Estimation of Parameters of Some Distribution Functions and its Application to Optimization Problem

Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

by

Ms. Satarupa Rath

Under the guidance of

Prof. M. R. Tripathy

Department of Mathematics
National Institute of Technology
Rourkela-769008
India

May 2012

Certificate

This is to certify that the thesis entitled **"Estimation of parameters of some discrete distribution functions and its application to optimization problem"**, which is being submitted by **Satarupa Rath** in the Department of Mathematics, National Institute of Technology, Rourkela, in partial fulfilment for the award of the degree of **Master of Science**, is a record of bonafide review work carried out by her in the Department of Mathematics under my guidance. She has worked as a project student in this Institute for one year. In my opinion the work has reached the standard, fulfilling the requirements of the regulations related to the Master of Science degree. The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree.

Place: NIT Rourkela
Date: May 2012

Dr. M. R. Tripathy
Assistant Professor
Department of Mathematics
NIT Rourkela-769008
India

Acknowledgement

It is a great pleasure and proud privilege to express my deep sense of gratitude to my guide, Prof. M.R.Tripathy. I am grateful to him for his, continuous encouragement and guidance throughout the period of my project work. Without his active guidance it would not have been possible for me to complete the work.

I acknowledge my deep sense of gratitude to the Head of the Department,all faculty members,all nonteaching staff members of the department of mathematics and the authorities of NIT Rourkela.

I also thank to all my friends for their support, co-operation and sincere help in various ways. I express my sincere thank with gratitude to my parents and other family members, for their support, blessings and encouragement. Finally, I bow down before the almighty who has made everything possible.

Place: NIT Rourkela

Date: May 2012

(Satarupa Rath)

Roll No-410MA2102

Abstract

The thesis addresses the study of some basic results used in statistics and estimation of parameters. Here we have presented estimation of parameters for some well known discrete distribution functions. Chapter 1, contains an introduction to estimation theory and motivation. Chapter 2, contains some basic results, definitions, methods of estimation and Chance constraint approach in linear programming problem. In chapter 3, we have estimated the parameters of well known discrete distribution functions by different methods like method of moments, method of maximum likelihood estimation and Bayes estimation. Further in Chapter 4, a Chance Constraint method is discussed which is an application of beta distribution function.

Contents

Chapter 1

Introduction

The theory of estimation was first studied by R. A. Fisher around 1930. Estimation theory traces its origin to the efforts of astronomers many years ago to predict the motion of our solar system. Estimation is a calculated approximation of the result which is given even if the input data is uncertain. This branch of statistic deals with estimating the value of parameter of measured data that has random component. Considering some practical data it is natural that it will follow certain distribution function. We may be interested in the characteristics of the distribution functions. Basically two types of estimation procedure is known. One is Point estimation and another is interval estimation. Here we mainly focus on point estimation. Suppose a random variable X follows $N(\mu, \sigma^2)$, where μ is unknown. let us take a sample $X_1, X_2, \ldots, Xn$ from $X \sim N(\mu, \sigma^2)$. The statistic $T = \sum_{i=1}^{n} x_i$ is the best estimate for μ. The value of T for $x_1, x_2, \ldots, xn$ is the estimate of μ and T is the estimator of μ. For both the unknown parameters.

For estimating the unknown parameters we follow some simple techniques. Specifically, method of moments, method of maximum likelihood estimation, known as classical techniques.

Further we consider a different approach where the unknown parameter is taken as an random variable. This technique is known as Bayesian approach. For this type of study we may consider informative prior or non-informative prior for the unknown parameter. For example: for binomial distribution taking prior as $g(p) = 1$. It is a noninformative prior.In this project work we have discussed some discrete distribution functions and estimation of its parameters.

Further, we have discussed the application of a distribution function to a stochastic linear programming problem. In 1950, stochastic linear programming is used as an application to linear programming problems. Here this is a special kind of linear programming problem in which the coefficients are treated as random variables with having some joint probability distributions.

Chapter 2

Some Definitions and Basic Results

In this chapter some definitions and basic results are given which are very much useful for the development of the consequence chapters. Below we start from a very basic concept known as random experiment or statistical experiment.

Definition 2.1 (Random experiment) *It is an experiment in which we are aware of all the outcomes. But the performance of the experiment is unknown. This experiment can be repeated under identical conditions.*

Definition 2.2 (Sample space) *Sample space of an random experiment is a pair* (Ω, S)*, where* Ω *is the set of all possible outcomes of the experiment and* S *is a* σ *field of a subset of* ω*.*

Definition 2.3 (Event) *A subset of a sample space* Ω *in which a statistician is interested is known as event.*

Definition 2.4 (Probability measure) *Let* (Ω, S) *be a sample space. A set function* P *defined on* S *is called probability if it satisfies the following conditions,*

(*i*) $P(A) \geq 0, \forall\ A \in S.$

(ii) $P(\Omega) = 1$.

(iii)$A_j \in S, j = 1, 2, ...$ *are disjoint sequence of sets.* $j \neq k$ *Then,*

$$P(\sum_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j).$$

Definition 2.5 (Random variable) *Random variable X is a function from sample space Ω to the set of real numbers such that the inverse image of a Borel measurable set in $\mathbb{R}$, under X is an event. That is $X : \Omega \to \mathbb{R}$, such that $X^{-1}(-\infty, a] \in S$, $a \in \mathbb{R}$.*

In this thesis we are suppose to study the discrete distribution function and its parameters, so below we are presenting some results related to this.

Definition 2.6 (Discrete random variable) *An random variable X which is define on (Ω, S, P) is called discrete type, if there exist a countable set $E \subseteq \mathbb{R}$, such that $PX \in E = 1$.*

Definition 2.7 (Cumulative distribution function) *Let X be an random variable defined on (Ω, S, P).A point function $F(.)$ on R defined by $F(x) = Q(-\infty, x] = P\omega : X(\omega) \leq x$ for all $x \in R$.Then the function F is known as the Cumulative distribution function of a random variable X.*

If X is a discrete random variable then we can define its probability mass function as below.

Definition 2.8 (probability mass function) *The collection of numbers p_i which satisfies $PX = x_i = p_i \geq 0$ for all i and $\sum_{i=1}^{\infty} p_i = 1$, is known as probability mass function of a random variable X.*

Definition 2.9 (Two dimensional discrete random variable) *a two dimensional random variable (X, Y) is known as discrete type if it takes on pair of values belonging to a countable set of pairs.*

Definition 2.10 (Joint probability mass function) *Let (X, Y) be a discrete random variable which takes a pair value $(x - i, y_j), i = 1, 2, ...$ and $j = 1, 2, ...$ we call*

$$p_{ij} = PX = x_j, Y = y_j$$

is the joint probability mass function.

Definition 2.11 (Marginal probability mass function) *Let (X, Y) be discrete random variables having distribution function F. Then the marginal distribution function of X is defined as*

$$F_1 = F(x, \infty) = \lim_{y \to \infty} (x, y) = \sum_{x_i \le x} p_i$$

Definition 2.12 (Conditional probability mass function) *Let (X, Y) be discrete random variable.If $PY = y_j > 0$. the function*

$$PX = x_i | Y = y_j = \frac{PX = x_i, Y = y_j}{PY = y_j}$$

for a fixed j, is known as the conditional probability mass function.

Definition 2.13 (Conditional expectation) *Let X and Y be random variables defined on a probability space (Ω, S, P) and let h be a Borel measurable function. The conditional expectation of $h(X)$, given Y, which can be written as $Eh(X)|Y$, is a discrete random variable that takes the value $Eh(X)|y$ defined as,*

$$Eh(X)|y = \sum_x h(x) PX = x | Y = y$$

when Y assumes the value y.

Next we discuss some characteristics of these distribution functions

Definition 2.14 (Moments) *Moments are parameters associated with the distribution of the random variable X. Let k be a positive integer and c be a any constant. If $E(X-c)^k$ exists, it is called the moment of order k about the point c. If we choose $c = E(X) = \mu$, then it is called central moments of order k. I particular if $k = 2$ we get variance.*

Further, we will discuss some terms related to the estimation of parameters of a discrete distribution function.

Definition 2.15 (Parameter space) *Let X be a random variable defined on a sample space Ω having probability mass function $f(x, \theta)$. Here θ is unknown and takes the values on a set called as parameter space Θ.*

Definition 2.16 (Statistic) *Let $X_1, X_2, ..., X_n$ be any sample taken from a distribution. Then any function of these say $T(X_1, X_2, ..., X_n)$ is called a statistic.*

Definition 2.17 (Estimator) *If this statistic is used to estimate an unknown parameter θ of the given distribution then it is known as an estimator.*

Definition 2.18 (Estimate) *A particular value of an estimator is called an estimate of θ.*

In this thesis we will only discuss the problem of point estimation. The process of estimating an unknown parameter is known as estimation.

When we estimate the unknown parameter θ of a distribution function $F_\theta(x)$, by an estimator $\delta(X)$ some loss is incurred. Hence we use some loss functions to know the amount of loss incurred as below.

Definition 2.19 (Linear Loss) *Linear loss is defined as*

$$\begin{aligned} L(\theta, \delta(x)) &= C_1(\delta(x) - \theta), \quad \delta \geq \theta \\ &= C_2(\theta - \delta(x)), \quad \delta < \theta, \end{aligned}$$

where C_1 and C_2 are constants.

Definition 2.20 (Absolute Error loss) *For this loss function the Bayes estimator is the posterior median.The function is defined as $L(\delta(x, \theta)) = |\delta(x) - \theta|$.*

Definition 2.21 (Quadratic Los) *This loss function is defined as $L(\delta(x), \theta) = C(\delta(x) - \theta)^2$.*

Definition 2.22 (Zero-One Loss) *Zero One Loss function is defined as*

$$L(\delta(x), \theta) = 0, \quad if|\delta(x), \theta| \leq k$$

2.1 Characteristics of Estimators

In this section we will discuss some characteristics of an estimators.

2.1.1 Unbiasedness

The estimator $T_n = T(X_1, X_2, ..., X_n)$ is said to be unbiased for θ if $E(T_n) = \theta$ for all values of parameter.

Remark 2.1 *If $E(T_n) > \theta$, T_n is said to be biased. If $E(T_n) < \theta$, T_n is said to be negative biased.*

Let $X_1, X_2, \ldots, X_n$ be identically and independently distributed random samples taken from Bernoulli population and θ be the parameter. Then the statistic $T = \sum_{i=1}^{n} X_i$ follows Binomial(n, θ). Here we can check that $E(T) = n\theta$, so $\frac{T}{n}$ is an unbiased estimate of θ. Further, $E\frac{T(T-1)}{n(n-1)} = \theta^2$, hence we can conclude that $\frac{T(T-1)}{n(n-1)}$ is an unbiased estimate of θ^2.

2.1.2 Consistency

The estimator $T_n = T(X_1, X_2, \ldots, X_n)$ for θ, based on a random sample of size n is said to be consistence if T_n converges to θ in probability.

Theorem 2.1 *Let X_n be a sequence of estimators such that for all $\theta \in \Omega$ $E_\theta(X_n) \to \gamma, n \to \infty$ $Var_\theta(X_n) \to 0, n \to \infty$, Then X_n is a consistent operator of (θ).*

Let $X_1, X_2, \ldots, X_n$ are i.i.d Bernoulli variants having parameter p. and the statistic T follows $B(n, p)$ and the expectation value is $E(T) = np$ and variance is $Var(T) = npq$. Then

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{T}{n}.$$

. such that $E(\bar{X}) = p.Var(\bar{X}) = \frac{pq}{n} \to 0$ as $n \to \infty$ As $E(\bar{X}) \to p$ and $Var(\bar{X}) \to 0$, for $n \to \infty$, then $\bar{X}$ is a consistent estimator of p.

2.1.3 Efficiency

If T_1 is any estimator with variance V_1 and T_2 is the most efficient estimator with variance V_2, then the efficiency E of T_1 is defined as $E = \frac{V_2}{V_1}$ and E can not exceed unity.

Remark 2.2 *In a class of consistent estimators of parameter, if there exists an estimator whose sampling variance is less than other estimators is called most efficient estimator.*

2.1.4 Sufficiency

Let $Y = (X_1, X_2, \ldots, X_n)$ be a sample from $G_\theta : \theta \in \Omega$.A static $T = T(x)$ is sufficient for θ , if and only if the conditional distribution of X, given T which does not depend on θ.

Theorem 2.2 (factorization criteria) *Let $X_1, X_2, \ldots, X_n$ be the discrete random variables with probability mass function $f_\theta(X_1, X_2, \ldots, X_n), \theta \in \Omega$. Then $T(X_1, X_2, \ldots, X_n)$ is sufficient for θ, if and only if*

$$f_\theta(X_1, X_2, \ldots, X_n) = g(X_1, X_2, \ldots, X_n)h_\theta(T(X_1, X_2, \ldots, X_n)).$$

where g is a non negative function of $Y_1, Y_2, \ldots, Y_n$ only and g does not depend on θ. h_θ is a non negative non constant function of θ and $T(X_1, X_2, \ldots, X_n)$ only, θ and $T(X_1, X_2, \ldots, X_n)$ can be multidimensional.

Let $X_1, X_2, \ldots, X_n$ be the i.i.d of $b(1, p)$ and the statistic $T = \sum_{i=1}^n X_i$. Then

$$\begin{aligned} P\{X_1 = x_1, \ldots, X_n = x_n | \sum_{i=1}^n X_i = t\} &= \frac{P\{X_1 = x_1, \ldots, X_n = x_n, T = t\}}{\binom{n}{t} p^t (1-p)^{n-t}}, \quad \text{if} \quad \sum_{i=1}^n X_i = t \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

Thus for $\sum_1^n x_i = t$ we have $P\{X_1 = x_1, \ldots, X_n = x_n | \sum_{i=1}^n X_i = t\} = \frac{1}{\binom{n}{t}}$. Therefore it is sufficient to concentrate on $\sum_1^n X_i$ only.

2.1.5 Completeness

Definition 2.23 *Let $\{g_\theta(x), \theta \in \Omega\}$ be a family of probability mass functions, then the family is complete if $E_\theta h(X) = 0$ for all $\theta \in \Omega$ implies that $P_\theta\{h(X) = 0\} = 1$ for all $\theta \in \Omega$.*

Definition 2.24 *A statistic $T(X)$ is said to be complete if the family of distributions of T is also complete.*

2.2 Methods of Estimation

2.2.1 Method of Moments

In this method we equate few moments of the population to the corresponding moments of the sample. Let θ, $\theta \in \Theta$ be a parameter to be estimated on the basis of random sample from a distribution function F. Then we calculate the sample moments and equate it to the population moments. Solving we will get the moment estimators.

2.2.2 Method of Maximum Likelihood

The main principle of maximum likelihood is that the sample is the representative of the population and taken as the estimator that value of parameter which maximizes the probability mass function $f_\theta(x)$.

Definition 2.25 (Likelihood function) *Let $(X_1, X_2, \ldots, X_n)$ be a random sample with probability mass function $f_\theta(X_1, X_2, \ldots, X_n)$, $\theta \in \Theta$. The function,*

$$L(\theta, x_1, x_2, \ldots, x_n) = f_\theta(x_1, x_2, \ldots, x_n) \tag{2.1}$$

considered as a function of θ, is called the likelihood function, where θ may be multiple parameter.

If $X_1, X_2, \ldots, X_n$ are identically and independently distributed random variable with probability mass function $f_\theta(x)$, then the likelihood function is

$$L(\theta, x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_\theta(x_i). \quad \theta \in \mathbb{R}^m.$$

The main principle of maximum likelihood estimator is to choose an estimator of θ, as $\hat{\theta}(x)$ such that it maximizes $L(\theta; x_1, x_2, \ldots, x_n)$. As we know that log is a monotonically increasing function we consider,

$$\log L(\theta; x_1, x_2, \ldots, x_n).$$

Hence instead of maximizing likelihood function we may maximize loglikelihood function with respect to the parameter. Further as $f_\theta(x)$ is positive, differentiable function of θ, we calculate

$$\frac{\partial \log L(\theta, x_1, x_2, \ldots, x_n)}{\partial \theta_j} = 0, \quad j = 1, 2, ..., m$$

and $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$. Solving these equations for θ we get the maximum likelihood estimators.

Remark 2.3 *Maximum likelihood estimators are always consistent, but may not be unbiased always.*

Theorem 2.3 *If maximum likelihood exists then it is most efficient in the class of such estimators. If a sufficient estimator exists, then it is the function of the maximum likelihood estimator.*

Theorem 2.4 (Invariance Property) *If T is a maximum likelihood estimator of θ and $\mu(\theta)$ is a one-to-one function of θ, then $\mu(T)$ is the maximum likelihood estimator of $\mu(\theta)$.*

2.2.3 Bayesian Estimation

Bayesian estimation is a different approach of probability which comes under decision theory. In this method the parameters are taken as random variables and assumes a distribution function which is known as a priori.

In Bayesian estimation we treat θ as a random variable distributed according to probability mass function $\pi(\theta)$ on Θ. Also π is called the priori distribution. Now $f(x|\theta)$ represents the conditional probability mass function of random variable X, given $\theta \in \Theta$ is held fixed. As π is the distribution of θ, it follows that the joint probability mass function of θ and X is given by

$$f(x, \theta) = \pi(\theta) f(x|\theta).$$

Here $R(\theta, \delta)$ is the conditional average loss, defined by $EL(\theta, \delta(x))|\theta$ given that θ is held fixed.

Definition 2.26 *The Bayes risk of an estimator δ is defined by*

$$R(\pi, \delta) = E_\pi R(\theta, \delta).$$

.

If θ is a continuous random variable and X is of continuous type then,

$$\begin{aligned} R(\pi, \delta) &= \int R(\theta, \delta)\pi(\theta)d\theta \\ &= \int\int L(\theta, \delta(x))f(x, \theta)dxd\theta. \end{aligned}$$

If θ is a discrete random variable and X is discrete type then,

$$R(\pi, \delta) = \sum_\theta \sum_x L(\theta, \delta(x))f(x, \theta).$$

Definition 2.27 *An estimator δ^* is known as a Bayes estimator if it minimizes the Bayes risk that is if $R(\pi, \delta^*) = \inf_\delta R(\pi, \delta)$.*

Definition 2.28 *The conditional distribution of random variable θ, given $X = x$, is called a posteriori probability distribution of θ, given the sample.*

Theorem 2.5 *Consider the problem of estimating a parameter $\theta \in \Theta \subseteq \mathbb{R}$ with respect to the quadratic loss function $L(\theta, \delta) = (\theta - \delta)^2$. A Bayes estimator is given by*

$$\delta(x) = E\theta \mid X = x.$$

2.3 Stochastic Linear Programming

In the stochastic linear programming, all the parameters of the problem that is the coefficients of the objective functions, the coefficients involve in inequalities are random. Hence all the parameters are treated as random variables.

The stochastic linear programming can be stated as

$$\min f(x) = C^T X = \sum_{j=1}^{n} c_j x_j$$

subject to

$$A_i^T X = \sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, i = 1, 2, \ldots, n, x_j \geqslant 0, j = 1, 2, \ldots, m,$$

Where a_{ij},c_j and b_j are random variables with known probability distribution, x_j assumed to be deterministic. The generic way of expressing chance constraint inequality is as follows.

$$P\{h(x, \xi) > 0\} \geqslant P$$

where x is decision, ξ random vector and P is the probability measure. We have taken $h(x, \xi) \geqslant 0$ is finite system of inequality. The classical linear programming problem is given as

$$\max Z(x) = \sum_{j=1}^{m} c_j x_j$$

$$\sum_{j=1}^{m} a_{ij} x_j \leqslant b_i, i = 1, 2, \ldots, n$$

$x_j \geq 0, j = 1, 2, \ldots, m$, coefficients are deterministic. Chance constraint technique is used to solve this problem.

Chance constraint programming deals with the random parameters in the optimization problems which is mainly used in engineering and finance sectors. The uncertainty arises due to uncertain state estimation as well as stochastic mode transition. This method is proposed by Chanes and Cooper which offers a powerful modelling stochastic decision and control system. The method mainly concerned with the problem that the decision makers must give the solution before random variables come true. The main difficulty of such model is due to optimal decision that to be taken prior to the observations of random parameter. The stochastic linear programming problem having chance constraint is formulated as

$$\min(\max)Z(x) = \sum_{j=1}^{m} c_j x_j$$

such that

$$P\{\sum_{j=1}^{m} a_{ij}x_j \leq b_i\} \geq 1 - u_i,$$

and $x_j \geq 0$, $j = 1, 2, \ldots, m$, $u_i \in (0, 1)$, $i = 1, 2, \ldots, n$.

Here a_{ij}, c_j and b_j are random variables. u_i's are probability taken. The k^{th} Chance constraint can be obtained as

$$P\{\sum_{j=1}^{m} a_{kj}x_j \leq b_k\} \geq 1 - u_k,$$

having lower bound $(1 - u_k)$ and x_j are deterministic. Where a_{kj}, c_j and b_k are random variables having known means and variance. If b_k is the random variable and F_a is the distribution function then the deterministic chance constraint is stated as

$$Pa_{kj}x_j \leqslant b_k \geqslant u_k.$$

If and only if

$$P\{b_k \geqslant a_{kj}x_j\} \geqslant u_k.$$

$$\Leftrightarrow 1 - F_a(a_{kj}x_j) \geqslant u_k.$$

$$\Leftrightarrow a_{kj}x_j < F_a^{-1}(1 - u_k).$$

Let us assume a_k be the random variable with normal distribution with the mean $E(a_{kj})$ and the variance $Var(a_{kj})$ and the covariance is zero between a_{kj} and a_{kl}.the random variable d_k is defined as $d_k = \sum_{j=1}^{n} a_{kj}x_j$ where a_{ki} are random variables with normal distribution and x_i are unknowns.the chance constraint gives the inequality as

$$\phi\{\frac{b_k - E(d_k)}{\sqrt{Var(d_k)}}\} \geqslant \phi(K_{uk}) \tag{2.2}$$

where K_{uk} are standard normal variable. Here

$$\phi(K_{uk}) = 1 - u_k.$$

The deterministic equivalent is stated as

$$E(d_k) + K_{uk}\sqrt{Var(d_k)} \leqslant b_k \tag{2.3}$$

Chapter 3

Estimation of Parameters of Some Well Known Discrete Distributions

In this chapter we consider the problem of estimation of parameters of different discrete distribution functions. First we consider the estimation problem for Binomial distribution.

3.1 Binomial Distribution Function

In this section the problem of estimation of the parameter of Binomial distribution considered. Basically the method of moments, maximum likelihood and the Bayes methods are considered.

3.1.1 Method of Moments Estimators

The probability mass function of a random variable X which follows Binomial distribution is given by,

$$f(x,p) = \binom{n}{x} p^x q^{n-x}. \tag{3.1}$$

We have to calculate the sample moments. The first moment is calculated as

$$\begin{aligned} E(X) &= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} \\ &= np. \end{aligned} \tag{3.2}$$

The method of moment estimator for p is

$$\hat{p} = \frac{X}{n}. \tag{3.3}$$

Here we assume that n is known.

3.1.2 Method of Maximum Likelihood

Let $(X_1, X_2, \ldots, X_n)$ identically and independently distributed random sample from Binomial(n, p). The joint probability mass function of these is given by

$$\begin{aligned} L(\underline{\mathrm{x}}, p) &= \prod_{i=1}^{n} f(x_i) \\ &= \prod_{i=1}^{n} \binom{n}{x_i} p^{\sum x_i} (1-p)^{n-\sum x_i}. \end{aligned}$$

Which is the likelihood function.

By taking log on both sides of the above equation we get,

$$\log L(\underline{\mathrm{x}}, p) = \log\{\prod_{i=1}^{n} \binom{n}{x_i} p^{\sum x_i} (1-p)^{n-\sum x_i}\}. \tag{3.4}$$

Now differentiating with respect to p and equating to 0 we have,

$$\frac{\partial}{\partial p}(\log L(p)) = \frac{\partial}{\partial p} \log \prod_{i=1}^{n} \binom{n}{x_i} p^{\sum x_i} (1-p)^{n-\sum x_i} = 0.$$

Which gives

$$\hat{p} = \sum_{i=1}^{n} X_i = \bar{X}.$$

Further we consider two Binomial distribution functions. Let $X \sim B(n, p_1)$ and $Y \sim B(n, p_2)$. We are interested to estimate the parameter $p_1 + p_2$.

Let $(X_1, X_2, \ldots, X_n) \sim B(n, p_1)$ and $Y_1, Y_2, \ldots, Y_n \sim B(n, p_2)$.

By using above maximum likelihood estimators for both p_1 and p_2 we can get the estimator of $p_1 + p_2$. Thus the estimator for p_1 is

$$\hat{p_1} = \bar{X} \tag{3.5}$$

and the estimator for p_2 is

$$\hat{p_2} = \bar{Y}. \tag{3.6}$$

Hence, the estimator for $p_1 + p_2$ can be obtained by adding above two estimator. This is possible because of the invariance property of maximum likelihood estimator. Therefor the estimator for $p_1 + p_2$ is given by

$$\hat{p_3} = \bar{X} + \bar{Y}.$$

Similarly we can estimate $\frac{p_1}{p_2}$ by

$$\hat{p_4} = \bar{X}/\bar{Y}. \tag{3.7}$$

.

3.1.3 Bayesian Estimation

The distribution for binomial distribution is given as

$$f(x) = \binom{n}{x} p^x q^{n-x}.$$

Let $X_1, X_2, \ldots, X_n$ be identically and independently distributed random variables taken from $X \sim B(n,p)$. Then the likelihood function is given as

$$\begin{aligned} L(\underline{x} \mid p) &= \prod_{i=1}^{n} \binom{n}{x} p^{x_i} q^{n-x_i} \\ &= p^s q^{nn-s} \prod_{i=1}^{n} \binom{n}{x_i}, \end{aligned}$$

where $\sum_{i=1}^{n} x_i = s$.

Consider the prior,

$$g(p) \sim p^{a-1}(1-p)^{b-1}, a > 0, b > 0$$

$$f(p \mid x) = \frac{f(p,x)}{\int_0^1 f(x,p)dp}. \tag{3.8}$$

The joint probability mass function is defined by

$$f(p,x) = \binom{n}{x} p^x q^{n-x} p^{a-1}(1-p)^{b-1}. \tag{3.9}$$

Marginal density is given

$$\begin{aligned} f(x) &= \int_0^1 f(p,x)dp \\ &= \int_0^1 Cp^{s+a-1}q^{nn-s+b-1}dp \\ &= C\int_0^1 p^{s+a-1}q^{nn-s+b-1}dp \\ &= B(s+a, nn+b-s). \end{aligned} \tag{3.10}$$

The conditional probability mass function is give as

$$f(p \mid x) = \frac{p^{s+a-1}(1-p)nn+b-s+1}{B(s+a, nn+b-s)} \tag{3.11}$$

Hence the Bayes estimator is obtained as

$$E(p \mid x) = \sum_{p=0}^{n} p \cdot \frac{p^{s+a-1}(1-p)^{nn+b-s+1}}{B(s+a, nn+b-s)} \tag{3.12}$$

$$= \frac{s+a}{nn+a+b}. \tag{3.13}$$

So the Bayes estimator is obtained as,

$$\hat{p} = \frac{s + a}{nn + a + b}. \tag{3.14}$$

For prior $g(p) \sim \frac{1}{\sqrt{p}}$ the joint density function is given as

$$f(p, x) = \binom{n}{x} p^x q^{n-x} \frac{1}{\sqrt{p}} = \binom{n}{x} p^{x-\frac{1}{2}} (1-p)^{n-x}. \tag{3.15}$$

The marginal density is given by

$$\begin{aligned} f(x) &= \int_0^1 f(x, p) dp \\ &= \int_0^1 \binom{n}{x} p^{x-\frac{1}{2}} (1-p)^{n-x} dp \\ &= \binom{n}{x} \int_0^1 p^{x-\frac{1}{2}-1} (1-p)^{n-x+1-1} dp \\ &= \binom{n}{x} B(x + \frac{1}{2}, n - x + 1). \end{aligned} \tag{3.16}$$

The conditional probability mass function is given by

$$\begin{aligned} f(p \mid x) &= \frac{f(p, x)}{\int_0^1 f(x, p) dp} \\ &= \frac{\binom{n}{x} p^{x-\frac{1}{2}} (1-p)^{n-x}}{\binom{n}{x} B(x + \frac{1}{2}, n - x + 1)} \\ &= \frac{p^{x+\frac{1}{2}-1} (1-p)^{n-x+1-1}}{B(x + \frac{1}{2}, n - x + 1)}. \end{aligned} \tag{3.17}$$

So the Bayes estimator is given by

$$\begin{aligned} E(p \mid x) &= \int_0^1 p \cdot \frac{p^{x+\frac{1}{2}-1} (1-p)^{n-x+1-1}}{b(x + \frac{1}{2}, n - x + 1)} \\ &= \frac{x + \frac{1}{2}}{n + \frac{3}{2}}. \end{aligned} \tag{3.18}$$

So the estimator

$$\hat{p} = \frac{x + \frac{1}{2}}{n + \frac{3}{2}}. \tag{3.19}$$

For noninformative prior that is

$$g(p) \sim f(p) = 1. \tag{3.20}$$

The probability mass function of $X \sim B(n, p)$ is given as

$$f(x) = \binom{n}{x} p^x q^{n-x}. \tag{3.21}$$

Joint probability mass function is given as,

$$f(p, x) = \binom{n}{x} p^x q^{n-x} \cdot p. \tag{3.22}$$

The marginal probability mass function is given by

$$\begin{aligned} f(x) &= \int_0^1 f(x,p)dp \\ &= \binom{n}{x} B(x+1, n-x+1). \end{aligned} \tag{3.23}$$

The conditional density is given as

$$\begin{aligned} f(p \mid x) &= \frac{f(p,x)}{\int_0^1 f(x,p)dp} \\ &= \frac{\binom{n}{x} p^x (1-p)^{n-x}}{\binom{n}{x} B(x+1, n-x+1)} \\ &= \frac{p^{x+1-1}(1-p)^{n-x+1-1}}{B(x+1, n-x+1)}. \end{aligned} \tag{3.24}$$

Then the Bayes estimation is obtained as

$$\begin{aligned} E(p \mid x) &= \int_0^1 p \cdot \frac{p^{x+1-1}(1-p)^{n-x+1-1}}{b(x+1, n-x+1)} \\ &= \frac{x+2}{n-x+1+x+2} = \frac{x+2}{n+3}. \end{aligned} \tag{3.25}$$

So the Bayes estimator is given by

$$\hat{p} = \frac{x+2}{n+3}. \tag{3.26}$$

Further we consider the prior $g(p) \sim B(\alpha, \beta)$. The joint probability mass function is defined as

$$f(x,p) = \frac{1}{B(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1}. \tag{3.27}$$

Then the marginal density is given by

$$\begin{aligned} f(x) = \int_0^1 f(x,p)dp \\ &= \frac{1}{B(\alpha,\beta)} B(\alpha,\beta). \end{aligned} \tag{3.28}$$

The the conditional density is given by

$$\begin{aligned} f(p \mid x) &= \frac{f(p,x)}{\int_0^1 f(x,p)dp} \\ &= \frac{\frac{1}{B(\alpha,\beta)} p^{\alpha-1}(1-p)\beta - 1}{\frac{1}{B(\alpha,\beta)} B(\alpha,\beta)} \\ &= \frac{p^{\alpha-1}(1-p)\beta - 1}{B(\alpha,\beta)}. \end{aligned} \tag{3.29}$$

The Bayes estimation is obtained as follows

$$\begin{aligned} \hat{p} &= E(p \mid x) \\ &= \int_0^1 p \cdot \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} \\ &= \int_0^1 \frac{p^{\alpha+1-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} \\ &= \frac{\alpha+1}{\alpha+1+\beta}. \end{aligned} \tag{3.30}$$

3.2 Poisson Distribution Function

In this section we consider a poison distribution function with unknown parameter λ. The problem is to estimate the parameter λ.

3.2.1 Method of Moments

Let $(X_1, X_2, \ldots, X_n)$ identically and independently distributed random sample from Poisson(λ). First moment is given as

$$\mu_1' = E(X) = \lambda. \tag{3.31}$$

The sample moment is given by $\sum_{i=1}^{n} X_i$. So, the estimator for $\lambda = \bar{X}$.

3.2.2 Method of Maximum Likelihood

The Probability mass function for Poisson distribution is given as

$$f(x, \lambda) = e^{-\lambda}\frac{\lambda^x}{x!}$$

for $x = 0, 1, \ldots$.

The likelihood function is given as

$$L = \prod_{i=1}^{n} f(x_i, \lambda) \quad = \quad e^{-n\lambda}\frac{\lambda^{\sum x_i}}{x_1!, x_2!, ..., x_n!}. \tag{3.32}$$

By taking log on both sides of above equation we have,

$$\begin{aligned} \log L &= -n\lambda + (\sum_{i=1}^{n} x_i)\log\lambda - \sum_{i=1}^{n}\log(x_i!) \\ &= -n\lambda + n\bar{x}\log\lambda - \sum_{i=1}^{n}\log(x_i!). \end{aligned} \tag{3.33}$$

Then the likelihood equation for λ is

$$\frac{\partial}{\partial\lambda}\log\lambda = 0.$$

$$\Rightarrow -n\lambda + \frac{n\bar{x}}{\lambda} = 0.$$

$$\frac{n\bar{x}}{\lambda} = x \Rightarrow \lambda = \bar{x}. \tag{3.34}$$

The variance of estimate is calculated by taking expectation over the likelihood equation.

$$\frac{1}{V(\hat{\lambda})} = \frac{n}{\lambda^2}E(\bar{x}) \quad = \quad \frac{n}{\lambda^2}\cdot\lambda = \frac{n}{\lambda}. \tag{3.35}$$

Further we consider the problem of estimating the parameter λ^α. If we want to estimate for λ^α then by using the invariance property of maximum likelihood method we can the result as follows. The maximum likelihood estimator of λ^α is given by $\bar{X}^\alpha$. Similarly, for $c\lambda$ the estimator is $c\bar{X}$.

3.2.3 Bayesian Estimation

The probability mass function for random variable which follows g Poisson(λ) is given by

$$f(x,\lambda) = \exp(-\lambda)\frac{\lambda^x}{x!}.$$

The prior is taken as

$$g(\lambda \mid a, b) = \exp(-b\lambda)\frac{b^a}{\Gamma(a)}\lambda^{a-1}. \tag{3.36}$$

Then the likelihood function is defined as

$$L(\lambda \mid \bar{x}) = \frac{\exp(-n\lambda)\lambda^{\sum_{i=1}^{n} x_i}}{x_1!, x_2!, \ldots.., x_n!}. \tag{3.37}$$

Let $\sum_{i=1}^{n} x_i = s$. Then the joint density is given as

$$\begin{aligned} f(x,\lambda) &= \frac{b^a e^{-\lambda(n+b)}\lambda^{s+a-1}}{\Gamma(a)x_1!, x_2!, ..., x_n!} \\ &= Ce^{-\lambda(n+b)}\lambda^{s+a-1}. \end{aligned} \tag{3.38}$$

Since

$$\Gamma(\alpha) = \int e^{-x}x^{\alpha-1}dx. \tag{3.39}$$

The marginal density can be obtained as,

$$\begin{aligned} f(x) &= \int_0^\infty f(x,\lambda) \\ &= C\int_0^\infty e^{-\lambda(n+b)}\lambda^{s+a-1}d\lambda \\ &= \frac{C\Gamma(s+a)}{(n+b)^{s+a-1}}. \end{aligned} \tag{3.40}$$

The conditional density is obtained by

$$\begin{aligned} f(\lambda \mid x) &= \frac{Ce^{-\lambda(n+b)}\lambda^{s+a-1}d\lambda}{\frac{C\Gamma(s+a)}{(n+b)^{s+a-1}}} \\ &= \frac{e^{-\lambda(n+b)}\lambda^{s+a-1}(n+b)^{s+a-1}}{\Gamma(s+a)}. \end{aligned} \tag{3.41}$$

The Bayes estimator can be calculated as,

$$\begin{aligned} E(\lambda,x) &= \int_0^\infty \frac{\lambda\cdot e^{-\lambda(n+b)}\lambda^{s+a-1}(n+b)^{s+a-1}}{\Gamma(s+a)}d\lambda \\ &= \frac{(n+b)^{s+a-1}}{\Gamma(s+a)}\frac{1}{(n+b)^{s+a}}\int_0^\infty e^{-\lambda(n+b)}(\lambda(n+b))^{s+a+1-1}d\lambda \\ &= \frac{(n+b)^{-1}}{\Gamma(s+a)}\Gamma(s+a+1) \\ &= \frac{s+a}{n+b}. \end{aligned} \tag{3.42}$$

For prior $g(\lambda) = 1$ that is for noninformative prior. The probability mass function is given as

$$f(x, \lambda) = \exp(-\lambda)\frac{\lambda^x}{x!}. \tag{3.43}$$

The likelihood function is

$$L(\lambda \mid \bar{x}) = \frac{\exp(-n\lambda)\lambda^{\sum_{i=1}^{n} x_i}}{x_1!, x_2!, \ldots.., x_n!}. \tag{3.44}$$

The joint density is given as

$$f(x, \lambda) = \frac{e^{-n\lambda}\lambda^s}{x_1!, x_2!, \ldots.., x_n!}. \tag{3.45}$$

Marginal density can be calculated as

$$\begin{aligned} f(x) &= \int_0^\infty f(x, \lambda) d\lambda \\ &= \int_0^\infty \frac{e^{-n\lambda}\lambda^s}{x_1!, x_2!, \ldots, x_n!} d\lambda \\ &= \frac{\Gamma(s+1)}{n^s C}. \end{aligned} \tag{3.46}$$

The posterior distribution is calculated as

$$\frac{e^{-n\lambda}\lambda^s n^s}{\Gamma(s+1)}.$$

The Bayes estimator $\hat{\lambda}$ is

$$\begin{aligned} E(\lambda \mid x) &= \int_0^\infty \frac{\lambda \cdot e^{-n\lambda}\lambda^s n^s}{\Gamma(s+1)} \\ &= \int_0^\infty \frac{e^{-n\lambda}\lambda^{s+1} n^s}{n\Gamma(s+1)} \\ &= \frac{\Gamma(s+2)}{n\Gamma(s+1)} \\ &= \frac{s+1}{n}. \end{aligned} \tag{3.47}$$

3.3 Geometric Distribution

The probability mass function of a random variable X which follows geometric distribution is given by $f(x,p) = (1-p)^{k-1}p$. Here we assume that the parameter p is unknown and we will try to estimate with the help of samples.

The method of moments estimator for geometric distribution is obtained as $\hat{p} = \frac{1}{X}$.

3.3.1 Bayesian Estimation

The probability mass function of X is given as

$$f(x \mid \theta) = \theta^x(1-\theta), 0 < \theta < 1. \tag{3.48}$$

The prior distribution for p is taken as $g(\theta \mid a) = a\theta^{a-1}$. The joint density is given as $f(x,\theta) = \theta^x(1-\theta)a\theta^{a-1}$. The marginal density is given as

$$\begin{aligned} f(x) &= \int_0^1 f(x,\theta)d\theta \\ &= \int_0^\infty \theta^{x+a-1}(1-\theta)d\theta \\ &= \int_0^1 \theta^{x+a-1}(1-\theta)^{2-1}d\theta. \end{aligned} \tag{3.49}$$

The conditional density can be obtained by

$$\begin{aligned} f(\theta \mid x) &= \frac{f(x,\theta)}{\int_0^1 f(x,\theta)d\theta} \\ &= \frac{\theta^{x+a-1}(1-\theta)}{\int_0^1 \theta^{x+a-1}(1-\theta)^{2-1}d\theta} \\ &= \frac{\theta^{x+a-1}(1-\theta)}{B(x+a,2)}. \end{aligned} \tag{3.50}$$

Then the Bayes estimator for θ is

$$\begin{aligned}
\hat{\theta} = E(\theta \mid x) & \\
&= \int_0^1 \frac{\theta \cdot \theta^{x+a-1}(1-\theta)d\theta}{B(a+x,2)} \\
&= \int_0^1 \frac{\theta^{x++1a-1}(1-\theta)^{2-1}d\theta}{B(a+x,2)} \\
&= \frac{B(a+x+1,2)}{B(a+x,2)}. \qquad (3.51)
\end{aligned}$$

If we want to estimate for θ^k then

$$\begin{aligned}
E(\theta^k \mid x) &= \frac{\int_0^1 \theta^k \theta^{x+a-1}(1-\theta)d\theta}{B(a+x,2)} \\
&= \frac{\int_0^1 \theta^{x+a+k-1}(1-\theta)^{2-1}d\theta}{B(a+x,2)} \\
&= \frac{B(x+a+k,2)}{B(a+x,2)}. \qquad (3.52)
\end{aligned}$$

Chapter 4

Application Of distribution Function

In this chapter we consider a stochastic linear programming problem. By using Chance constraint approach we have derived the deterministic model when the coefficients follow a beta distribution.

4.1 Stochastic Linear Programming

The stochastic linear programming model is considered in Chapter 2. We refer the same model for present discussion.

Let $X_1, X_2, ..., X_n$ be the random variables having $EX_j = 0$ and $E|X_j|^3 < \infty$ and $j = 1, 2, \ldots, n$, where

$$\begin{aligned} \sigma^2 = EX_j^2, ..., B_n &= \sum_{j=1}^{n} \sigma_j^2, ..., F_n(x) \\ &= PB_n^{\frac{1}{2}} \sum_{j=1}^{n} X_j < x\}, ..., L_n \\ &= B_n^{\frac{-3}{2}} \sum_{j=1}^{n} E|X_j|^3. \end{aligned}$$

Then

$$\underset{x}{Sup}|F_n(x) - \Phi(x)| \leq SL_n. \tag{4.1}$$

Then the above statement can be proved as follows. The above inequality for large value of n can be defined as

$$P[B_n^{\frac{-1}{2}}\sum_{j=1}^{n} X_j - E(\sum_{j=1}^{n} X_j) < x] = \Phi(x) + \frac{\sum_{j=1}^{n} E(X_j - E(X_j))^3 e^{\frac{-x^2}{2}}(1-x^2)}{6\sqrt{2\pi}B_n^{\frac{3}{2}}} + O(n^{\frac{-1}{2}}) \tag{4.2}$$

By using equation (4.2) we can explain the beta distribution for chance constraint method.I n linear programming the constraints are as follows:

$$Ax \leq b \text{ if and only if } \begin{bmatrix} a_{11}a_{12}.....a_{1n} \\ \\ a_{xj}a_{k2}.....a_{kn} \\ \\ a_{m1}a_{m2}.....a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ x_k \\ . \\ x_n \end{bmatrix} \leqslant \begin{bmatrix} b_1 \\ . \\ b_k \\ . \\ b_m \end{bmatrix} \tag{4.3}$$

Matrix A represent the coefficient matrix. Suppose $d_k = a_k x, k = 1, 2, ..., m$ then the k^{th} row of equation [4.3] can be written as $d_k \leq b_k$ if and only if

$$[a_{k1}.a_{k2}, ..., a_{kn}]. \begin{bmatrix} x_1 \\ . \\ x_k \\ . \\ x_n \end{bmatrix} \leq b_k \tag{4.4}$$

where a_{kj}'s are k^{th} row of matrix A which are independent beta random variable.Then the chance constraint method define in CHAPTER-2 is stated as

$$P(d_k \leq b_k) \geq 1 - u_k, k = 1, \ldots, m. \tag{4.5}$$

We have assumed each random variable a_{kj} has beta distribution having parameters $(\alpha_{kj}, \beta_{kj})$. By statement given in equation (4.1) the random variable $r_j = a_{kj}x_j - E(a_{kj}x_j, j = 1, 2, ..., n)$ are considered.

The expected value and variance of each random variable a_{kj} is

$$E(a_{kj} = \alpha_{kj}\beta_{kj})Var(a_{kj}) = \alpha_{kj}\beta_{kj}^2.$$

Hence the expected value and variance of random variable r_j are given as,

$$E(r_j) = E(d_j^2 - E(a_{kj}x_j)) = x_j[\alpha_{kj}\beta_{kj} - \alpha_{kj}\beta_{kj}] = 0.$$

$$Var(r_j) = E(d_j)^2 - [E(d_j)]^2 = x_j^2 Var(a_{kj}) = x_j^2 \alpha_{kj}\beta_{kj}^2.$$

Then the absolute third moment of d_j can be obtained as

$$E|r_j|^3 = E|a_{kj} - E(a_{kj}x_j)|^3 = x_j^3 E|a_{kj} - \alpha_{kj}\beta_{kj}|^3. \tag{4.6}$$

The expected value in (4.6) can be written as

$$\begin{aligned} E|a_{kj} - \alpha_{kj}\beta_{kj}|^3 &= \int_0^\infty |a_{kj} - \alpha_{kj}\beta_{kj}|^3 f(a_{kj}) da_{kj} \\ &= \int_0^{\alpha_{kj}\beta_{kj}} |a_{kj} - \alpha_{kj}\beta_{kj}|^3 f(a_{kj}) da_{kj} \\ &\quad + \int_{\alpha_{kj}\beta_{kj}}^\infty |a_{kj} - \alpha_{kj}\beta_{kj}|^3 f(a_{kj}) da_{kj} \\ &= I_{xj} + II_{xj}. \end{aligned}$$

Where

$$I_{xj} = \int_0^{\alpha_{kj}\beta_{kj}} -|a_{kj} - \alpha_{kj}\beta_{kj}|^3 f(a_{kj}) da_{kj}.$$

$$= \frac{1}{B(\beta_{kj}\alpha_{kj})} \int_0^{\alpha_{kj}\beta_{kj}} (a_{kj}^3 - 3a_{kj}^2\alpha_{kj}\beta_{kj} + 3a_{kj}\alpha_{kj}^2\beta_{kj}^2 - \alpha_{kj}^3\beta_{kj}^3) a_{kj}^{\beta_{kj}-1}(1 - a_{kj}^{\alpha_{kj}-1}).$$

Take $\frac{1}{B(\beta_{kj}\alpha_{kj})} = K$, then the integral can be written as

$$\begin{aligned} I_{xj} = K \int_0^{\alpha_{kj}\beta_{kj}} & a_{kj}^{\beta_{kj}+2}(1-a_{kj})^{\alpha_{kj}-1} da_{kj} \\ -K(\alpha_{kj}^3\beta_{kj}^3) \int_0^{\alpha_{kj}\beta_{kj}} & a_{kj}^{\beta_{kj}-1}(1-a_{kj})^{\alpha_{kj}-1} da_{kj} \\ -K3(\alpha_{kj}\beta_{kj}) \int_0^{\alpha_{kj}\beta_{kj}} & a_{kj}^{\beta_{kj}+1}(1-a_{kj})^{\alpha_{kj}-1} da_{kj} \\ +K3(\alpha_{kj}^2\beta_{kj}^2) \int_0^{\alpha_{kj}\beta_{kj}} & a_{kj}^{\beta_{kj}}(1-a_{kj})^{\alpha_{kj}-1} da_{kj}. \end{aligned}$$

$$= KM_1 + K(3\alpha_{kj}\beta_{kj})M_2 + K(3\alpha_{kj}^2)M_3 + K(\alpha_{kj}^3\beta_{kj}^3)M_4. \tag{4.7}$$

This integrals can be evaluated to get the d_k as stated in Chapter 2. Further modifications can be done and converting to standard normal random variable we may formulate the deterministic model. The rest of the research work will be carried out as my future work.

Conclusions and Scope of Future Works

In this project work I have studied a new area namely estimation of parameters of different distribution function. Particularly Binomial(n, p), Poisson (λ) and Geometric (μ) distribution are discussed. I have learned different techniques for estimating of parameters. Finally, an application of distribution functions in optimization problems by chance constraint method is briefly studied.

The work of the thesis can be extended in different directions.Some of the future work are discussed bellow. Numerical methods can be adopted to find the closed form of some estimators. Risk of different estimators can be compared. A real world problem can be considered to apply our method in chance constraint problem.

Bibliography

Johnson, N.L., Kotz, S. and Balakrishnan, N.(1992): Univariate Discrete Distributions, Wiley-interscience, New York.

Lehman, E.L. and Casella, G.(2008): Theory of Point Estimation, Springer, New York.

Miller, I. and Miller, M.(2003): John E.Freund's Mathematical Statistics,Prentice-Hall of India, New Delhi.

Rohatgi, V. K. and Saleh, A. K. Md. E.(2009): An Introduction to Probability and Statistics, Wiley-interscience, New York.

Sinha, S. K.(2004): Bayesian Estimation, New Age International, New Delhi.

Tripathy, M. R.(2009): Estimation of Parameters Under Equality Restriction, Ph.d.Thesis, IIT, Kharagpur.

Atalay, K. D. and Apaydin, A.(2011): Gamma distribution approach in chanceconstrained stochastic programming model.*Atalay and Apaydin Journal of Inequalities and Applications*, 2011:108.