Turk J Math
29 (2005) , 193 – 200.

An Exponential Model for Treatment Effects Without a Control Group

Kemal Gürsoy

Abstract

It may be desirable to estimate the behaviour of a pair of random variables and their functions through the information acquired by utilizing only one of them and its functions. In this work, such an approach has been used. Motivated by the need to provide treatment to every patient in a new drug trial, an exponential model was considered. This approach provides sufficient information to make inferences about the effect of a treatment without using a control group who will be otherwise denied treatment, as an alternative method to the commonly used controlled clinical trials.

Key Words: Controlled clinical trials, exponential probability distribution.

1. Introduction

It is one of the major human activities to acquire information from the observation of experiments. In one approach, Fisher [3] proposed a method in which random samples are selected from a population under study, until sufficient statistics are obtained and characteristics of the population are estimated by utilizing the sample information.

In the medical field, design of experiments are widely based on Fisher's method, known as controlled clinical trials model. It refers to comparing the behaviour of two groups, both randomly selected from the same ill population: one, a control group which does not receive treatment but instead a placebo (usually some harmless nontreatment); the other, which receives the treatment being evaluated. Then the effect of the treatment is inferred

2000 *AMS Mathematics Subject Classification:* 62C12, 62P10

by comparing the average behaviour of these two groups. A brief exposure to diverse utilization areas of controlled clinical trials may be acquired from references [1], [2], [5], [10] and [11]. There has been an alternative method to acquire the desired information about the effect of treatments, without utilizing a control group (where treatment is provided to all individuals participating in the trial), introduced by Robbins [8], in the form of statistical learning applied to the field of medical treatments and their evaluation. This work follows in that direction.

2. Formulations

Let X and Y be a pair of nonnegative random variables which are exponentially distributed and parametrized by an unknown θ. Also assume that there are two functions of X, u and v, both are σ-finite (a set is said to be σ-finite if it is a countable union of sets of finite measure. For example, real numbers with Lebesgue measure is σ-finite but not finite), such that the following equation holds,

$$\int v(\mathrm{X})dF(\mathrm{X}|\theta) = \theta \int u(\mathrm{X})dF(\mathrm{X}|\theta), \text{ for all } 0 < \theta < \infty \tag{1}$$

Assume that there is a finite number (say, n) of observations about the pair of (X_i, Y_i), with their corresponding distinct parameter θ_i, for $i = 1, 2, \ldots, n$. Here, θ_i is a finite, positive, identical and independently distributed (i.i.d.) random variable coming from an unknown distribution $G(\theta)$. Also, X_i is the pre-treatment value and Y_i is the post-treatment value of an observed state, for example, body temperature.

Suppose that the illness effects people from time to time with various degrees of symptoms. Hence, we can assume that there is a reoccurence process of the illness, captured and expressed by an exponential distribution. Consequently, let the assumptions (model) be as follows.

(A_1): θ is the independent positive random variable, with a finite expected value ($E[\theta] < \infty$).

(A_2): $X_i|\theta_i \sim \exp(1/\theta_i)$ is the independent random variable.

(A_3): $Y_i|X_i, \theta_i \sim \exp(1/c\theta_i)$ is a conditionally independent random variable.

In this model construction, c denotes a positive multiplicative treatment effect factor.

The objective is to find an unbiased estimator [4] of the treatment effect on the

population, as a relationship of X_i and Y_i, such as $S_n = \sum_{i=1}^{n} u_i(X_i)\, Y_i$, only by a function of X_i, say $T_n = \sum_{i=1}^{n} v_i(X_i)$, at a desired confidence level.

Let $u_i(X_i)$ be the indicator function of the event $\{X_i > a\}$, denoted by

$u_i(X_i) = I\{X_i > a\} = \begin{cases} 1, & \text{if } X_i > a \\ 0, & \text{otherwise} \end{cases}$, for $i = 1, 2, \ldots n$, where, a is a nonnegative constant for selecting the treatment group, say it is a treshold value (cutoff point) for the acceptable level of X_i, which indicates the state of a person's health. Thus, if $X_i \leq a$, then the i^{th} person observed is healthy and does not need treatment. On the other hand, if $X_i > a$, then apply the treatment and observe its effect by Y_i, on this person. The biased random selection criterion, $X > a$, may be based on historical data, or any prior medical knowledge. Since it is desired to estimate S_n, by utilizing T_n, such that $E[S_n] = cE[T_n]$, then let $v_i(X_i) = \int_0^{X_i} I\{t > a\}dt$, for any i.

Lemma 1 The pair of functions, (u_i, v_i), satisfies the equation (1), that is, $E[v_i(X_i)|\theta_i] = \theta_i\, E[u_i(X_i)|\theta_i]$.

Proof.

Pick an $i \in \{1, 2, \ldots, n\}$. Since $v_i(X_i) = \int_0^{X_i} I\{t > a\}dt$ and
$u_i(X_i) = I\{X_i > a\}$, then
$E[v_i(X_i)|\theta_i] = \int_0^{\infty} v_i(x)dF(x|\theta_i) = \int_0^{\infty} \int_0^{x} I\{t > a\}dt(1/\theta_i)e^{-x/\theta_i}dx$
$= \int_0^{\infty} \int_0^{x} I\{t > a\}(1/\theta_i)e^{-x/\theta_i}dtdx$
$= \int_0^{\infty} I\{t > a\} \int_t^{\infty}(1/\theta_i)e^{-x/\theta_i}dxdt$, by changing the order of integration,
$= \int_0^{\infty} I\{t > a\}e^{-t/\theta_i}dt$
$= \theta_i \int_0^{\infty} I\{t > a\}(1/\theta_i)e^{-t/\theta_i}dt$
$= \theta_i \int_0^{\infty} u_i(x)dF(x|\theta_i)$
$= \theta_i\, E[u_i(X_i)|\theta_i]$. □

Lemma 2 $E[S_n] = cE[T_n]$.

Proof.

Since $E[T_n] = E[\sum_{i=1}^{n} v_i(X_i)] = \sum_{i=1}^{n} E[v_i(X_i)] = \sum_{i=1}^{n} E[E[v_i(X_i)|\theta_i]]$,
and $E[v_i(X_i)|\theta_i] = \theta_i\, E[u_i(X_i)|\theta_i]$, by the Lemma 1, then

$$E[T_n] = \sum_{i=1}^{n} E[\theta_i \, E[u_i(X_i)|\theta_i]] \tag{2}$$

On the other hand, $E[S_n] = E[\sum_{i=1}^{n} u_i(X_i)\, Y_i] = \sum_{i=1}^{n} E[E[u_i(X_i)\, Y_i|\theta_i]]$
$= \sum_{i=1}^{n} E[E[u_i(X_i)|\theta_i]\, E[Y_i|\theta_i]]$, by (A_3).
Since $E[Y_i|\theta_i] = c\theta_i$, then

$$\begin{aligned} E[S_n] &= \sum_{i=1}^{n} E[E[u_i(X_i)|\theta_i] c\theta_i] = \sum_{i=1}^{n} cE[E[\theta_i u_i(X_i)|\theta_i]] \\ &= c\sum_{i=1}^{n} E[\theta_i E[u_i(X_i)|\theta_i]] \end{aligned} \tag{3}$$

Therefore, $E[S_n] = cE[T_n]$, by equations (2) and (3). □

3. Estimation of S_n

Let $S_n = \sum_{i=1}^{n} I\{X_i > a\}\, Y_i$, and $T_n = \sum_{i=1}^{n} \int_0^{X_i} I\{t > a\} dt$.
Since $E[T_n] = \sum_{i=1}^{n} E[\int_0^{X_i} I\{t > a\} dt]$ and
$\int_0^{X_i} I\{t > a\} dt = (X_i - a)\, I\{X_i > a\}$, then
$E[T_n] = \sum_{i=1}^{n} E[(X_i - a)\, I\{X_i > a\}]$
$= \sum_{i=1}^{n} E[E[(X_i - a)\, I\{X_i > a\}|\theta_i]]$,
where, $E[(X_i - a)\, I\{X_i > a\}|\theta_i] =$
$\int_0^{\infty} (x - a)\, I\{x > a\}\, (1/\theta_i) e^{(-x/\theta_i)} dx = \theta_i \, e^{(-a/\theta_i)}$.
Hence, $E[E[(X_i - a)\, I\{X_i > a\}|\theta_i]] = E[\theta_i \, e^{(-a/\theta_i)}] = \int_0^{\infty} \theta_i e^{(-a/\theta_i)} \, dG(\theta_i)$.

Since $E[X] < \infty$ and as $n \to \infty$, $\frac{1}{n} \sum_{i=1}^{n} (X_i - a)^+ \to E[T_n]$, by the law of large numbers [4], then $E[T_n] = \sum_{i=1}^{n} E[\theta_i e^{(-a/\theta_i)}] = \sum_{i=1}^{n} \int_0^{\infty} \theta_i e^{(-a/\theta_i)} \, dG(\theta_i)$ is estimated by $\hat{T_n} = \frac{1}{n} \sum_{i=1}^{n} (X_i - a)^+$.

In this formulation, $(Z)^+ = \max\{0, Z\}$, for any Z. Therefore, $E[S_n]$ is estimated by $c\hat{T_n} = \frac{c}{n} \sum_{i=1}^{n} (X_i - a)^+$, as a consequence of Lemma 2.

3.1. Confidence interval for the estimator

Since $E[Y_i|\theta_i, X_i] = c\theta_i$ and $Var(Y_i|\theta_i, X_i) = c^2\theta_i^2$, then $E[Y_i^2|\theta_i, X_i] = 2c^2\theta_i^2$. Also, since $E[S_n] = cE[T_n]$ and $S_n - cT_n =$

$\sum_{i=1}^{n}(I\{X_i > a\}\, Y_i - c\int_0^{X_i} I\{t > a\}dt)$, then as $n \to \infty$,

$\frac{S_n - cT_n}{\sqrt{n}} \to \mathcal{N}(0, \sigma^2)$, in distribution. Here, $\mathcal{N}(\mu, \sigma^2)$ denotes the *Normal* probability distribution with mean μ and variance σ^2.

Hence, $\sigma^2 = E[(I\{X_i > a\}\, Y_i - c\int_0^{X_i} I\{t > a\}dt)^2]$

$=E[I^2\{X_i > a\}\, Y_i^2 - 2cI\{X_i > a\}\, Y_i \int_0^{X_i} I\{t > a\}dt + (c\int_0^{X_i} I\{t > a\}dt)^2]$

$=E[Y_i^2\, I\{X_i > a\}] - 2cE[Y_i\,(X_i - a)\, I\{X_i > a\}] + c^2E[(X_i - a)^2\, I\{X_i > a\}]$

$=E[2c^2\theta_i^2\, E[I\{X_i > a\}|\theta_i] - 2c^2\theta_i\, E[(X_i - a)\, I\{X_i > a\}|\theta_i]$

$+c^2E[(X_i - a)^2\, I\{X_i > a\}|\theta_i]]$, by conditioning [9], (A_2) and (A_3).

Therefore, we can find an interval to estimate the value of the unknown parameter at any desired precision. Without loss of generality, fix the Type-I error probability [4] to be 0.05, then at 95% confidence level, an estimate of the σ^2 (say, $\hat{\sigma}^2$) would be as follows. We have $E[X|\theta] = \theta$ by (A_2), and then an unbiased estimator of θ would be $\hat{\theta}$, with $E[\theta] = E[E[X|\theta]] = E[X]$. Thus, $\hat{\theta} = \overline{X}$ is an unbiased estimator of $E[X]$, where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Let $w_i(X_i) = 2c^2\theta_i^2\, I\{X_i > a\} - 2c^2\theta_i\,(X_i - a)^+ + c^2[(X_i - a)^+]^2$, for any positive and finite θ_i, for all i's. Consequently for any positive and finite θ we have, $w(X, \theta) = 2c^2\theta^2 I\{X > a\} - 2c^2\theta(X - a)^+ + c^2[(X - a)^+]^2$.

Since $Var(X|\theta) = \theta^2$, by (A_2), then $E[Var(X|\theta)] = E[\theta^2]$. The θ^2 is estimated by $\hat{\theta}^2$, while $E[Var(X|\theta)]$ itself can be estimated by $(\frac{n}{n-1})S^2$, based on the available empirical evidence, where,

$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n}\{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\} = \frac{1}{n}\sum_{i}^{n} X_i^2 - \overline{X}^2$. Therefore, as $n \to \infty$, $E[\theta^2]$ can be estimated by $(\frac{n}{n-1})S^2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\overline{X}^2$.

Also, we have $E[w(X, \theta)] = E[2c^2\theta^2 I\{X > a\} - 2c^2\theta(X - a)^+$

$+ c^2\{(X - a)^+\}^2] = 2c^2E[\theta^2 E[I\{X > a\}|\theta]] - 2c^2E[\theta E[(X - a)^+|\theta]]$

$+ c^2E[E[\{(X - a)^+\}^2|\theta]] = 2c^2E[\theta^2]E[I\{X > a\}] - 2c^2E[\theta]E[(X - a)^+]$

$+ c^2E[\{(X - a)^+\}^2]$, by conditioning and (A_1).

As $n \to \infty$, we have $\frac{1}{n}\sum_{i=1}^{n} w_i(X_i) \to E[w(X, \theta)]$, by the law of large numbers [4]. Then $\hat{w}(X) = 2c^2S^2\,(\frac{1}{n-1})[\#\{i|i \in \{1, \ldots, n\}, X_i > a\}] - 2c^2\overline{X}\,[\frac{1}{n}\sum_{i\in\{1,\ldots,n|X_i>a\}}(X_i - a)] + c^2[\frac{1}{n}\sum_{i\in\{1,\ldots,n|X_i>a\}}(X_i - a)^2]$, is an estimator of $w(X, \theta)$, based on the observed X_is, such as when $n \to \infty$, $\hat{w}(X) \to E[w(X, \theta)]$. Here, $\#\{A\}$ denotes the cardinality of a set A.

Also, let $W_n = \frac{1}{n}\sum_{i=1}^{n} w_i(X_i)$. While, as $n \to \infty$, $E[W_n] \to \sigma^2$, with probability one (w.p. 1); thus, $\hat{\sigma}^2 = \hat{W}_n$ would be an estimator of σ^2. Here, $\hat{W}_n = \hat{w}(X)$. Also, as $n \to \infty$, we have $\frac{S_n - cT_n}{\sqrt{W_n}} \to \mathcal{N}(0,1)$, in distribution. Therefore, at the desired level of confidence, $2\Phi(z) - 1 = 0.95 \Rightarrow \Phi(z) = 1.95/2 \Rightarrow z = \Phi^{-1}(1.95/2) \Rightarrow z = 1.96$. Here, $\Phi(.)$ denotes the cumulative distribution function (c.d.f.) [9] of the *Standard Normal* distribution.

Consequently, the confidence interval of S_n (at the desired level of precision), centered about $E[S_n]$, would be $[E[S_n] - z\sigma, E[S_n] + z\sigma]$ and it is asymptotically estimated by $[c\hat{T}_n - 1.96\sqrt{\hat{W}_n}, c\hat{T}_n + 1.96\sqrt{\hat{W}_n}]$, at the 95% level of confidence.

4. Estimating the treatment effect

The unknown value of the treatment effect could be found by the relationship, $S_n = cT_n$, as follows. Let c_n be such that $S_n = c_n T_n$. Since $T_n = \sum_{i=1}^{n}(X_i - a)^+$ and $S_n = \sum_{i=1}^{n}(X_i - a)^+ Y_i$, based on the pre- and post-treatment observations and together with the selection criterion, $X_i > a$, for $i = 1, \ldots, n$, then we can find $c_n = S_n/T_n = \sum_{i=1}^{n}(X_i - a)^+ Y_i / \sum_{i=1}^{n}(X_i - a)^+$, by utilizing the available data. Hence, as $n \to \infty$, we have $c_n \to c$, almost surely (a.s.), as a consequence of Lemma 2.

5. Conclusions

In this work, a method for forming inferences about the characteristics of a pair of random variables and their functions, by utilizing information generated by only one of them and its functions, has been investigated for the exponential probability distribution, with example application as an alternative method for clinical trials, in the field of biomathematics. Therefore, the relevant information about the effect of a treatment can be acquired by using the above mentioned sequential method.

In this construct, the unknown distribution of the parameter G becomes an irrelevant factor to the estimation procedure, thus this approach may be called a semi-parametric method. The model may also explain the time to become ill (for a population), such that the mean time to illness is θ. Also, from the perspective of public health policy, a decision maker may wish to assess the risk, or the likelihood of being ill, by $P\{X > a\} = 1 - F_{X|\theta}(a) = e^{-\frac{a}{\theta}}$ (similarly, the probability of being healthy, $P\{X \leq a\} = 1 - e^{-\frac{a}{\theta}}$, which may be the degree of health of a population). The model constructed here assumes that the driving force behind the stochastic events is primarily time, similar to the assumption

used in quantum statistics (such as in the Schrödinger's equations) [6]. Furthermore, one can construct arbitrary nonnegative random variables from exponential random variables, by utilizing smooth enough inverse function relationships, to approximate varieties of other distributions [12]. However, a nonparametric model, which does not rely on any particular type of probability distribution assumption, would provide much more robust explanations than the semi-parametric model employed in this work.

Acknowledgement

The author is grateful for the productive suggestions made by an anonymous referee, and the help provided by Professor Ahmet Feyzioglu.

References

[1] Bernstein L., Hanisch R., Sullivan-Halley J., Ross R. K.: *Treatment with human Chorionic Gonadotropin and risk of breast cancer*, Cancer Epidemiology Biomarkers & Prevention, Vol. 4, No. 5, pp. 437-440 (1995).

[2] Fidler P., Loprinzi C. L., O'Fallon J. R., Leitch J. M., Lee J. K., Hayes D. L., Novotny P., Schutjer D. C., Bartel J., Michalak J. C.: *Prospective evaluation of a chamomile moutwash for prevention of 5-FU induced oral mucositis*, Cancer, Vol. 77, No. 3, pp. 522-525 (1996).

[3] Fisher R. A.: *On the mathematical foundations of theoretical statistics*, Philosophical Transactions of the Royal Society of London, Series A, Vol. 222, pp. 309-368 (1922).

[4] Hogg R. V., Craig A. T.: *Introduction to Mathematical Statistics*, Prentice Hall, Inc., Upper Saddle River, New Jersey, 1995.

[5] Hornberger J. C., Brown B. WM. Jr.: *Designing a cost-effective clinical trial*, Statistics, No. 20, pp. 2249-2 (1995).

[6] Khinchin A. Y.: *Mathematical Foundations of Quantum Statistics*, Dover Publications, Inc., Mineola, New York, 1998.

[7] Moher D., Jadad A. R., Nichol G., Penman M., Tugwell P., Walsh S.: *Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists*, Controlled Clinical Trials, Vol. 16, No. 1, pp. 62-73 (1995).

[8] Robbins H.: *The U,V method of estimation*, in Statistical Decision Theory and Related Topics IV, Vol. 1, Editors; Gupta S. S., Berger J. O., p.p. 265-270, Springer-Verlag, New York, 1988.

[9] Ross S.: *A First Course in Probability*, Macmillan Publishing Co., Inc., New York, 1976.

[10] Schmoor C., Olschewski M., Schumacher M.: *Randomized and non-randomizes patients in clinical trials: Experiences with comprehensive cohort studies*, Statistics in Medicine, Vol. 15, No. 3, pp. 263-271 (1966).

[11] Senn S.: *A personal view of some controversies in allocating treatments to patients in clinical trials*, Statistics in Medicine, Vol 14, No. 24, pp. 2661-2674 (1995).

[12] Szekli R.: *Stochastic Ordering and Dependence in Applied Probability*, Springer-Verlag, New York, 1995.

Kemal GÜRSOY
Department of Mathematics,
Bogazici University,
225 Anderson Hall,
Bebek, İstanbul-TURKEY
e-mail: gursoyk@boun.edu.tr

Received 01.03.2004