THE BASIC PROBLEMS OF THE PROBABILITY USED IN MANAGERIAL PLANNING

RYŪEI SHIMIZU

I. PREFACE

The choosing principles used in the model of managerial planning with a single object may be classified into three in accordance with the the decision theory classification of the Choosing Principle under Certainty, the Choosing Principle under Risk, and the Choosing Principle under Uncertainty. The Choosing Principle under Certainty is the principle for the maximization or the minimization of object according to the marginal unit analysis. The Choosing Principle under Risk includes the principle for the maximization or the minimization of the expected value of object, the comparative principle of posterior density, the level of significance principle, etc. The Choosing Principle under Uncertainty is the principle for the maximization or the minization of object under some condition.

The object to be chosen under risk is a probability event. In this case, it is assumed that the density or the distribution of probability is known beforehand. But it is rare for the occurence probability of a social event to be known beforehand as objectively as a natural affair. The study of the probability in this risk, especially as applied to managerial planning, in other words, the elucidation of the process of managerial planning involving subjective element, is the aim of this essay.

II. THE DEFINITION OF PROBABILITY AND ITS PROBLEMS

The prevalent definitions of probability may be classified roughly into three. The first starts with the assumed axioms defining the characters of probability, and deducts from them the general theory of probability. This was advocated for the first time by A. Kolmogorov and has been the main current of the pure theory on probability to the present. The second theory is one which obtains probability from the relative frequencies of many observations performed under the same condition. This was advocated by R. von Mises, H. Reichenbach and others. The third is the theory which develops probability for each event after the "combination" theory, granting that its occurrence will be "equally probable or equiprobable." This has been the most primitive of the various probability theories since the time of P. S. Laplace and D. Bernoulli.

To be useful for the later discussion, we shall consider here the probability theory based on some assumed axioms.

First, we shall examine H. G. Tucker's[1] theory on the probability derived from an axiom, where an elementary event is ω, the collection of some elementary events, that is, the subset is A, the collection of all the elementary events is Ω and the collection of all the A's is $\mathfrak{A}$. In this case, $\mathfrak{A}$ is called a fundamental probability set, which is generally designated the "sample space."

Given the sets above, there will be among them the following relations:

(a) If A is included in $\mathfrak{A}$, A^c will be included in $\mathfrak{A}$, where A^c is a complement event of A, that is $A \in \mathfrak{A} \rightarrow A^c \in \mathfrak{A}$.

(b) When $A_1, A_2, \cdots, A_n$ are the denumerable sequence of events in $\mathfrak{A}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}$.

(c) $\phi \in \mathfrak{A}$, where ϕ is an empty set.

The group of events with the three characters above is called sigma field.

Here is the definition of probability for the sigma field: "The probability P is the function which gives the value $P(A)$ to all the A events included in $\mathfrak{A}$. $P(A)$ is called the probability of the event A, for which the following axioms are assumed:

(a) For all the $A \in \mathfrak{A}$, $P(A) \geqq 0$

(b) $P(\Omega) = 1$

(c) When any two of $A_1, A_2, \cdots, A_n$ are mutually exclusive events, $P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$

From these axioms:

(1) $P(\phi) = 0$

(2) If any two of $A_1, A_2, \cdots, A_n$ are mutually exclusive events, $P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$

(3) For all the events of A and B, $P(A \cup B) = P(A) + P(B) - P(AB)$

(4) If $A \subset B$, $P(A) \leqq P(B)$

(5) $P(A^c) = 1 - P(A)$ or $P(A) = 1 - P(A^c)$

These five are the fundamental theorems of probability derived from the axioms above.

What we should notice of the axioms above is that they do not provide for the "equally probable." Also, we should know that even the so-called addition theorem is contained in the axioms. Since the assumption "equally probable" is not included in these axioms, it may be assumed that in case of two tosses of a coin, $P(A) = P\{$both the first and the second toss turn out tail} $= P(T.T.) = 1/5$, $P(A) = P$\{at least the first or the second toss turns out head\} $= 1 - P(T.T.) = 4/5$. For $P(A)$ is a probability function for Ω which satisfies the axioms: (a), (b), (c). In other words, "to determine the probability function for the sample space S, is to conceive the mathematical model for a practical problem."[(2)] Hence, $P(A) = 1/4$, $P(A^c) = 3/4$ signify that they represent the model of "regular" coin, while $P(A) = 1/5$, $P(A^c) = 4/5$ are examples of the "irregular" coin.

Next, we shall consider the probability based on the relative frequency according to Von Mises's concept.[(3)] In the first place, we take up the finite label space S, equipped with k kinds of label space, $a_1, a_2, \cdots, a_k$, and the sequence $\{X_j\}$, each of those X_js being a_i, and considering a_i to a specific label. Also, it is assumed that among the first n elements of the sequence $\{X_j\}$ there are n_i elements which bear the label a_i. In this case n_i depends on n, a_i and $\{X_j\}$. The ratio n_i/n is designated the frequency or the relative frequency of a_i in the first n elements of $\{X_j\}$.

Here Von Mises's assumption is introduced. The sequence may be indefinitely extended, and the frequency n_i/n approaches a limited value as n gets to infinity. In other words, $\lim_{n\to m} n_i/n = P_i$, $i = 1, 2, \cdots, k$. This sequence $\{X_j\}$ is one whose elements are randomly distributed, that is, it is "insensible to place selection." This indefinite frequency and the random distribution are the two characteristics of the definition of probability which is based on relative frequency.

Next, we shall examine the third theory, which develops probability after the "combination" theory. It is assumed that all the occurrence number of events are N, none of them is duplicated, and every one of them is equiprobable. If the occurrence number of the event E is a, the probability for E to occur is a/n. This is the definition of the probability based on the "Combination" theory. These thoughts of approach, that is, "all the possible number of events" and "none of them is duplicated" assume that it is possible at the initial stage to analyze an event to its probable ultimate elements (ω, the previously mentioned sign of the axiomatic probability), and define the total set (Ω, another sign of the same) by gathering all the elements observed.

Besides, as we are unable to obtain the further knowledge of these elements, it seems quite legitimate to consider them "as equiprobable," according to Laplace's principle. In spite of the assumption that "it is possible at this initial stage to analyze an event to it probable ultimate elements and define the total set by gathering all the elements observed," here are difficult issues how to decide transcendentally the elements of what level be taken as the ultimate ones, and whether or not we are qualified to assume them as "equiprobable", even when we may grant the elements at a certain level as the ultimate ones.

The first two out of the three probability theories mentioned above call for no subjective judgment in the exercise of their axioms and assumptions as such. Even these two, however, need subjective judgment, when they are applied to a social event. The first probability, for example, which is calculated according to an axiom, depends on subjective judgment in deciding a total set and what probability be given to each event, when it is applied to a social event.

For example, in order to estimate the probability distribution of demand for a specific good in small scale investment planning, we first, on the basis of the axiomatic probability, make an assumption of the prospective maximum and minimum limit of demand, and assign a probability value for each demand so as to ascertain the estimation for distribution.

The second approach for probability, which resorts to frequency, also depends upon subjective judgment, when it is applied to a social event, in determining to what standard probability process the social event be simulated. For example, the sample inspection at random in the qualitative control of product takes the distribution of its population to be analogous to the standard probability process of the normal distribution.

Coming to the third type of probability which theorizes with the "combination" theory, we find it embraces subjective judgment even in its own axioms or assumptions such as the "equally probable." Especially when it is applied to a social event, it depends upon subjective judgment in determining, say, what kind of social event be taken as the ultimate one, hence, what type of total set or sub-set be constructed. For example, concerning the probability distribution of demand for bread in inventory control, the daily demand (or the weekly demand) is assumed to be the ultimate element. On the assumption that the daily demand (regardless of its actual amount) is equally probable, the equal probability is given, whereby distribution is calculated. In this case, if the extent of daily demand (of an element) is determined, and the aggregated probability for the demand which falls within the determined extent, the sub-set probability should be considered and determined subjectively.

Thus it is clear now that, even definition of probability, which was objectively contrived, is unable to disregard the indispensability of subjectivity, when applied to a social event, or managerial planning in this particular case.

III. THE CALCULATION of DIFFERENT KINDS of PROBABILITY AND THE INTERPOLATION OF SUBJECTIVE ELEMENT

In this Chapter are examined the probability calculation methods of different scholars, and how the subjectivity of managerial planners comes into play. The probabilities discussed here are Von Neumann and Morgenstern's subjective probabilily; Fellner's sub-probability; the post-probability by Baysian method as adapted by Raiffa and Schlaifer, and Chernoff and Moses; and the author's "assigned probability." Here we exclude the probability which can easily we estimated from the comparatively stable frequency calculated by the objective data of past.

III 1. Von Neumann and Morgenstern's[4] subjective probability

There are various methods in calculating the subjective probability. Here is taken up the Von Neumann and Morgenstern's method. The essence of this method is found in the establishment of the subjective probability of the event of non-standard process on the testified assumption of human preference for consistency, the utility function derived from it, and another assumption on the indiscriminatory character of the expected utility, seeking the subjective probability of the event of non-standard process.

Concerning the human preference for consistency, the axioms below have been contrived, and the function of linear utility is obtained from them.

In the first place, the system U consisting of the entities u, v, w, is formed; it is assumed that within this system, there exists a relation $u > v$, and the applicability of the operation, $\alpha u + (1-\alpha)v = w$ to any value of $(0 < \alpha < 1)$. In this case, the elements of the system U satisfy the following axioms:

Axiom A: $u > v$ is a complete ordering of U. In other words,

(A, a) For any two elements u, v, one of the three relations below, and at least one relation is possible. $u = v, u > v, u < v$.

(A, b) If $u > v$ and $v > w$, then $u > w$.

Axiom A, taking into consideration the human attitude toward preference, attempts to give ordering and transivity to utility.

Axiom B: Ordering and combining.

(B, a) If $u < v$, then $u < \alpha u + (1 - \alpha)v$.

(B, b) If $u > v$, then $u > \alpha u + (1 - \alpha)v$.

These two, (B, a) and (B, b), are the so-called "independent axioms." They signify: when v is independently preferred to u, the combination event which brings u and v respectively at the rate of α and $(1 - \alpha)$

is preferred to u.

(B, c) If $u < w < v$, there will be α which satisfies $\alpha u + (1 - \alpha)v < w$.

(B, d) If $u > w > v$, there will be α which satisfies $\alpha u + (1 - \alpha)v > w$.

These axioms, (B, c), (B, d), intend to give continuity to utility. In other words, they, (B, c) and (B, d) assume that if a proper value is given to α, any two utilities can be so combined with the probability that the utility by the probabilistic combination will fall between the two utilities.

Axiom C: The algebra of combining.

(C, a) $\alpha u + (1 - \alpha)v = (1 - \alpha)v + \alpha u$.

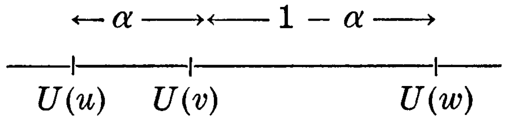
(C, b) If $\gamma = \alpha\beta$, then $\alpha\{\beta u + (1 - \beta)v\} + (1 - \alpha)v = \gamma u + (1 - \gamma)v$

(C, a) assumes that the composing elements u and v, may be arranged in any order. In other words, the mathematical commutative theorem is taken for granted. (C, b) is the axiom which was contrived to apply the conditional probability rule to utility function.

If the system U satisfies the axioms above, it is possible to give numerical value to the utility differences and their utilities, that is, it becomes possible to provide utility functions. Suppose now a person evaluates the elements u, v, w (in this case, they are taken as analogous to the results of some acts), and sets their preference order to be $u < v < w$. He would be considered to be judging a compound event,—that is, simple result u occurs at probability α, and simple result w occurs at probability $(1 - \alpha)$,—and simple event v to be equally preferable. In other words, if the preference order of $u < v < w$ is clarified, the existence of α which will satisfy the subjective indiscrimination of the compound element and the simple element,—that is, satisfy $U(u)\alpha + U(w)(1 - \alpha) = U(v)$,—will be recognized.

Thus there will be assigned a numerical value to the utilities of u, v, w, so that the ratio $(U(v) - U(u))/(U(w) - U(v))$ will be equal to $\alpha/(1 - \alpha)$. If U in III: 1: Chart 1 be taken as the original point, in other words, if $U(u) = 0$, and $U(w) = M$, then the distance of

III: 1: Chart 1



$U(v)$ from the original point will be $M\alpha$. As can be seen in this instance, if the minimum point of preference degree, that is, the original point (utility = 0), and the maximum point of preference degree (utility = M in this case), are set, $M\alpha$ that is, an utility value for each of the elements v_i (or results) falling between the maximum point and the minimum point will be uniquely determined according to the indiscriminatory principle of preference. This is the linear utility function of $u, v_1, v_2, \cdots, v_i, \cdots, v_n, w$.

Now, let these elements be the results ($f_i(S_j)$) in the Payoff Matrix as presented below in III: 1: Chart 2:

III: 1: Chart 2

	S_1	S_2
f_1	$f_1(S_1)$	$f_1(S_2)$
f_2	$f_2(S_1)$	$f_2(S_2)$

So long as these results satisfy the axioms above, the linear functions of the results will be established. With a assumed conditions described in III: 1: Chart 2, the event S_1, that is, the average annual increase rate of sales for the coming 5 years is less than 10%; the event S_2, that is, the average annual increase rate of sales for the coming 5 years is no less than 10%; act f_1, that is, the 500,000,000 yen equipment investment is made; act f_2, that is, the 800,000,000 yen equipment investment is made; result $f_1(S_1)$, that is, the average annual profit is 70,000,000 yen; result $f_1(S_2)$, that is, the average annual profit is 100,000,000 yen; result $f_2(S_1)$, that is, the average annual profit is 30,000,000 yen; result $f_2(S_2)$, that is, the average annual profit is 200,000,000 yen, the linear utility function of these results will be established.

Let us now try to seek the subjective (occurrence) probability of events out of the linear utility function of the results above; the subjective probability of the state S_1 which is likely to happen in future. See III: 1: Chart 3.

III: 1: Chart 3

	S_1	S_2
f_1	$f_1(S_1)$	$f_1(S_2)$
f_2	$f_2(S_1)$	$f_2(S_2)$
$\vdots$	$\vdots$	$\vdots$
f_n	$f_n(S_1)$	$f_n(S_2)$

Here we take up f_i and f_j, the expected values of which are equal, out of the set of acts or plans $F = (f_1, f_2, \cdots, f_n)$ that are conceivable at present, that is, there will be $U(f_i(S_1))\cdot P(S_1) + U(f_i(S_2))\cdot P(S_2) = U(f_j(S_1))\cdot P(S_1) + U(f_j(S_2))\cdot P(S_2)$, where $P(S_1)$ represents the subjective probability of S_1. If n is enlarged, such acts, the expected utilities of which will be equal, are sure to be found. On the assumption, therefore, of $P(S_1) + P(S_2) = 1$, out of the formula above is derived:

$$P(S_1) = \frac{U(f_j(S_2)) - U(f_i(S_2))}{U(f_i(S_1)) - U(f_i(S_2)) - U(f_j(S_1)) + U(f_j(S_2))}\,.$$

Even when n is small, it is possible to find acts, the expected utilities of which will be equal, so long as the repetition of f may be possible to some extent, and thus the compound acts of f can be conceived, and hence to seek the subjective probability of S_1.

Using the example above, we shall now consider various questions in the calculation of subjective probability. In this case, since S_1; the average annual increase rate of sales for the coming 5 years, is set to be less than 10%, and S_2; the average annual increase rate of sales for the coming 5 years is set to be no less than 10%, it is always easy to assume $P(S_1) + P(S_2) = 1$. But if S_1; the average annual increase rate of sales for the coming 5 years is 10%, and S_2; the average annual increase rate of sales for the coming 5 years is 20%, it is very difficult to assume $P(S_1) + P(S_2) = 1$.

Besides, most of the practical cases in managerial planning will have to take into consideration several events for future. To conceive a matrix like this requires a large n, making the construction of such a matrix itself very difficult. Especially, it is very difficult for an individual to judge the indiscriminatory character of two acts, while examining each result of numerous events simultaneously.

Thus this Von Neumann and Morgenstern's method of calculating subjective probability can not be considered very useful on such an occasion when the number of events (states) requiring their future perspective are likely to increase, and also on such a "business phenomenon" for the analysis of which subjective utility may not be advantageously utilized.

III 2. Fellner's(5) Semi-probability.

To calculate the semi-probability, Fellner explains first the standard process[1] postulate, which serves as the standard for the calculation

1) Suppose a set of events occurred in a physical process which he intended to make strictly analogous to such an event as the toss of coin or random sampling card, such a physical process is called a standard process of an individual.

of "general probability," and then the semi-probability act postulate which serves as the standard for the calculation of semi-probability.

In other words, according to the "general probability theory", an individual act is required not only to satisfy the utility principle of Von Neumann and Morgenstern, but also to meet the postulate of equivalent degree of belief. This means that a rational human being must give to a non-standard process event an equivalent probability and weight as he gives to a standard process event, when he is concerned about monetary gain. In other words, if E' stands for non-standard process event, and E for standard process event, there should be $P(\sim E') = 1 - P(E)$.

On the other hand, the act of the semi-probabilistic individual who appears in the semi-probability theory may satisfy the utility principle of Von Neumann and Morgenstern, but may not necessarily meet the postulate of equivalent degree of belief. In other words, the degree of belief is rather ambiguous in a non-standard process, and is not quite stabilized even at a particular point of time psychologically. So semi-probability theorizers adopt the slanted probability as the determined weight, instead of the postulate of the equivalent degree of belief. This slant signifies the instability of the judgment for probability, and the width of slant is in accordance with the size of gain bidded for a probability event.

As the conditions for defining the semi-probability act, Fellner mentions three: (1) ordering of possible transitivity, (2) the principle of "probabilization," (3) the common limits in which true probability falls. The characteristics of this semi-probabilitic act consist in the inapplicability of the so-called independent axiom to the semi-probability, the value of semi-probability varying with the size of bid, and the sum of the slanted semi-probability, as the determined weight, not necessarily amounting to 1.

The construction of a corrected probability from a semi-probability as the determined weight is made through the following procedure:

In the case when all the uncorrected probabilities, which belong to a set, are equal, all the corrected probabilities will also be equal, and their sum will amount to 1. For example, when the gain from a bid is 100 dollars, a person may give a non-standard process event E', $\sim E'$, the same determined weight, that is, uncorrected probability, $Pu(E') = Pu(\sim E') = 0.2$. Suppose in this case, the gain from the bid varies to a value other than 100 dollars, the values of the two uncorrected probabilities will be the same, although they may differ from 0.2. It seems, therefore, correct to consider the values of all the corrected probabilities to be equal. If the two events E' and $\sim E'$ only are considered, there will be $Pc(E') = Pc(\sim E') = 0.5$, where Pu denotes an uncorrected probability, and Pc, a corrected one.

When the uncorrected probabilities which belong to a set are different, only the limits for the corrected probabilities will be determined, instead of their corrected probabilities being uniquely determined. Suppose a person puts up the uncorrected probabilities for a certain gain of a bid: $Pu(E') = 0.6$, $Pu(\sim E') = 0.2$, his corrected ones, $Pc(E')$ and $Pc(\sim E')$ will fall between 0.6 and 0.8, 0.2 and 0.4 respectively. If this definite amount of gain from a bid changes, being followed by the changes of value for $Pu(E')$ and $Pu(\sim E')$, the limits for Pc will also be affected. But if this person should act consistently, and satisfy the third conditions mentioned above, the limits for Pc will hold on the common element, although the gain from the bid may very. For example,

limits in which $P_{c,10}$ falls $\subset$ limits in which $P_{c,100}$ 100 falls $\subset$
limits in which $P_{c,1000}$ falls

The $P_{c,i}$ above denotes the corrected probability in case of the gain being i.

As can be seen from the above, corrected probabilities can not be uniquely calculated when uncorrected probabilities differ. But even in this case, if it is possible to assume that "the uncorrected probability was obtained from a true probability by slanting down or up on same rate," a corrected probability can be constructed uniquely from a different uncorrected one. For example, when the uncorrected probabilities are $Pu(E') = 0.6$ and $Pu(\sim E') = 0.2$, and if the assumption above is satisfied, $x = 5/4$ will be derived from $0.6x + 0.2x = 1$, hence $Pc(E') = 0.6 \times 5/4 = 3/4$, $Pc(\sim E') = 0.2 \times 5/4 = 1/4$.

In the practical application, however, of the corrected probability in the study of environmental conditions (the states of an event), as well as in the choice of plan for an enterprise, this corrected probability should be multiplied by the conditional probability obtained by the objectively observed data, so as to seek the posterior probability.

As is clear from the above, Fellner's theory is characterized by its denial of the possibility to set up the fixed linear utility function, as devised by Von Neumann and Morgenstern, from which the subjective probability is derived through the application of the indiscrimination principle of the expected value, as well as by its denial of the possibility to bring "automatically" the sum of the semi-probabilities, that have been sought as the determined weights, to 1, since the determined weights get disturbed by the size of bid in a non-standard process event.

So, Fellner's theory tried to seek the subjective probability by modifying the semi-probabilities, although the method was somewhat awkward, to obtain such subjective probabilities, the sum of which would be 1. Then he tried to seek such posterior probabilities which would be effective in the choice of plan, by applying the Baysian method to the previously obtained subjective probabilities. It is true, as Fellner claims, that the distribution (needless to say, the density) of posterior probabilities (assumed) can be calculated, if some objective, exact conditional probabilities are secured, even though the corrected probabilities are rather strong in their subjectivity.

Is it, however, possible to obtain conditional probabilities for many events? The business events which call for the application of semi-probabilities are the ones that repeat themselves rather less frequently. Fellner's method for the calculation of probability also seems to have a great difficulty because of the lack of the objective data essential in the search of conditional probabilities. It is especially difficult to have objectively testified conditional probabilities, when some change in economic structure takes place. The result will be an inevitable dependence on subjective elements in the calculation of conditional probabilities.

III 3. The calculation of probability by the Baysian method as adopted by Raiffa and Schlaifer[6], or by Chernoff and Moses

The Baysian method primarily attempts to seek posterior probability (density or distribution), first by setting intuitively the prior probability (density or distribution), and then multiplying this prior probability by the conditional probabilities (densities) obtained through all sorts of information.

Let us now examine in outline the Baysian method as adopted by Raiffa and Schlaifer. In the first place, an assumption is made about the family of conjugate distributions, and then its parameter is estimated somewhat in line of the subjective judgment of a decision maker. The prior probability thus obtained will be next modified by the newly calculated observed value, to estimate the posterior probability distribution. In order to make an efficient use of the newly calculated observed values one by one a kernel of the likelihood of observed values and a kernel of conjugate prior densities are expressed with the sufficient statistic peculiar to the family of conjugate distributions of the prior probability distributions, as parameter. In other words, this is an attempt to seek the posterior probability density by a simple mathematical exercise using the sufficient statistic which expresses the prior probability distribution, based mainly on subjective judgment, and the likelihood of observed values. This method is basically derived from the Bayse's theorem that the posterior probability density is proportionate to the product of the kernel of prior probability density multiplied by the kernel of the likelihood of observed values.

As the conjugate prior distribution which makes the Baysian treatment like this possible, only the specific density functions such as the normal distribution, the beta distribution, the gamma distribution, and their combined distribution are used.

To illustrate this method, let us now consider the true average influence which a design on packing sheet exercises on the sale to the customer per head. This true influence is expressed by $\tilde{\theta} = \tilde{\mu}$. First, let us explain the process seeking the prior probability density by subjective judgment.

Assumption: The sale to the customer per head is assumed to take the normal distribution with the known precision. Consequently, the variance $1/h$ is also known. In other words, the family of conjugate distributions is assumed to be the normal distribution.

Examination of assumptions: (1) Ask the decision maker of the best estimate value ($\hat{\mu}$) concerning the true value of $\tilde{\mu}$, (2) Ask if the distribution of $\tilde{\mu}$ for this true is asymmetrical, (3) If the question (2) is answered affirmatively, ask the decision maker if he would bid the same amount of money, regardless the proposition whether the true value of μ falls within the asymmetrical interval or outside of it, in order to have the exact knowledge of the asymmetric interval about $\hat{\mu}$.

Estimation of the prior probability: Next, estimate this assumed parameter of normal distribution, using the decision maker's subjective judgment, that is, determine quantitatively the prior probability by the following procedure: according to the assumed noraml distribution, calculate the probability, $P\{\tilde{\mu} > \mu\}$ for some value of μ, and also calculate μ which will satisfy $P\{\tilde{\mu} > \mu\} = \alpha$ for some values of α. Ascertain whether the results of these calculations, and these probabilities, or the subjective judgments for μ agree or not, by asking the decision maker by trial and error if he is willing or refuses to make the bid, so as to estimate their parameters and determine the prior probability.

Expressed by a mathematical formula, this density of prior probability is:

$$f_N(\mu \mid m', hn') = (2\pi)^{-1/2}(hn')^{1/2}e^{-(1/2)hn'(m'-\mu)^2}$$

Where:

$$-\infty < \mu < \infty\,, \qquad h > 0$$
$$-\infty < m' < \infty\,, \quad n' > 0$$

Hence:

Sufficient statistic n', m' is estimated.

Consequently:

$f_N(\mu \mid m', hn')$ can be estimated.

And the kernel for this:

$$e^{-(1/2)hn'(m'-\mu)^2}$$

Let us now explain the process for calculating the posterior probability using conditional probability. First, calculate the sufficient statistic, affecting the posterior probability, from the objectively observed new sample value. Seek the statistic: n=number of observed value x and $m = 1/n\,\Sigma x$, as the sufficient statistic in the case of the family of conjugate normal distributions, for the likelihood of the observed sample value is proportionate to a kernel, $e^{-(1/2)hn(m-\mu)^2}$.

From the newly observed n and m can be easily derived the posterior probability as follows:

$$f_N(\mu \mid m'', hn'') = (2\pi)^{-1/2}(hn'')^{1/2}e^{-(1/2)hn''(m''-\mu)^2}$$

Where:

$$n'' = n' + n$$
$$m'' = \frac{1}{n''}(n'm' + nm)$$

If another new value $\tilde{x}'$ has been observed, derive therefrom new n and m to repeat the operation above, so that the modification by the conditional probability of the observed value can be easily performed. Thus the posterior probability will be obtained by multiplying the prior probability by the conditional probability.

The above is the Baysian method as adopted by Raiffa and Schlaifer. This method has the merit of being able to utilize the objective data gathered in the process of managerial planning one by one mathematically, but at the same time it has the disadvantage that its use is limited only to the direct and objective information about $\tilde{x}$, because of its being too precise mathematically.

Also, it has the drawback that it is capable of treating only such specific distributions as the beta distribution, the gamma distribution, and the normal distribution mentioned before.

In order to do away with these weaknesses, Chernoff and Moses contrived their own interpretation of the Baysian method, making use of the concept of conditional probability as it is (that is, without paying any attention to such concepts as the kernel and the sufficient statistic) to obtain posterior probabity, although the result would be somewhat rough mathematically.

Let us explain it by an example. Suppose we build an automobile assembly factory in South Eastern Asia. There, the automobile market can be conceived as divided into two: one inhabited by the prosperous Chinese businessmen who, being rational and practical, want small-sized cars, say, 300 cc type, and the other mainly occupied by the farming landowner class who, being vain of their appearance, like to have showy, large or medium-sized cars, say, 1,500 cc type. It is important, therefore, to find out which type of car a district in South-Eastern Asia is likely to demand, for there is a great difference in the building conditions of an automobile factory according to the type of car produced.

Now, we will see the process of applying the Baysian method in this particular case. The district, which demands the unpractical medium-sized car, is designated θ_1, and the district which prefers the practical small-sized car, θ_2. Also, suppose you have obtained the set of information below:

E_1: The district where the main crop is rice and canals are used for its distribution;

E_2: The district where the earners of large income are found among landowners

E_3: The district which sends a large number of youth to Japan for study, the majority of them being the sons of prosperous Chinese businessmen.

Under the assumption that the district in question is one which demands unpractical medium-sized cars, that is, θ_1, the following conditional probabilities are set by subjective judgment:

$$P\{E_1 \mid \theta_1\} = 0.6 \qquad P\{E_2 \mid E_1, \theta_1\} = 0.8$$
$$P\{E_3 \mid E_1, E_2, \theta_1\} = 0.3$$

Under another assumption that the district in question is one which demands practical small-sized cars, that is, θ_2, the following conditional probabilities are set by subjective judgment:

$$P\{E_1 \mid \theta_2\} = 0.3 \qquad P\{E_2 \mid E_1, \theta_2\} = 0.5$$
$$P\{E_3 \mid E_1, E_2, \theta_2\} = 0.9$$

Now, bring E_1, E_2, E_3 together with the sign Z. The conditional probabilities, that Z is gained under the condition θ_1 or θ_2, are respectively:

$$f(Z \mid \theta_1) = 0.144 \qquad f(Z \mid \theta_2) = 0.135$$

Further, the prior probabilities of θ_1 and θ_2, taking into consideration the historical conditions of the country, are assumed to be $w_1 = 0.45$, $w_2 = 0.55$.

Hence their posterior probabilities are respectively:[2]

$$P\{\theta = \theta_1\} = \frac{w_1 \cdot f(z \mid \theta_1)}{w_1 \cdot f(z \mid \theta_1) + w_2 \cdot f(z \mid \theta_2)} = 0.466$$
$$P\{\theta = \theta_2\} = \frac{w_2 \cdot f(z \mid \theta_2)}{w_1 \cdot f(z \mid \theta_1) + w_2 \cdot f(z \mid \theta_2)} = 0.534$$

This Chernoff and Moses type of Baysian method also has the following weakness: Since it uses conditional probability as a multiplied product, it is capable of only making use of indirect information in the same direction which makes it possible to grasp the causational relationship between the combined event of $E_1, \cdots, E_i, \theta_i$ and the event E_{i+1}. In other words, it is unable to use indirect information of different directions, the causational relationships of which can not be followed.

In order to make up this weakness, the "assigned probability" which makes no use of conditional probability as a product has been devised, although it further recedes in exactness as a theory.

In spite of the weakness mentioned above, the Baysian method is credited with the general advantage that it, applying the conditional probability and likelihood, succeeds in making a systematic use of the information which is very close to the decision maker's value structure and perception structure,[3] so that the calculation of probability

2) In a practical choice by the Baysian method, each of these posterior probabilities is multiplied by the regret (that is, a factory was built according to the misjudgment of the real value of θ, derived from this information) to calculate an expected regret. The construction plan with the lesser expected regret is adopted.

3) A person's value structure means the system of value based on his inborn traits and the education in a broad sense which he has received. A person's perception structure means the collection of data that have become his consciousness and of the data that are related to them as consciousness. It is not a mere physical filing of information. See the reference (8).

will be systematically carried.

III 4. The "assigned probability"[8]

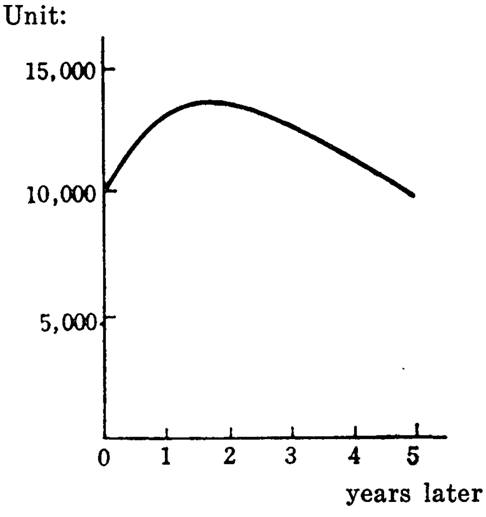
As we see with the definitions of probability so far discussed, we find that they always comprise such a subjective judgment as the assumption of "total set", "sub-set", "convergency", or "equiprobable" when probability is applied to a social event. And in an actual calculation of the density or the distribution of probability, the subjective judgment as mentioned above is taken for granted. Coming to the "assigned probability", we rather make a positive use of subjective judgment.

In the first place, we make an assumption of the "total set" and the "sub-set" of various events in the non-standard process with the exception of the events that do not appear like to bring about an important result at the time of planning, or some very elusive events. A distribution of probability is assigned intuitively and directly at the events thus selected.

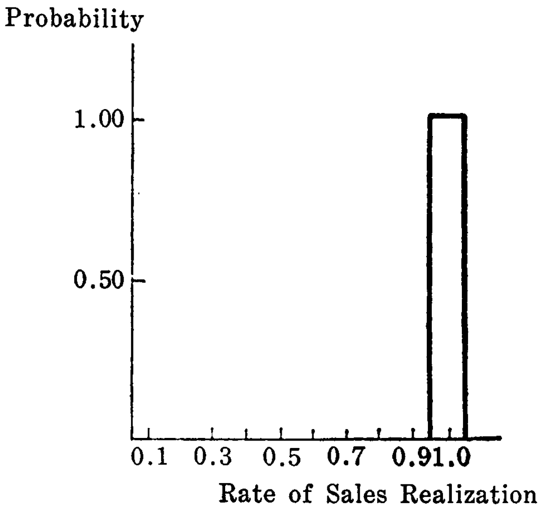
Next, the direct or the indirect data which are organically related to them are gathered, and they are used through the trial and error procedure so as to effect the necessary modification of the assigned probability little by little. In this case, the modification is made without "being particular about the product of conditional probability."

In the last place, this modified distribution of probability is once again examined by analogously treating it as a standard process event so that the obtained probability of distribution will not be grossly amiss, although the non-standard process event may become too abstracted. If there is found no great error, the distribution of

III: 4: Chart 1
Trend of Estimated Sales



III: 4: Chart 2
Probability Distribution (1)



probability in question will be taken as final.

The calculation process of the "assigned probability" will be illustrated by the simulation model of a simple managerial planning presented below:

Here is a medium-sized company *A* manufacturing precision machines. *E*, one of its customers and a large electric apparatus maker, inquired of *A* whether *A* would make a part of a newly-devised good. *A* was willing to accept the order, since *E* had been a good customer of *A* for many years in the past. The trouble was, however, that *A* had to procure immediately some machines necessary to make the part in question, in addition to the ones they had. Also, the new good for which the part was essential was a portable ice box, something new to the Japanese, and *A* was not sure if this new apparatus would be popular with the Japanese. Therefore *A* felt rather uncertain as to what estimate he should make for the prospective order.

So, *A* made an estimate of the probable profit expected of the prospective order. Using the trend of estimated order for the coming 5 years as depicted by *E* (III: 4: Chart 1), *A* set the occurrence probability of the sales pattern as depicted by *E* to be 1. (The occurrence probability of the sales, the rate of realization[4] of which will be 1.0, was made 1. See III: 4: Chart 2.). Making a detailed simulation of the labor cost, the raw material cost, and other definite costs, as well as the equipment investment and the constant amount of stored raw material, *A* found the rate of return on invested capital to be 60%. Further, *A* made another calculation on the basis taht he purchased all the pieces of equipment anew, instead of using the old ones, and found that in that case the rate of return on the invested capital would be 50%.

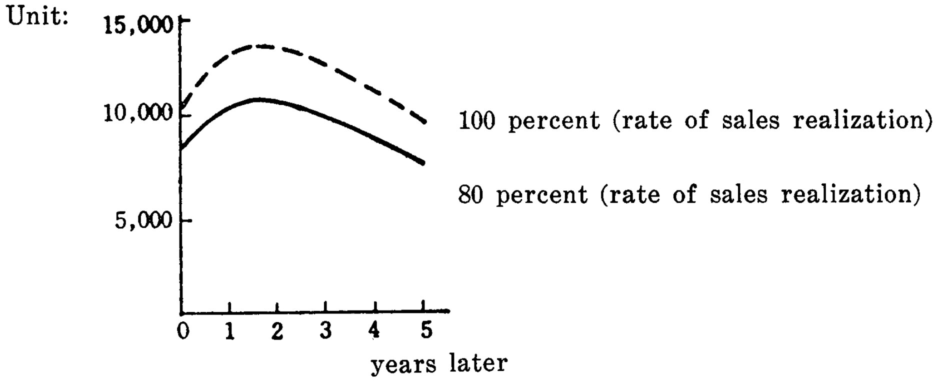
As a results of this simulation, *A* became aware that *E*'s order was not very exact. For *E* might have increased the number of technicians so that they could make the part themselves, if they had known that the rate of return on investment would amount to that much.

So, taking the pattern of the life-cycle of the new product to be the same as the trend of estimated order shown by *E*, *A* assumed the probability of the event that the prospect of the *A*'s sales estimate

4) $$\text{Rate of sales realization} = \frac{\text{Estimated amount of order from E}}{\text{Amount of order placed by E}}$$
$$= \frac{\text{Amount of order expected by A}}{\text{Amount of order placed by E to A}} = \frac{\text{Estimated sales of A}}{\text{Estimated sales of A placed by E}}$$

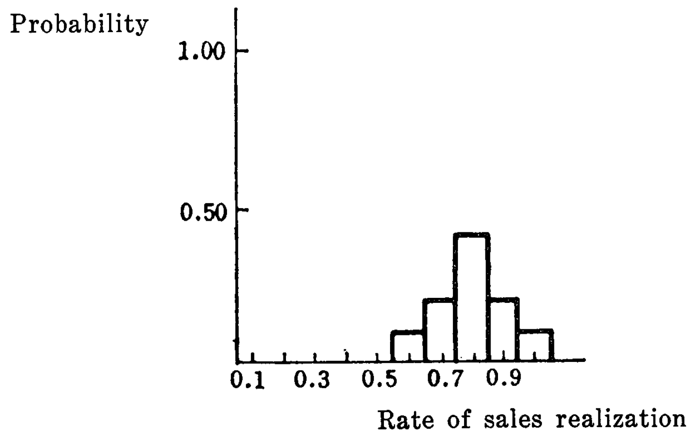
would be on the average 80 per cent of the amount which E had estimated. (III: 4: Chart 3)

III: 4: Chart 3. Prospect of *A*'s Sales Estimate



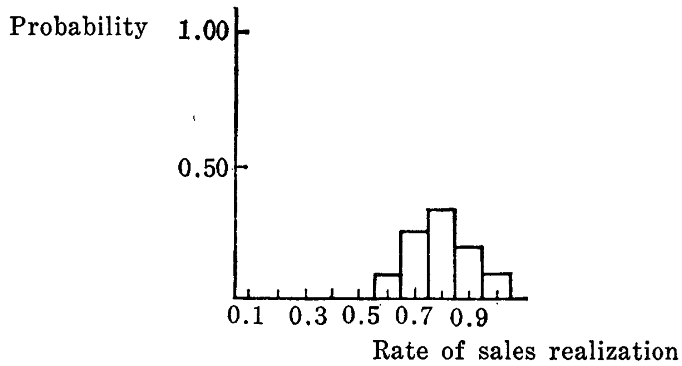
In other words, assuming the probability distribution (normal distribution, the horizontal axis of which designates the rate of sales realization, III: 4: Chart 4), the expectation of which would be 80 per cent, and conducting a simulation on the basis of the probability obtained, we found the estimated rate of return on investment (mathematical expectation) to be 35 per cent.

III: 4: Chart 4. Probability Distribution (2)



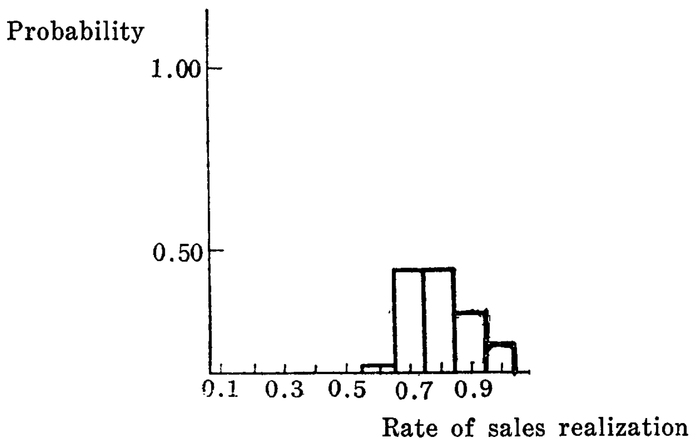
While this order was under consideration in A, the information revealed that B, another precision machine maker and A's competitor, declined the E's order. A knew that B had accepted many orders in the past, if the estimated rate of return on investment was more than 30 per cent. In other words, the decline this time by B signified that the rate of sales realization estimated by B had proved to be less than 80 per cent. So, the probability density of the rate of sales realization 80 per cent was reduced by 0.05, and this 0.05 was added to the probability density of the rate of sales realization 70 per cent. (III: 4: Chart 5)

III: 4: Chart 5. Probability Distribution (3)



Next, the simulation model was perturbed inversely, so that the rate of return on investment would be 30 per cent, and it was found that the E's resultant amount of order would be considerately less, and A's rate of sales realization would be as low as 60 per cent. If such should be the case, the sales of the new product would be very small in amount. As it was inconceivable that E, which has a central research institute for the study of demand forecasting, would dare to set forth for the sale of a new product having such a low sales forecast, the probability density of the rate of sales realization 60 per cent, was reduced by 0.08, and that 0.08 was added to 70 per cent, as is shown in the probability distribution chart, III: 4: Chart 6.

III: 4: Chart 6. Probability Distribution (4)



The next thing which A was very much concerned was the conceivable minimum sales, and decided to look into the case of the rate of sales realization of 50 or 40 per cent. So, A made a private inquiry of the marketing research conducted by the sales section of *E*. It was revealed that *E* had sent out a series of enquete to the owner drivers through car makers to obtain information of the prospective demand for the portable ice-box. *A* conducted a research of the owner drivers among its own employees asking the same questions as were found in the enquete. They learned that 40% of the respondents wanted the portable ice-box at the price of 10,000 yen. Also, they secured a rough knowledge as to what income levels or personal environmental conditions require such an apparatus through a cross analysis study. Taking further into consideration the increasing trend of owner drivers' number, we estimated the potential cumulative demand for the coming 5 years. This estimated cumulative demand proved possible to bring in the total amount of sales which would be larger than that calculated on the basis of the realization rate 50 per cent, for the *E*'s prospective order for the coming 5 years. So, it is hardly conceivable that the rate of sales realization would be 40 or 50 per cent, if one should consider the *E*'s efforts for the promotion of sales by such a method as advertisement or propaganda hereafter. Consequently the probability distribution, especially its lower limit (III: 4: Chart 6) was left as it was.

Receiving no further conceivable information which could indirectly relate to the *E*'s inexact estimate, and seeing no possible way of adjusting it, *A* decided to examine it for the last time by analogizing it to the standard probility process. But generally it is a common sense[5)] that the standard probability process event which should be analogized to an error in demand be taken as an accidental erroneous event. Its distribution shows the normal distribution, being high in the centre and low at both tails. Thus III: 4: Chart 6 was judged not to be very erroneous, and it was determined that this probability distribution be taken as final.

Currently it seems that under this theory of "assigned probability" there are some functional relations which are generally recognized

5) Since the distribution of probability which is commonly held as true often proves entirely different from the scientifically ascertained truth, it is important that the so-called "common sense" knowledge of this sort be subjected to scrupulous investigation.

Suppose the two companies of the same size are in competition for market share. It is important to consider the probability that one of them continues to take lead against the other. In common sense, it is quite possible that such a probability can happen equally to each of them. In other words, it is conceivable that, in a graphic presentation, the probability distribution is highest in its density at the centre. But when we analogize the event of this competition to the standard probability process, it has been discovered that the probability distribution would be lowest at the centre.[(9)]

among various factors in and outside of an enterprise, and that there exist certain rough but admissible simulation models of those phases; especially, it seems, assuming that the establishment of fairly stable models among the factors within the enterprise, and the examination of those internal factors by the trial and error method such as the computer, are possible. On the condition, and only on the condition that the examination of internal factors becomes possible by the trial and error method such as the computer, a rough knowledge, that is, the indirect information of the relations prevalent among the various phases of an enterprise proves to be an useful agency in the calculation of probability distribution.

CONCLUSION

It is clear from the discussion above that subjective judgment is indispensable in the estimation of the density or the distribution of probability for a future event. Subjective judgment, however, should not be taken as something, the function of which is only reluctantly admitted in the process of probability estimation. It is essential, therefore, that the probability concept be effectively utilized in order to apply *subjective judgment positively* and *systematically in any managerial planning*. So, the method for the calculation of the density or the distribution of probability should be selected in accordance with whether it is convenient for a comprehensive and systematic exercise of subjective judgment. In other words, the probability calculation method is closely related to the decision maker's value and perception structure.

Since the probability calculation method is selected by the process described above, the choosing principle of managerial planning, which makes use of the probability, also should be not mechanically applied. For example, to an event which is not very important and which occurs repeatedly in an enterprise, the principle of mathematical expectation or significance level may be applied, while to an event which is relatively important to an enterprise and which is rarely repeated, such a principle which enables to examine the probability density of the least advantage or the principle of mathematical expectation is to be applied. Further, to the business phenomenon which involves very few assumed events, such a principle as the principle of comparing posterior probability densities is applied.

Having surveyed the various managerial planning models and having seen the importance of the rational application of probability, one finds it is essential that the calculation process of the density or the distribution of probability, especially the process where subjective judgment plays a significant role, be definitely clarified.

REFERENCES

[1] H. G. TUCKER: *An Introduction to Probability and Mathematical Statistics*, pp. 9-11, Academie Press Inc., 1963.
(2) Ichirō TAJIMA: *Statistics* (*A New Book on Modern Mathematics*), p. 24, Shibundō, Tōkyō, 1964.
(3) R. Von MISES: *Mathematical Theory of Probability and Statistics*, pp. 5-10, Academie Press Inc., 1964.
(4) J. Von NEUMANN and O. MORGENSTERN: *Theory of Games and Economic Behavior*, pp. 26-27, John Wiley & Sons Inc., 1964.
(5) W. FELLNER: *Probability and Profit*, pp. 57-169, Richard D. Irwsn Inc., 1965.
(6) H. RAIFFA and R. SCHLAIFER: *Applied Statistical Decision Theory*, pp. 28-76, Division of Research Graduate School of Business Administration, Harvard Univ., 1961.
(7) H. CHENOFF and L. E. MOSES: *Elementary Decision Theory, Parts VI-VII*, John Wiley & Sons Inc., 1959.
(8) Ryūei SHIMIZU: *Theories of Managerial planning* (*Keiei-keikaku-settei-riron*) pp. 121-133, Chūō-keizai-sha, Tōkyō, 1966.
(9) W. FELLER: *An Introduction to Probability Theory and Its Applications*, Vol. I, pp. 77-81, John Wiley & Sons, 1957.