Frequentist Methods and the Base Rate Fallacy

Elizabeth Rodriguez

November 21, 2022

Statisticians and data scientists have long fought over the methodology in regard to modeling or analyzing data, many picking the side of history, known as Frequentism. Frequentist statistics rely on assigning probabilities to events, using a sample space of n points. On the other side, Bayesian statistics is a newer methodology used to gain a distribution for unknown parameters based on likelihood and prior knowledge of the population being tested. Bernoulli's fallacy is one of the main objectives Bayesians make against using Frequentist methods. In Aubrey Clayton's Bernoulli's Fallacy, he defines it as the idea that one can make inferences off of data that was created only using sample probabilities (in reference to Frequentist methods). A sub-category of this idea is another fallacious belief, the base-rate fallacy. In this paper, I will explore the base-rate fallacy. First by defining what it means. Then, I will present the argument Aris Spanos made in his paper 'Is Frequentist Testing Vulnerable to the Base-Rate Fallacy?'. I claim that his argument is valid, and furthermore, his argument brings to light a problem with the classification of false positive and false negative probabilities.

1 The Base-Rate Fallacy

The base-rate fallacy refers to the ignorance of the base probability of the population being tested. This concept comes directly from the Bayes theorem, which states

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \tag{1}$$

$P(X)$ is the percentage of the population being tested that is ignored. Therefore, the conditional probability $P(X|Y)$ is mistaken for being equal to the conditional probability $P(Y|X)$. There are some who believe that this fallacy takes place in Frequentist methods, largely due to the use of type I and type II errors. This is false, the comparison of these errors and conditional probabilities is as misconstrued as the mistaken equivalence of said conditional probabilities.

2 Conditional Probabilities vs Error Rates

The bickering between Bayesians and Frequentists is a long path in history. Some Bayesians argue one reason to not use Frequentist methods is the base-rate fallacy. This is due to an idea that type I and type II errors are equal to the conditional probabilities used in Bayes theorem, $P(X|Y)$ and $P(Y|X)$. Therefore, they are committing large errors because the base rate is not being taken into account. Spanos argues that type I and type II errors are not the same as false positive/negative probabilities, take for example the Harvard Medical School test. The Harvard medical test used is not committing the base-rate fallacy by using false positive and false negative rates because it is not analogous to using type I and type II errors. Instead, I claim that the fallacy being committed is using these false positive/negative probabilities in place of error probabilities. For clarity, we shall define a type I error as the error obtained when your result falls in the critical region (obtained by alpha), if the p-value obtained is below alpha you may reject your null hypothesis. Type II error is the error associated with accepting a false null hypothesis. There is a bijection between both type errors, meaning decreasing alpha (type I error) will result in a larger type II error, and vice versa.

Spanos claims that error probabilities cannot be conditional, hence already separating the type I and II probability terms with being equal to false positive and false negative rates. I claim that Spanos is correct, error rates are based upon a chosen alpha, a distribution, a test statistic, and a population of size n. Taking into account these reasons, this directly contradicts what frequentist testing is. Frequentist testing is based almost entirely on sample size (n), and hypothesis testing. To say that false positive and false negative rates are a fallacy in regard to frequentist testing is a contradiction in itself, as false positive and false negative rates should not be considered frequentist methods at all. Therefore, frequentist methods do not commit the base-rate fallacy due to the absence of conditional probability in frequentist testing.

Also going back to sample sizes, computing a false negative/positive rate not only uses conditional probability but is only using one sample ($n = 1$). Frequency (it's in the name!) is a large part of frequentist testing, probabilities are turned into frequencies due to large sample sizes. Take, for example, the probability that person A does not have disease (Y) but obtained a positive test (X). This conditional probability, $P(X|Y)$, is based only on one person, A. Therefore we cannot obtain any hypotheses, due to the lack of data. The given probabilities to obtain said conditional probability is not related to a test statistic, a distribution, or proper errors.

Spanos also claims that false positive and false negative probabilities are not equivalent to prior or posterior probabilities. For context, a posterior probability is calculated as

$$\pi(\theta|z_0) \propto \pi(\theta) * f(z_0|\theta) \tag{2}$$

where $\pi(\theta|z_0)$ is the posterior distribution, which is proportional to the prior probability, $\pi(\theta)$, multiplied the likelihood function $f(z_0|\theta)$. Spanos' argument is that the conditional probability, say from the example above, $P(X = 1|Y = 1)$, is "nothing more than a deductive calculation within the context of a known statistical model"(Spanos, 580). The larger issue is, again, the problem with the sampling size. It is not possible to run tests about hypotheses if the data set is only one observation. This leaves us to wonder what is a false negative and false positive methodology, if not Frequentist or Bayesian.

A report like the Harvard Medical School test using Bayes theorem to calculate a conditional probability like the example above is making the mistake of using said probability with other frequentist methods (confidence intervals, maximum likelihood methods, etc.), and is a dangerous slope to place research on. Frequentist statistics are used to test data (plural) and observe what that data can tell us. This is impossible for a conditional probability based on one observation, and therefore should not be applied/related to a statistic for a population. Frequentist testing also does not allow for probabilities to be assigned to certain events, instead, chooses a distribution to match the data points. The biggest fallacy regarding these medical types of reports is describing these conditional probabilities as frequentist methods.

3 Discussion and Conclusions

Although ignorance of the base rate in certain applications does classify as a miscalculation, the true fallacy is the concept that the base-rate fallacy takes place in frequentist testing via type I and type II errors. Furthermore, due to the criteria for what constitutes as a frequentist test, as well as a Bayesian test, false positive/negative probabilities do not fit in either of these categories. Therefore, I claim that the base-rate fallacy is indeed a fallacy, but cannot be applicable to frequentist testing via type I and type II errors. The true fallacy here is in regards to the concept of false positive and false negative rates being equivalent to type I and type II errors. This, of course, begs the question of what is a false positive/negative rate if not a Frequentist or Bayesian method? Due to only using one observation to compute this error, the only way to truly classify it is to say that it is just a conditional probability. There is no methodology that is based around having one sample only used, therefore, it is in a class of its own.

4 Citations

Clayton, A. (2022). Bernoulli's Fallacy: Statistical illogic and the crisis of modern science. Columbia University Press.

Koehler, J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. Behavioral and Brain Sciences, 19(1), 1-17. doi:10.1017/S0140525X00041157

Romeijn, Jan-Willem, ”Philosophy of Statistics”, The Stanford Encyclopedia of Philosophy (Fall 2022 Edition), Edward N. Zalta Uri Nodelman (eds.), URL = ¡https://plato.stanford.edu/archives/fall2022/entries/statistics/>.

Spanos, Aris (2010). Is frequentist testing vulnerable to the base-rate fallacy? Philosophy of Science 77 (4):565-583.