An Adaptive Design For Discrete Responses Of Patients In Clinical Trials

V Lakshmi Naryana[1],
S.P. Kishore[2], K. Santosh Reddy[3]
Vardhaman College of Engineering, Hyderabad, India

ABSTRACT

Adaptive design is a trial design that allows modifications to one or more specified aspects of the study design and hypothesis based on the analysis of data after its initiation without undermining the validity and integrity of the trial.. It has become very attractive to the pharmaceutical industries. In this paper, discrete adaptive designs for clinical trials survival responses are studied using computer simulations. The greatest interest in adaptive design clinical trials has been in the adequate and well controlled study setting intended to support marketing a drug. The discrete adaptive design (DAD) described here decides the immediate future allocation of treatment to patients based on the analysis of data .At each phase of analysis the decision to terminate or continue the trial is based on the expected loss function .

Keywords: Adaptive design, robustness, decision making, allocation pattern, loss function.

1. INTRODUCTION

The objective of clinical trials is saving or curing the maximum number of patients and of several alternative therapies, treatment plans or drug formulations the physician has to decide on adopting one that would yield the maximum cure rate or minimum mortality. The statistician has a vital role to play in these sorts of testing plans.

IJERT

Take for instance a clinical trial where in subjects (human patients) arrive sequentially desiring to receive, adaptively, the better of two competing treatments A and B in which, A is a drug that is already administered widely, and B is the new introduction. Those subjected to such a trial are human beings and are proposed in this paper as a DAD that would allocate less efficient treatment therapy to fewer numbers of patients. This scheme of allocation does not just dichotomize the response of the patients, but used for the next allocation, all the information gathered from previous discrete responses.

Various data-dependent adaptive allocation rules are available for a sequential chain of patients undergoing clinical trials. There is the concept of play-the- winner (PW) rule introduced by Zelen (1969). Subsequently Wei and Durham (1978) and Wei (1979) proposed the randomized PW rule. Durham and Yu (1990) designed a two arm urn model that ignores failure and in which the urn composition changes only as and when a success occurs.

The contributions of Ware (1989) and Rout et al. (1993), Anderson et al. (1994) Faries et al. (1995), Bandyopadhyay and Biswas (1996, and 1997), and Durham et al.(1998) have made available other adaptive designs. However over the past two decades, the contributed number of real life response adaptive designs, is not very large. Statisticians can actively influence the design and analysis of clinical trials since recent advancement in computer technology could make clinical trials more adaptive in all respects. A new type of urn design called a birth and death urn of Ivan ova et al. (2000) has enabled them to derive the maximum likelihood estimators of success probabilities and their limiting distribution. All designs listed above are urn designs where responses of the patients are assumed to be dichotomous (success or failure).

2. DISCRETE ADAPTIVE DESIGN (DAD)

The DAD, now being described, decides the immediate future allocation of treatment to patients based on the quantum of updated information. The population of responses under treatment A is represented by X and that of

B by Y. The real valued random variables X and Y are assumed to be of the discrete type. Let X_i or Y_i denote the response of treatment A or B by the i^{th} patient depending on the one that has been allocated to him/her. Consider δ_i to be the indicator of assignment that takes the value 1 (or 0) if the i^{th} patient is allocated to treatment A (B).

The first patient is allocated to treatment A and the second to treatment B and their responses are measured as $\left(\delta_1=1, X_1, Y_1=0\right)$, and $\left(\delta_2=1, X_2, Y_2=0\right)$ respectively. Starting from the third up to the n^{th} allocation, incoming patients are allocated to either treatment A or B according to an adaptive allocation scheme (AAS), which is detailed below. When 'n' is the maximum number of patients participating in the trial, this AAS used the statistics of observed sample based on the discrete responses of triplets $\left\{\left(\delta_i, X_i, Y_i\right); i=1,2,\ldots..k\right\}$ of the previous 'k < n' patients to decide the value of δ_{k+1} without dichotomizing them.

The stage at which the number of patients N_{AK} and N_{Bk} treated by treatment A and B respectively just after the entry of k^{th} patient is

$$N_{AK} = \sum_{i=1}^{k} \delta_i \text{ and } N_{BK} = \sum_{i=1}^{k} (1-\delta_i) \qquad (1)$$

The observed sample means for the two populations X and Y can respectively be denoted by

$$\hat{\mu}_{AK} = \frac{\sum_{i=1}^{K} \delta_i X_i}{N_{AK}} \qquad \hat{\mu}_{BK} = \frac{\sum_{i=1}^{K} (1-\delta_i) Y_i}{N_{BK}} \qquad (2)$$

The steps leading to the decision rule for allocating $\left(k+1\right)^{th}$ patient to either treatment A or B as per AAS or equivalently for the determination of $\left(k+1\right)^{th}$ value are as follows:

Step 1: A symmetric and continuous cumulative distribution function G (.) is chosen i.e G(0)=0.5, G (-x) =1-G(x).

Step2: The observed sample means $\hat{\mu}_{AK}$ and $\hat{\mu}_{BK}$ are determined as in (2) based on the data $\left\{(\delta_i, X_i, Y_i), i=1,2,.....k\right\}$ of the previous 'k < n' patients.

Step 3: The 'k < n' patient is allocated to treatment A with probability G (.) and to treatment B With probability 1-G (.), where c is some appropriate scaling constant.

This kind of an allocation will always be biased towards the treatment that has led on average to larger number of favorable responses on average in the past, and that is what the AAs aims to achieve. (See Biswas and Basu (2001)).

3. THE PROBLEM OF DECISION MAKING

The decision-making problem addressed in (3) below declares that the sample evidence supports one of the following three hypothesis (H):

$$H_0 : \mu_A = \mu_B , H_1 : \mu_A > \mu_B , H_2 : \mu_A < \mu_B \qquad (3)$$

If the observed difference $d = \hat{\mu}_{An} - \hat{\mu}_{Bn}$ of sample means obtained on the basis of 'n' sample observations $\left\{\left(\delta_i, X_i, Y_i\right); i=1,2,\ldots..n\right\}$ and sample means $\hat{\mu}_{An}$ and $\hat{\mu}_{Bn}$, is greater than a pre-specified cut-off point 'v' or less than '-v' then we accept H_1 or H_2 otherwise we tend to accept H0

IJERT

There is the opportunity for identifying a few outlying observations. Each outlier can exert undue influence on the values of the sample means $\hat{\mu}_{An}$ and $\hat{\mu}_{Bn}$. Special attention has been paid in recent years to the statistical property called robustness which is an indicative of the extent to the estimation procedures that are adversely affected by violations of underlying assumptions.

Making use of appropriate robustness considerations as Biswas and Basu (2001) have in the estimates $\hat{\mu}_{An}$ and $\hat{\mu}_{Bn}$, the influence of outlier may be minimized. Accommodation methods such as trimming data and Winsorization are other commonly used procedures for dealing with outliers. The computer simulations alone can solve most of the questions of robustness that are difficult to answer.

3.1 NUMERICAL ILLUSTRATIONS

In most of the applications of adaptive designs available in literature, the responses are either assumed to be dichotomous (success or failure) or intentionally dichotomized (if continuous only) setting appropriate threshold value. But the design being considered, measured the responses of patients in discrete units after the allocation of treatment to patients.

For our numerical study, responses X and Y are assumed to follow binomial distributions $B(x;m,p_0)$ and $B(y;m,p_1)$ respectively so that mean μ_A = mean of $X = mp_0$, μ_B = mean of $Y = mp_1$, σ^2_A = variance of $X = mp_0(1-p_0)$, and σ^2_B = variance of $Y = mp_1(1-p_1)$,.

Normal distribution can provide a close approximation of the binomial $B(x;m,p)$ distribution when the number of trials m, is very large and p, the probability of success on an individual trial is close to 0.5. However, normal distribution is often applied to approximate binomial probabilities even when m is fairly small. A good rule of thumb is to use this approximation only when mp and $m(1-p)$ values are greater than 5. Algorithm of drawing a sample from binomial distribution $B(x;m,p)$ is outlined below subject to the condition $mp > 5$ and $m(1-p) > 5$:

Step 1: A point 'u' at random is drawn from the uniform population $U[0,1]$.

Step2: The value z_1 of $Z = \dfrac{X - mp}{\sqrt{mp(1-p)}} \sim N(0,1)$ is computed such that $P[Z \le z_1] = u$

Where $N(0,1)$ stands for the standard normal distribution.

Step3: Computed is X=number of successes = $z_1(\sqrt{mp(1-p)}) + mp$

Step 4: Steps 1 to 3 are repeated until the sample desired is selected from the binomial population $X \sim B(x;m,p)$

Detailed below is data on (δ_i, X_i, Y_i) obtained for a fixed number 'n' of patients $(i.e.\ i = 1,2,\ldots n)$. The first patient is allocated to treatment A and second to treatment B and random observation X_1 from $B(x;m,p_0)$ and observation Y_2 from $B(y;m,p_1)$ populations are drawn so that $(\delta_1 = 1, X_1, Y_1 = 0)$, and $(\delta_2 = 0, X_2 = 0, Y_1)$ are completely known. For each of successive allocations, the current estimate $\hat{\mu}_{Ak}$ and $\hat{\mu}_{Bk}$ can be computed and the allocation determined according to the algorithm of AAS. And so the next allocation will either be to treatment

A $\left(\text{i.e. } \delta_{k+1} = 1\right)$ with probability $G\left(\frac{\mu_{Ak} - \mu_{Bk}}{c}\right)$ or treatment $B\left(\text{i.e. } \delta_{k+1} = 0\right)$ with probability $1 - G\left(\frac{\mu_{Ak} - \mu_{Bk}}{c}\right)$.

For the present numerical illustration we use the symmetric distributions $G_1(.)$ and $G_2(.)$. $G_1(.)$ has been defined as symmetric over the real valued –r to r as follows:

$$G_1(x) = \int_{-r}^{x} g(x)dx = \frac{x + r}{2r} \qquad \ldots (4)$$

So that $g(x) = 1/(2r))$ for $-r \le x \le r$ where r is a positive and finite real number. Quantities '0' and '$r^2/3$' and will respectively be the mean and variance values of $G_1(x)$.

The performance of $G_1(.)$ as allocation probability in the DAD is then compared with another cumulative distribution function (CDF) say $G_2(.)$ that has been selected as the CDF $\Phi(.)$ of the standard normal distribution $N(0,1)$.

The estimates μ_{Ak} and μ_{Bk} are simple means of 'k' observations and for each method, the final estimates μ_{An} and μ_{Bn} and the number of allocations T_A and T_B to the treatment groups A and B respectively ($T_A+T_B=n$) have been calculated.

The functions listed below will help estimate loss consequent on problem (3) of the final decision-making.

(i) The loss is zero for choosing H_i when H_i is true for i=0,1 and 2.

(ii) The loss is 1 for choosing H_0 when H_1 or H_2 is true.

(iii) The loss L is greater than zero for choosing H_1 when H_2 is true or vice versa.

The reader may refer to Ferguson (1967) for details on a loss function.

IJERT

The above binomial $B(m, p)$ trials have been carried out with n=20 over a large number of samples say N=200. The following sample estimates have been made from 200replications of the sampling distribution: (i)$E(T_A)$=Expected value of T_A, (ii)$V(T_A)$=Variance of T_A ,(iii) Prob1=estimate of the probability of H_1 of (3) and Prob2= estimate of the probability of H_2 of (3) and (iv)$E(p_0)$and $E(p_1)$.

Numerical calculations carried out for different choices of input parameters such as m, p_0, p_1,r, v and c of the above DAD, and the corresponding results presented in tables 1 through 4 will help gain a better understanding of the performance of the design.

TABLE 1: Simulation results for Binomial model. Data (responses) from populations B(m=17, p_0=0.45)i.e. A and B (m=17, p_1=0.45) i.e. B are generated for n=20 trails which is replicated N=200times with cut off v=2.

CDF	c	$E(T_A)$	$V(T_A)$	Prob1	Prob2	$E(\hat{p}_0)$	$E(\hat{p}_1)$
$G_1(.)$=U(-r,r)	5	10.070	7.3651	0.010	0.985	0.4480	0.4493
r=1.7321	10	10.135	5.4768	0.005	0.995	0.4496	0.4510
	20	10.110	5.2579	0.000	1.000	0.4504	0.4528
$G_1(.)=\Phi(.)$	5	9.9750	9.8642	0.015	0.975	0.4453	0.4465
	10	10.125	6.1494	0.005	0.995	0.4494	0.4500
	20	10.110	5.4418	0.000	1.000	0.4502	0.4517

TABLE 2: : Simulation results for Binomial model. Data (responses) from populations B(m=17, p_0=0.41)i.e. A and B (m=17, p_1=0.52) i.e. B are generated for n=20 trails which is replicated N=200 times with cut off v=2

CDF	c	$E(T_A)$	$V(T_A)$	Prob1	Prob2	$E(\hat{p}_0)$	$E(\hat{p}_1)$
$G_1(.)=U(-r,r)$ r=1.7321	5	8.1050	7.2140	0.260	0.740	0.4062	0.5198
	10	9.1550	5.0510	0.265	0.735	0.4106	0.5222
	20	9.6350	4.8918	0.260	0.740	0.4116	0.5231
$G_2(.)=\Phi(.)$	5	7.4600	8.4584	0.280	0.72	0.4018	0.5198
	10	8.8450	5.3710	0.245	0.775	0.4096	0.5215
	20	9.4200	4.8236	0.255	0.745	0.4120	0.5228

It could be noted that when r =$\sqrt{3}$=1.7321, variance of $G_1(.)$ equals one so that both $G_1(.)$ and $G_2(.)$ have the same mean (i.e. zero) and variance(i.e. unity). A scrutiny of the results presented in the above two tables, will show that both $G_1(.)$ and $G_2(.)$ produce close sample values but $V(T_A)$ of $G_1(.)$ is smaller. Further, if the probability of success of finding a positive symptom 'p' for the two treatments are same as in Table 1 or differs moderately by less than 0.1 unit between p_0 and p_1 as in Table 2, any one real value between c= 5and c=20, can be preferred as scale value in the allocation process.

Secondly, Prob2 (which should fall very close to 1 if $p_0 = p_1$ values and close to zero if $p_0 \neq p_1$) value under $G_1(.)$ estimation is greater than or equal to the corresponding Prob2 value under $G_2(.)$. This pattern of calculation has been repeated in two more combinations with vast differences between the proportions of p_0 and p_1 and the corresponding results have been tabulated and presented in Table3 below for further verification.

TABLE 3: Simulation results for Binomial model. Data (responses) from populations B(m=17, p_0=0.42)i.e. A and B (m=17, p_1=0.67) i.e. B are generated for n=20 trails which is replicated N=200 times with cut off v=2.

IJERT

CDF	c	$E(T_A)$	$V(T_A)$	Prob1	Prob2	$E(\hat{p}_0)$	$E(\hat{p}_1)$
$G_1(.)=U(-r,r)$ r=1.7321	5	5.450	6.2775	0.955	0.045	0.4007	0.6698
	10	7.985	5.3548	0.965	0.035	0.4203	0.6706
	20	9.09	4.7719	0.955	0.045	0.4213	0.6713
$G_2(.)=\Phi(.)$	5	4.500	4.6600	0.960	0.040	0.3975	0.6709
	10	7.150	5.4975	0.945	0.055	0.4186	0.6703
	20	8.740	4.8324	0.955	0.045	0.4210	0.6717

TABLE4: Simulation results for Binomial model. Data (responses) from populations B(m=25, p_0=0.42)i.e. A and B (m=25, p_1=0.67) i.e. B are generated for n=20 trails which is replicated N=200 times with cut off v=2.

CDF	c	$E(T_A)$	$V(T_A)$	Prob1	Prob2	$E(\hat{p}_0)$	$E(\hat{p}_1)$
$G_1(.)=U(-r,r)$ r=1.7321	5	3.175	4.3144	1.000	0.000	0.3817	0.6706
	10	6.775	5.4664	1.000	0.000	0.4179	0.6702
	20	8.600	4.8100	1.000	0.000	0.4203	0.6712
$G_2(.)=\Phi(.)$	5	2.990	3.1299	0.995	0.005	0.3917	0.6713
	10	5.790	4.8159	1.000	0.000	0.4129	0.6698
	20	7.985	5.0548	1.000	0.000	0.4197	0.6700

It could again be noticed from a study of the results presented in Table3 and Table 4 that each expected value E(.) from the simulated samples are found to be so close under $G_1(.)$ and $G_2(.)$ corresponding to each scaling factor c= 5,10, and 20. This ensures that allocation between treatments A and B according to both probabilities $G_1(.)$ and $G_2(.)$ as specified by DAD, not only produces similar results but also favors the treatment that performs better with reference to the response of the patients. Secondly, the estimate of Prob1 (which should be very close to 1 if $p_0 \neq p_1$) value under $G_1(.)$ is greater than or equal to that of the corresponding Prob1 value under $G_2(.)$.

4. RESULTS AND DISCUSSION

It could be seen that (i) the scaling factor should increase with an increase in each value of 'm' and (ii) any one of the allocation processes from the family $\{G_1(.), G_2(.)\}$ may be selected for allocating patients between the competing treatments A and B when the number of successes out of 'm' symptoms from patient follow the binomial $B(m, p)$ law. For each treatment, 'p' the proportion of success is such that $mp > 5$ and $m(1-p) > 5$. The value of 'r' can be selected from the neighborhoods of 1.7321, in the allocation process under G1 (.) =U(-r, r), depending upon the input values of DAD.The proposed DAD will perform as expected if there is no delayed response. But the allocation type of next incoming patient to the type of treatment in adaptive trails is decided by the summary of responses up to the current trials. Better estimates may be obtained on the basis of available responses before the next entry of a patient if the number of delayed responses is very few.

The allocation pattern of patients is governed by the scaling factor 'c' and the value of 'r' of the distribution $G_1(.) = U(-r, r)$ and they are allocated to either treatment, according to the state at any stage of the data on responses. Small values of 'c and r', for example, will have sensitive impacts on outliers. On the other hand, large values tend to influence information and say the allocation ratio of patients to treatments A and B eventually towards a 50:50 pattern. The choice of c and r and the standard deviation of the estimated difference in mean responses of samples selected from A and B populations are factors that influence the decision of allocation of patients to the better of the two lines of treatments and hence they should always be viewed in conjunction.

IJERT

The allocation probabilities will be different for different choices of G (.), but this difference can be minimized by choosing c depending on G(.). in general the selection of c is a more difficult task than that of the choice of G(.).

CONCLUSIONS

In most of the applications of adaptive designs, the responses are either assumed to be dichotomous or intentionally dichotomized setting appropriate threshold value, but the design being considered here measured the responses of patients in discrete units after allocation of treatments to patients. Adaptive design can increase the number of responses in a trial and provide more benefits to the patient in comparison to the classic design.

REFERENCES

1. Andrews, D. F. and Herzberg, A. M. (1985). Data: A Collection of Problems from Many Fields for the Student and Research Worker, New York: Springer-Verlag.
2. Barlow, W. E. and Prentice, R. L. (1988). Residuals for relative risk regression. Biometrika IS, 65-74.
3. Crowley, J and Hu, M. (1977). Covariance anlaysis of heart transplant survival data. Journal of the American Statistical Association 72, 27-36.
4. Cox, D.R. (1972). Regression models and life- tables (with discussion). Journal of the Royal Statistival Society, Scries B 34, 187-220.

5. Cox, D.R. and Oakcs, D. (1984). Analysis of Survival Data. London: Chapman and Hall. David, H. A. and Mocschbcrgcr, M. L. (1978). The Theory of Competing Risks. London: Griffin. Ilancll, F. and Lee, K. (1986). Verifying assumptions of the proportional hazards model. Proceeding of the Eleventh Annual SAS User's Group International, pp.823-828.
6. Kalbflcisch, J.D. and Prentice, R.L. (1980). The Statistical Analysis of Failure Time Data. New York: John Wiley 7 Sons.
7. Kav, R. (1986). Treatment effects in competing- risks analysis of prpstate cancer data, Biometrics 42, 203-211.
8. Breslow, N. (1972). Contribution to the discussion of paper by D.R.Cox journal of the Royal Statistical Society, Series b 34, 216-217.
9. Fleming, T.R. and Harrington, D. P. (1991). Counting Processes and Survival Analysis New York: Wiley.
10. Friedman, J.H. (1991). Multivariate adaptive refression splines (with dicussion) Annals of Statistics 19, 1-141.
11. Gentleman, R. and Crowley J. (1991). Local full likelihood estimation for the prpportional hazards model. Bimetrics 47, 1283-1296.
12. Grambsch, P.M. and Therneau, T.M (1994). Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81, 515-526.
13. Grambsch, P.M. and Therneau, T.M., and Fleming, T.R. (1995). Diagnostic plots to reveal functional form for covariates in multiplicative intensity models. Biometrics 51, 1469-1482.
14. Gray, R.J. (1992). Flexible mothods for analyzing survival data using splines with applications to breast cancer prpgnosis. Journal of the American Statistical Associatin 87, 942-951.
15. Hastie, T. and Tibshirani, R. (1990). Generalized Additive Models. London: Chapman and Hall.
16. Kooperberg, C., Stone, C.j., and Truong, Y.K. (1995). Hazards and regfression Journal of the American Statistical Association 90, 78-94.
17. O' Sullival, F. (1988). Non parametric estimation of relative risk using splines and cross- validation. SIAM, Journal on Scientific and Statistical Computing 9, 531-542.
18. Saeed, S.M., Stock-Novack, D., Pohold, R., Crowley, J. and Salmon, S.E.(1993). Prognostic correlation of plasma cell acid phosphatase and beta-glucoronidase in multiple myeloma. Blood 81, 869-870.
19. Therneau, T., Grambasch, P., and Fleming, T. (1990). Martingale based residuals for survival models. Biometrics 77, 147-160.
20. Kuk, A.Y.C. (1992). A semiparametric mixture model for the analysis of competing risks data. Australian Journal of Statistics 34, 169-180.
21. Larson, M.G. and Dinse, G.E. (1985). A mixture model for the regression analysis of competing risks data. Applied Statistics 34,201-211.
22. Lin, D.Y. and Wei, L.J. (1989). The robust inference for the cox proportional hazards model. Journal of the American Statistical Association 84, 1074-1079.
23. Wei, L.J.,Lin, D.Y. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distrihbutions. Journal of the American Statistical Association 84, 1065-1073.
24. Covariance analysis in generalized linear measurement error models. Statistics in Medicine 8, 1075-1093.
25. Carroll, R.J., Ruppert, D., and Stefanski, L.A.(1995). Measurement Error in Nonlinear Models. London: Chapman and Hall.

IJERT