for the observed x. Note that the likelihood ratios for the two experiments are also the same when 2 is observed, and also when 3 is observed. Hence, no matter which experiment is performed, the *same* conclusion about θ should be reached for the given observation.

This example clearly indicates the startling nature of the Likelihood Principle. Experiments E_1 and E_2 are very different from a frequentist perspective. For instance, the test which accepts $\theta=0$ when the observation is 1 and decides $\theta=1$ otherwise is a most powerful test with error probabilities (of Type I and Type II, respectively) 0.10 and 0.09 for E_1, and 0.74 and 0.026 for E_2. Thus the classical frequentist would report drastically different information from the two experiments.

The above example emphasizes the important distinction between initial precision and final precision. Experiment E_1 is much more *likely* to provide useful information about θ, as evidenced by the overall better error probabilities (which are measures of initial precision). Once x is at hand, however, this initial precision is no longer relevant, and the Likelihood Principle states that whether x came from E_1 or E_2 is irrelevant. This example also provides a good testing ground for the various conditional methodologies that were mentioned in Subsection 1.6.3. For instance, either of the conditional frequentist approaches has a very hard time in dealing with the example.

So far we have not given any reasons why one *should* believe in the Likelihood Principle. Examples 15 and 16 are suggestive, but could perhaps be viewed as refutations of the Likelihood Principle by die-hard classicists. Before giving the axiomatic justification that exists for the Likelihood Principle, we indulge in one more example in which it would be very hard to argue against the Likelihood Principle.

EXAMPLE 17 (Pratt (1962)). "An engineer draws a random sample of electron tubes and measures the plate voltages under certain conditions with a very accurate voltmeter, accurate enough so that measurement error is negligible compared with the variability of the tubes. A statistician examines the measurements, which look normally distributed and vary from 75 to 99 volts with a mean of 87 and a standard deviation of 4. He makes the ordinary normal analysis, giving a confidence interval for the true mean. Later he visits the engineer's laboratory, and notices that the voltmeter used reads only as far as 100, so the population appears to be 'censored'. This necessitates a new analysis, if the statistician is orthodox. However, the engineer says he has another meter, equally accurate and reading to 1000 volts, which he would have used if any voltage had been over 100. This is a relief to the orthodox statistician, because it means the population was effectively uncensored after all. But the next day the engineer telephones and says, 'I just discovered my high-range voltmeter was not working the day I did the experiment you analyzed for me.' The statistician ascertains that the engineer would not have held up the experiment until the meter was fixed, and informs him that a new analysis will be required. The engineer is astounded. He says, 'But the experiment turned out just the same as if the high-range meter had been working. I obtained the precise voltages of my sample anyway, so I learned exactly what I would have learned if the high-range meter had been available. Next you'll be asking about my oscilloscope.'"

In this example, two different sample spaces are being discussed. If the high-range voltmeter had been working, the sample space would have effectively been that of a usual normal distribution. Since the high-range voltmeter was broken, however, the sample space was truncated at 100, and the probability distribution of the observations would have a point mass at 100. Classical analyses (such as the obtaining of confidence intervals) would be considerably affected by this difference. The Likelihood Principle, on the other hand, states that this difference should have no effect on the analysis, since values of x which did not occur (here $x \geq 100$) have no bearing on inferences or decisions concerning the true mean. (A formal verification is left for the exercises.)

Rationales for at least some forms of the Likelihood Principle exist in early works of R. A. Fisher (cf. Fisher (1959)) and especially of G. A. Barnard (cf. Barnard (1949)). By far the most persuasive argument for the Likelihood Principle, however, was given in Birnbaum (1962). (It should be mentioned that none of these three pioneers were unequivocal supporters of the Likelihood Principle. See Basu (1975) and Berger and Wolpert (1984) for reasons, and also a more extensive historical discussion and other references. Also, the history of the concept of "likelihood" is reviewed in Edwards (1974).)

The argument of Birnbaum for the Likelihood Principle was a proof of its equivalence with two other almost *universally* accepted natural principles. The first of these natural principles is the sufficiency principle (see Section 1.7) which, for one reason or another, almost everyone accepts. The second natural principle is the (weak) conditionality principle, which is nothing but a formalization of Example 14. (Basu (1975) explicitly named the "weak" version.)

The Weak Conditionality Principle. *Suppose one can perform either of two experiments E_1 or E_2, both pertaining to θ, and that the actual experiment conducted is the* mixed *experiment of first choosing $J=1$ or 2 with probability $\frac{1}{2}$ each (independent of θ), and then performing experiment E_J. Then the actual information about θ obtained from the overall mixed experiment should depend only on the experiment E_j that is actually performed.*

For a proof that sufficiency together with weak conditionality imply the Likelihood Principle in the case of discrete $\mathscr{X}$, see Birnbaum (1962) or Berger and Wolpert (1984); the latter work also gives a similar development