Note on an Application of Bayes' Rule in the Classification of Hazards in Experience Rating.

BY

ARNE FISHER.

In a recent review in the *Journal* of the Royal Statistical Society of my treatise on "The Mathematical Theory of Probabilities" under the discussion of the sixth chapter, dealing with Bayes' Theorem, the reviewer states:

"Upon the whole we agree with Mr. Fisher's conclusions respecting the theorem, but we should need much space to define our exact measure of agreement. We may, however, remark that the real value of Bayes' principle seems to reside in its application to the study of consistency, that is to say to the problem of determining whether two or more samples can properly be regarded as having been derived from one and the same 'universe.' This aspect of the matter has not engaged Mr. Fisher's attention."

This suggestion of a further—although not new—application of the famous and often misused Rule of Bayes is of some value in classifying hazards inside various industries in workmen's compensation, and I gladly take the opportunity to extend my discussion of the principle as originally developed in my book, especially in view of the fact that this renowned theorem has received very little attention among actuaries, mostly due to a completely false conception of the true Rule of Bayes as given in the usual discussions under what is known by the ambiguous name of "inverse probability."

Starting from first principles we have observed a certain event, E, the probability of which is unknown, to have happened m and failed $s - m$ times in s total trials (samples). Using the principle of equal distribution of ignorance as the basis of our calculations, merely assuming that all possible events are, in the absence of any grounds for inference, equally likely, the probability that the event, E, will occur in a following trial (i. e., in the $s + 1$ trial) is expressed by the integral:

$$P = \left[\int_0^1 y^{m+1}(1-y)^{s-m}dy \right] \div \left[\int_0^1 y^m (1-y)^{s-m}dy \right].$$

(See Fisher, "Probabilities," pages 72–74.)

The probability that the event, E, will occur n times and fail $t-n$ times in a second series of t total trials (order of happening of the individual events being immaterial) may then be expressed as follows:

$$P_{(t,n)} = \frac{\lfloor t}{\lfloor n \lfloor t-n} \frac{\int_0^1 y^{m+n}(1-y)^{s-m+t-n}dy}{\int_0^1 y^m(1-y)^{s-m}dy}.$$

Letting n assume all integral values from $n=0$ to $n=t$, we get the various probabilities that E will happen 0, 1, 2, 3 . . . or t times in the second series of t trials. The sum of all those probabilities must necessarily equal unity as some one of those combinations is bound to occur. Hence we have:

$$\sum_{n=0}^{n=t} P_{(t,n)} = 1.$$

The values of $P_{(t,n)}$ for various integral values of n are easily computed from the table of Degen. (See Fisher, "Probabilities," page 101.)

The great practical value of the formula lies in its application to test whether two samples may be regarded as belonging to the same type or universe. A few illustrations will better serve to illustrate this statement.

Example 1.—The Danish physician and biologist, Dr. Permin, in his "Tetanusstudier" gives the following observations on treatment of tetanus (lockjaw) by means of serum. One hundred ninety-nine cases of tetanus were not treated with the serum and only 42 or 21 per cent. were cured. Another sample of 189 cases were treated with the serum and 80 or 42 per cent. were cured. The question is now: Is the variation due to sampling, or would it be reasonable to assume that the serum has been favorable?

Here $s=199$, $m=42$, $t=189$, $n=80$.

Substituting these values in the formula we have, using Degen's Table:

$\log \underline{\lvert 122} =$	202.9945390	$\log \underline{\lvert 389} =$	840.2439992
$\log \underline{\lvert 266} =$	531.1078500	$\log \underline{\lvert 42} =$	51.1476782
$\log \underline{\lvert 200} =$	374.8968886	$\log \underline{\lvert 157} =$	278.0692820
$\log \underline{\lvert 189} =$	349.7071362	$\log \underline{\lvert 80} =$	118.8547277
	1458.7064138	$\log \underline{\lvert 109} =$	176.1595250
			1464.4752121

or $\log P_{(189, 80)} = \bar{6} \cdot 2312192$, $P_{(189, 80)} = \cdot 000001703$.

Hence the probability that the two samples are identical is about 2 in a million, or we may say with certainty that the serum has been beneficial.

Example 2.—A certain tannery with a payroll of 1,008,000 has shown a loss during the year of 8,000. Another tannery with a payroll of only 251,000 has in the same year shown a loss of 3,000. Would it be reasonable to assume that the second plant was inferior to the first in safety protection? I have no doubt that many of our so-called "practical" safety experts would jump to the conclusion that on the strength of those figures the second plant had shown a safety standard of 50 per cent. less than the first plant, the loss ratio being 12 per 1,000 as against 8 per 1,000 of the first plant. Now let us see how the same problems look in the light of the theory of probabilities. Choosing 10,000 as the unit of payroll and 1,000 as the unit of losses, we have here a neat little problem in chance, worded as follows. A first sample of 100 observations showed 8 successes, what is the probability that a second sample of 25 observations will give 3 successes?

The formula gives

$$(s = 100,\ m = 8,\ t = 25,\ n = 3).$$

$\log \underline{\lvert 11} =$	7.6011557	$\log \underline{\lvert 126} =$	211.3751464
$\log \underline{\lvert 114} =$	186.4054419	$\log \underline{\lvert 8} =$	4.6055205
$\log \underline{\lvert 101} =$	159.9743250	$\log \underline{\lvert 92} =$	142.0947650
$\log \underline{\lvert 25} =$	25.1906457	$\log \underline{\lvert 3} =$	0.7781513
	379.1715683	$\log \underline{\lvert 22} =$	21.0507666
			379.9043498

Hence $\log P_{(25, 3)} = \bar{1} \cdot 2672185$, or $P_{(25, 3)} = \cdot 185020$.

In other words, we may expect that the loss will be 12 per 1,000 in about 19 in 100 cases, by no means a rare occurrence. "Safety experts" please take notice and don't make rash conclusions, as it is quite probable that the apparent increase in hazard simply is due to random sampling.

I give below a complete tabulation of the probabilities of $P_{(25,n)}$ for various values of n from 0 and upwards.

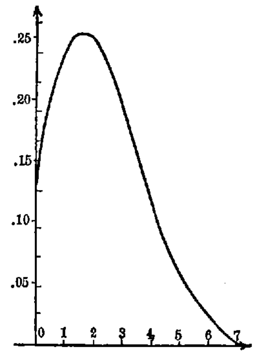
When $s = 100$ and $m = 8$

n	$P_{(25,n)}$
0	.126823
1	.243890
2	.252300
3	.185020
4	.107117
5	.051757
6	.021565
7	.007910
8	.002589
9	.000763
10	.000203
11	.000049
12	.000011
13	.000002
14–25	.000000

The above table shows that we can expect a loss of 3,000 or more in a sample of 250,000 in about 39 out of 100 cases, such excessive loss being due entirely to random sampling (chance) and not due to other influences.

Fitting the above data to a Charlier B curve (Poisson-Charlier Frequency Curve) we obtain, as will be seen from the accompanying figure, a decidedly skew distribution, indicating once more how careful we must be in using a normal Gaussian distribution in compensation work.

I could go on and quote number upon number of fallacies of medical health officers and actuaries who with truly procrustean efforts attempt to verify a pet theory of their own by samples too small to be representative. It is, I am sure, only the mathematically trained statistician who will be able to tell whether deviations from standard rates are the result of random sampling or due to truly representative causes. When preferential rate-making just now is in such vogue in American assurance circles, I can but apply a friendly warning to the statisticians and actuaries to be extremely careful and, before making a final decision, to submit the data to a painstaking mathematical analysis, which again should be undertaken only by the properly trained expert.



Mr. Mowbray, as well as Mr. Woodward, have mentioned the importance of chance variation in compensation rate making. Unfortunately only a few members of this Society seem to recognize the important bearing this has upon the whole subject of rate making, as well as the fact that such variation due to random sampling can be treated by mathematical methods only. Although the above application of some of the most elementary theorems in the theory of probabilities constitutes only a modest attempt to show what can be accomplished by such methods, I have the impression that my deductions will from many sides be viewed as having no "practical" bearing on compensation rates. Personally, I feel that this little word "practical" has been greatly abused by many statisticians and the gibe—alas only too common—that mathematical statistics is of theoretical interest only, is not justified. The engineer and the chemist use mathematics in nearly every branch of their work. Yet, nobody accuses them of being impractical, not even when the telephone engineer employs the higher criterions of probabilities in estimating the future revenues of a specified group of subscribers. Modern electrical and chemical engineering rest on essential mathematical foundations. Where would electrical engineering be to-day without the aid of the mathematical researches of a Fourier, a Helmholtz, a Hertz, a Maxwell or a Kelvin? Lord Kelvin once said that "there is no part of mathematics the engineer might not apply." I for one believe that this holds true to a still greater extent for statistics. Many a time I have had occasion to feel the limitations of my elementary mathematical training in certain statistical problems where a thorough knowledge of the higher methods of modern mathematical analysis would have carried me over the difficulties.

The day may not be so very far off when the practical statistician will be required to have a thorough mathematical training. By this I do not mean that the statistician must be a pure mathematician. Statistics must be handled *with* mathematics, not *as* mathematics. Herein lies often the danger of the pure mathematicians who often lose sight of the fact that mathematics is only a tool—although a very powerful one—in statistical analysis. This danger has been shown in the conventional and absolutely erroneous method of presenting and applying Bayes' Rule in most of those American universities I have had occasion to visit. Only through a mutual understanding between the statisticians and the mathematicians such errors in application of method to practical problems may be avoided. It is up to the statisticians to take a more conciliatory view towards the introduction of mathematical methods in statistics instead of taking a suspicious, if not actually acrimonious and ignoring attitude towards the lonely little band of students who attempt to reach a mutual understanding with the mathematicians. Such an understanding does not exist and seemingly there is a wide, and so far unbridged, gap between the mathematicians and the statisticians. What we actually need is an "entente cordiale."