principle, elicit your own initial probability density $\alpha(\lambda)$ for the unknown weight of the potato, but in practice the self-interrogation may not work very well and you may be very vague about $\alpha(\lambda)$. There is a temptation, under these circumstances, to say that $\alpha(\lambda)$ does not exist or that you know nothing about the weight of the potato. Actually, it has proved impossible to give a satisfactory definition of the tempting expression 'know nothing'. Still more, you do know a good deal about your opinions about the weight of the potato, and these can be quite well expressed in terms of partial specifications of $\alpha(\lambda)$. If, for example, it were necessary to mail the potato without weighing it, you could put on enough postage to be reasonably sure that it would not be returned to you nor be 500 per cent overpaid.

More important, you are almost sure to have a certain kind of knowledge about $\alpha(\lambda)$, which has seldom been mentioned explicitly, but which will be very useful after you have a chance to weigh the potato on a good balance. To illustrate, suppose that you found out, as a result of some experiment, that the weight of the potato to the nearest gram was either 146 or 147 gm. Given this knowledge, you would probably be willing to accept odds only slightly more favourable than one-to-one in favour of either of the two possibilities, 146 gm or 147 gm. This may be interpreted to mean that, for you, the average value of $\alpha(\lambda)$ near 146 is almost the same as its average value near 147. Continuing along this line, you might arrive at the conclusion that $\alpha(\lambda)$ varies by at most a few per cent in any 10-gm interval, included between, say, 100 and 300 gm. You might also conclude that $\alpha(\lambda)$ is nowhere enormously greater, say 1000 times greater, than even the smallest value that it attains between the bounds of 100 gm and 300 gm.

Armed with such knowledge, what could you conclude after weighing the potato on a tried and true balance known to have a normally distributed error with a standard deviation of 1 gm? Bayes's theorem, in this context, can be written

$$\alpha(\lambda \mid x) \propto \phi(x-\lambda)\, \alpha(\lambda), \tag{2}$$

where

$$\phi(z) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}z^2}. \tag{3}$$

At first sight, (2) may seem inapplicable, because you know so little about $\alpha(\lambda)$. But suppose, for definiteness, that $x = 174{\cdot}3$ gm. As Fig. 1 shows, the function $\phi(174{\cdot}3-\lambda)$ is almost zero outside the interval $174{\cdot}3 \pm 5$; the function $\alpha(\lambda)$ varies by at most a few per cent inside that interval and is never enormously larger outside the interval than it is inside of it. Under these circumstances, the product on the right side of (2) is well approximated for many purposes by

$$\phi(174{\cdot}3-\lambda)\,\alpha(174{\cdot}3).$$

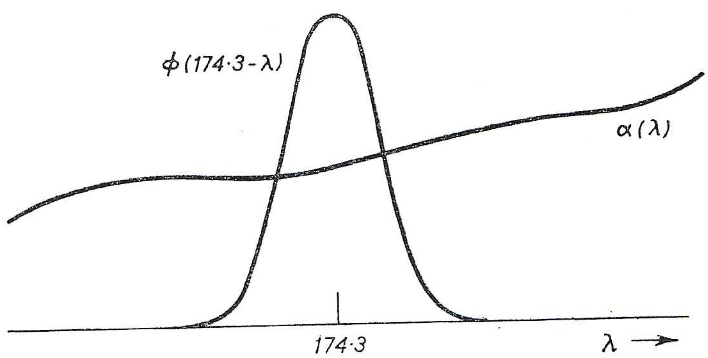

Fig. 1. Prior probability density $\alpha(\lambda)$ and likelihood ϕ $(174{\cdot}3-\lambda)$.

The probability density $\alpha(\lambda)$ and the likelihood $\phi(174{\cdot}3-\lambda)$ are not drawn to the same vertical scale. Such quantities need not generally even be of the same dimension; with Poisson data, for instance, one would be probability per unit frequency, the other simply probability. Therefore, $\alpha(\lambda|x)$ is a probability density in λ that is well approximated by some constant multiple of $\phi(x-\lambda)$, but the only such multiple that is a probability density in λ, that is, the one that is suitably normalized, is clearly $\phi(x-\lambda)$ itself. Thus, after the weighing, your opinion about λ is expressed to a good degree of approximation by saying that λ is normally distributed around 174·3 gm with a standard deviation of 1. Though this is much the kind of conclusion that is usually ridiculed in the statistics classroom, I hope you now feel that, in the presence of reasonable assumptions about your own initial subjective probability, it is not ridiculous but true.