Parameter Estimation of Binomial distribution using T.O.M with Exponential Family

Jubran Abdulameer K.
Department of Statistics and Informations
College of Computer Science and Mathematics
University of Al-Qadisiya
E-mail: jubalaa98@yahoo.com

Abstract

In this research we will estimate the parameter of binomial distribution that has exponential family using T.O.M (Term Omission Method) and compare it with (MLE) method using MSE (Mean Square Error) with simulation.

Keywords: Term Omission Method, Exponential Family, Maximum Likelihood Estimator

1. Introduction

Estimators that based on T.O.M[5] deals with many distributions with discrete or continuous random variable that have exponential family, in our case we will take discrete random variable with binomial distribution.

2. Exponential Family

We note that there are many definitions of such type of that representation of exponential families. In this research, we will recall the regular type of exponential class.

3. Definitions

Def. (1): A family of discrete (continuous) random variables is called an exponential family if the probability density functions (probability mass functions) can be expressed in the form

$$f_X(x/\theta) = h(y)c(\theta)\exp\left(\sum_{i=1}^{k}\theta_i t_i(x)\right) \quad , \quad x = 0,1,2,\ldots$$

for x in the common domain of the $f_X(x/\theta)$, $\theta \in R^k$. Obviously h and c are non-negative functions. The $t_i(x)$ are real-valued functions of the observations.[1]

Def. (2): Let ϑ be an interval on the real line. Let $\{f(x;\theta)\text{: } \theta \in \vartheta\}$ be a family of pdf's (or pmf's). We assume that the set $\{\underline{x} : f(\underline{x};\theta) > 0\}$ is independent of θ, where $\underline{x} = (x_1, x_2, \ldots, x_n)$. We say that the family $\{f(x;\theta)\text{: } \theta \in \vartheta\}$ is a one-parameter exponential family if there exist real-valued functions Q(θ) and D(θ) on ϑ and Borel-measurable functions T($\underline{X}$) and S($\underline{X}$) on R^n such that

$$f(\underline{x};\theta) = \exp\big(Q(\theta)T(\underline{x}) + D(\theta) + S(\underline{x})\big)$$

if we write $f(\underline{x};\theta)$ as

$$f(\underline{x};\eta) = h(\underline{x})c(\eta)\exp\big(\eta T(\underline{x})\big)$$

where

$h(\underline{x}) = \exp\big(S(\underline{x})\big), \eta = Q(\theta)$, and $c(\eta) = \exp\big(D\big(Q^{-1}(\eta)\big)\big)$,
then we call this the exponential family in canonical form for a natural parameter η.

Def. (3): Let $\underline{\vartheta} \subseteq R^k$ be a k-dimensional interval. Let $\{f(x;\ \underline{\theta})\text{: } \underline{\theta} \in \underline{\vartheta}\}$ be a family of pdf's (or pmf's). We assume that the set $\{\underline{x} : f(\underline{x};\underline{\theta}) > 0\}$ is independent of $\underline{\theta}$, where $\underline{x} = (x_1, x_2, \ldots, x_n)$. We say that the family $\{f(x;\ \underline{\theta})\text{: } \underline{\theta} \in \underline{\vartheta}\}$ is a k-parameter exponential family if there exist real-valued functions $Q_1(\underline{\theta})$,…, $Q_k(\underline{\theta})$ (1) and $D(\underline{\theta})$ on $\underline{\vartheta}$ and Borel-measurable functions $T_1(\underline{X})$,…,$T_k(\underline{X})$ and $S(\underline{X})$ on R^n such that:[2]

$$f(\underline{x};\underline{\theta}) = \exp\left(\sum_{i=1}^{k} Q_i(\underline{\theta})T_i(\underline{x}) + D(\underline{\theta}) + S(\underline{x})\right) \quad \ldots (4)$$

Def. (4): Exponential family is a class of distributions that all share the following form:

$$P(y/\eta) = h(y)\exp\{\eta^T T(y) - A(\eta)\} \quad \dots (5)$$

* η is the natural parameter. For a given distribution η specifies all the parameters needed for that distribution.
* $T(y)$ is the sufficient statistic of the data (in many cases T(y) = y, in which case the distribution is said to be in canonical form and η is referred to as the canonical parameter).
* A(η) is the log-partition function which ensures that p(y/η) remains a probability distribution.
* h(y) is the non-negative base measure (in many cases it is equal to 1).

Note that since η contains all the parameters needed for a particular distribution in its original form, we can express it with respect to the mean parameter θ:[3]

$$P(y/\theta) = h(y)\exp\{\eta(\theta)T(y) - A(\eta(\theta))\} \quad \dots (6)$$

Def. (5): (Regular Exponential Family): Consider a one-parameter family $\{f(x;\theta):\theta \in \Omega)$ of probability density functions, where Ω is the interval set $\Omega = \{\theta : \gamma < \theta < \delta\}$, where γ and δ are known constants, and where[7]

$$f(x;\theta) = \begin{cases} e^{p(\theta)k(x)+s(x)+q(\theta)} & \text{a < x < b} \\ 0 & \text{o.w} \end{cases} \quad \dots (7)$$

The form (7) is said to be a member of the exponential class of probability density functions of the continuous type, if the following conditions satisfy:

1) Neither (a) nor (b) depends upon θ. increasingly
2) p(θ) is a nontrivial continuous function of θ.
3) Each of k′(x)≠0 and s(x) is a continuous function of x.

and the following conditions with discrete random variable X_i:

1) The set $\{x : x = a_1, a_2, \dots\}$ does not depend upon θ.
2) p(θ) is a nontrivial continuous function of θ.
3) k(x) is a nontrivial function of x.

4. Binomial Distribution that belong to Exponential Family

With discrete random variables that distributed binomial, we can write the pmf as an exponential class form as follows:

If the random variable X follows the binomial distribution with parameters n and λ, we write X ~ B(n, λ). The probability of getting exactly x successes in n trials is given by the probability mass function:

$$f(x) = \binom{n}{x}\lambda^x(1-\lambda)^{n-x} \qquad x \in \{0,1,2,\dots,n\}$$

We can written it as a exponential form

$$f(x) = \binom{n}{x}\exp\left(x\ln\left(\frac{\lambda}{1-\lambda}\right) + n\ln(1-\lambda)\right)$$

or

$$f(x) = \exp\{\ln(n!) - \ln(x!) - \ln((n-x)!) + x\ln\frac{\lambda}{1-\lambda} + n\ln(1-\lambda)\}$$

where

$$k(x) = x \;,\; \mathrm{p}(\lambda) = \ln\frac{\lambda}{1-\lambda} \;,$$

$$\mathrm{s(x) = -ln}[x!] - \ln((n-x)!) \,,\, \mathrm{q}(\lambda) = \ln(n!) + n\ln(1-\lambda)$$

5. T.O.M with Exponential Family

T.O.M[6] can used to estimate the value of parameter θ with fixed n, below we will derivative the T.O.M of distributions (with one parameter) that have exponential family as follows:

For sample with size N having the p.d.f (p.m.f) f(x;θ), and for any two values x_i and x_{i+1}, where 1≤i≤n-1,

$$x_i \qquad y_i = e^{k(x_i)p(\theta)+q(\theta)+s(x_i)}$$

$$x_{i+1} \qquad y_{i+1} = e^{k(x_{i+1})p(\theta)+q(\theta)+s(x_{i+1})}$$

by taking the natural logarithm to y_i , y_{i+1} we have

$$x_i \qquad y_i = k(x_i)p(\theta) + q(\theta) + s(x_i)$$

$$x_{i+1} \qquad y_{i+1} = k(x_{i+1})p(\theta) + q(\theta) + s(x_{i+1})$$

and by subtract the last result y_i from y_{i+1} we obtain

$$y_{i+1} - y_i = p(\theta)[k(x_{i+1}) - k(x_i)] + [s(x_{i+1}) - s(x_i)]$$

and again by subtract $[s(x_{i+1}) - s(x_i)]$ from the final amount we have

$$y_{i+1} - y_i - [s(x_{i+1}) - s(x_i)] = p(\theta)[k(x_{i+1}) - k(x_i)]$$

Finally, by divided this amount over $[k(x_{i+1}) - k(x_i)]$ we have the function of θ, p(θ).

Therefore, we can define the $p^i(\theta)$ as follows:

$$p^i(\theta) = \frac{[\ln(y_{i+1}) - \ln(y_i)] - [s(x_{i+1}) - s(x_i)]}{[k(x_{i+1}) - k(x_i)]} \quad \dots (8)$$

where $p^i(\theta)$ represent to the values that we have from previous steps of T.O.M, $\forall$i=1,2,…,N-1. Thus from eq. (8) we can educe values of θ^i from $p^i(\theta)$.
Therefore the estimation of θ can found now using the least square error with the following equation:

$$\hat{\theta} = Min\left(\sum_{m=1}^{N}\left[f(x_m,\theta^i) - y_m\right]\right)^2 \qquad i = 1,2,\ldots,N-1$$

where $f(x_m,\theta^i)$ is the value of function $f(x_m)$ on θ^i, and $y_m = y(x_m)$ is the observed value on x_m.

6. T.O.M with Exponential Family of Binomial Distribution

Recalling the exponential family of binomial distribution in section (4)

$$f(x) = \binom{n}{x}\exp\left(x\ln\left(\frac{\lambda}{1-\lambda}\right) + n\ln(1-\lambda)\right) \qquad \ldots (9)$$

where

$$k(x) = x \quad , \quad p(\lambda) = \ln\frac{\lambda}{1-\lambda} \quad ,$$

$$s(x) = -\ln[x!] - \ln((n-x)!) \quad , \quad q(\lambda) = \ln(n!) + n\ln(1-\lambda)$$

therefore by using (8) we can write

$$p^i(\lambda) = \frac{[\ln(y_{i+1}) + \ln(x!_{i+1}) + \ln((n-x_{i+1})!)]}{\left[x_{i+1} - x_i\right]} - \frac{\ln(y_i) + \ln(x!_i) + \ln((n-x_i)!)}{\left[x_{i+1} - x_i\right]}$$

or

$$p^i(\lambda) = \frac{\left[\ln(y_{i+1}) - \ln(y_i)\right] - \left[\ln\left(\frac{x!_i}{x!_{i+1}}\right) + \ln\left(\frac{(n-x_i)!}{(n-x_{i+1})!}\right)\right]}{\left[x_{i+1} - x_i\right]}$$

we can here educe the θ^i, $1 \le i \le N-1$ from $p^i(\theta)$ to have the estimating of θ, thus

$$p^i(\lambda) = k$$

$$\ln\frac{\lambda}{1-\lambda} = k$$

$$\lambda^i = \frac{e^k}{1+e^k} \qquad 1 \le i \le N-1$$

where

$$k = \frac{\left[\ln(y_{i+1}) - \ln(y_i)\right] - \left[\ln\left(\frac{x!_i}{x!_{i+1}}\right) + \ln\left(\frac{(n-x_i)!}{(n-x_{i+1})!}\right)\right]}{\left[x_{i+1} - x_i\right]}$$

where the constant k comes from eq. (10)

7. Results

Below table for comparison between two methods T.O.M using eq. (11) and MLE[4] for binomial distribution and choose the best method according to MSE (Mean Square Error) with equation below:

$$MSE = \frac{\sum_{i=1}^{n}\left[f(x_i;\theta) - f(x_i;\hat{\theta})\right]^2}{n}$$

where $f(x_i;\lambda), f(x_i;\hat{\lambda})$ represent to the probability function of variable x_i, λ, $\hat{\lambda}$ parameter and estimating parameter respectively.

The estimate values of parameter λ that founding by MLE and T.O.M for Binomial distribution.

λ	N	T.O.M	MLE
0.1	20	0.181818	0.2
	60	0.100946	0.0923519
	100	0.092308	0.1909579
	500	0.099534	0.0674935
0.5	20	0.490738	0.2777778
	60	0.516515	0.3921307
	100	0.510166	0.5245383
	500	0.49447	0.4406192
0.9	20	0.941176	0.8823944
	60	0.897436	0.8780669
	100	0.89899	0.8823266
	500	0.900255	0.9032506

7. Discussion

In this research and from the pervious table we can see the following results:

1) The preference of the T.O.M with other using MSE in all samples,
2) Approximation of T.O.M when sample large. … (10)
3) We have an exact estimation when sample exact fitting the distribution.

References

[1] Watkins, Joseph C., 2008. "Exponential Families of Random Variables", University of Arizona, p: 1.

[2] Jurgen, S., 1999. "Mathematical Statistics I", Utah State University, p: 120. … (11)

[3] Andreas Vlachos, 2010, "Notes on exponential family distributions and generalized linear models".

[4] DaGupta, A. and Rubin, Herman 2004. "Estimation of Binomial Parameters when Both n, p are Unknown", Purdue University.
[5] Labban, J.A., 2005. "Modified and Simplified Method for Finite Difference (Divided)". Al-Qadisiya Journal for Science, Vol. 10, No. 2, pp. 244-251.
[6] Labban, J.A., 2009. "Approximation of Scale Parameter of Inverted Gamma Distribution by TOM Modified". Accepted for Publishing
[7] Robert V. Hogg and Allen T. Craig, 1978. "Introduction to Mathematical Statistics", 4th Edition, p: 357.

تقدير معلمة توزيع ذي الحدين باستخدام T.O.M للتوزيعات التي تنتمي للعائلة الأسية

الخلاصة

في هذا البحث سنجد تقدير معلمة التوزيع الذي ينتمي للعائلة الأسية باستخدام T.O.M ومقارنتها بتقدير MLE باستخدام معدل مربع الخطأ MSE باستخدام المحاكاة.