Parson Bayes Meets the CCP Virus

Craig Pirrong
Bauer College of Business
University of Houston

March 22, 2020

Public debate and public policy regarding Covid-19, AKA the CCP virus, have been driven by data on the prevalence of the illness derived from testing. In this note I show why this is an extremely dubious basis for policy. Due to the nature of testing, it does not measure what we really want to measure.

In the first instance, what we want to measure is the likelihood of contracting the virus, as a function of time. This can be represented as a probability at any time (measured from the timing of "case zero") or as a hazard rate (the probability of contracting the virus conditional on not having contracted it at a previous date).

The basic problems with test-based data are (a) the error rates in the test, and (b) the non-random nature of the testing. Basic probability theory, and in particular Bayes' Law, help understand these problems.

Denote $\mathcal{P}(C|+)$ the probability of having CCP virus conditional on testing positive; $\mathcal{P}(+|C)$ the probability of testing positive if you are indeed infected (a "true positive," or TP); $\mathcal{P}(-|C)$ the probability of testing negativve if you are infected (a false negative, or FN); $\mathcal{P}(+|NC)$ the probability of testing positive if in fact you are not infected (a false positive, or FP); $\mathcal{P}(-|NC)$ the probability of testing negative when not infected (a true neg-

ative, or TN), and $\mathcal{P}(C) = \hat{p}$, the (unconditional) probability of infection. Note that if testing was random, $\hat{p}$ is what we are interested in: the actual probability of being infected: momentarily I will show that one *cannot* derive this probability from the test data we have, because testing is not random.

Note that $\hat{p}$, whether viewed as a cumulative probability or a hazard rate, should depend on time. Indeed, for forecasting in order to determine policy, this time dependence is essential. If the hazard rate declines over time, due to biological factors or endogenous behavioral responses (a point emphasized by Richard Epstein), forecasts based on exponential extrapolations exaggerate wildly the risks. Despite this importance, I suppress that time dependence to simplify the analysis. Just note that all of the problems apparent in the simplified analysis are even more acute when one tries to condition on time.

Bayes' Rules says:

$$\mathcal{P}(C|+) = \frac{\mathcal{P}(+|C)\hat{p}}{\mathcal{P}(+|C)\hat{p} + \mathcal{P}(+|NC)(1-\hat{p})}$$

That is, the probability of having CCP virus, conditional on a positive test result, depends on the unconditional probability of being infected, and the rates of false and true positives. Similarly:

$$\mathcal{P}(C|-) = \frac{\mathcal{P}(-|C)\hat{p}}{\mathcal{P}(-|C)\hat{p} + \mathcal{P}(-|NC)(1-\hat{p})}$$

If there are high rates of error, the foregoing kinds of calculations are conventionally used to demonstrate that tests overstate the true probability that someone with a positive test is infected, especially if the underlying unconditional probability of an infection is small. That is of use when determining how to treat (or not) someone with a positive test. But our main interest is in estimating the unconditional probability of infection.

To get at this, note that the probability of testing positive is:

$$\alpha = \mathcal{P}(+|C)\hat{p} + \mathcal{P}(+|NC)(1-\hat{p})$$

and the probability of testing negative is:

$$1 - \alpha = \mathcal{P}(-|C)\hat{p} + \mathcal{P}(-|NC)(1 - \hat{p})$$

One tests either positive or negative, so:

$$1 = \mathcal{P}(+|C)\hat{p} + \mathcal{P}(+|NC)(1 - \hat{p}) + \mathcal{P}(-|C)\hat{p} + \mathcal{P}(-|NC)(1 - \hat{p})$$

With a set of test results, and knowledge of the error rates, it would be possible to infer $\hat{p}$. This is problematic in current conditions for a variety of reasons. For example, how do you determine the rate of false positives for a disease that may not result in symptoms? Or may result in symptoms that are indistinguishable from other illnesses (e.g., flu)? Perhaps the former issue could be addressed by defining condition C as "symptomatic coronavirus," which is what we are mainly interested in anyways. The latter issue is more difficult to address. Moreover, different tests have been used, and testing protocols differ across countries, making it difficult to pool data.

If one defines C as being symptomatic, asymptomatic positive tests are false positives. There is anecdotal evidence that is a serious matter. For instance, several NBA players tested positive, but are asymptomatic. The same is true of other public figures.

All this means that we are unlikely to have good estimates of test error rates, and hence are unlikely to be able to correct test results to extract $\hat{p}$.

Even if we could adjust the frequency of positive and negative test results for test error rates to recover $\hat{p}$, this would not reveal $\mathcal{P}(C)$, because $\hat{p}$ is the unconditional probability of infection *among the tested population.* But if tests are not random–and it is clear that they are not, the tested group is not representative of the population at large, and hence $\hat{p}$ is a biased measure of the probability of infection in the population at large.

To expand the notaton, $\hat{p} = \mathcal{P}(C|T)$, i.e., it is the probability that someone has the virus, conditional on being tested. Again applying Bayes' Rule:

$$\hat{p} = \mathcal{P}(C|T) = \frac{\mathcal{P}(T|C)\mathcal{P}(C)}{\mathcal{P}(T|C)\mathcal{P}(C) + \mathcal{P}(T|NC)(1 - \mathcal{P}(C))}$$

What we really want to know is $p^* = \mathcal{P}(C)$. If testing is random, $\mathcal{P}(T|C) = \mathcal{P}(T|NC)$, which would mean that $p^* = \hat{p}$. That is, if testing were random, error rate corrected test frequencies could provide an unbiased estimate of what we really want to know.

But we know that testing is not random. The symptomatic and exposed are more likely to have been tested. Given that CV19 symptoms are similar to those for influenza, influenza sufferers are more likely to be tested. Furthermore, the symptomatic are more likely to have other health conditions that makes them more likely to express symptoms, and hence are unrepresentative of the population at large. Also, the influential are more likely to have been tested, and the influential are not representative. The hypochondriacial are more likely to have been tested. They are also unrepresentative.

Do you see the problem here? What we want to know is p^*–the probability of contracting the virus. (We actually want to know this probability conditional on a plethora of factors, including domicile (and the climate in that domicile), age, health condition, and on and on.) The results of tests are many steps removed from what we want to know. The reliablity of the tests, and crucially, the fact that testing is not random, stand between the test results that are reported obsessively and what we really want to know.

My conclusion from this is that test results are more likely to mislead than provide helpful guidance. The foregoing analysis implies that *if* the tests are completely accurate (no false positives or false negatives) and *if* testing is random, than test rates measure accurately the rate of coronovirus incidence. But we know neither condition is true, and not even remotely

correct.

So if the test data are so defective, what should we look at? What we care about is acute illness and death from this disease. Such outcomes will be reflected in aggregate data on hospital admissions for acute respiratory conditions, death rates overall, and death rates from respiratory illness. That is, if this is a prevalent, deadly virus, death rates should be elevated.

Moreover, these data should reflect crucial conditioning variables, such as age, location, date of death (respiratory ailments and deaths being strongly seasonal), and pre-existing conditions (obesity, blood pressure, heart conditions, renal conditions, respiratory health, smoking status, etc.). Such data exist on a historical basis. Use those data to benchmark current conditions to see whether indeed we are experiencing something outside the realm of ordinary experience–an in particular, so far outside historical experience to justify incurring *trillions* of cost in terms of lost income. Moreover, by conditioning on variables related to expected lifespan, productivity, etc., we can develop more targeted policy responses.

So forget the tests. Don't look at those. Demand data on what we really need to focus on: illnesses and death of the type that the CCP virus can cause.