事象も存在するが、それらが偶然とは言い難いほど高い相関かどうかを確かめることが統計的な検定である。

では、偶然か偶然とは言えないかをどのように判断したらよいのだろうか。あいこは数えず、じゃんけんを100回繰り返すことを想像してもらいたい。じゃんけんの勝ち負けは、いかさまや相手の出し手のパターンが分かってしまう場合を除けば、五分五分である。そうであれば、約50回勝ち、約50回負けるのが平均的である。しかし、時には55回も勝ったり45回しか勝てないことも偶然起こり得るだろう。この100回のじゃんけんを、1,000組のペアで実施してみると、50回前後勝つ（負ける）ペアが最も多くなり、40回勝つ（負ける）ペアや30回勝つ（負ける）ペアは急速に少なくなることが想像できるだろう。これを図示すると図15-1のようになる。図15-1の横軸が100回じゃんけんの勝利数、縦軸はそれが起きたペアの数を示している。

図15-1 100回じゃんけんを1,000ペアで実施した場合の勝利数とペア数

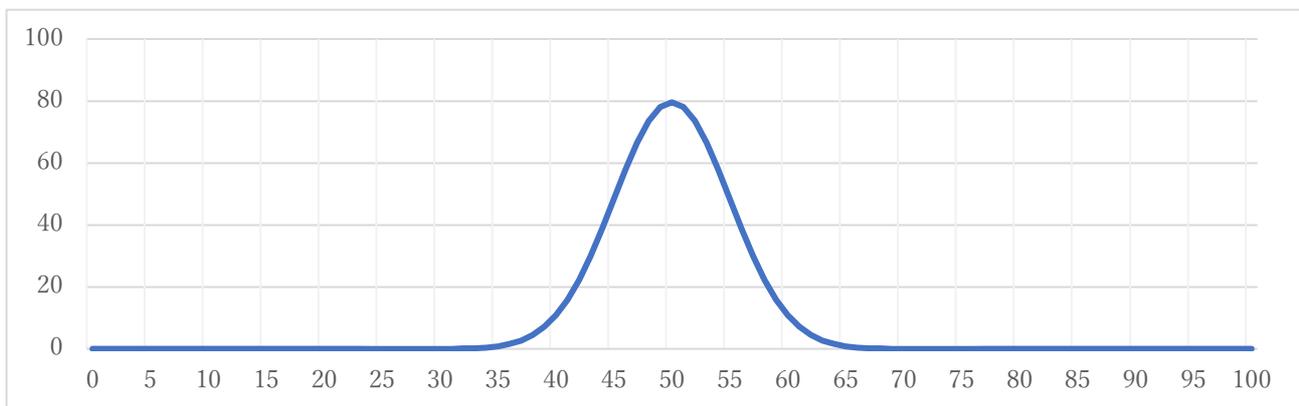


図15-1から分かるように、50勝のペアが最も多くなるものの、55勝（または45勝）は多くはないが珍しくはない、60勝（または40勝）はかなり珍しい、65勝（または35勝）はほとんど起こらないということが分かるだろう。ここで偶然には起こりそうにもない水準を慣例的に5%とすると、60回以上勝

つ場合と 40 回以下しか勝たない場合が 5%に含まれる。つまり、ペアで 100 回じゃんけんをしたときに、60 回以上勝ったり、40 回以下しか勝てなかったりする場合には、偶然では 5%以下でしか起きない事象であるため、偶然とは言い難いと推測する。これは統計的に 5%で有意と表現される。反対に、95%の確率に含まれてしまう 41 勝以上 59 勝以下であれば、勝率が 5 分 5 分のじゃんけんでも偶然起こり得る範囲であると判断される。

偶然に 5%の確率で生じるのであればまだ偶然である可能性を捨てきれないという場合には、さらに厳しい水準で考えることもできる。慣例的に 1%の水準で考えると、63 回以上勝ったり、37 回以下しか勝てない場合が含まれる。ここまできると 5 分 5 分の勝率のじゃんけんではほとんど起きることがないため、相手のパターンを読んでいる（あるいは相手に読まれている）か、いかさまが疑われることになる。

なお、社会現象の中には、じゃんけんのような勝ち負けとは違う散らばりになる場合がある。例えば正規分布では、平均値の他に散らばりの大きさを表す標準偏差を指定する必要がある。(図 15-2)。平均は同じでも散らばりが大きくなることで、グラフの形が変化する。標準偏差については 3 節で説明する。

図 15-2 平均 50、標準偏差 20 の正規分布

