6th International Conference on Industrial Engineering and Industrial Management.
XVI Congreso de Ingeniería de Organización. Vigo, July 18-20, 2012

Stochastic Demand Model for Parcel Logistics Companies

Oscar Rioja San Martín[1], Jesús María Larrañaga Lesaca

Abstract

One of the factors that most considerably influences a Parcel Logistics Company (hereby referred to as PLC) is the randomness of the goods its' clients entrust it for distribution. This randomness considerably complicates any decision making related to the goods received by a PLC client. This article will focus on two aspects that are affected by the uncertainty of the goods received from a client, and will be analysed. These are: the revenue of a PLC and, the delivery and collection costs.

To reduce the risk associated with decision making, the variables that affect the revenue and the delivery and collection costs will be fitted by probability distribution functions and how to create a stochastic model with these probability distribution functions will be studied and analysed. The three variables that significantly affect the nature of the goods given by a client to a PLC for its transportation and distribution are the number of shipments per day., the weight of each shipment, and finally, the distance to the delivery point of each shipment. The uncertainty of these variables has a critical impact on the revenue and the delivery and collection costs of a PLC.

Keywords: Demand, stochastic model

1.1 Introduction

One of the factors that most considerably influences a Parcel Logistics Company (hereby referred to as PLC) is the randomness of the goods its' clients entrust it for distribution. This randomness complicates considerably any decision making related to the goods received by a PLC client. This article will focus on two aspects that are affected by the goods received from a client, and will be analyzed. These are: the revenue of a PLC and, the delivery and collection cost.

To reduce the risk associated with decision making, the variables that affect noteworthy the revenue and the delivery and collection costs are going to be fitted by probability distribution functions and how to create a stochastic model with these probability distribution functions is going to be studied and analysed in this research. The behaviour of the revenue and the costs will be predicted by this model, and this prediction will lead to reliable and accurate information about them. Generally, reliable and accurate information reduces the risk of any management decision; in this particular case, this information will help in the decisions related to the assignment of a tariff to each client with the objective of guarantee a minimum margin that satisfies the necessary profitability requirements.

For this research, the production of five PLC clients will be studied and modelled, with the aim of simulating them, and in this way, being able to predict the revenue and delivery and collection costs of each client in a period of time larger than the one in the study sample.

[1] Oscar Rioja San Martín (✉)
Technological Univesity of Catalonia
Barcelona, Spain
e-mail: errioxa007@hotmail.com

1.2 Case Study

When making a decision, to have at one's disposal reliable and accurate information, reduces the risk of a wrong decision and increases the chance that the chosen decision is the correct one. One of the most used tools in the information obtainment is model creation. These models will behave in an analogous way as the real system that is analysed, and by the observation and study of the models, desired information will be obtained. After the study of the information, the pertinent conclusions will be drawn, and these conclusions can be inferred into the real system. (Shannon, 1988)

Therefore, through this model the reality is represented to understand it in a better way, and put into practice this knowledge in the decision making management in the real system.

In this particular case, once the revenue and delivery and collection costs of a client are known valuable information is available to assign a correct tariff to each client. With a correct tariff, a minimum margin is secured, and in this way, an optimum profitability with each client.

The production of every client will be modelled with a stochastic model, and this model will be created by probability distribution functions. The steps to create this kind of models are:

- Variable analysis and sampling
- Fit a probability distribution function to each variable
- Model validation
- Model simulation

Horn, in 1981, recommended the use of computers to study markets' behaviour and forecast customers demand. The most suitable models he proposed for this kind of study were stochastic methods based on probability distributions.

Since then, this technique has been used in different transport fields for demand analysis. Recent examples of the use of this technique are:
Traffic forecasts in transportation and passenger demand in the rail system were modelled with statistical methods by Flyvberjerg, Holm, et all, in 2005.
Powell and Topaloglu, in 2003, demonstrated that stochastic methods solve some problems of the deterministic methods' weaknesses caused in freight transportation.

1.3. Variable Analysis

One fundamental characteristic that allows classifying a simulation model as a stochastic one is the use of random variables to formalize the evolution of the system represented (Keijnen, 1974). Therefore, the first step is to determine which variables become involved in the system, how they evolve in the system and how they influence its outputs.

In this case study, it is important to consider that the number and the nature of the shipments that a PLC client gives to the PLC for their delivery and distribution vary from day to day, being feasible that the goods entrusted are completely different on consecutive days.

The variation in the number and the nature of the shipments on different days is common to be of considerable magnitude. This variation is due to particular conditions, these being unique for each client, and therefore, exclusive. The exclusivity of the conditions for every customer means that each customer must be analysed in a particular and separate way, to be able to study the randomness of its goods.

The aforementioned variety is a direct consequence of three variables that significantly affect the nature of goods given by a client of a PLC for its transportation and distribution. The first of these variables is the number of shipments per day. The second is the weight of each shipment. And finally, the third one is the distance of the delivery point of each shipment.

A PLC is more interested in clients with high nominal value of every variable. The reason is that the higher the number of services, the shipment's weight and the delivery distance, the higher the PLC's revenue due to the transportation and delivery of the shipments.

The values that these three variables can take are uncertain. For instance, the number of shipments of a client on a certain day does not have to match the number of shipments that the same client gives to the PLC the next day. This variable will depend on the number of orders the PLC's client has.

With the variable of shipment's weight, exactly the same happens. The weight of all shipments given can be completely different on two consecutive days. Each shipment's weight is also random, and different shipments' weight given on the same day does not have to coincide one with the other. This variable is the reflection of the different nature of each order the PLC's client has.

Finally, the variable distance of the delivery point is analogous to the previous two. The delivery destination will be different for each shipment, according to the place where the person or the company who orders to the PLC's client is.

System Output: Revenue and Delivery and Collection Costs

The randomness of these three variables has a high influence on two key aspects for a PLC. The first is the revenue from the transportation and delivery of goods. The second is the cost related to the delivery and the collection of each shipment.

The revenue, as has been previously indicated, depends on the weight and the distance to be travelled to deliver the shipment. The higher the weight and the distance, the higher the income will be for to the delivery of the shipment.

The delivery and collection costs can be described as follows: the PLCs have a number of vehicles for the capillary distribution at their disposal, with the aim of delivery of all the shipments that are within its hub's sphere of influence. Whilst they are delivering, these vehicles are also responsible of the collection of other clients' goods, and take them to their hub. There, these goods are labeled, separated and classified according to their final destinations where they will have to be delivered, and loaded in the lorries that are in charge of taking them to the destination hub. Once the goods are in the destination hub, the cycle starts again, the capillary distribution vehicles are loaded with the shipments that have to be delivered in the destination hub's sphere of influence, and while they are delivering, they will do new collections.

Every PLC has these distribution vehicles in the regime most beneficial for the company, but in most cases it is a production regime: the more deliveries and collections they make, the more they earn. The vehicle drivers do not have fixed incomes for going to their job and their revenues depend on the number of deliveries and collections they do. Therefore, the higher the number of deliveries and collections has done, the higher their revenue. These delivery and collection revenue are tabulated in different strata, the strata being based on the weight of each delivery or collection.

A fixed price is assigned to the drivers of the vehicles due to the delivery of each shipment or each collection done. Another variable price is added to the fixed one depending on the weight of each shipment they deliver or collect done. Both, the fixed and variable part, depend on the weight of every delivered shipment or collection. The income for a delivery does not have to be the same as for a collection, and the variation among the prices of the different strata does not have to be lineal. Every PLC regulates these prices in the way it believes best suits its interests. An example of this kind of table can be seen in the figure table 1:

Table. 1.1. Stratified delivery and collection cost table

COLLECTION			DELIVERY		
kg	**Fixed Part**	**Variable part**	**kg**	**Fixed Part**	**Variable part**
1	1,9433	0,0179	1	1,8332	0,0181
2	1,9433	0,0179	2	1,8525	0,0181
3	1,9433	0,0179	3	1,8718	0,0181
4	1,9433	0,0179	4	1,8911	0,0181
5	1,9433	0,0179	5	1,9104	0,0181
10	2,3918	0,0178	10	2,1832	0,0179
20	2,9218	0,0165	20	2,5637	0,0175
30	3,4857	0,0160	30	3,0190	0,0173
40	3,4857	0,0160	40	3,0876	0,0173
50	3,4857	0,0160	50	3,1562	0,0173
60	3,7575	0,0159	60	3,2199	0,0166
70	3,7575	0,0158	70	3,3098	0,0166
80	3,7575	0,0158	80	3,4404	0,0166
90	3,7575	0,0158	90	3,5807	0,0166
100	3,7575	0,0158	100	3,6284	0,0166
200	4,0633	0,0147	200	3,9138	0,0151
500	4,3826	0,0139	500	4,2332	0,0144
1000	4,7360	0,0134	1000	4,5661	0,0137
2.000	4,9534	0,0131	2.000	5,1509	0,0133
3.000	16,0565	0,0077	3.000	6,9966	0,0124
24.000	24,4271	0,0049	24.000	8,8525	0,0118

As can be seen in table 1, not only the PLC revenue is directly related to the clients' goods, but also the delivery and collection costs depend on them.

1.4 Sampling

Data collection is the most laborious stage in a stochastic model construction, yet it is of utmost importance for achieving an efficient model. Any model is only as good as the data on that it is based (Vincent, 1998).

The different steps followed to obtain the data and samples which the probability distributions are fitted with are these: first of all, all the data related to a client's shipments will be obtained from the computer system of the PLC. This data will be broken down so a probability distribution function can be fitted to every random variable.

The variables that the revenue of a PLC depends on are the number of shipments, each shipment's weight and the distance travelled to deliver the shipments. As for the distance variable, the way every PLC prices each destination has to be considered. In the particular case of this PLC all the destinations are grouped into 4 scales, which are radial to the location where the PLC is located, the metropolitan area of Barcelona (Fig. 1.1).

Fig.1.1. *Scale map*

To obtain a sample of the variables that affect the collection costs, the data has to be broken down in the following way: how many kilograms are collected for a client each day, independently to which scale they are sent or the number of shipment it gives. As noted earlier, the collection costs are a fixed and variable one, which both are based on the weight of the collection.

Finally, to obtain a sample of the variables that determine the delivery costs, the data needed is the same that has been obtained to fit probability distribution function for the calculation of the revenue. In the case of the revenue, the data has been broken down in the number of shipments for every scale, and the weight of each shipment. The delivery costs are a combination of a fixed part due to the delivery of the shipment, plus a variable part due to shipment's weight. Both parts are based on the weight of the shipment.
In this particular case, the table values of delivery and collection costs is the same in all scales, but it could be that for a different PLC these values would be different depend on the scale. In that case, different values should be put into practice for each scale.

Therefore, the first step to obtain an appropriate sample is to put all the shipments of a client (with their correspondent weight) in order according to the date of receipt in the PLC. In this way, it is possible to achieve the total weight that is collected per day for each client.

Once the values of the collected weight per day variable are known, the next step is to take again all the original data and separate the data into the different scales, and then all the shipments have to be put in order according to the date of receipt in the PLC, so, it is possible to know how many shipments are given by client for each scale per day. When the values of the shipments per day variable are known, the next data that needs to be known is each shipment's weight per scale. With all these data, a sample of the different variables that affects the system is at disposal: weight collected per day, the number of shipments per day and each shipment's weight. As the delivery data is disaggregated in scales, the destination variable has also been considered, and all needed data to fit a probability distribution function of all the variables is already obtained.

Once all the data is broken down, all the samples are duly prepared to be fitted a probability distribution function to create the model and to run a simulation with it.

1.5 Model Development and Validation

As already stated, the variables that affect the received goods by a PLC (number of shipments, each shipment's weight, each shipment's destination and collected weight per day) are random variables.

Once the appropriate samples have been obtained, the next step is to create the model from them. In this research, this model is going to be created by probability distribution functions, where a probability distribution function is the relative frequency of occurrence of a random variable (Guasch et al, 2003).

To do that, it is necessary to know if the observations of the samples are independent, i.e. that they are not correlated. Once the independence of the observations of the samples is proved, the next step is to know if all the observations are identically distributed. If they are, it means that all these observations follow the same probability distribution, and this probability distribution and its parameters can be known. To know the probability distribution function and its parameters, several computer tools exist that made this process fast and accurate. In this particular case the statistical software used is Minitab.

To ensure the adequacy of the chosen probability distribution, and in this way guarantee that all the observations from the sample are identically distributed, two extras test have been undertaken in this research. These tests are known as nonparametric goodness of fit tests, and they are Kolmogorov-Smirnov and Chi Square hypothesis tests (Webster, 2001).

The goodness fit Chi Square test can be used for discrete and continuous distribution, while Kolmogorov-Smirnov can only be used for continuous distributions (Law and Kelton, 1991).

1.6 Model Simulation

Once all the variables have been fitted by a probability distribution function, and this and its parameters are both known, the model is created. After the model is created a simulation of the model can be run with the aim of see how the model evolves, and inference these observations and the conclusions drawn from them into the system.

In this particular case, the values taken of each variable obtained from the model can be observed, their behaviors studied and finally predict how the goods given by each client will evolve, and consequently, how much the revenue and the delivery and collection costs related to these goods will be.

This simulation is going to be run by the repetition method (Guasch et al, 2003). This method consists of the use of random number series (o Monte Carlo method), and from them, using inverse probability distributions values of the variables are generated.

This provides an independent sample that allows the use of classical statistics with sufficient guarantee that the results are reliable. Monte Carlo methods analyse stochastic systems way probability distribution and random within a mathematical model of the system operation (Rios, 2004).

This technique has been used to prove and simulate several models in logistics sector. Among all of them Ravula in 2007 designed five possible alternatives for bio-mass transport to the ethanol processing centres, and he determined which one was the best by simulating each route with Monte Carlo method.

Another example is Daeki in 1994 who simulated the lorry road transport network by this technique. Limpaitoon in 2005, simulated the costs of the goods delivery postponement for a logistics company by Monte Carlo method.

Once an independent sample is obtained from the simulation, the characteristics of the system cannot be clearly specified, but a probabilistic valuation of the rank of these characteristics is possible to be given. For that a significance level has to be determined, that indicates the error probability or the probability that a determined interval does not include the unknown population mean. This interval is named as confidence interval and indicates the percentage of sample means of size n that are into the limits marked by this confidence interval, due to the central limit theorem.

In this study, for each scale of each client the shipments of 1000 days are going to be simulated. Afterwards, the weight for each one of those shipments will also be simulated. Obviously, for each variable previously calculated probability distribution function will be used.

Once this information is available, the revenue will be obtained by multiplying the simulated weight of each shipment with the price assigned to the corresponding tariff. For the delivery cost, the correspondent strata according to the weight of each shipment will be looked for in table 1, and the product of the weight by the correspondent variable part will be added to the fixed part. The collection cost will be calculated in an analogous manner, with the exception of the simulated collected weight per day will be looked for in table 1 instead of the simulated weight of a shipment.

To forecast the rank where the revenue and delivery and collection costs will be, the mean value of the sample for each random variable will be obtained, and the corresponding confidence interval will be developed.

1.7 Conclusion

This research has proved that it is possible to model the goods given by a client of a PLC, by a stochastic model that is developed by fitting each random variable to a probability distribution function.

More specifically, the number of shipments variable of each scale has been modeled by Poisson distribution function. However, for the weight of each shipment and the weight of each collection the lognormal distribution function has been used. In the case of the lognormal distribution, after checking that it was possible to fit these variables for some clients with this distribution, the weight of each shipment and the weight of each collection have been fitted for all the clients with the lognormal distribution. It is true that there were some other distributions that fitted the data with a little more accuracy. Despite this, for a proper significance level, it has been verified that all the variables can be modeled by the lognormal distribution function.

Therefore, it can be concluded that all the goods given by a PLC client can be modeled by Poisson and lognormal distribution functions. Once this particularity is known, it is easy to calculate the parameters of the probability distribution function for other clients: to know the parameters of the Poisson distribution only it is needed to calculate the mean value of the number of shipments of each scale, and this value will be Poisson distribution's mean. For the case of the weight of each shipment and weight of each collection variables, the natural logarithm of each of the values of a sample is needed to be calculated, and from these new data the mean and variance has to be calculated. With this mean and this variance the values of the parameters of the lognormal distribution can be obtained.

When all the parameters of the different probability distribution functions are known, by the repetition method, a simulation has been done and the margin for each client has been obtained after removing the delivery and collection costs.

With the observation of the simulation output, accurate and reliable information of the evolution of the goods given by a PLC client, the revenue obtained by the transportation of these goods and the costs related to their delivery and their collection is at disposal of the managers. With this information, appropriate decisions for each client's tariff can be undertaken to ensure a minimum return and the viability of the company. Through the use of a model, and its simulation, the error in the assignment of client's tariff is significantly reduced, facilitating the managers' job.

1.8 References

Banks, J. (1998) Handbook of simulation: principles, methodology, advances, applications, and practice. John Wiley & Sons, New Jersey.

Caliskan D. O. (2007) Demand management in decentralized logistics. Georgia Institute of Technology

Coss Bu, R. (1993) Simulación: Un enfoque práctico. Noriega Editores, México D.F.

Daeki, K. (1994) Routing models and solution procedures for regional less-than-truckload operations problems. Massachusetts Institute of Technology

Flyvbjerg, B et al (2005) How (in)accurate are demand forecast in public works projects?. Journal of the American Planning Association, Vol. 71, No. 2. Chicago.

Fu, H. (2004) Development of dynamic travel demand models for hurricane evacuation. Louisiana State University

Gordon, G. (1978) System simulation. Prentice Hall, New Jersey

Guasch, A et al (2002) Modelado y simulación. Aplicación a procesos logísticos de fabricación y servicios. Edicions UPC, Barcelona.

Horn, M. J. (1891) Marketing models of consumer behaviour. Massachusetts Institute of Technology

Kleijnen, J. P.C. (1974) Statistical techniques in simulation. Maral Dekker inc., New York.

Law, A. M., Kelton, D. W. (1991) Simulation modelling and analysis. Mc Graw Hill, New York

Limpaitoon, T. (2008) Real time multi-period truckload routing problems. Massachusetts Institute of Technology

Montgomery, D. et al (2002) Probabilidad y estadística aplicada a la ingeniería. Editorial Limusa, Grupo Noriega Editores, México D.F.

Navidi, W (2006) Estadística para ingenieros y científicos. Mac Graw Hill, México D.F.

Powell W. B.; TOPALOGU H. Stochastic programming in transportation and logistics; Princeton University; 2003.

Ravula, P. P (2007) Design, Simulation, Analysis and Optimization of transportation system for a biomass to ethanol conversion plant. Virginia Polytechnic Institute

Ríos Isuna, SIXTO at al (2004) Investigación operativa. Modelos deterministicos y estocásticos. Editorial Centro de Estudios Ramón Areces, Madrid.

Shannon, R. E. (1988) Simulación de sistemas. Diseño, desarrollo e implantación. Prentice Hall, México D.F.

Sobol, I.M. (1976) Lecciones populares de matemáticas. Método de Montecarlo. Editorial MIR, Moscú.

Vincent, S. (1998) Input data Analysis. Handbook of simulation. Ed. Banks, J., John Wiley & Sons, New Jersey.

Webster, A.L. (2001) Estadística aplicada a los negocios y la economía. Irwin McGraw-Hill, Mexico D.F.