

偽陽性にご用心：新型コロナウイルス検査とベイズ定理

どのような病気の検査も誤りは避けられない。新型コロナウイルス感染の有無を調べる検査に関しても同様である。まず、感染しているのに陰性とする誤り、そして感染していないのに陽性とする誤り（偽陽性）である。

それでは、無差別に検査して、ある人の検査結果が陽性だった場合、この人が実際に感染している確率はどれくらいになるのか。

感染していて陽性とする確率と感染していないのに陽性とする確率はそれぞれ

$$P(\text{陽性}|\text{感染}) = \frac{P(\text{陽性} \cap \text{感染})}{P(\text{感染})}$$

$$P(\text{陽性}|\text{非感染}) = \frac{P(\text{陽性} \cap \text{非感染})}{P(\text{非感染})}$$

なので、検査結果が陽性だった人が実際に感染している確率は

$$\begin{aligned} P(\text{感染}|\text{陽性}) &= \frac{P(\text{陽性} \cap \text{感染})}{P(\text{陽性})} = \frac{P(\text{陽性} \cap \text{感染})}{P(\text{陽性} \cap \text{感染}) + P(\text{陽性} \cap \text{非感染})} \\ &= \frac{P(\text{陽性}|\text{感染}) \times P(\text{感染})}{P(\text{陽性}|\text{感染}) \times P(\text{感染}) + P(\text{陽性}|\text{非感染}) \times P(\text{非感染})} \end{aligned}$$

となる。

1 万人に 1 人が感染しており、感染しているのに誤って陰性とする確率が 10 パーセント、感染していないのに誤って陽性とする確率が 0.1 パーセントであれば、

$$P(\text{感染}) = \frac{1}{10,000}, \quad P(\text{非感染}) = \frac{9,999}{10,000}$$

$$P(\text{陰性}|\text{感染}) = \frac{1}{10}, \quad P(\text{陽性}|\text{感染}) = \frac{9}{10}$$

$$P(\text{陽性}|\text{非感染}) = \frac{1}{1,000}, \quad P(\text{陰性}|\text{非感染}) = \frac{999}{1,000}$$

を上式に代入すればよい。結果は

$$\begin{aligned} P(\text{感染}|\text{陽性}) &= \frac{P(\text{陽性}|\text{感染}) \times P(\text{感染})}{P(\text{陽性}|\text{感染}) \times P(\text{感染}) + P(\text{陽性}|\text{非感染}) \times P(\text{非感染})} \\ &= \frac{\frac{9}{10} \times \frac{1}{10,000}}{\frac{9}{10} \times \frac{1}{10,000} + \frac{1}{1,000} \times \frac{9,999}{10,000}} = \frac{900}{900 + 9,999} \\ &= 8.3\% \end{aligned}$$

である。陽性という結果が出ても、実際に感染している確率はわずか 8.3 パーセント。逆

にいえば、感染していない偽陽性の確率が 91.7 パーセントである。

新型コロナウイルスに限らず、人口に占める感染者の割合が低い状況で無差別に検査を行った場合、かなり精度の高い検査であったとしても、陽性という検査結果が出た人が実際に感染している確率はかなり低い。

このベイズ定理の簡単な応用で得られる、偽陽性の割合の高さについて、残念ながら、今回の検査の是非をめぐる議論では、ほとんど言及されることがない。このノートが検査に関する建設的な議論に多少とも貢献できれば幸いである。