This theorem is a consequence of Theorems 1′ and 4′ and the result of Sierpinski, used by Professor Moore in the proof of Theorem 5.

NEW YORK UNIVERSITY

# NOTE ON A SCHOLIUM OF BAYES

BY F. H. MURRAY

In his fundamental paper on a posteriori probability,* Bayes considered a certain event $M$ having an unknown probability $p$ of its occurring in a single trial. In deriving his a posteriori formula he assumed that all values of $p$ are equally likely, and he recommended this assumption for similar problems in which nothing is known concerning $p$. In the corollary to proposition 8 he derives the value

$$\int_0^1 \binom{n}{x} p^x(1-p)^{n-x}dp = \frac{1}{n+1}$$

for the probability of $x$ successes in $n$ trials. This result is independent of $x$; in a scholium he observes that this consequence is what is to be expected, on common sense grounds, from complete ignorance concerning $p$, and this concordance is considered to justify the assumption that all values of $p$ are equally likely.†

In order to complete the argument of the scholium it is necessary to show that no other frequency distribution for $p$ has the same property.

More precisely, given that a cumulative frequency function $f(p)$ has the property that for $0 \leqq x \leqq n$, $x$, $n$ being integers,

$$\int_0^1 \binom{n}{x} p^x(1-p)^{n-x}df(p) = \frac{1}{n+1},$$

* Bayes, *An essay towards solving a problem in the doctrine of chances*, Philosophical Transactions of the Royal Society, vol. 53 (1763), pp. 370–418.

† In other words, the assumption "all values of $p$ are equally likely" is *equivalent* to the assumption "any number $x$ of successes in $n$ trials is just as likely as any other number $y$, $x \leqq n$, $y \leqq n$." It has been suggested verbally by Mr. E. C. Molina that this proposition has a possible importance in certain statistical questions.

it is required to determine $f(p)$ from this equation. Now if $n=x$, the equation becomes

$$\int_0^1 p^x df(p) = \frac{1}{x+1}$$

consequently the moments of $f(p)$ are known. The function $f(p)$ can be completely calculated from these moments with the aid of a theorem of Stieltjes.*

Let

$$F(z) = \int_0^1 \frac{df(p)}{p+z} = \frac{1}{z}\int_0^1 \frac{df(p)}{1+\dfrac{p}{z}} \qquad (|z|>2)$$

$$= \frac{1}{z}\left[\int_0^1 df - \frac{1}{z}\int_0^1 p\,df + \frac{1}{z^2}\int_0^1 p^2 df - \frac{1}{z^3}\int_0^1 p^3 df + \cdots\right].$$

If $f$ is the function already discussed, this becomes

$$F(z) = \frac{1}{z} - \frac{1}{2z^2} + \frac{1}{3z^3} - \frac{1}{4z^4} + \cdots$$

$$= \log\left(\frac{z+1}{z}\right).$$

Consequently the function $f$ satisfies the equation (for $|z|>2$)

$$\log\left(\frac{z+1}{z}\right) = \int_0^1 \frac{df(p)}{p+z}.$$

From the theorem of Stieltjes, if $\psi(x)$ is a non-decreasing function of $x$, and

$$F(z) = \int_{-\infty}^{\infty} \frac{d\psi(x)}{z+x},$$

then

$$\frac{\psi(\xi-0)+\psi(\xi+0)}{2} - \frac{\psi(a-0)+\psi(a+0)}{2}$$

$$= \lim_{\eta=+0} R\left(\frac{1}{\pi i}\int_{-\xi-i\eta}^{-a-i\eta} F(z)dz\right).$$

* Stieltjes, *Récherches sur les fractions continues*, Annales de Toulouse, vol. 8 (1894), pp. 172–175. Also, Perron, *Die Lehre von den Kettenbrüchen*, p. 372.

Now the function $F(z)=\log\{(z+1)/z\}$ can be defined on the real axis by continuation, hence the limits above and below the real axis are uniquely determined. Suppose $\xi$, $a$ on the segment $0<a<\xi<1$.

Then

$$\int_{-\xi-i\eta}^{-a-i\eta}[\log(z+1)-\log z]dz$$
$$=\int_{-\xi}^{-a}[\log(1+x-i\eta)-\log(x-i\eta)]dx$$
$$=\left[\begin{matrix}(1+x-i\eta)\log(1+x-i\eta)-(1+x-i\eta)\\ -(x-i\eta)\log(x-i\eta)+(x-i\eta)\end{matrix}\right]_{-\xi}^{-a}.$$

Now

$$(1+x-i\eta)\log(1+x-i\eta)-(1+x-i\eta)$$

approaches real limits, for $x=-a$, $x=-\xi$, as $\eta\to 0$, hence makes no contribution to the sum required. We have only to consider

$$-(-a-i\eta)\log(-a-i\eta)+(-\xi-i\eta)\log(-\xi-i\eta).$$

Now as $\eta\to 0$, $-\xi-i\eta\to-\xi$. Since the approach is from below the axis of reals, and since the argument of $\log z$, like that of $\log(1+z)$, is zero for a real positive $z$, the argument here is $-i\pi$. Hence this sum becomes

$$(a+i\eta)[-\pi i+\log(a+i\eta)]-(\xi+i\eta)[-\pi i+\log(\xi+i\eta)].$$

This approaches the limit, as $\eta\to 0$,

$$\pi i(\xi-a)+a\log a-\xi\log\xi.$$

Hence

$$\lim_{\eta\to 0} R\left[\frac{1}{\pi i}\int_{-\xi-i\eta}^{-a-i\eta}F(z)dz\right]=\xi-a.$$

Substituting in the identity, we find

$$\frac{\psi(\xi-0)+\psi(\xi+0)}{2}-\frac{\psi(a-0)+\psi(a+0)}{2}=\xi-a,$$

or

$$\frac{\psi(\xi - 0) + \psi(\xi + 0)}{2} = \xi + \text{const}.$$

Consequently $\psi$ itself is continuous, $0<\xi<1$.

Now if $a>1$, $\xi>1$, the integral

$$\int_{-\xi}^{-a} [\log (z + 1) - \log z] dz$$

is seen to be real, hence

$$\tfrac{1}{2}[\psi(\xi - 0) + \psi(\xi + 0)] - \tfrac{1}{2}[\psi(a - 0) + \psi(a + 0)] = 0.$$

The same is true if both $a$ and $\xi$ are negative.

There are three additive constants yet to be determined, one on each of the intervals $(-\infty, 0)$, $(0, 1)$, $(1, \infty)$. If it is assumed that $\psi(-\infty)=0$, $\psi(+\infty)=1$, and $\psi$ is a non-decreasing function,

$$\begin{aligned} \psi(+\infty) - \psi(-\infty) = 1 = \psi(+0) &- \psi(-0) \\ + \psi(1 - 0) &- \psi(+0) \\ + \psi(1 + 0) &- \psi(1 - 0). \end{aligned}$$

The central term being one, the two remaining terms vanish. Hence $\psi(-0)=\psi(+0)=0$, $\psi(1+0)=\psi(1-0)=1$. Finally

$$\psi(\xi) = \begin{cases} 0, & \text{if} \quad \xi = 0, \\ \xi, & \text{if} \quad 0 < \xi < 1, \\ 1, & \text{if} \quad 1 < \xi. \end{cases}$$

AMERICAN TELEGRAPH AND TELEPHONE COMPANY