What is a Probability Anyway? T. Parent

What does it even *mean* to say that a claim has such-and-such probability??? Let p be any claim or statement, and let $P(p)$ be the probability of that statement. We are thus asking what it *means* to say that $P(p) = 50\%$ or 60% or what have you.

1. *The Classical View*

The classical view says that $P(p)$ represents the number of cases where p is true ("p-cases") over the total number of possible cases. For example, suppose p is "I will get the ace of spades in one draw from this ordinary deck of 52 cards." Then, when we say $P(p) = 1/52$, we simply mean that there is one ace of spades out of the 52 possible cards that you could get in one draw.

<u>Objections</u>

- Not every statement has a definite number of possibilities associated with it, e.g., 'I will have a rewarding career.'
- Often, some possibilities are more likely than others, meaning that classical probabilities will often not be the true probability. Suppose we are estimating the probability that a newborn baby girl will live until age 80. We do not consider this one case over the total number of ages that females live to, in order to get something like 1/122. For that would mean it is equally likely our newborn will live up to 80 as up to age 122.

2. *The Frequency View*

This is a view developed in response to the previous objection. The idea is that the probability of p is not the p-cases over the total number of *possible* cases, but rather the p-cases over the total number of *actual* cases. Thus, to know the likelihood that a female will live to 80, we take a representative sample of females, and see how many live up to 80. We then take this number (the "frequency" of females living up to 80) over the total number of females in our sample.

<u>Objections</u>

- Sometimes our samples are not representative, meaning that there are difficulties in knowing probabilities. The frequentist might respond that if we take a large enough sample, we can avoid having a skewed one. Still, there is no guarantee at any point that our sample isn't skewed.
- The frequentist view only gives us "generic" probabilities, e.g., the proportion of females (generically speaking) who live up to 80. But this is not to give us the probability that *this particular newborn girl* will live up to 80, and it is possible that this "individualized" probability will be different from the generic one.

3. *The Propensity View*:

Do not think of probabilities as *mere* frequencies, but rather as frequencies that mirror *propensities in nature* (as described by scientific laws).

- ➢ So instead of relying on a (possibly) skewed sample, we can use laws about human biology or sociology to estimate the probability that a girl will live to 80.

- Also, if we consider the facts about *this particular newborn girl*, we can use scientific laws in estimating her particular "propensity" to live this long.

Objections:

- Sometimes we don't know enough to say what the natural tendency of something is.
- Sometimes it doesn't make sense to talk of probabilities as natural propensities. For instance, we can talk about the probability of God existing, but it does not make sense to talk of the "natural propensity" for God to exist.

4. *Bayesianism*

The previous two views of probability are all *objectivist* in the sense that probability is determined by objective facts about the world (propensities, frequencies, or bare possibilities). Traditional Bayesianism, however, is a *subjectivist* view in that probabilities are personal assessments of how likely something is.

In going subjective, Bayesians avoid previous troubles about knowing probabilities, since those were troubles concerning knowledge of *objective* facts. Also, we can assess probabilities for both single cases and generic cases, as well as assess a probability for the truth of God existing.

Objections:

- There is still a problem in knowing probabilities, since it is hard to assign a *number* that represents your precise assessment of how likely something is.
- Personal assessments can sometimes be irrational. When that's the case, they shouldn't represent the true probability of something.
- What determines the prior probability of p, that is, the likelihood of p before *any* evidence is in? Is it based on a hunch? Or is it always .5? [A Bayesian might opt for the latter, under the assumption that the evidence (eventually) will correct the initial assignment to something more accurate.]

- The Bayesian response to the first objection is to look at your *betting behavior*. In particular, we consider what is the riskiest bet you would accept on the truth of p, as opposed to $\sim p$. Those odds then indicate the degree to which you believe that p is true (i.e., your subjective probability that p). For instance, if p = 'The Earth revolves around the sun,' you may be willing to accept very risky bet. Suppose the riskiest bet you would accept requires you to pay \$99 and only pays \$100 if you win. This would reflect a very high subjective probability that 'The Earth revolves around the sun' is true. Indeed, it would reflect a subjective probability of .99, since your confidence in its truth is so high that you see it as worth risking \$99 to gain \$1.

  This requires us, however, to make some idealizations. We must suppose that you don't have any problem with betting *per se*, so that you will accept any bet that's worth the risk (based on your subjective probabilities). Also, we must assume that your acceptance of a bet *really is* based on your personal assessment of likelihoods, rather than a whim. But if these conditions are met, the Bayesian says we *can* know exactly someone's subjective probabilities, contra the first objection.

- The Bayesian response to the second objection is to propose some necessary conditions on a set of *rational* personal assessments, i.e., on a *rational* credence function.

Definition $C_t(p)$ is the *credence function* at time $t \geq 0$, which takes ANY p as input and assigns the subjective probability for p at time t [or in short, $C_t(p)$ assigns $P_t(p)$ for any p].

Bayesian Constraints on Rational Credence Functions

Bayesians Agree on:
1. (Coherence) My credence function $C_t(p)$ is rational only if the probability-assignments it produces are consistent with the Axioms of Probability, which is to say, only if no Dutch Book can be made from these probability-assignments.

Definition: A Dutch Book can be made against a set of probability assignments iff a series of bets can be made (in which risk is proportional to the probability assignments), where you end up losing money, no matter what. For example, if you accept the earlier bet of \$99 to gain \$100—yet you also accept a bet of \$2 to gain \$100 on the *falsity* of 'The Earth revolves around the sun'—then you would lose \$1 either way. For regardless of the truth of this statement, you've paid \$101 for the two bets, but only gained \$100 from whichever bet you won, meaning that either way you'll suffer a net loss of \$1.

2. (Conditionalization) A subjective probability $P_{t+1}(p)$ is rational only if: $P_{t+1}(p) = P_t(p/q)$, where $P_{t+1}(q) = 1$.[*]

- $P(p)$, after discovering that q is true, should be the same as $P(p/q)$ immediately *before* the discovery.

Other Constraints:
Bayesians offer further constraints on rationality, but most of these are not agreed upon. Here are two examples:

3. (Strict Coherence) A subjective probability $P_n(p)$ is rational only if: $P_n(p) = 1$ only if p is necessarily true.

- You should only be certain about *necessary* truths.

4. (Reflection) A (prior) probability $P_n(p)$ is rational only if: $P_n(p/P_{n+1}(p) = k) = k$.

- Knowing that p will have a future probability of k should be enough to make k its present subjective probability.

5. *Pluralism*
Many statisticians now accept *pluralism* about probability. This is the view that our talk about "probability" does not concern one thing all the time, but rather concerns classical probability in some cases, frequentist probability in other cases, Bayesian probability in still other cases, and son on. Moreover, this pluralism is often seen as a *good* thing. The objections to the various views suggest that sometimes it is better to focus on one kind of probability over another, but no single type of probability seems optimal for all theoretical purposes.

---

[*] Note that, of course, my posterior credence function $C_{t+1}(p)$ is rational only if the posterior probabilities it assigns are rational.