

●特集 / 数学の論理・社会の論理・日常の論理

PCR 検査の統計と論理

谷 克彦

COVID-19 パンデミックは、実効再生産数を 1 より低下させると鎮まります。この対策は、次の 3 つの数値を下げることです：

- ① 感染者が感染力をもつ期間
- ② 感染者が接触する人数
- ③ 感染者との接触で感染する確率

そして、それぞれに対する施策は以下のようです：

- ① 感染者を早期に発見し隔離する。このために PCR 検査の拡充が必要。
- ② 効果的なロックダウンの期間、地域、方法を、シミュレーションで予測し戦略的に介入する。
- ③ ワクチン接種で、感染感受性のある人の割合を減らす。

COVID-19 に感染すると、次のような経過になります。

感染(陽性) → 潜伏無症状期
 → 発症期(無症状もありうる)
 → 回復(陰性)or 死亡

陽性の期間 [潜伏無症状期 + 発症期(無症状もありうる)] は、「罹患者」が感染源となる有病状態なので「有病罹患」と呼ぶことにします。感染源となる「有病罹患」(症状の有無にかかわらず)を早く発見し隔離する必要があります。

有病罹患者の発見は PCR 検査でなされます。検査の目的は蔓延率の推定だけでなく、感染源となる有病罹患者をできるだけ早期に探し出し隔離するという緊急な役割があります。検査対象を限定し、検査の陽性的中率を上げることが検査の目的になってはいけません。

実世界の現象は、多数の原因と結果が複雑に絡んだ因果関係をなし、数学(統計)で論理的に推論するのは、その一部を切り取った世界です。その範囲で得た数学(統計)的推論の結果を、系全体

中で解釈できる論理的な思考が必要です。

正しい数学(統計)推論で得られた結果でも、複雑な全体系で非論理的に利用されるとしたら、社会を誤った方向へ導く主張に、数学が加担してしまうことになります。PCR 検査の規模拡大は有病率の低い集団ではすべきではないというのは正しい主張ではありますが、その主張の根底にある PCR 検査の性能から見直し、これを論理的に考察してみましょう。

有病率とは

日本感染症学会の定義によると、有病率とは、「その疾患をもっている人数の全人口に対する割合」ですが、日々発表される厚労省の新規陽性者数と検査数のデータから、日本の有病率を計算できるように、有病率 x_0 の解釈を次のようにします。

$$(\text{定義}) \quad \text{有病率} = \frac{\text{有病罹患数}}{\text{累積検査数}}$$

$$\begin{aligned} \text{有病罹患数} &= \text{累積陽性者数} \\ &\quad - \text{累積退院} \cdot \text{回復者数} \\ &\quad - \text{累積死亡者数} \end{aligned}$$

これにより、日時 T (= 5 月 15 日) の日本の有病率を求めると $x_0 = 0.58\%$ になります。ここでは、陽性者 = 罹患者と見なしています。

この検査集団の陽性率(T)と罹患者率(T)は、7 日平均(9~15 日)を用いて、

$$\text{陽性率}(T) = \frac{\text{陽性者数}(T)}{\text{検査数}(T)} = \frac{6288}{92167} = 0.068$$

$$\begin{aligned} \text{罹患者率}(T)^{*1} &= \frac{\text{陽性者数}(T)}{\text{累積検査数}(T)} \\ &= \frac{6288}{13015244} = 0.0005 \end{aligned}$$

*1 感染症学会の定義では、罹患者率の分母は、「集団の感受性のある人数」ですが、ここでは「集団の人数」としています。

累積 PCR 検査数	累積陽性者数	累積退院・回復者数	累積死亡者数	有病罹患数
13,015,244	671,841	584,770	11,365	75,706

累積 PCR 検査数	PCR 検査数(7日平均)	陽性者数(7日平均)
13,015,244	92,167	6,288

国内の発生状況。厚労省(2021.5.15 現在)

日時 T の有病罹患数(T)を別の定義で表現してみましょう。有病罹患状態は、14日くらいで回復(陰性になる)します(運悪く死亡の場合もありますが、少数のため無視します)。今日 T の有病罹患患者(感染源となる罹患)の中には、 t 日前に罹患した者も残っています。そこで、次の定義が成り立ちます：

$$\text{有病罹患数}(T) = \int_0^T \text{罹患数}(t) p_a(T-t) dt$$

ここで、 $p_a(t)$ は、陽性保持確率(病気の減衰関数のような性質)で、[付録 3] の大規模調査の結果に報告されています。病気の感染初日を推定するのは大変難しいのです。PCR 検査で見つかった日が感染日というわけではないでしょう。感染から5日目あたりが、感染者が最も多くウイルスを放出するので、その頃が最も発見されやすいのではないかと思います。

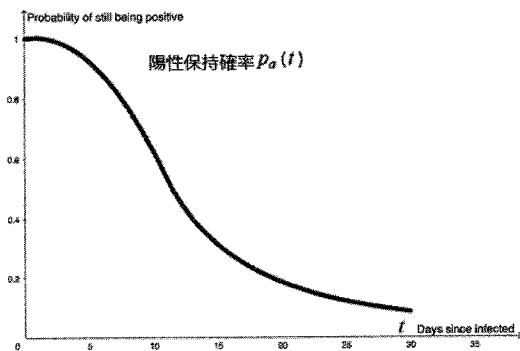
病気減衰関数の半減期は約10日ですので、上の積分は次のように近似できます：

$$\text{有病罹患数}(T) = 10 \times \text{罹患数}(T).$$

この集団の累積検査数で規格化すると、

$$\text{有病率}(T) = 10 \times \text{罹患率}(T)$$

が得られます。



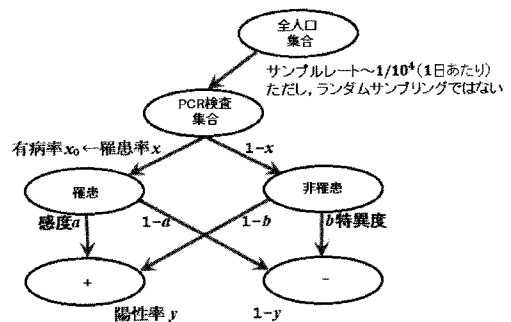
PCR 検査の感度と特異度

PCR 検査の感度 a とは、罹患患者を PCR 検査で陽性(+)と正しく判定する確率のことで、真の罹患患者でも PCR 検査が陰性(-)(偽陰性)と判定される確率は $1-a$ 程度あります。検査の特異度 b とは、非罹患患者を正しく陰性(-)と判定する確率のことで、非罹患患者を陽性(+) (偽陽性)と判定する確率は $1-b$ 程度です。

確率表示	+判定	-判定
罹患	a (真陽性)	$1-a$ (偽陰性)
非罹患	$1-b$ (偽陽性)	b (真陰性)

確率 a, b は1に近いほど、優秀な検査になります。従来の議論に用いられてきたこれらの数値は、 $a = 0.70, b = 0.99$ ですが、昨年の英国 ONS による大規模調査 [付録 3] で判明した数値は、 $a = 0.95 (0.85 \sim 0.98), b = 0.9992$ です。

低い有病率の集団で PCR 検査対象を拡大すると、莫大な偽陽性が出て医療崩壊につながるので、有症状者や濃厚接触者に限定して検査を行っているとの主張がありますが、英国 ONS 調査の感度と特異度を採用すると、この主張の根拠が覆えることを検証します。



ベイズ推定による有病罹患の内訳

この集団の罹患率を $p(\text{罹患}) = x$ とします。この集団で、PCR 検査が + 判定のとき罹患患者である確率 $p(\text{罹患} | +)$ 、および、- 判定のとき罹患患者である確率 $p(\text{罹患} | -)$ 、などを推定しましょう。

$$p(\text{罹患} | +) = \frac{p(+ | \text{罹患}) \cdot p(\text{罹患})}{p(+)}$$

$$= \frac{a \cdot x}{a \cdot x + (1-b)(1-x)},$$

$$p(\text{罹患} | -) = \frac{p(- | \text{罹患}) \cdot p(\text{罹患})}{p(-)}$$

$$= \frac{(1-a)x}{(1-a)x + b(1-x)},$$

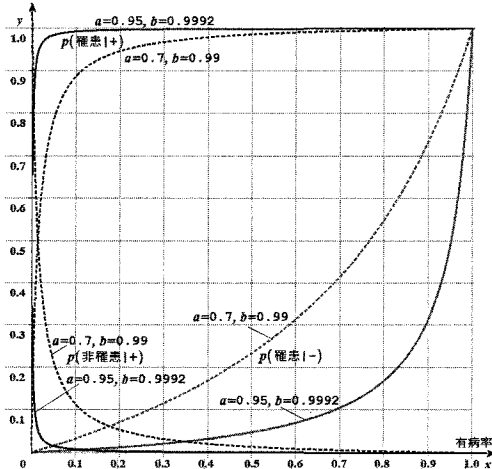
$$p(\text{非罹患} | +) = \frac{p(+ | \text{非罹患}) \cdot p(\text{非罹患})}{p(+)}$$

$$= \frac{(1-b)(1-x)}{(1-b)(1-x) + a \cdot x},$$

$$p(\text{非罹患} | -) = \frac{p(- | \text{非罹患}) \cdot p(\text{非罹患})}{p(-)}$$

$$= \frac{b(1-x)}{b(1-x) + (1-a)x},$$

下に x を変数とするグラフを示します。ベイズの定理は線形システムなので、重ね合わせができ、罹患状態を束ねた状態で定義される有病率 x_0 を入力にしてもよいでしょう。



集団の有病率 $x = 0.006$ に対して、感度 a 、特異度 b を変えて比較

(1) $a = 0.95, b = 0.9992$ (英国 ONS) を用いた場合

$$p(\text{罹患} | +) : p(\text{非罹患} | +) = \text{真陽性} : \text{偽陽性}$$

$$= 0.9 : 0.1$$

$$p(\text{罹患} | +) : p(\text{罹患} | -) = \text{真陽性} : \text{偽陰性}$$

$$= 0.9 : 0.0003$$

(2) $a = 0.7, b = 0.99$ (従来) を用いた場合

$$p(\text{罹患} | +) : p(\text{非罹患} | +) = \text{真陽性} : \text{偽陽性}$$

$$= 0.3 : 0.7$$

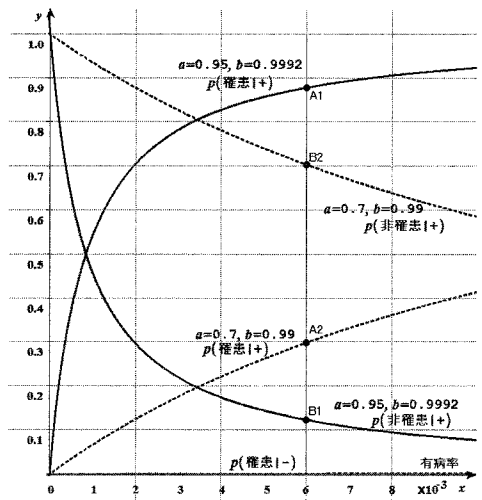
$$p(\text{罹患} | +) : p(\text{罹患} | -) = \text{真陽性} : \text{偽陰性}$$

$$= 0.3 : 0.002$$

PCR 検査が + 判定であるときに、真陽性の確率 $p(\text{罹患} | +)$ と、偽陽性の確率 $p(\text{非罹患} | +)$ を比較すると、従来用いられている感度、特異度では、偽陽性確率が真陽性確率の 2.3 倍もありますが、英国 ONS の数値では、逆に 1/10 になり、医療崩壊を懸念する根拠にはなりません。

いずれにしろ、陽性確定までに PCR 検査は 2 度行われ、さらに、抗原検査の併用もありますから、偽陽性の誤判定リスクは回避可能です。積極的に PCR 検査の対象を拡大し、感染源となる無症状の罹患患者を拾い出し早期隔離する道を閉ざすべきではありません。

次に、真陽性の確率 $p(\text{罹患} | +)$ と偽陰性の確

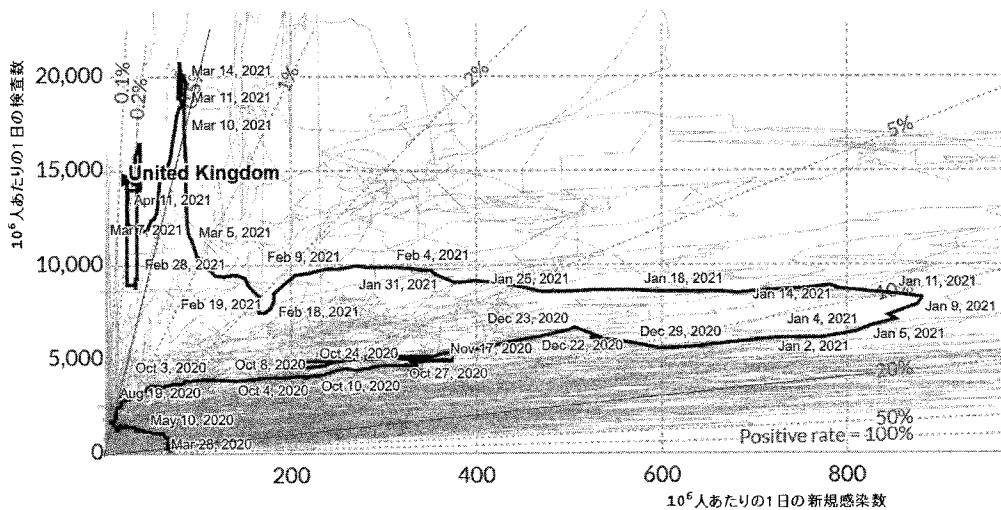
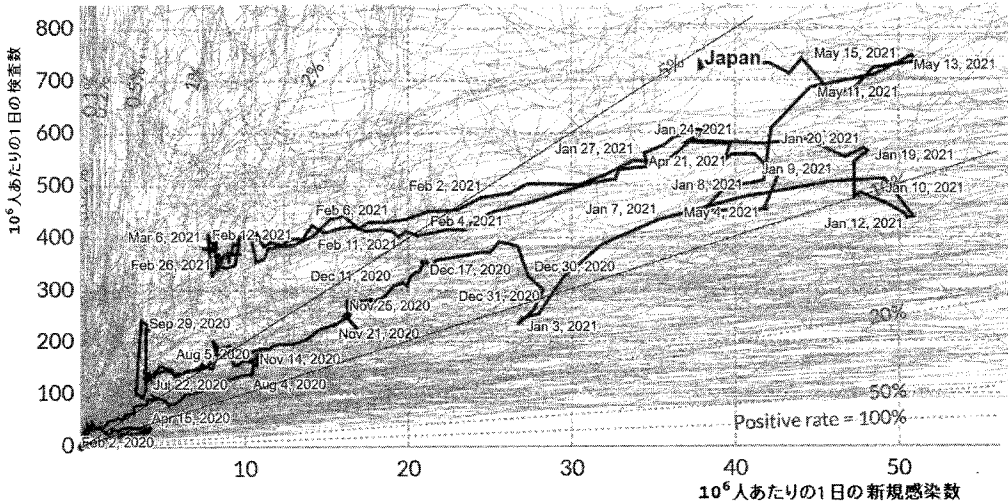


率 $p(\text{罹患}| -)$ を比較すると、(1)では、 $1:3 \times 10^{-4}$ 、(2)では、 $1:7 \times 10^{-3}$ でいずれも偽陰性は小さい確率です。人数で比較するには、陽性率 $y = 0.07$ 、陰性率 $1 - y = 0.93$ を、それぞれ、 $p(\text{罹患}| +)$ と $p(\text{罹患}| -)$ に乗じます。(1)では、真陽性人数：偽陰性人数 = $1:4 \times 10^{-3}$ 、(2)では、真陽性人数：偽陰性人数 = $1:1/9$

この集団の陽性率は7%と低いので、陰性集団が大多数で、偽陰性の確率が小さくても偽陰性者数は多いとの主張もあります。確かに、従来の数値(2)を用いると、真陽性者の1/9の偽陰性者がいます。しかし、数値(1)を用いると、偽陰性者の数は1/250で、PCR検査対象を拡大しない理由にはなりません。検査を拡大すれば、拾い出せる陽性者は検査数に比例して確実に増加します。

PCR検査数は十分か

ここに引用した2つの図は、1日の 10^6 人あたりの(横軸)検査陽性数：(縦軸)検査数の散布図で、日本(上)と英国(下)の例です。英国の散布図スケールは日本のものより縦軸で25倍、横軸で16倍大きいのでご注意ください。散布図パターンを比較すると特徴的な違いがあります。時間経過とともに、右横あるいは右下がりに伸びる部分では、検査数が足りず陽性者の増加傾向を頭打ちにしている可能性があり、縦に伸びる部分では、陽性者を拾い出す十分な検査が行われているようです。英国の例を見ると、始めは、陽性者が多く検査数が間に合わないほどでしたが、現在みられる検査数を増しても陽性者が一定となる状態は、



十分な検査数が確保されている証拠です。日本の例で、検査数と陽性者数の増減の比例が見られる傾向は、検査数を増やせば、陽性者数も増加する可能性があります。

これらの図は、2020.1.25～2021.5.25の期間のもので、直線は、日本の図では、陽性率5%と10%の勾配、英国の図では、0.5%と20%の勾配を示しています。

これらのグラフは、<https://ourworldindata.org/coronavirus-testing> から引用しました。

[付録1] ベイズの定理

条件付き確率についての「ベイズの定理」とは次のように説明できます。

$$p(Y|X)p(X) = p(X \cap Y) = p(X|Y)p(Y)$$

ただし、記号の意味は例えば以下のようです。

$p(X)$: X が起こる確率

$p(Y|X)$: X が起こる条件下で、 Y が起こる確率

$p(X \cap Y)$: X かつ Y が起こる確率

ベイズの定理は、 X (原因)が起こる条件下で Y (結果)が起こる確率 $p(Y|X)$ と、 X と Y を入れ替えた確率 $p(X|Y)$ を結び付ける定理です。

[付録2] PCR 検査の原理

PCR(ポリメラーゼ連鎖反応)を利用して、わずかに数分子のターゲット DNA から数ミリグラムの DNA に増幅する技術を、1983年にマリスが発見し、1993年のノーベル化学賞を受賞しました。

RNA に対しても逆転写酵素により DNA を合成し、DNA の複製反応が利用できます。ターゲット DNA 鎖全体の複製ではなく、ウイルスを特徴づける断片の複製をします。これが、パウエル社が発表(1987)した RT-PCR 技術です。PCR 検査は、検体に含まれる COVID-19 ウイルスの微量な RNA から、逆転写酵素を用いて、cDNA を合成し、温度サイクル処理を 40 回程度繰り返して、cDNA を検出可能な濃度まで複製するというものです。このような PCR 検査の原理から、検体にわずかでもターゲット RNA が含まれていればいくらかでも増幅できますが、含まれていなければ誤混入がない限り増幅物はまったく生じません。

すなわち、特異度の高い検査です。

5 時間ほどかかった検査が 75 分に短縮でき、精度も確保できる自動化が普及しています。

[付録3] 英国 ONS の COVID-19 感染の大規模調査

英国国家統計局(ONS)による COVID-19 の大規模感染調査は、パイロット調査としてイングランド(2020.4 末)で開始されました。7 月以降、規模を拡大し、住所リストから世帯のランダム・サンプリングを行い、8 月には、2 週間ごとにテストされるサンプルを、イングランドで 15 万人、ウェールズ、スコットランド、北アイルランドで各 1.5 万人の規模に上げました。調査は、小児の症候性および無症候性感染の有病率と発生率を知るため、2 歳以上の子供を含みます(参照: ONS の COVID-19 学校感染調査)。16 歳以上の成人の 20% は血液サンプルの提供を求められました。2020.7.31～9.10 の 6 週間で、208,730 の検体調査を行いました。偽陽性は症状のない人で発生すると予想されましたが、調査研究のデータからは、症状の有無によらず同じでした。血液サンプルは、オックスフォード大学で、抗体検査、IgG 免疫グロブリンのテストをしました。

PCR 検査に関する知見は、その感度が、85%～98%(95% である可能性が最も高い)にあり、特異度は、99.92% でした。

綿棒スワブの検体採取は、適切な監督下で行えば、自己採取でも医療従事者が直接採取した場合と同様の正確な結果になることも報告されています。

●引用データ……………

- 1) 厚生労働省ホームページ
- 2) Coronavirus(COVID-19) Infection Survey, UK Statistical bulletins
- 3) Keeping up with COVID-19, Rachel Thomas;
<https://plus.maths.org/content/keeping-covid-19>
- 4) <https://ourworldindata.org/coronavirus-testing>

(たに・かつこ / NPO 法人数学月間の会理事)