İstanbul Üniv. Fen Fak. Mat. Der. **51** (1992), 97-105

THREE PRINCIPLES OF DATA REDUCTION IN STATISTICAL INFERENCE

İ. H. ARMUTLULU

Sufficiency, likelihood and invariance are three basic principles of statistical inference. An investigator uses the information in a sample to make inferences about unknown parameter θ. Any statistics, $T(\underline{X})$, defines a form of data summary. A sufficient statistics for a parameter θ is a statistics that, in a certain sense, cover all the information about θ contained in the sample. This consideration leads to the sufficiency principle. Consider experiment $E = \{X_1, X_2, \ldots, X_n, \theta, \{f(\underline{x}; \theta)\}\}$ and suppose $T(\underline{X})$ is a sufficient statistics for θ. If $\underline{x}$ and $\underline{y}$ are sample points satisfying $T(\underline{x}) = T(\underline{y})$, then evidence of $(E, \underline{x})$ is equal to evidence of $(E, \underline{y})$. The likelihood principle states that the same conclusion about θ should be drawn for any two sample points satisfying $T(\underline{x}) = T(\underline{y})$. Consider two experiments E_1 and E_2, where the unknown parameter θ is the same in both experiments. Suppose $\underline{x}_1$ and $\underline{x}_2$ are sample points from E_1 and E_2, respectively, such that $L(\theta; \underline{x}_1) = CL(\theta; \underline{x}_2)$ for all θ and for some constant C which may depend on $\underline{x}_1$ and $\underline{x}_2$ but not θ. Then evidence of $(E_1, \underline{x}_1)$ is equal to evidence of $(E_2, \underline{x}_2)$. This is the formal likelihood principle. The invariance principle describes inference technique in slightly different way. There are two different invariance considerations. The first type of invariance might be called measurement invariance. The second type of invariance is called formal invariance. It states that if two inference problems have the same mathematical structure then the same inference procedure should be used in both problems.

All three techniques restrict the set of allowable inferences and, in this way, simplify the analysis of the problem.

1. THE SUFFICIENCY PRINCIPLE

A sufficient statistics for a parameter θ is a statistics that, in a certain sense, implies all the information about θ contained in the sample [4]. Considering this definition, if $Y(\underline{X})$ statistics is sufficient for θ, any inference related to $\underline{X}$ sample about θ is made by $Y(\underline{X})$. In other words, if $\underline{x}$ and $\underline{z}$ are two sample points so that $Y(\underline{x}) = Y(\underline{z})$, then it makes no difference to observe $\underline{X} = \underline{x}$ or $\underline{X} = \underline{z}$ in any inference about θ. This thought is called the sufficiency principle. The formal definition of sufficient statistics is as follows:

Definition. A statistics $Y(\underline{X})$ is a sufficient statistics for θ if the conditional distribution of the sample $\underline{X}$ given the value of $Y(\underline{X})$ does not depend on θ [9].

Beginning from this definition, being $P_\theta(Y(\underline{X}) = y) > 0$, let's think of the conditional probability $P_\theta(\underline{X}=\underline{x} \mid Y(\underline{X})=y)$. If at the point $\underline{x}$ $Y(\underline{x})\neq y$ then $P_\theta(\underline{X} = \underline{x} \mid Y(\underline{X}) = y) = 0$. Considering the definition, if $Y(\underline{X})$ is a sufficient statistics, the conditional probability mentioned here is the same for ah values θ.

Suppose there are two researchers. First researcher is observing $\underline{X} = \underline{x}$ and calculating $Y(\underline{X}) = Y(\underline{x})$. This researcher will make inference about θ by using information of $\underline{X} = \underline{x}$ and $Y(\underline{X}) = Y(\underline{x})$. The second researcher, on the other hand, knows without observing the value of $\underline{X}$ that $Y(\underline{X}) = Y(\underline{x})$ and $P(\underline{X} = \underline{z} \mid Y(\underline{X}) = Y(\underline{x}))$ is the probability distribution defined on $A_{y(\underline{x})} = \{\underline{Z} \mid Y(z) = Y(\underline{x})\}$. Because this probability can be calculated by using the model without knowing the true value of θ. Therefore, the second researcher by using this distribution, can get the z value which maintains $P(\underline{Z} = \underline{z} \mid Y(\underline{X}) = Y(\underline{x})) = P(\underline{X} = \underline{Z} \mid Y(\underline{X}) = Y(\underline{x}))$ equality from a random number generator. Inversely thinking, for each value of θ, $\underline{X}$ and $\underline{Z}$ have the same probability distribution. As a result, the first researcher knows the value of $\underline{X}$ that he has observed, the second researcher knows the value of $\underline{Z}$ that he has derived, and both of them have the same degree amount of information for θ. To complete this claim we must show that for each $\underline{x}$ and θ values, $\underline{X}$ and $\underline{Z}$ have the same unconditional distribution as $P_\theta(\underline{X} = \underline{x}) = P_\theta(\underline{Z} = \underline{x})$.

The events $\{\underline{X} = \underline{x}\}$ and $\{\underline{Z} = \underline{z}\}$ are both subsets of the event $\{Y(\underline{X}) = Y(\underline{x})\}$. Their conditional probabilities do not depend on θ. Thus we have

$$P(\underline{X} = \underline{x} \mid Y(\underline{X}) = Y(\underline{x})) = P(Z = \underline{x} \mid Y(\underline{X}) = Y(\underline{x}))$$

and so,

$$P_\theta(\underline{X} = \underline{x}) = P_\theta(\underline{X} = \underline{x} \text{ and } Y(\underline{X}) = Y(\underline{x}))$$

and by conditional probability definition, it can be seen that

$$\begin{aligned} P_\theta(\underline{X} = \underline{x}) &= P(\underline{X} = \underline{x} \mid Y(\underline{X}) = Y(\underline{x}))\, P(Y(\underline{X}) = Y(\underline{x})) \\ &= P(\underline{Z} = \underline{x} \mid Y(\underline{X}) = Y(\underline{x}))\, P(Y(\underline{X}) = Y(\underline{x})) \\ &= P(\underline{Z} = \underline{x} \text{ and } Y(\underline{X}) = Y(\underline{x})) \\ &= P(\underline{Z} = \underline{x}). \end{aligned}$$

Beginning from the last definition, when proving that $Y(\underline{X})$ is a sufficient statistics for θ, it must also be proved that for any two constant sample points $\underline{x}$ and $\underline{z}$, the probability of $P_\theta(\underline{X} = \underline{x} \mid Y(\underline{X}) = \underline{z})$ is the same for all values of θ. Therefore, only it must be proved that the probability $P_\theta(\underline{X} = \underline{x} \mid Y(\underline{X}) = Y(\underline{x}))$ does not depend on θ. But, because the event $(\underline{X} = \underline{x})$ is a subset of $\{Y(\underline{X}) = Y(\underline{x})\}$,

$$P_\theta(\underline{X} = \underline{x} \mid Y(\underline{X}) = Y(\underline{x})) = \frac{P_\theta(\underline{X} = \underline{x} \text{ and } Y(\underline{X}) = Y(\underline{x}))}{P_\theta(Y(\underline{X}) = Y(\underline{x}))}$$

$$= \frac{P_\theta(\underline{X} = \underline{x})}{P_\theta(Y(\underline{X}) = Y(\underline{x}))}$$

$$= \frac{p(\underline{x} \mid \theta)}{q(Y(\underline{x}) \mid \theta)}$$

where $p(\underline{x} \mid \theta)$ is the joint probability density function of the sample $\underline{X}$ and $q(Y(\underline{x}) \mid \theta)$ is the joint probability density function of $Y(\underline{X})$.

Thus, $Y(\underline{X})$ is a sufficient statistics for θ if and only if for every $\underline{x}$ the above ratio of probability density functions is constant as a function of θ.

Factorization Theorem [10]. Let $f(\underline{x} \mid \theta)$ denote the joint pdf of a sample $\underline{X}$. A statistics $Y(\underline{X})$ is a sufficient statistics for θ if and only if there exist functions $g(y \mid \theta)$ and $h(\underline{x})$ such that, for all sample points $\underline{x}$ and all parameter points θ,

$$f(x \mid \theta) = g(y \mid \theta)\, h(\underline{x}).$$

This theorem is very useful in proving the existence of sufficient statistics.

Minimal sufficient statistics. In any problem more than one sufficient statistics can be found for the unknown parameter θ. Being as a sufficient statistics if $Y(\underline{X})$ is a function of any other sufficient statistics $Y'(\underline{X})$, then $Y(\underline{X})$ is called the minimal sufficient statistics [14].

Ancillary Statistics. The statistics $S(\underline{X})$ whose distribution does not depend on θ is called an ancillary statistics. This definition has been brought by Basu in 1959. Robert J. Buehler defined ancillary statistics for many distributions on its own article in 1982. An ancillary statistics contains no information about θ. But it gives important information for inferences about θ. For example, if $\hat{\theta}$ is the sufficient statistics and $S(\underline{X})$ is an ancillary statistics then the statistics $(\hat{\theta}, S(\underline{X}))$ becomes minimal sufficient statistics. Besides, Var $(\hat{\theta} \mid S(\underline{X}); \theta)$ only depends on $S(\underline{X})$, does not depend on θ [4].

A minimal sufficient statistics is a statistics that contains all the informations about θ and reduces maximum data. On the other hand, ancillary statistics is a statistics that does not contain information about θ. In this case it can be expected that they are independent of each other. But in independency concept completeness and Basu's theorem are very important.

Definition. Let $\{f(y; \theta), \theta \in \Omega\}$ be a family of pdf for a statistics $Y(\underline{X})$. The family of probability distributions is called complete if $E_\theta g(Y) = \theta$ for all θ implies $P_\theta(g(Y) = \theta) = 1$ for all θ. Equivalently, $Y(\underline{X})$ is called a complete statistics [15].

Basu's Theorem. If a statistics $Y(\underline{X})$ is minimal sufficient and completes statistics then it is independent of every ancillary statistics [4].

Inverse of this theorem is not true, but Lehmann (1981) gave a further definition about ancillary statistics and with respect to this definition inverse of Basu's theorem can be improved. Lehmann's definition is as follows:

Definition. A statistics $S(\underline{X})$ is called first-order ancillary if $E_\theta(S(\underline{X}))$ is independent of θ.

Lehmann then proves the following theorem, which is somewhat converse to Basu's Theorem:

Theorem. Let T be a statistics with Var $T < \infty$. A necessary and sufficient condition for T to be complete is that every bounded first-order ancillary S is uncorrelated (for all θ) with every bounded real-valued function of T.

Lehmann also notes that a type of converse is also obtainable if, instead of modifying the definition of ancillary, the definition of completeness is modified [14].

2. THE LIKELIHOOD PRINCIPLE

The likelihood principle in data reduction is explained starting from the hkelihood function. Let $f(\underline{x}; \theta)$ denote the joint pdf of sample $\underline{X}=(X_1, X_2, \ldots, X_n)$. Then given that $\underline{X} = \underline{x}$ is observed, the function of θ defined by

$$L(\theta; \underline{x}) = f(\underline{x}; \theta)$$

is called the likelihood function. If $\underline{X}$ is a discrete random vector then $L(\theta; \underline{x}) = P_\theta(\underline{X} = \underline{x})$. If we compare the likelihood function at two parameter points and find that

$$P_{\theta_1}(\underline{X} = \underline{x}) = L(\theta_1 ; \underline{x}) > L(\theta_2 ; \underline{x}) = P_{\theta_2}(\underline{X} = \underline{x})$$

then the sample we actually observed is more likely to have occurred if $\theta = \theta_1$ then if $\theta = \theta_2$ which can be interpreted as saying that θ_1 is a more plausible value for the true value of θ than is θ_2. It can be reasonable to examine the probability of the sample we actually observed under various possible values of θ. If X is a continuous, real-valued random variable and if the pdf of X is continuous in x then, for small ε, $P_\theta(x - \varepsilon < X < x + \varepsilon)$ is approximately $2\varepsilon f(x; \theta) = 2\varepsilon L(\theta; x)$ with respect to the definition of derivative. Thus,

$$\frac{P_{\theta_1}(x - \varepsilon < X < x + \varepsilon)}{P_{\theta_2}(x - \varepsilon < X < x + \varepsilon)} \approx \frac{L(\theta_1 ; x)}{L(\theta_2 ; x)}.$$

Definition. If $\underline{x}$ and $\underline{y}$ are two sample points such that $L(\theta; \underline{x})$ is proportional to $L(\theta; \underline{y})$, that is, there exists a constant $C(\underline{x}, \underline{y})$ such that

$$L(\theta; \underline{x}) = C(\underline{x}, \underline{y})\, L(\theta; \underline{y})$$

for all θ, then the conclusions drawn from $\underline{x}$ and $\underline{y}$ should be identical. This principle is called the likelihood principle [5].

Note that the constant $C(\underline{x}, \underline{y})$ may be different for different $(\underline{x}, \underline{y})$ pairs but $C(\underline{x}, \underline{y})$ does not depend on θ. In the special case if $C(\underline{x}, \underline{y})=1$ then the likelihood principle states that if two sample points result in the same likelihood function then they contain the same information about θ. If $L(\theta_2 ; \underline{x})=2L(\theta_1 ; \underline{x})$ then, in some sense, θ_2 is twice as plausible as θ_1 and it is also true that $L(\theta_2 ; \underline{y})=2L(\theta_1 ; \underline{y})$. Thus, whether we observe $\underline{x}$ or $\underline{y}$ we conclude that θ_2 is twice as plausible as θ_1. Furthermore, although $f(\underline{x}; \theta)$, as a function of $\underline{x}$, is a pdf, there is no guarantee that $L(\theta; \underline{x})$, as a function of θ, is a pdf. On this point one form of inference, called fiducial inference, explicitly interprets likelihoods as probabilities for θ [16]. That is, $L(\theta ; \underline{x})$ is multiplied by $M(\underline{x}) = \left(\int_{-\infty}^{\infty} L(\theta ; \underline{x})\, d\theta \right)^{-1}$ and then $M(\underline{x})\, L(\theta; \underline{x})$ is interpreted as a pdf for θ. Most statisticians do not subscribe to the fiducial theory of inference but it has a long history, dating back at least to the work of Fisher in the 1930 [16].

For example let $X_1, X_2, \ldots, X_n$ be a random sample from $n(\mu, \sigma^2)$, σ^2 known. The fiducial distribution of unknown parameter μ is $n(\bar{X}, \sigma^2/n)$ and calculations are done with assuming that the distribution of $(\mu - \bar{X}) / (\sigma / \sqrt{n})$ is $n(0, 1)$. In this distribution our random variable is μ.

Let's suppose a three-dimensional experiment or observation model by considering likelihood principle and sufficiency principle alltogether : Being E as experiment, $\underline{X}$ as sample, θ as parameter to be estimated and $f(\underline{x};\theta)$ as density of $\underline{x}$ observations defined on a subset of parameter space Ω, let $E=(\underline{X}, \theta, \{f(\underline{x}; \theta)\})$. Starting from $X=\underline{x}$ observation, the result obtained for θ is $Ev(E, \underline{x})$. In this connection, the definition of sufficiency principle, for two different $\underline{x}$ and $\underline{y}$ observations, becomes $Ev(E, \underline{x}) = Ev(E, \underline{y})$. From this point the formal likelihood principle is as follows [2], [16], [5] :

Definition. Let $E_1 = (\underline{X}_1, \theta, \{f_1(\underline{x}_1, \theta)\})$ and $E_2 = (\underline{X}_2, \theta, \{f_2(\underline{x}_2, \theta)\})$ be two experiments. C being constant for all θ values and if $\underline{x}_1^*$ and $\underline{x}_2^*$ are sample points in E_1 and E_2 respectively so that

$$L(\theta, \underline{x}_2^*) = CL(\theta; \underline{x}_1^*)$$

then

$$Ev(E_1, \underline{x}_1^*) = Ev(E_2, \underline{x}_2^*).$$

The reason for the difference between formal likelihood and likelihood principles defined at the beginning is that the formal likelihood principle is for two

different experiments. If $E_1 = E_2$, then likelihood principle and formal likelihood principle is the same. In this case E_2 is called shadow (dummy) experiment. For supplementary reasons this can be mentioned: If $E = (\underline{X}; \theta; \{f(\underline{x}; \theta)\})$ is an experiment, the result $Ev(E, \underline{x})$ which is obtained from this experiment depends on E and $\underline{x}$ only by $L(\theta\ ;\ \underline{x})$.

Many statistical studies violate the formal likelihood principle. With these studies, different concluisons would be reached for the same parameter in two different experiments.

3. THE INVARIANCE PRINCIPLE

In data reduction, two different invariance can be defined. First type of invariance is measurement invariance. According to this, let a researcher measure the same objects by using English measuring units and the other researcher measure them by using metric measuring units. If the measuring tools, which they use, have the same sensitivity, the results will be the same when the measuring units are transformed to each other. More remarkable, when using a measuring tool whose one side has mm. scale and the other side inch scale, the result doesn't change when either of side are used.

The second type of invariance is called formal invariance. If in two inference problems, the same formal structure is used on the basis of mathematical model then the results of both problems are tied to the same process. Formal invariance is about mathematical inputs and rejects the physical definitions of the experiment. Let the parameter space $\Omega = \{\theta; \theta > 0\}$ be used in two different problems. Suppose, one of the problems is about the weights of people in Turkey and the other is about the heights of giraffes in Africa. The same real numbers set has been defined for θ in both of the problems. Casella, G . and Berger, R.L. (1990) has defined the invariance principle as follows :

Definition. If $\underline{Y} = g(\underline{X})$ is a change of measurement scale such that the model for $\underline{Y}$ has the same formal structure as the model for $\underline{X}$, then an inference procedure should be both measurement invariant and formally invariant.

Under the invariance principle, a set of transformation functions, which is defined on the sample space, must be a group (see Lehmann (1990), pp. 19-26).

Definition. A set of functions $\{g(\underline{x}) : g \in G\}$ from the sample space S onto S is called a group of transformations of S if

(i) For every $g \in G$ there is a $g' \in G$ such that $g'(g(\underline{x})) = \underline{x}$ for all $\underline{x} \in S$,

(ii) For every $g \in G$ and $g' \in G$ there exists a $g'' \in G$ such that $g'(g(\underline{x})) = g''(\underline{x})$ for all $\underline{x} \in S$,

(iii) The identity, $e(\underline{x})$, defined by $e(\underline{x}) = \underline{x}$ is an element of G,

Definition. Let $F=\{f(\underline{x};\theta):\theta\in\Omega)$ be a set of pdf for $\underline{X}$ and let G be a group of transformations of the sample space **S**. Then F is invariant under the group G if for every $\theta\in\Omega$ and $g\in G$ there exists a unique $\theta'\in\Omega$ such that $\underline{Y}=g(\underline{X})$ has the distribution $f(\underline{y};\theta')$ if $\underline{X}$ has the distribution $f(\underline{x};\theta)$.

If we explain the last definition by a simple example; in a Bernoulli trial with the size n, because of $\underline{X}\sim$ binomial (n, p), $g_1(\underline{X})=n-\underline{X}\sim$ binomial $(n, 1-p)$, $g_2(\underline{X})=\underline{X}\sim$ binomial (n, p), in the inference made for $p, p'=n-\underline{x}$ can also be used and the result does not change. Because under the group $G=(g_1, g_2)$, the set of binomial distributions has the characteristics of invariance.

4. CONCLUSION

The three principles presented in this work are the principles not ceasable on the basic subject matters of statistical inference, namely, point estimation, interval estimation and hypothesis testing. Hogg and Craig (1970) built in their book the subject matter of point estimation on the principle of sufficiency. Lehmann (1990) has studied the matter as a whole by considering the three principles in point estimation.

Most analysts perform some sort of "model checking" when analyzing a set of data. Most model checking is, necessarily, based on statistics other than a sufficient statistics. It is common practice to examine residuals from a model, statistics that measure variation in the data not accounted for by the model. Such a practice immediately violates the sufficiency principle, since the residuals are not based on sufficient statistics. Of course, such a practice directly violates the likelihood principle also. Thus, it must be realized that before considering the sufficiency principle or the likelihood principle, we must be comfortable with the model.

All three principles prescribe similar relationships between inferences at different sample points. Thus, all three data reduction techniques restrict the set of allowable inferences and, in this way, simplify the analysis of the problem.

REFERENCES

[1] AGRESTI, A. : Categorical Data Analysis, Wiley, New York, 1990.

[2] BROWN, L.D. and FARRELL, R.H. : *Complete Class Theorems For Estimation of Multivariate Poisson Means and Related Problems*, The Annals of Statistics, **13** (1985), No. 2, 706-726.

[3] BRUNK, H.D. : Mathematical Statistics, Xerox, Massachusetts, 1975.

[4] BUEHLER, R.J. : *Same Ancillary Statistics and Their Properties*, Journal of The American Statistical Association, 77 (1982), No. 379, 581-594.

[5] CASELLA, G. and BERGER, R.L. : Statistical Inference, Wadsworth, California, 1990.

[6] DAVIS, L.J. : *Consistency and Asymptotic Normality of The Minimum Logit Chi-Squared Estimator When The Number of Design Points is Large*, The Annals of Statistics, 13 (1985), No. 3, 947-957.

[7] DeGROOT, M.H. : Probability and Statistics, Addison-Wesley, California, 1975.

[8] FRASER, D.A.S. : Probability and Statistics: Theory and Application, Duxbury Press, Massachusetts, 1976.

[9] FREUND, J.E. and WALPOLE, R.E. : Mathematical Statistics, Prentice-Hall, New Jersey, 1987.

[10] GERTSBAKH, I. and WINTERBOTTOM, A. : *Point and Interval Estimation of Normal Tail Probabilities*, Commun. Statist.-Theory and Meth., 20 (4) (1991), 1497-1514.

[11] HOGG, R.V. and CRAIG, A.T. : Introduction to Mathematical Statistics, Macmillan Publishing Co., New York, 1970, p. 219-220.

[12] KALE, B.K. : *Essential Uniqueness of Optimal Estimating Functions*, Journal *of Statistical Planning and Inference*, 17 (1987), 405-407.

[13] LEHMANN, E.L. : Testing Statistical Hypotheses, Wadsworth, California, 1991, p. 141, 172-173.

[14] LEHMANN, E.L. : Theory of Point Estimation, Wadsworth, California, 1990, p. 36-47.

[15] LEHMANN, E.L. : *An Interpretation of Completeness and Basu's Theorem*, Journal *of The American Statistical Association*, 76 (1981), p. 335-340.

[16] SAVAGE, L.J. : *On Rereading P.A. Fisher*, The Annals of Statistics, 4 (1976), 441-500.

ASSISTANT PROFESSOR OF NUMERICAL METHODS
MARMARA UNIVERSITY
ISTANBUL-TURKEY

ÖZET

Yeterlilik, olabilirlik ve değişmezlik, istatistiksel vardamanın üç temel ilkesidir. Bir araştırmacı, örneklemdeki bilgileri, bilinmeyen parametre θ hakkında vardamada bulunmak için kullanır. Herhangi bir istatistik $T(\underline{X})$, verilerin özet formunu tanımlar. θ parametresi için bir yeterli istatistik, örneklemdeki θ ile ilgili tüm bilgileri kapsayan bir istatistiktir. Bu düşünce, yeterlilik ilkesine ışık tutar. $E=\{X_1, X_2, \ldots, X_n, \theta, \{f(\underline{x}; \theta)\}\}$ deneyi ve θ için bir yeterli istatistik $T(\underline{X})$ düşünülsün. Eğer $\underline{x}$ ve $\underline{y}$ örneklem noktaları $T(\underline{x})=T(\underline{y})$ şartını sağlıyorsa $(E, \underline{x})$ kanıtı ile $(E, \underline{y})$ kanıtı birbirine eşittir. Olabilirlik ilkesi de $T(\underline{x}) = T(\underline{y})$ şartını sağlayan herhangi iki örneklem noktası $\underline{x}$ ve $\underline{y}$ için θ hakkında aynı sonucu söyler. E_1 ve E_2, aynı θ parametresi için iki deney olsun. Bütün θ lar için ve θ ya bağlı olmayan, fakat E_1 den gelen $\underline{x}_1$ ve

E_2 den gelen $\underline{x}_2$'ye bağlı olabilen C sabit fonksiyon olmak üzere $L(\theta; \underline{x}_1) = C L(\theta; \underline{x}_2)$ oluyorsa $(E_1, \underline{x}_1)$ kanıtı ile $(E_2, \underline{x}_2)$ kanıtı birbirine eşittir. Bu, formel olabilirlik ilkesidir. Değişmezlik ilkesi oldukça farklı vardama tekniğini açıklar. İki farklı değişmezlik ilkesi düşünülmektedir. Birincisine ölçme değişmezliği, ikincisine formel değişmezlik denir. Buna göre, eğer iki vardama problemi aynı matematiksel yapıya sahip ise iki problemde de aynı vardama yöntemi kullanılmıştır.

Bütün bu üç teknik, kabul edilebilir sonuçları kısıtlayarak, bu yolla, problemin analizini basitleştirirler.