

ベイズ統計 超入門

YPC 2020年5月例会Zoom
横浜物理サークル 2020年5月17日(日)17:20~17:30

新型コロナウイルス感染検査への対応におけるベイズ統計を考える

検査で得られた「陽性判定」「陰性判定」をどうとらえるか。

現実の信頼度は元の集団の感染率に依って変わる。検査のエラーが1%で識別度70%での「陽性判定」でも、その93%が感染していない場合もある。

検査で得た情報の推定活用が重要

陽性判定だが感染している
陰性判定だが非感染

陽性判定で感染している
陰性判定で感染していない

一人の人間が感染症の検査を受ける場合はこれだけの情報

検査の識別度は70%とする。

0.7	感染している	0.01	感染していない
0.3	0.99	0.01	0.99

非感染者を陽性と間違えるエラー度を0.01とする。

あらゆる検査において、陰性判定だが感染している部分と陽性判定だが感染していない部分もある。

一人の人間が感染症の検査を受ける場合はこれだけの情報

検査の識別度は70%とする。

0.7	感染している	0.01	陽性陰性判定で非感染
0.3	陽性判定で感染	0.99	陰性判定で非感染
	陰性判定で感染		

非感染者を陽性と間違えるエラー度を0.01とする。

陽性判定だった！ 陰性判定だった！

一人の人間が感染症の検査を受ける場合はこれだけの情報

検査の識別度は70%とする。

0.7	陽性判定で感染	0.01	陽性陰性判定で非感染
0.3	陰性判定で感染	0.99	陰性判定で非感染

非感染者を陽性と間違えるのを0.01とする。

「私は、どの程度の推定確率で感染しているのでしょうか？」という問いかけに答えたい；

一人の人間が感染症の検査を受ける場合はこれだけの情報

検査の識別度は70%とする。

0.7	陽性判定で感染0.35	0.01	陽性陰性判定で非感染0.005
0.3	陰性判定で感染0.15	0.99	陰性判定で非感染0.495

非感染者を陽性と間違えるのを0.01とする。

他に何の情報も無いと感染可能性と非感染の可能性を等価0.5対0.5にするしかない。
陽性判定の場合 $0.5 \times 0.7 + 0.5 \times 0.01 = 0.355$ なので、本当に感染している可能性は $0.35 / 0.355 = 0.986$ となる。

陰性判定の場合 $0.5 \times 0.3 + 0.5 \times 0.99 = 0.645$ なので非感染可能性は $0.495 / 0.645 = 0.767$ になるが？ 同様な陰性判定者の4人に1人は感染している……「検査の意味が不明」という印象を与える。★数学としてシンプルだが、**実用として適切**だろうか？

一人の人間が感染症の検査を受ける場合はこれだけの情報

検査の識別度は70%とする。

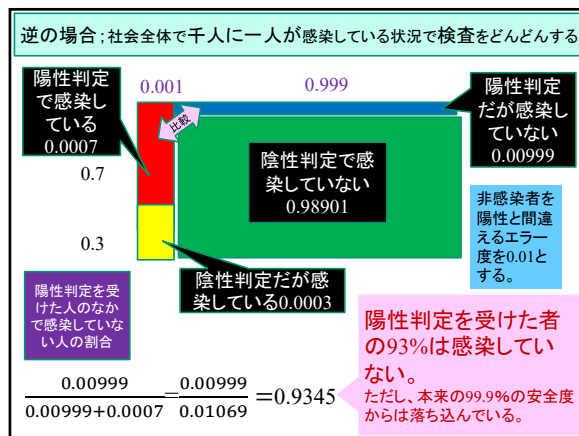
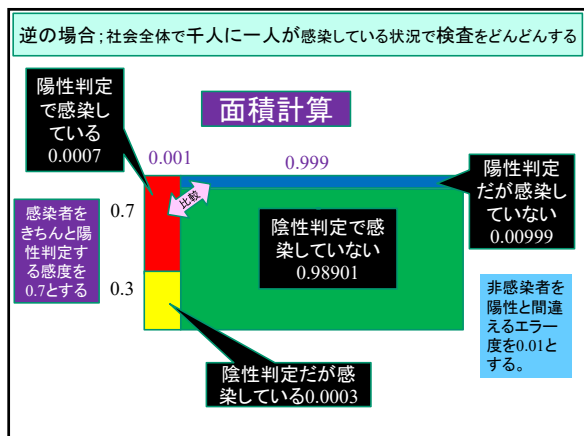
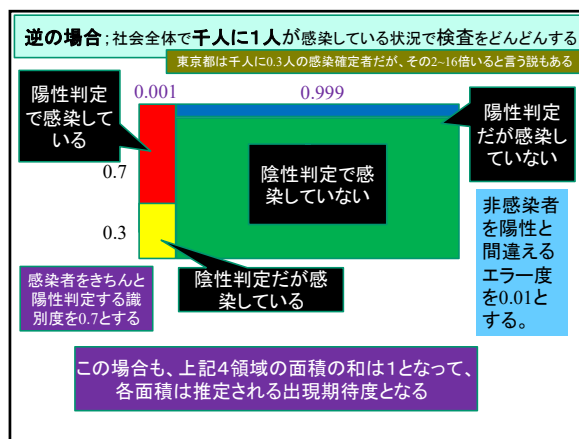
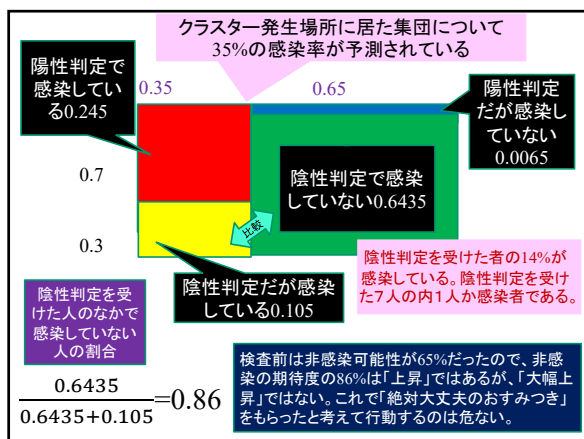
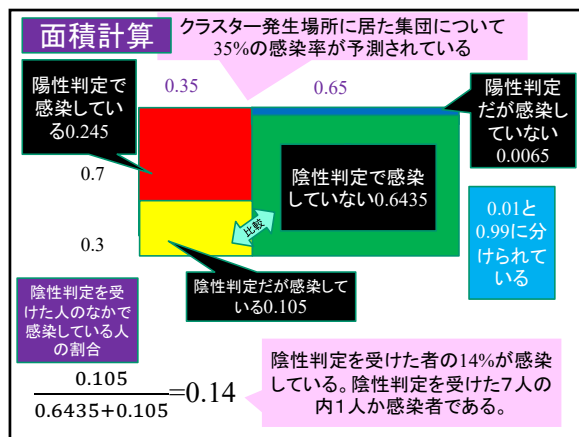
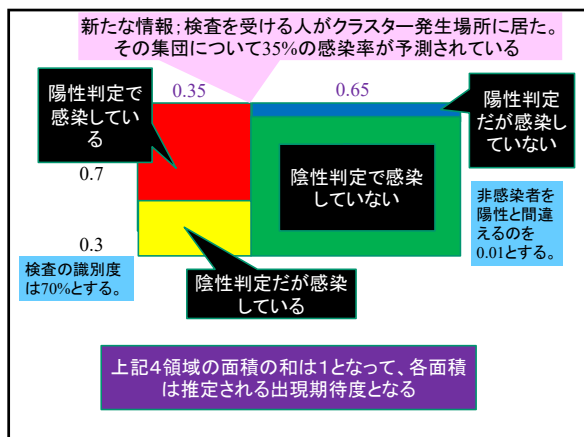
0.7	陽性判定で感染	0.01	陽性陰性判定で非感染
0.3	陰性判定で感染	0.99	陰性判定で非感染

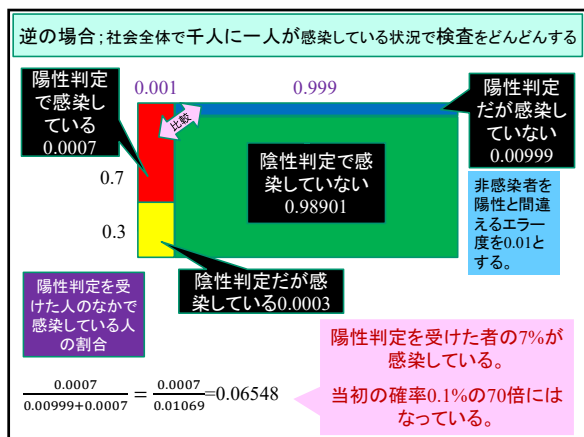
非感染者を陽性と間違えるのを0.01とする。

「私は、どの程度の推定確率で感染しているのでしょうか？」という問いかけに答えたい；

他に何の情報も無いと感染可能性と非感染の可能性を等価にするしかない。

ベイズ統計：さらに得られる情報を使って、その情報を取り込んだ確率(期待度)を求めてみる。ある意味では「追加された見通し=『主観』」であるが、「根拠を説明出来て方法論計算式が提示できる」という意味で「科学」である。





このような推定は正しいか？

ひとえに新しい情報である「周囲という母集団(統計学用語)における感染の状況」が正しく把握されているかどうかによって異なる。「正しい」かどうかより、「より妥当になる方法を与えている」ことが重要と考える。

ベイズ推定統計

多くのテキストでは「新たな情報」を「事前確率」と言っている。

一人の人間の一回の検査において、それ以外の周囲から得られる新たな情報を使って、その情報を取り込んだ確率(期待度)を求めてみる。ある意味では「追加された見通し=『主観』」であるが、「根拠を説明出来て方法論計算式が提示できる」という意味で「科学」である。

以上はベイズ統計の入門であって、さらに深い世界が待っています。