MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 17/1976.

ON THE ESTIMATION OF A PARAMETER OF A CONVOLUTION WITH AN APPLICATION TO QUEUEING THEORY

by Jacob Eshak Samaan

INTRODUCTION

Suppose that the observed random variable X is the sum of two independent random variables Y and Z, where Y is uniformly distributed in the interval $(0, 1)$ and Z is exponentially distributed with unknown parameter λ. We consider here the problem of estimating λ from a sample $X_1, X_2, \dots, X_n$.

Such a problem may arise, if we consider, for example, the single server queueing system for which the arrival process is Poisson with unknown parameter λ, while the service time is uniformly distributed in the interval $(0, 1)$. The customer which arrives when the server is busy is rejected. Let us observe the departure times $\tau_1, \tau_2, \dots, \tau_n$ of the served customers. Let:

$$X_0 = 0,$$

$$X_i = \tau_{i+1} - \tau_i \qquad i = 1, 2, \dots, n.$$

Then each X_i is the sum of a service time and an interarrival time (or a part of interarrival time which is again exponentially distributed), and our purpose is to estimate the arrival rate λ from the sample $X_1, X_2, \dots, X_n$.

Conditional sufficient statistic:

If we have n independent observations $y_1, y_2, \dots, y_n$ from a population with density function $p_\Theta(y)$, where Θ is an unknown parameter, then a necessary and sufficient condition for a statistic $T(Y) = T(y_1, \dots, y_n)$, to be sufficient for Θ is that, the joint density function:

$$L_n(Y, \Theta) = L_n(y_1, \dots, y_n; \Theta) = p_\Theta(y_1), \dots, p_\Theta(y_n)$$

is factorizable in the form:

$$(1) \qquad L_n(Y; \Theta) = g_\Theta[T(Y)].\ K(Y)$$

where the first factor may depend on Θ but depends on Y only through $T(Y)$, whereas the second factor is independent of Θ.

It may happen that the joint density function of the observations is factorizable only on a subset of the whole sample space.

This means that, sufficient statistics does not exist on the whole space, but there is a sufficient statistic on a subset of the whol space, such sufficient statistic may be called "conditional sufficient statistic" for Θ.

Such situation occurs in the problem of estimating λ.

Conditional sufficient statistic for λ:

The density function of X is given by:

$$h(x) = \begin{cases} 1 - e^{-\lambda x} & 0 \leqslant x < 1 \\ (e^{\lambda} - 1)e^{-\lambda x} & x \geqslant 1. \end{cases}$$

Since any sample $x_1, x_2, \ldots, x_n$ may contain observations which have values less than one, so it is clear that the joint density function is, generally, not factorizable in the whole sample space.

However in the subset $X \geqslant 1$, it is factorizable and has the form:

$$L_n(X;\lambda) = (e^{\lambda} - 1)^n e^{-\lambda \sum_{i=1}^{n} X_i}$$

This means that the statistic::

$$T_n = \sum_{i=1}^{n} X_i$$

is a sufficient statistic for λ on the subset $X_i \geqslant 1$.

Conditional likelihood estimator for λ :

Suppose that we continue sampling until we get n observations all of which have value greater than one.

Let:

$A \equiv$ denotes the event that $X_i \geqslant 1$ for all $i = 1, 2, \ldots, n$.

Then

$$P(A) = [P(x \geqslant 1)]^n = [\int_1^{\infty} (e^{\lambda} - 1)e^{-\lambda x} dx]^n = (\frac{1}{\lambda^n})e^{-\lambda n}(e^{\lambda} - 1)^n$$

Thus, the conditional likelihood function of the sample $(X_1, X_2, \ldots, X_n)$ is:

$$L_n(X;\lambda \mid A) = \frac{L_n(X;\lambda)}{P(A)} = \lambda^n e^{n\lambda} e^{-\lambda \sum_{i=1}^{n} x_i},$$

i.e. the conditional likelihood equation is:

$$\frac{n}{\lambda} + n - \sum_{i=1}^{n} x_i = 0.$$

The solution of this equation is:

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i - n},$$

which we refer to it as the conditional likelihood estimator for λ.

If: $A_i \equiv$ denotes the event that $x_i \geqslant 1$, then:

$$P(X_i \mid A_i) = p\{x_i \leqslant x \mid x_i \geqslant 1\} = 1 - e^{-\lambda(x-1)}$$

which shows that under the condition A_i, the random variable $(X_i - 1)$ has an exponential distribution with parameter λ. This means that under the condition A, the random variable:

$$W = \sum_{i=1}^{n} X_i - n$$

has a gamma distribution with parameters (n, λ). So

$$E(\hat{\lambda}) = \frac{n}{n-1}\lambda$$

and consequentely the estimator $\hat{\lambda}$ is a biased estimator for λ.
To correct this biasedness we can consider the estimator

$$\hat{\lambda} = \frac{n-1}{\sum_{i=1}^{n} X_i - n}.$$

Now, let us have a sample of fixed size n, which contains observations greater than one and others less than one, and we want to find an estimator for λ which bases only on the observations greater than one in this sample. Let X_i^* denotes the observations greater than one, and n^* denotes the number of such observations in a general sample with fixed size n.

Then n^* is a random variable with binomial distribution:

$$P(n^* = k) = \binom{n}{k} p^k q^{n-k}, \quad \text{where}$$

$$P = \int_1^{\infty} (e^{\lambda} - 1)e^{-\lambda x} dx = \frac{1}{\lambda}(1 - e^{-\lambda}), \qquad q = 1 - p.$$

Thus, we may introduce the estimator

$$\lambda^* = \frac{n^* - 1}{\sum_{i=1}^{n^*} x_i^* - n^*}, \qquad n^* > 0,$$

as an estimator for λ, which depands only on the observations X_i^* in the sample. Since

$$p\{(x_i^* - 1) \leqslant x\} = 1 - e^{-\lambda x}, \qquad x \geqslant 0.$$

Then, we have:

$$p\{Y^* \leqslant x \mid n^* = k\} = \int_0^x \lambda \frac{(\lambda y)^{k-1}}{\Gamma(k)} e^{-\lambda y} dy,$$

where

$$Y^* = \sum_{i=1}^{n^*} X_i^* - n^*.$$

i.e.

$$\frac{p\{Y^* \leqslant x, n^* = k\}}{p\{n^* = k\}} = \int_0^x \lambda \frac{(\lambda y)^{k-1}}{\Gamma(k)} e^{-\lambda y} dy.$$

Thus, the joint density function of Y^* and n^* is given by:

$$f(x, k) = \lambda e^{-\lambda x} \frac{1}{\Gamma(k)} \binom{n}{k} p^k q^{n-k} (\lambda x)^{k-1}, \qquad x \geqslant 0.$$

Since

$$P\{n^* > 0\} = 1 - q^n$$

then:

$$E\{\lambda^* \mid n^* > 0\} = \lambda.$$

So, λ^* may be called conditionally unbiased estimator for λ.

Direct estimator for λ :

Assume that we have a general sample of size n, which contains observations greater than one and others less than one. Then:

$$\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{2}$$

is an unbiased estimator of $\frac{1}{\lambda}$.

Thus, we may introduce the statistic:

$$\bar{\lambda} = \frac{n}{\sum_{i=1}^{n} X_i - \frac{n}{2}}.$$

as an estimator for λ, and call it a direct estimator. A disadvantage of this estimator is that, there is a positive probability that it have negative values.

Let $\bar{n}$ denotes the number of times we get negative estimator for λ out from one hundred samples with size n. The following tables show the values of $\bar{n}$ corresponding to different values of λ for: $n = 1000$, $n = 100$.

λ	$\bar{n}$
100	9
70	7
50	2
30	–
15	–

$n = 1000$

Table 1.

λ	$\bar{n}$
25	7
20	5
15	2
10	–
5	–

$n = 100$

Table 2.

All the estimators $\hat{\lambda}$, λ^* and $\bar{\lambda}$ are consistent estimators for λ. But because of the difficulties in calculating the variances of λ^* and $\bar{\lambda}$ we can not compare their variances.

But to see the advantages and disadvantages of using estimates based on conditional sufficient statistics, we consider $\frac{1}{\hat{\lambda}}$ and $\frac{1}{\bar{\lambda}}$ as estimators for $\frac{1}{\lambda}$ and compare their variances.

Since:

$$p\{(X_i^* - 1) \leqslant x\} = 1 - e^{-\lambda x}, \qquad x \geqslant 0,$$

then we have:

$$E(\frac{1}{\hat{\lambda}}) = E(x^* - 1) = E((x_i - 1) \mid A_i) = \frac{1}{\lambda},$$

$$E(\frac{1}{\bar{\lambda}}) = E(x_i - \frac{1}{2}) = \frac{1}{\lambda}.$$

And

$$\text{Var}\,(\frac{1}{\hat{\lambda}}) = \frac{1}{n}\,\text{Var}\,(X_i^*) = \frac{1}{n\lambda^2},$$

$$\text{Var}\,(\frac{1}{\bar{\lambda}}) = \frac{1}{n}\,\text{Var}\,(X_i) = \frac{1}{n}\,(\frac{1}{\lambda^2} + \frac{1}{12}).$$

Thus from the point of view of minimum variance estimators, we see that the estimators which based on conditional sufficient statistics are better than the direct estimator for $\frac{1}{\lambda}$. But, on the other hand, in order to continue the observations upto obtaining n values all of which greater or equal one, we need to have a random number ν_n of observations. The random variable ν_n has the so–called negative binomial distribution given by:

$$P\{\nu_n = N\} = \begin{cases} \binom{N-1}{N-n} \quad p^n q^{N-n} & \text{if} \quad N \geqslant n \\ \\ 0 & \text{if} \quad N < n \end{cases}$$

Thus

$$E(\nu_n) = \frac{n}{P}, \qquad P = \frac{1}{\lambda}(1 - e^{-\lambda}),$$

which is more greater than n for the large values of λ.

To avoid the needness of such a large number of observations in such cases, we introduce the statistic

$$\frac{1}{\lambda^*} = \frac{1}{n^* + 1}\left[\sum_{i=1}^{n^*} (X_i^* - 1)\right],$$

which depends on the observations X_i^* contained in a sample of fixed size n, as an estimator for $\frac{1}{\lambda}$ and compare the variances of $\frac{1}{\overline{\lambda}}$ and $\frac{1}{\lambda^*}$.

For $\frac{1}{\lambda^*}$ we have:

$$E(\frac{1}{\lambda^*}) = \sum_{k=0}^{n} \int_0^\infty \frac{x}{k+1} \cdot \frac{1}{\Gamma(k)} \binom{n}{k} p^k q^{n-k} \lambda e^{-\lambda x} (\lambda x)^{k-1} dx$$

$$= \frac{1}{\lambda} \sum_{k=0}^{n} \frac{k}{k+1} \binom{n}{k} p^k q^{n-k}.$$

Noting that:

$$\sum_{k=0}^{n} \frac{1}{k+1} \binom{n}{k} p^k q^{n-k} = q^n \sum_{k=0}^{n} \frac{1}{k+1} \binom{n}{k} (\frac{p}{q})^k$$

$$= \frac{q^{n+1}}{p} \cdot \frac{(1 + \frac{p}{q})^{n+1} - 1}{n+1} = \frac{1 - q^{n+1}}{(n+1)p}$$

$$= \frac{1 + q + \ldots + q^n}{n+1}.$$

Then:

$$E(\frac{1}{\lambda^*}) = \frac{1}{\lambda}[1 - \frac{1 + q + \ldots + q^n}{n+1}].$$

Also:

$$E(\frac{1}{\lambda^{*2}}) = \frac{1}{\lambda^2} \sum_{k=0}^{n} \frac{k}{k+1} \binom{n}{k} p^k q^{n-k}.$$

Thus:

$$\text{Var}\,(\frac{1}{\lambda^*}) = E(\frac{1}{\lambda^{*2}}) - E^2(\frac{1}{\lambda^*})$$

$$= \frac{1}{\lambda^2}(1 - \frac{1 + q + \ldots + q^n}{n+1})(\frac{1 + q + \ldots + q^n}{n+1}).$$

Now, for the variance of $\frac{1}{\lambda^*}$, to be less than that of the statistic $\frac{1}{\bar{\lambda}}$ we must have:

$$\frac{1}{(n+1)\lambda^2}(1 - \frac{1 + q + \ldots + q^n}{n+1})(1 + q + \ldots + q^n) < \frac{1}{n\lambda^2}(1 + \frac{\lambda^2}{12})$$

i.e.

(2) $$\frac{n}{n+1}(1 - \frac{1 + q + \ldots + q^n}{n+1})(1 + q + \ldots + q^n) < 1 + \frac{\lambda^2}{12}.$$

Since:

$$1 - \frac{1 + q + \ldots + q^n}{n+1} < \frac{n}{n+1}.$$

Then (2) will be true if:

(3) $$(\frac{n}{n+1})^2 \cdot \frac{1 - q^{n+1}}{p} < 1 + \frac{\lambda^2}{12}.$$

Thus for any value of n, we can say that the variance of $\frac{1}{\lambda^*}$ is less than the variance of $\frac{1}{\bar{\lambda}}$ if:

$$\frac{1}{\lambda}(\frac{\lambda^2 + 12}{12})(1 - e^{-\lambda}) > 1.$$

This is true for the large values of λ (nearly $\lambda \geqslant 11$).

This fact may be shown also as follows:

For the large values of λ we have

$$P = \frac{1}{\lambda}(1 - e^{-\lambda}) \simeq \frac{1}{\lambda}.$$

Thus (3) may be written in the form:

$$\frac{1}{\lambda}(1+\frac{\lambda^2}{12}) > (\frac{n}{n+1})^2[1-(1-\frac{1}{\lambda})^{n+1}].$$

This relation is also true if:

$$\frac{1}{\lambda}(1+\frac{\lambda^2}{12}) \geqslant 1$$

i.e. for large values of λ (nearly $\lambda \geqslant 11$).

This means that for such large values of λ and for any n, we have:

$$\text{(4)} \qquad \operatorname{Var}(\frac{1}{\lambda^*}) < \operatorname{Var}(\frac{1}{\bar{\lambda}}).$$

In fact, the condition $(\lambda \geqslant 11)$ deduced above is more than needed, because for some smaller values of λ, we can determine a value of n such that (4) is true.

For values $\lambda \geqslant 4$ we can carry out the approximation.

$$p = \frac{1}{\lambda}(1-e^{-\lambda}) \simeq \frac{1}{\lambda}.$$

The error caused by using this approximation is equal to $\frac{1}{\lambda}e^{-\lambda}$ which is less than 0.005 for $\lambda = 4$, and decreases when λ increases.

Thus for any given $\lambda \geqslant 4$, the value of n is determined from (3) after substituting $p = \frac{1}{\lambda}$ and $q = 1 - \frac{1}{\lambda}$, i.e. from the relation

$$\text{(5)} \qquad (\frac{n}{n+1})^2\,[1-(1-\frac{1}{\lambda})^{n+1}] < \frac{1}{\lambda}(1+\frac{\lambda^2}{12}).$$

For $\lambda = 4$, the maximum value of n satisfying (5) is $n = 5$, and for $\lambda = 7$ we have $n = 11$.

Also for $\lambda = 10$, relation (5) gives $n = 28$, while (5) is true for any n when $\lambda \geqslant 11$.

From the preceeding discussion, it is clear that the estimators which have smaller variances need a greater number of observations than the number needed for the other estimators with larger variances. Thus the costs of the observations must be taken into consideration when we compare the estimators. This leads us to consider the sum of the costs of the observations needed for each estimator and its variance, as an objective function and compare these objective functions. So the objective functions corresponding to the estimators $\frac{1}{\hat{\lambda}}$, $\frac{1}{\bar{\lambda}}$, $\frac{1}{\lambda^*}$ are:

$$\hat{f}_n = cE_n + \mathrm{Var}(\frac{1}{\hat{\lambda}}) = \frac{cn}{p} + \mathrm{Var}(\frac{1}{\hat{\lambda}}),$$

$$\bar{f}_n = cn + \mathrm{Var}(\frac{1}{\bar{\lambda}}),$$

$$f_n^* = cn + \mathrm{Var}(\frac{1}{\lambda^*}),$$

where c is the cost of each observation.

Now the values of n for which:

(6) $$\hat{f}_n < \bar{f}_n$$

will be determined.

It is to be observed that comparing the functions $\bar{f}_n$ and f_n^* is exactly the same as comparing the variances of the estimators $\frac{1}{\bar{\lambda}}, \frac{1}{\lambda^*}$.

So it is required to determine the values of n which satisfies the relation:

$$\frac{cn}{p} + \frac{1}{n\lambda^2} \leqslant cn + \frac{1}{n}(\frac{1}{\lambda^2} + \frac{1}{12})$$

i.e.

$$n^2 \leqslant \frac{p}{12qc} \cdot$$

This relation has meaning only when:

$$p > 12qc.$$

Acknowledgement

I would like to thank my supervisor Dr. J. Tomkó for suggesting the problem and the help he has given by way of many discussions.

Összefoglaló

Egy konvolució paraméterének becsléséről és annak alkalmazása
a sorban-állási problémák körében

Jacob Eshak Samman

A konvolució paraméterének becslését az ugynevezett "feltételes elegendő statisztika" felhasználásával végezzük. Az érkezési paraméter becslését egy veszteséges sorban-állási rendszerben a feltételes elegendő statisztikára alapozzuk, amelyet itt bemutatunk és összehasonlitunk a direkt becslésekkel.

Резюме

Оценка параметра конволюции, и
применение его в решении проблемы очередей

Йасоб Ешхак Шамаан

Для оценки параметра конволюции применяется так называемая "условно достаточная статистика". Оценка параметров прибытия проведется в одной системе обслуживания с потерями на основе условно достаточной статистики, которое показывается и сравнивается другими непосредственными методами оценки.