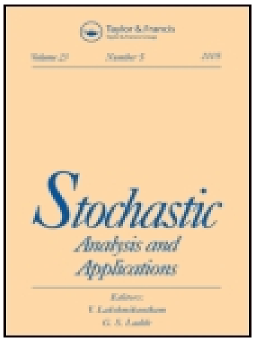


Stochastic Analysis and Applications

ISSN: 0736-2994 (Print) 1532-9356 (Online) Journal homepage: www.tandfonline.com/journals/lsaa20

An Application of Stochastic Programming with Weibull Distribution–Cluster Based Optimum Allocation of Recruitment in Manpower Planning

M. Jeeva, Rajalakshmi Rajagopal, V. Charles & V.S.S. Yadavalli

To cite this article: M. Jeeva, Rajalakshmi Rajagopal, V. Charles & V.S.S. Yadavalli (2004) An Application of Stochastic Programming with Weibull Distribution–Cluster Based Optimum Allocation of Recruitment in Manpower Planning, Stochastic Analysis and Applications, 22:3, 801-812, DOI: 10.1081/SAP-120030457

To link to this article: https://doi.org/10.1081/SAP-120030457

Published online: 15 Feb 2007.

STOCHASTIC ANALYSIS AND APPLICATIONS
Vol. 22, No. 3, pp. 801–812, 2004

An Application of Stochastic Programming with Weibull Distribution–Cluster Based Optimum Allocation of Recruitment in Manpower Planning

M. Jeeva,[1] Rajalakshmi Rajagopal,[2] V. Charles,[3] and V. S. S. Yadavalli[4,*]

[1]Department of Mathematics, The Ethiraj College for Women, Chennai, India
[2]Department of Mathematics, Loyola College, Chennai, India
[3]Department of Mathematics and Humanities, National Institute of Technology, A.P. Warangal, India
[4]Department of Industrial and Systems Engineering, University of Pretoria, Pretoria, South Africa

ABSTRACT

Recruitment of persons for various assignments with required talents in an organization is an important feature, since it plays a vital role in the growth of the organization. To achieve the required expertise in

*Correspondence: V. S. S. Yadavalli, Department of Industrial and Systems Engineering, University of Pretoria, Lynwood Road, Pretoria 0002, South Africa; Fax: 27-12-3625103; E-mail: yadavalli@postino.up.ac.za.

DOI: 10.1081/SAP-120030457 0736-2994 (Print); 1532-9356 (Online)

recruitment, in this paper Linear Stochastic Programming (LSP) is applied along with cluster analysis technique. The aim of this paper is to obtain an optimal allocation of persons to different jobs, so that the time taken to complete all the jobs is minimum. The time taken for a person to complete a job is assumed to follow Weibull distribution. The parameters of Weibull distribution is obtained through Maximum Likelihood Estimator (MLE) approach, along with Cohen's iterative process.

Key Words: Cluster analysis; Linear stochastic programming; Weibull distribution; Sequential programming; Nonlinear probabilistic constraints; Maximum likelihood estimators.

1. INTRODUCTION

Consider an organization with n jobs for which persons are recruited from a large population of applicants using Cluster analysis technique[5] based on some similarity conditions. We are interested in obtaining the Optimal number of persons to be selected from various clusters to various jobs so as to minimize the time required to complete all jobs, that is to maximize the overall efficiency of persons selected. Here the time taken by each person to complete a given job is assumed to be random in nature and at least one of the constraints involved in this problem are probabilistic. Hence the problem under consideration is a Linear Stochastic programming problem (LSPP), where the parameters follow certain empirical distributions. In our case it is assumed to follow Weibull distribution.

Section 2, deals with the formation of groups from the population under study, using Cluster Analysis Techniques.[6] Section 3 explains three parameter Weibull distribution where the parameters are estimated by MLE method after applying Cohen's iterative process.[8] Sections 4 and 5 describes LSPP model (Models I & II) developed in this context and also the conversion of the probablistic constraints into the deterministic constraints[1,2,4] in two different models. In Sec. 6, the numerical solution of examples of LSPP models I and II are obtained along with some observations. Section 7, as conclusion, highlights the applications of LSPP in Cluster based recruitment.

2. CLUSTER ANALYSIS TECHNIQUE

Among the well known Cluster Analysis techniques[5] that are available, we have chosen K-means method to divide the population of N eligible applicants into $K = n$ homogeneous clusters (say) $G_1, G_2, \ldots, G_n$, each cluster is formed with the similarity of specialization in one of the jobs as stated by them in their applications. Based on the efficiency level defined by the organization, each cluster is examined (Interviewed with Practicals) and is divided into binary clusters, using selection criteria. The selected clusters are denoted by $C_1, C_2, \ldots, C_n$, where $C_i \subset G_i$ $(i = 1, 2, \ldots, n)$. Ranking is made based on the performance skills of persons within the selected clusters.

3. WEIBULL DISTRIBUTION

The three parameter Weibull distribution is described by the following probability density function

$$f(x, \beta, \eta, \gamma) = (\beta/\eta)\left(\frac{x-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{x-\gamma}{\eta}\right)^{\beta}}$$

where η is a scale parameter and β is a shape parameter and γ is a location parameter. $x > \gamma$, $\gamma > 0$, $\eta > 0$ and $\beta > 0$. The mean of the distribution is given by $\bar{x} = \hat{\gamma} + \hat{\eta}\Gamma(1/\beta + 1)$ and the variance is given by $s_{ij}^2 = \hat{\eta}^2[\Gamma(2/\hat{\beta} + 1) - \Gamma^2(1/\hat{\beta} + 1)]$ The parameters $\hat{\beta}$, $\hat{\eta}$ *and* $\hat{\gamma}$ are estimated by the Maximum likelihood method by applying a modified version of Cohen's iterative process.[3,8]

3.1. Estimation of Parameters (MLE Method Using Cohen's Iterative Process)

Logarithm of Likelihood function of three parameter Weibull distribution is given by

$$\log L = n \log \beta - n\beta \log \eta + (\beta - 1)\sum_{i=1}^{n} \log(x_i - \gamma) - \sum_{i=1}^{n}\left(\frac{x_i - \gamma}{\eta}\right)^{\beta} \quad (1)$$

$\log L/\partial\eta = 0$ gives

$$\eta = \left[\frac{\sum_i (x_i - \gamma)^{\beta}}{n}\right]^{1/\beta} \quad (2)$$

Elimination of β and η from equations $\partial \log L/\partial\beta = 0$, $\log L/\partial\eta = 0$, $\log L/\partial\gamma = 0$ yields

$$\Psi_\beta(\gamma) \equiv \sum_i \frac{1}{x_i-\gamma}\left[1+\frac{\sum_i \log(x_i-\gamma)}{n} - \frac{\sum_i (x_i-\gamma)^\beta \log(x_i-\gamma)}{\sum_i (x_i-\gamma)^\beta}\right] - \frac{n\sum_i (x_i-\gamma)^{\beta-1}}{\sum_i (x_i-\gamma)^\beta}$$

The Cohen's[3] iterative process is given below:

(1) For an arbitrary $\beta = \beta_0$, use a starting value $\hat{\gamma} < x_{(1)}$, (where $x_{(1)}$ is the first value of the ordered sample) in the sample and decrease $\hat{\gamma}$ subsequently to γ' such that $\Psi(\gamma') > 0$ and $\Psi(\gamma' + \epsilon) < 0$ (or vice-versa) and then interpolate for $\bar{\gamma}$ such that $\Psi(\hat{\gamma}) = 0$.
(2) Substitute $(\beta_0, \hat{\gamma})$ in (2) to estimate η.
(3) Evaluate $\log L$ in (1) with $(\beta_0, \hat{\gamma}, \hat{\eta})$.
(4) Repeat 1 to 3 with a new set of $\beta = (\beta_1, \beta_2, \ldots)$ and evaluate (1) for each $\beta = \beta_i$.
(5) Choose $\beta = \hat{\beta}$ and the corresponding $(\hat{\gamma}, \hat{\eta})$ which maximizes $\log L$.

4. LINEAR STOCHASTIC PROGRAMMING MODEL

The mathematical formulation of the chance constrained Linear Stochastic Programming problem[7] is as follows:

Obtain the vector $X = (x_1, x_2, \ldots, x_n)$ which optimizes $f(X) = \sum_j^n c_j x_j$ subject to

$$P\left[\sum_{j=1}^{n} a_{ij}x_j \le b_i\right] \ge (1-\alpha_i), \quad x_j \ge 0, \quad (0 < \alpha_i < 1) \quad i = 1, 2, \ldots, m$$

where c_j, a_{ij} and b_i are random variables, α_i, is the level of significance, and $(1-\alpha_i)$ is the least probability with which ith constraint is satisfied.

In this model our goal is to decide on the optimal number of persons to be allocated from various clusters to various jobs with the objective of minimizing the total time taken to complete all the jobs.[2] We assume that the number of persons required for the jobs from jth cluster should not exceed the number of persons available in the jth cluster, to avoid infeasible solution. Let x_{ij} be the number of persons selected for ith job from jth cluster. Let t_{ij} be independent random variables denoting the time taken for completing ith job by a person from the jth cluster $(i, j = 1, 2, \ldots, n)$. Assume that the upper bounds for completion time of jobs and the upper bounds for manhours expected for each cluster are fixed and predefined. Assume also that the organization fixes a target on the number of employees for each job and for each cluster with the assumption that number of persons in the clusters put together, equals the number of persons required to complete all the jobs. The above details are furnished in Table 1.

The mathematical formulation design of the LSPP is as follows:

Determine x_{ij}, for which

$$\sum_{j=1}^{n}\sum_{i=1}^{n} t_{ij}x_{ij} \quad \text{is minimum subject to}$$

$$P\left[\sum_{j=1}^{n} t_{ij}x_{ij} \leq a_i\right] \geq (1-\alpha_i) \quad \text{for all } i = 1, 2, \ldots, n \tag{3}$$

$$P\left[\sum_{i=1}^{n} t_{ij}x_{ij} \leq b_j\right] \geq (1-\alpha_j) \quad \text{for all } j = i, 2, \ldots, n \tag{4}$$

$$\sum_{j=1}^{n} x_{ij} = x_i \quad \text{for all } i = 1, 2, \ldots, n \tag{5}$$

Table 1. Cluster matrix.

	Clusters								
	C_1		C_2		...	C_n			
Jobs	Mean	S.D.	Mean	S.D.		Mean	S.D.	Completion time	Persons required
J_1	$\overline{t_{11}}$	s_{11}	$\overline{t_{12}}$	s_{12}	...	$\overline{t_{1n}}$	s_{1n}	a_1	x_1
J_2	$\overline{t_{21}}$	s_{21}	$\overline{t_{22}}$	s_{22}	...	$\overline{t_{2n}}$	s_{2n}	a_2	x_2
–	–	–	–	–	...	–	–	–	–
–	–	–	–	–	...	–	–	–	–
J_n	$\overline{t_{n1}}$	s_{n1}	$\overline{t_{n2}}$	s_{n2}	...	$\overline{t_{nn}}$	s_{nn}	a_n	x_n
Expected man hours	b_1		b_2		...	b_n			
Expected persons	y_1		y_2		...	y_n			

$$\sum_{i=1}^{n} x_{ij} = y_j \quad \text{for all } j = 1, 2, \ldots, n \tag{6}$$

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{n} y_j \tag{7}$$

$$x_{ij} \geq 0 \quad \text{for all } (i, j = 1, 2, \ldots, n) \tag{8}$$

where $(1 - \alpha_i)$ and $(1 - \alpha_j)$, $[(0 < \alpha_i < 1), (0 < \alpha_j < 1)]$, are the least probabilities with which ith and jth constraints of Eqs. (3) and (4) are satisfied respectively and at least one among a_i, b_j and t_{ij} are random in nature.

Note. The above model reduces to the usual Transportation problem when we ignore inequalities (3) and (4).

5. LSPP MODEL

Model-1

Assume a_i b_j be fixed constants and t_{ij} is the only random variable. Let t_{ij} be independent random variable following Weibull distribution with three parameters, that is, $t_{ij} \sim W(\beta, \eta, \gamma)$ with the known mean $\overline{t_{ij}} = \hat{\gamma} + \hat{\eta} \quad \Gamma(1/\beta + 1)$ and standard deviation $s_{ij} = \hat{\eta}\left(\sqrt{\Gamma(2/\beta + 1) - \Gamma^2(1/\beta + 1)}\right)$. Let $l_i = \sum_{j=1}^{n} t_{ij} x_{ij}$ then

$$E(l_i) = \sum_{j=1}^{n} \overline{t_{ij}} x_{ij} \quad \text{and} \quad \text{Var}(l_i) = \sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 \quad \forall i = 1, 2, \ldots, n$$

(since t_{ij} are independent random variables, the covariances are zero). The ith probabilistic constraint of (3) is

$$P[l_i \leq a_i] \geq (1 - \alpha_i) \quad \text{or} \quad P[Z_i \leq z_i] \geq (1 - \alpha_i) \quad \forall i = 1, 2, \ldots, n$$

which could be rewritten as

$$\Phi(z_i) \geq \Phi(K_{(1-\alpha_i)}), \quad \forall i = 1, 2, \ldots, n$$

where

$$Z_i = \frac{[l_i - E(l_i)]}{\sqrt{\text{Var}(l_i)}} \quad \forall i = 1, 2, \ldots, n$$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

is the Standard Normal variate of l_i by Liapounovff Central Limit Theorem, and Φ is the cumulative distribution function of Standard Normal variate $Z_i \sim N(0, 1)$ with $(1-\alpha_i) = \Phi(K_{(1-\alpha_i)})$. Since Φ is a non decreasing continuous function

$$z_i \geq K_{(1-\alpha_i)} \quad \forall i = 1, 2, \ldots, n$$

which implies

$$E(l_i) + K_{(1-\alpha_i)}\sqrt{\mathrm{Var}(l_i)} \leq a_i \quad \forall i = 1, 2, \ldots, n$$

This implies

$$\sum_{j=1}^{n} \overline{t_{ij}} x_{ij} + K_{(1-\alpha_i)} \sqrt{\sum_{j=1}^{n} s_{ij}^2 x_{ij}^2} - a_i \leq 0 \quad \forall i = 1, 2, \ldots, n \tag{9}$$

This is the ith deterministic nonlinear constraint which replaces the given ith probabilistic constraint of Eq. (3). Similarly the jth deterministic nonlinear equation which replaces the jth probabilistic constraint of Eq. (4) is

$$\sum_{i=1}^{n} \overline{t_{ij}} x_{ij} + K_{(1-\alpha_j)} \sqrt{\sum_{i=1}^{n} s_{ij}^2 x_{ij}^2} - b_j \leq 0 \quad \forall j = 1, 2, \ldots, n \tag{10}$$

Thus the deterministic nonlinear programming problem equivalent to the given Stochastic Linear Programming problem is as follows

Determine x_{ij} for which

$$\sum_{j=1}^{n} \sum_{i=1}^{n} \overline{t_{ij}} x_{ij}$$

is minimum, subject to constraints (9), (10), (5), (7) and (8) for all $i, j = 1, 2, \ldots, n$.

Model-2

Assume t_{ij}, a_i and b_j are all random variables following Weibull distribution. Let $t_{ij} \sim W(\beta, \eta, \gamma)$ with mean $\overline{t_{ij}}$ and standard deviation s_{ij}. Let $a_i \sim W(\beta_1, \eta_1, \gamma_1)$ with mean $\overline{t_{a_i}}$ and standard deviation $s_{a_i}^2$ and

$b_j \sim W(\beta_2, \eta_2, \gamma_2)$ with mean $\overline{t_{b_j}}$ and standard deviation $s_{b_j}^2$, which are known by substituting the estimated values of the parameters. Let

$$l_i = \sum_{j=1}^{n} t_{ij}x_{ij} - a_i, \quad \forall i = 1, 2, \ldots, n$$

then

$$E(l_i) = \sum_{j=1}^{n} \overline{t_{ij}}x_{ij} - \overline{t_{a_i}} \quad \text{and} \quad \text{Var}(l_i) = \sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 + s_{a_i}^2 \quad \forall i = 1, 2, \ldots, n$$

(since t_{ij} are independent random variables, the covariances are zero). The ith probabilistic constraint of (3) is

$$P[l_i \le 0] \ge (1 - \alpha_i) \quad \forall i = 1, 2, \ldots, n$$

Following the lines of Model-1, we have

$$\Phi\left[\frac{-E(l_i)}{\sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 + s_{a_i}^2}\right] \ge (1 - \alpha_i) \quad \forall i = 1, 2, \ldots, n$$

that is

$$\Phi\left[\frac{E(l_i)}{\sqrt{\sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 + s_{a_i}^2}}\right] \le \alpha_i \quad \forall i = 1, 2, \ldots, n$$

which in turn implies

$$\Phi\left[\frac{E(l_i)}{\sqrt{\sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 + s_{a_i}^2}}\right] \le \Phi(K_{ai}) \quad \forall i = 1, 2, \ldots, n$$

where $\Phi(\cdot)$ is the cumulative function of $N(0, 1)$ which is non decreasing continuous function. Hence

$$E(l_i) \le K_{ai}\sqrt{\sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 + s_{a_i}^2}$$

that is

$$\sum_{j=1}^{n} \overline{t_{ij}}x_{ij} - \overline{t_{a_i}} - K_{\alpha_i}\sqrt{\sum_{j=1}^{n} s_{ij}^2 x_{ij}^2 + s_{ai}^2} \le 0 \quad \forall i = 1, 2, \ldots, n \tag{11}$$

which is the ith deterministic nonlinear constraint which replaces the ith probabilistic constraint (3). Similarly the jth deterministic nonlinear constraint which replaces the jth probabilistic constraint of Eq. (4) is given by

$$\sum_{i=1}^{n} \overline{t_{ij}}\, x_{ij} - \overline{t_{b_j}} - K_{\alpha_j}\sqrt{\sum_{i=1}^{n} s_{ij}^2 x_{ij}^2 + s_{b_j}^2} \leq 0 \quad \forall j = 1, 2, \ldots, n \tag{12}$$

Hence the model takes the following form: Determine x_{ij} which minimizes $\sum_{i=1}^{n}\sum_{j=1}^{n} \overline{t_{ij}} x_{ij}$ subject to the constraints (11), (12), (5), (6), (7) and (8) for all $i, j = 1, 2, \ldots, n$. Linear Stochastic Programming Problem can be solved using various techniques like Separable Programming, Genetic Algorithms, Sequential Programming and Khun–Tucker Condition Method. In this paper we have incorporated the techniques of Sequential Linear Programming.[7]

6. EXAMPLES

Model-1

Consider an organization which has (say) three types of jobs related to electronics for which it called for posts with the condition that every one must be able to do all the three jobs. First the applicants are grouped into 3 clusters G_1, G_2 and G_3 using K-means method by applying some similarity conditions on the efficiency in each job required by the organization on the details stated in the relevant applications. Then the persons from the clusters are examined in the interview by the organization based on it's designed norms on efficiency in each job and each cluster is divided into binary clusters of selected and non selected ones. The persons from the selected clusters are ranked based on their efficiency. Let the three final selected cluster be C_1, C_2 and C_3 and the jobs be J_1, J_2 and J_3 with the data given in Table 2.

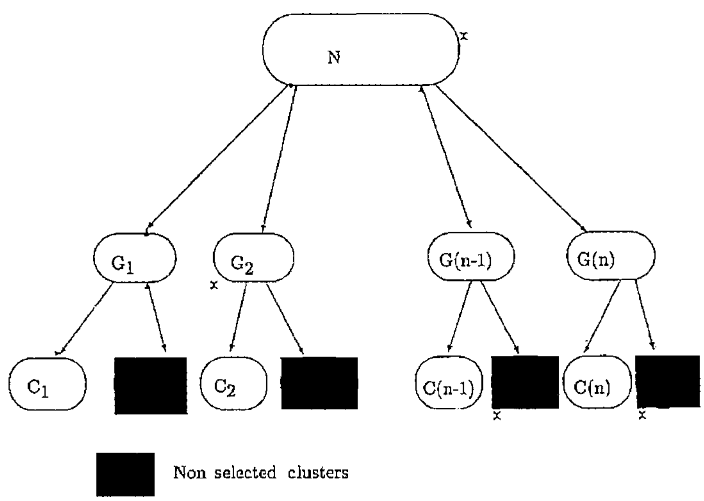
When $\alpha_i = \alpha_j = 0.025$ for all $i, j = 1, 2, \ldots, n$, the above nonlinear model have been linearized using Sequential Conversion Algorithm. Then the solution to the linearized model is obtained using Simplex algorithm or by using some standard Operations Research packages like TORA, or the solution to the problem could be directly obtained from LINGO packages for nonlinear equations. Thus the solution obtained is as follows, with the minimum objective value of 453 h.

$x_{11} = 18 \quad x_{12} = 1 \quad x_{13} = 1$
$x_{21} = 6 \quad x_{22} = 18 \quad x_{23} = 1$
$x_{31} = 6 \quad x_{32} = 6 \quad x_{33} = 18$

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table 2. Numerical illustration of cluster matrix.

	Clusters							
	C_1		C_2		C_n			
Jobs	Mean	S.D.	Mean	S.D.	Mean	S.D.	Completion time	Persons required
J_1	6	1	8	2	7	4	500	20
J_2	8	2	5	1	6	3	700	25
J_3	10	3	9	2	4	1	1000	30
Expected man hours	1000		700		500			
Expected persons	30		25		20			



Observations

(1) It is to be noted that maximum number of persons from clusters are allotted to jobs for which they posses minimum completion time. No one is left unemployed in each cluster because of constraint (5).

(2) If Eq. (2) is deleted the solution remains the same objective value of 453 h. which shows that the constraint (2) is a redundant one in this example. Similar analysis could be made by deleting one or more of constraints in the model and various results can be obtained.

Model-2

In the same illustration if the S.D. for completion time for jobs J_1, J_2 and J_3 are taken as 5, 4 and 6 and the S.D. for manhours for clusters C_1, C_2 and C_3 are 6, 4 and 3 respectively, then the solution is obtained as follows along with the objective function value of 455 h.

$$\begin{array}{lll} x_{11} = 18 & x_{12} = 1 & x_{13} = 1 \\ x_{21} = 7 & x_{22} = 17 & x_{23} = 1 \\ x_{31} = 5 & x_{32} = 7 & x_{33} = 18 \end{array}$$

Analysis similar to the one discussed in example of Model-1, can also be done here.

7. CONCLUSION

In this paper a detailed study of recruitment based on cluster analysis technique with an application of Linear Stochastic Programming Problem is considered. Though the formulation tends to Stochastic nature, it has been drastically brought down to the form of LPP through Sequential linear programming technic. The theory of Cluster Analysis enables us to sort out the best group of persons from the population based on the efficiency in completion of jobs, it yields a more accurate result when Cluster Analysis technique is applied with LSPP. The model developed can also be modified with more probabilistic constraints.

REFERENCES

1. Birge, J.; Louveaux, F. *Introduction to Stochastic Programming*; Springer Series in operations Research, Springer-Verlag: New York, 1997.
2. Charls, V.; Dutta, D.; Appal Raju, K. Linear stochastic fractional programming problem. International Conference on Mathematical Modelling, Univ. of Roorkee, Jan. 2001, Article 37, 211–217.
3. Cohen A.C., Jr. Estimating parameters of logarithmic-normal distribution by maximum likelihood. J. Amer. Statist. Assoc. **1951**, *46*, 206–212. Eibull distribution based on complete and on censored samples. Technometrics **1965**, *7*, 559–588.
4. Dantzig, G.B. Linear programming under uncertainty. Manage. Sci. **1995**, *41*, 197–206.

5. Gorden, A.D. *Classification*; Chapman and Hall: London, New York, 1981.
6. Jeeva, M.; Rajalakshmi, R. An application of cluster analysis in manpower planning–recruitment policies. A Paper Presented at the XXXIII Covention of Operational Research Society of India, IAM, Ahmedabad, India, 2000.
7. Rao, S.S. *Optimization Theory and Applications*; Wiley Estern Limited: New Delhi, 1986.
8. Sinha, S.K. *Reliability and Life Testing*; Wiley Estern Limited: New Delhi, 1986.