Markdown policies for optimizing revenue, towards optimal pricing in retail

Rudi Meijer and Michael Pieffers
VU University Amsterdam Dep. of Mathematics/Bijenkorf Departmental Stores
Netherland

Sandjai Bhulai
VU University Amsterdam, Department of Mathematics, Netherland

Keywords
Survival analysis; sell through; regression; convolution; markdowns; optimal revenue.

Abstract
Due to inadequate initial allocation of supply, retail organizations are usually confronted with markdowns during a sales season. In this article the apparent paradox between the traditional Law of Demand and retailing pricing is solved. Through the use of Survival Analysis, as is commonly used in Biomedical Research, a framework is developed for determining optimal price paths in a retail context. When thought of in analogy to birth and death processes, sell through curves can be seen as life cycles of retail goods and can be applied in Revenue Optimization. We show that two significant events in calendar time exist: the moment a probable markdown occurs and the instance the underlying good is sold. Next, we apply convolution to define the markdown point in calendar time and estimate the price elasticity with an exhaustive search yielding the markdown moment that optimizes revenue

1. Introduction

In optimizing revenue retailers in the fashion industry face a major challenge. During the selling period due to unforeseen discrepancies between actual demand and supply, markdowns have to be applied in an adequate and timely manner. The dominant source of the need for markdowns stems from the probability of misallocation of supply over the different stock keeping units. The main goal of this article is to develop a theoretical framework that can be practically applied for identifying optimal pricing strategies in terms of markdown management.

Although a sophisticated statistical and mathematical apparatus involving risks and pricing has been developed in the field of finance, in retail pricing no such generally accepted theoretical framework exists. Retail prices are the dominant part of all consumer prices though. Current retail pricing still heavily relies on the traditional cost price theory of the 19th century. In theory of the price setting, traditionally the law of demand should hold. When projecting this principle on the field of retailing, it will lead to an apparent inconsistency though. When the traditional law of demand holds, at the beginning of the selling period, when supply is abundant, prices should be low. At the end of the sale season when supply becomes scarce, prices should be high.

In retailing practice the opposite situation can be witnessed though: during the selling period prices will decline rather than go up, even though supply is depleted. This paradoxical phenomenon can only be rationalized when one realizes that it is not the quantity (amount) of the supply that determines the retail price but the allocation and risks involved of the supply that determine the retail price. By taking the risks in to account, as is addressed in survival analysis, different price levels during a selling period can be justified. Given the law of demand the estimates of the price elasticities should still be negative.

In retail it is common to measure performance in terms of sell through. Related to calendar time we form sell through curves which are guided by the same statistical principles as survival curves used in biomedical research. The careful approach in biomedical research to derive unbiased estimates is also applied to the sell through curves used in retail. Through the use of Cox regression, hazard rates for different characteristics of the products are identified. Next, spline regression is used to find the price elasticity parameter related to the markdown moment. Once estimates for the parameters are found, we develop a revenue formula. Given the revenue formulation it is possible to optimize the revenue for a stock keeping unit by an exhaustive search for the optimal markdown moment (Gosavi, 2003). In the univariate case we do not need the more sophisticated approach like Markov Decision Processes (MDPs) where we have competitive risks and we cannot use an exhaustive search for the optimum. A retail selling period covers about 200 days at most and therefore poses no real challenge for today's computing power.

Both survival analysis as pricing in finance are founded in martingale theory and related stochastic calculus. (Aalen, et al., 2009)(Flemming & Harrington, 2005)(Gill, 1984)(Gjessing, et al., 2008)(Kalbfleisch & Prentice, 2002). We will make extensive use of the results but in this article no mathematical rigor is attempted, rather the chain of arguments will be intuitive and results are graphically presented to elucidate the results. Martingale theory is rooted in the idea of a fair game and is as such a model, a guarantee for the unbiasedness of the estimates.

2. The revenue formulation

We start with a revenue formulation for optimal markdown policies. In the following paragraphs the different elements of the revenue formula are clarified and we demonstrate how to derive the optimal markdown moment given the price elasticity that gives rise to the optimal revenue.

$$R = \int_0^T \left\{ \left[1 - S_0{}^{\exp(\beta x)}\right] \left[\sigma LN\left[\frac{\widehat{H}_0}{H_0}\right] + \bar{p} + \frac{1}{2}\alpha(t-\tau)t \right] \right\} dt - cp.$$

Here, the notation is given by:

R	Retail revenue
$\left[1 - S_0{}^{\exp(\beta x)}\right]$	The sell through curve
$\exp(bx)$	β stock keeping unit allocation regression parameter (color, size)
σ	Supply price variation
$\frac{H_0}{\widehat{H}_0}$	change of measurement of base hazard
α	Price elasticity parameter
$t-\tau$	convolution in calendar time, τ optimal moment for markdowns
t	calendar time
$\bar{p}$	Average initial full retail price
cp	Cost price

3. The sell-through curve

Sell through is used in a rather naïve way in retail practice. In retail no attempt is made to use sell through rates in a statistical way to get unbiased results as is done in biomedical research. Estimates of sell through are mostly pessimistic in event time and mostly optimistic in calendar time. Sell through is defined as the ratio of the cumulative products sold during calendar time and the inventory at the start of the season. The relation with survival analysis is clear when product life cycles are thought of in analogy to the life cycle of people. A 'birth' of a product is analogous to the moment the item becomes available for sale. Contrarily, the sale of a good can be considered as a 'death' of the concerning item. Hence by applying survival analysis to retail goods, one can determine the 'expected life span' of the concerning items. Survival analysis treats careful attention to censored cases and the so-called left truncation. In retail, left truncation is primarily caused by delayed entries of the products on the shopping floor. Because of the delayed entries every product has its own event time than can be different from calendar time. The sell through is measured in calendar time. The difference between event and calendar time while related gives a change of measurement when we derive our revenue curve.

Let us start with an example that explains why it is important to follow survival analysis to estimate the sell through for retail products. The example is inspired by the survival lecture notes from Daowen Zhang(Zhang, 2005).

Table1

Interval Start Time	*Number Entering Interval*	*Number Withdrawing during Interval*	*Number Exposed to Risk*	*Number of Terminal Events*	*Proportion Terminating*	*Proportion Surviving*	*Cumulative Proportion Surviving at End of Interval*	*SellThrough*
0	10.315	204	10.213	2.182	0,21	0,79	0,79	0,21
20	7.929	380	7.739	1.254	0,16	0,84	0,66	0,34
40	6.295	386	6.102	707	0,12	0,88	0,58	0,42
60	5.202	0	5.202	1.660	0,32	0,68	0,40	0,60
80	3.542	82	3.501	670	0,19	0,81	0,32	0,68
100	2.790	0	2.790	537	0,19	0,81	0,26	0,74
120	2.253	107	2.200	521	0,24	0,76	0,20	0,80
140	1.625	168	1.541	385	0,25	0,75	0,15	0,85
160	1.072	183	981	221	0,23	0,77	0,11	0,89
180	668	374	481	119	0,25	0,75	0,09	0,91
200	175	140	105	35	0,33	0,67	0,06	0,94

Life Table

The data have been grouped in 20 days intervals and the time is measured in event times.
We start with 10.315 in the product group 'Women's Fashion T-shirts 'at 68 different non-discounted retail price levels. The numbers withdrawing means that we have no follow up for the article, it could be lost for administrative reasons, could be stolen, could be damaged and at the end of the season the article could still not be sold. These represent the censored cases in the analysis.
We could ask the question: What is de sell through after 100 days?
Two naïve and incorrect answers are given by

1. $F(100) = P[T < 100] = \frac{7.010\ sold\ in\ 100\ days}{10.315\ products} = 68.0\%,$

2. $F(100) = P[T < 100] = \frac{7.010\ sold\ in\ 100\ days}{10.315 - 1.052\ withdrawn\ in\ 100\ days} = 75.7\%$.

1. The first answer would be correct if all censoring occurred after 100 days. This is not the case so the answer is too pessimistic; the real sell through rate after 100 days is 74%.
2. The second estimate would be correct if all the observations censored in 100 days were censored immediately upon entering the study. This estimate is too optimistic. This situation arises when we measure in calendar time.

Given the retail practice, the naïve estimated sell through rate is mostly too pessimistic measured in event time and measured in calendar time the naïve estimate sell through rate is too optimistic.

We want to measure the influences on the sell through for different stock keeping units like color and size and related price elasticities. In order to determine accurate and unbiased estimates for sell through, we use survival analysis in our research.

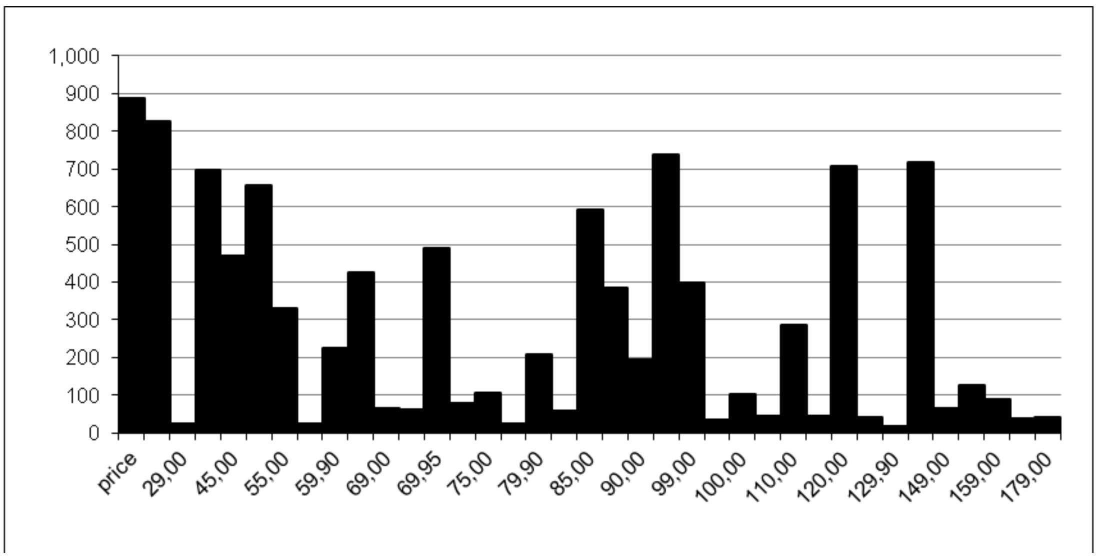

Figure1

The analyzed data set contains 37 different price levels, 41 different colors and 5 sizes.

3.1. Cox regression

We are interested in the time it takes to sell a product. We want to characterize the distribution of "time to event". Specifically, we are interested in the variations within the stock keeping units like color and size. The random variable T denotes the time to event. In survival analysis the distribution of T can be described in a number of equivalent ways. Notice that we start in event time as is usual in survival analysis.

- The cumulative distribution function $F(t) = P[T \leq t], t \geq 0$, $f(t) = \frac{dF(t)}{dt}, F(t) = \int_0^t f(u)du$;
- The survival function $S(t) = P[T \geq T] = 1 - F(t^-)$;
- The hazard function $\lambda(t) = \frac{\lim_{h \to 0} \frac{P[t \leq T < t+h]}{h}}{P[T \geq t]} = \frac{f(t)}{S(t)}$.

The hazard rate and the survival function are related by:

$S(t) = e^{-\int_0^t \lambda(u)du}$.

A proportional hazards model looks like
$S(t|z) = [S_0(t)]^{\exp(z^T\beta)}$.

This is the Cox-regression model in which we have a basic survival curve, the **z** is a p x 1 vector of covariates of the stock keeping units, and **β** is a p x 1 vector of regression factors to estimate. With the Cox-regression model we measure the influence of the covariate or factor on the survival function. For the stock keeping units there is just a factor with the value 1 if it is applied and 0 if it is not applied. Cox regression is a semi-parametric approach to survival analysis. The estimation of the parameters is done by partial maximum likelihood. Notice that no constant is estimated, the constant is absorbed in the baseline of the hazard. The base line is unspecified in the estimation procedure and that explains the use of partial information. For the estimation of the base line Breslow's procedure is used (SPSS Inc. , 2010).

In Figure 2, four sell through curves for four different colors are given. We immediately notice that the curves are proportional. The proportionality is a characteristic of the Cox-regression model. For proportionality we have independence of time. However, we are interested in the price markdowns; they are per definition time dependent covariant. The sell through curves should not be proportional. Cox regression gives the opportunity to incorporate time-dependent covariance but we have chosen to convolute the integral and so making explicit this time dependence of the markdowns on the sell through.

3.2. Martingale residuals

Please note that the role of the martingale residuals and the measured noise in normal regression differs. Given the Doob-Meyer condition, it is tempting to interpret the martingales residuals equally to the residuals in normal regression (Therneau & Grambsch, 2000). When measured noise in regression parameter estimation is not random, it will still capture some structure and can be used to find covariates that will further improve our model. This is not the case for the martingale residuals found in the Cox-regression though. Due to the counting process, the martingale residuals are inherent to the stochastic process. We will find the same outcomes as when deriving the formula for the supply price during the selling period.

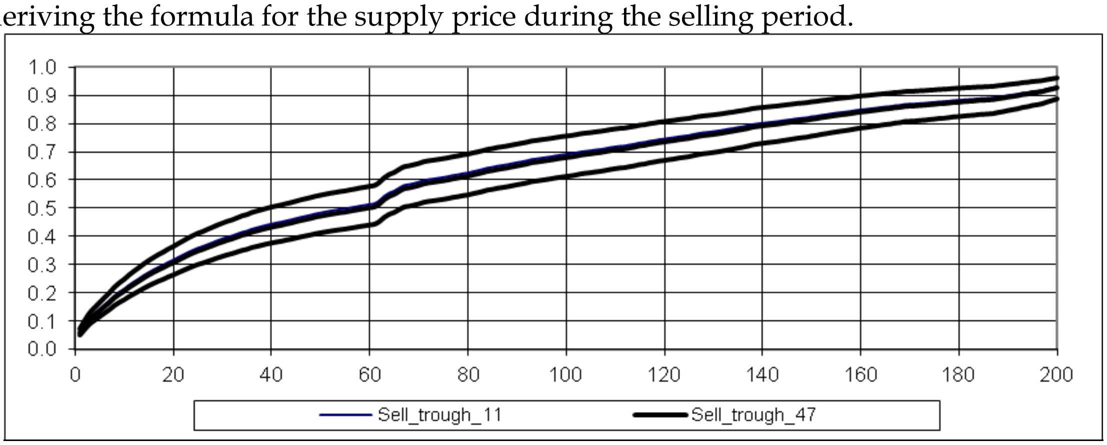

Figure 2

Table2

	B	SE	Wald	df	Sig.	Exp(B)
Color			28,497	3	0,000	
Color 11	-0,177	0,062	8,107	1	0,004	0,838

Color 47	-0,397	0,075	27,698	1	0,000	0,673
Color 85	-0,211	0,068	9,660	1	0,002	0,809
Size			50,270	4	0,000	
Size(1)	-0,059	0,093	0,399	1	0,528	0,943
Size(2)	0,128	0,067	3,720	1	0,054	1,137
Size(3)	0,268	0,061	18,964	1	0,000	1,307
Size(4)	0,358	0,062	33,299	1	0,000	1,431

Variables in the Equation

In the table, B are the estimated coefficients. They are interpreted as the predicted change in the log hazard for the factorized values (0,1) in color and size. The variables SE are the standard errors of the estimated coefficients B, and Wald are the Wald statistics. If the df=1, the Wald statistic can be calculated as $\left({}^{B}/_{SE}\right)^2$. It is used to test whether the estimated coefficients are significantly different from 0 . Sig is the significance level for the Wald statistic, Exp(B) is the relative risk (i.e., the ratio of the risk for the different factors, color and size). The β coefficients of the colors 11, 47 and 85 have all negative values. The color 93 has the reference value 0. The negative values suggest that for these colors the survival is relatively high meaning the sell through is low. The relative risks for the colors are below 1 suggesting that the sell through for these colors is expected to be low. For these colors we should expect possible markdowns.

4.Convolution

We start with a graphical presentation to create an intuitive understanding of what is necessary to shift from the allocation influence of stock keeping units to the related price domain.

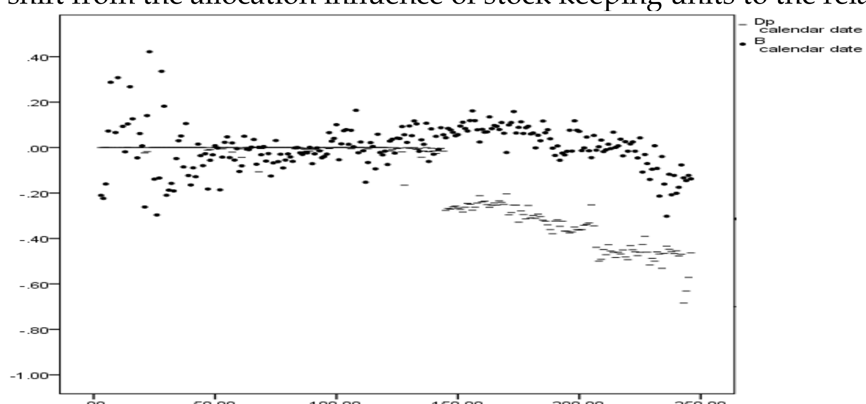

Figure 3

We find in this graph the mean estimated β's for all the stock keeping units over calendar time and the mean price differentiation over calendar time. Clearly, there is a relation between the markdowns in percent and the estimated risks of the stock keeping units after about 120 days. How can we incorporate this relation in the survival analysis and address the duality in allocation, quantity and prices?

Convolution formula:

Let X_1 and X_2 be two independent, non-negative random variables with respective probability distribution functions $F_1(x)$ and$F_2(x)$. Assume that $F_2(x)$ has a probability density$f_2(x)$. Then, by a direct application of the law of total probability, we have the convolution formula

$P\{X_1 + X_2 \leq x\} = \int_0^x F_1(x-y) f_2(y)dy, \quad x \geq 0$. (Tijms, 2003)

A good can have two events in calendar time: first, a probable markdown and secondly, the moment that it is sold. With this knowledge, we can apply convolution to define the markdown point in calendar time. The sum of the time to markdown and the time for the article to be sold is the total event time. The convolution applies.

4.1.Spline regression

We formulate a regression equation to give form to the convolution and find the markdown moment and the related price elasticity coefficient. Spline regression can be used for our problem, it gives the 'knot', i.e., the markdown moment as a result.

$\beta x = \dot{p} \; = \alpha x(t-\tau)$.

$x$is just the indicator of the stock keeping unit applied and equals 1 if applied. β is the regression parameter of the influence on the sell through of the stock keeping unit and $\dot{p}$ is the price differential in time$t$ and τ the markdown moment convolution applied. The estimated α gives the price elasticity.

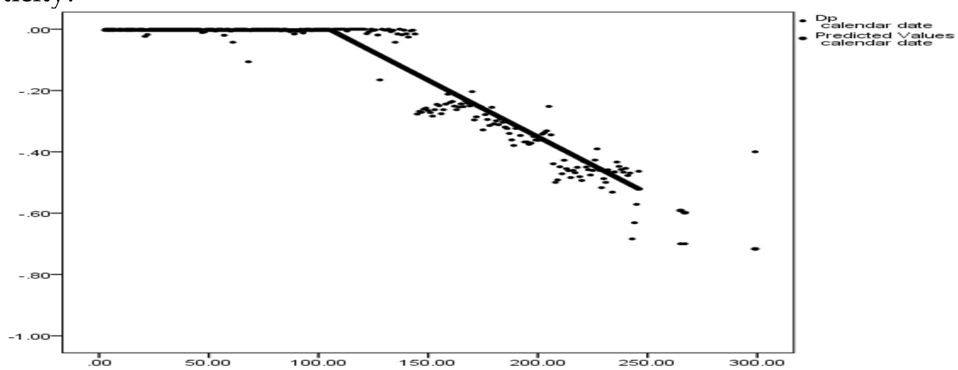

Figure 4

As an overall estimate for the price elasticity parameter we found $\alpha = \; -0.004$ and the 'knot' is the 105.69 day in the selling period.

As mentioned before, the constant in Cox-regression is absorbed in the base line. Clearly, from Figure 4 there is a constant to estimate if we wanted relate the results from the spline regression equation to the estimated β's.

The hazard is defined as:

$H(t|x) = H_0(t)\exp(x'\beta)$.

We want to preserve these estimated hazards when we change to the price domain without having a constant in the exponential. As a result we apply a change in measurement for the base line hazard. The relation of the hazard in price domain can be formulated as:

$H(t|p) = \; \ddot{H}_0(t)\exp(\dot{p})$.

So we can define our constant as$\ln\left[\frac{H_0}{\ddot{H}_0}\right]$.

Define $\dot{p} \; = \alpha(t-\tau) + \; \lvert\hat{\beta} - \hat{p}\rvert$. $\lvert\hat{\beta} - \hat{p}\rvert$is the absolute difference between the average β's and average $\dot{p}$'s. $H_0\, e^{\beta} = \; H_0\, e^{\alpha(t-\tau)+\lvert\hat{\beta}-\hat{p}\rvert}$.

$$\hat{H}_0 = H_0 e^{|\ddot{\beta}-\hat{p}|}.$$
$$Ln\left[\frac{\hat{H}_0}{H_0}\right] = |\hat{\beta} - \hat{p}|.$$

4.2. Stochastic calculus

We defined the constant as a change in measure but the constant is really stochastic in nature. The average price of a product which the retailer observed is the superposition of a weighted average of the supply prices at the shopping floor. The average price follows a Brownian motion by the delayed entry process caused by the consumer demand in time. This is the abstract formulation to say that observed supply prices and so their differentials fluctuate randomly as we can see in Figure 3. It is Î to calculus that gives the recipes to define the inherent volatility of the supply price.

$$dP(t) = [\alpha(t-\tau)]dt + \sigma dW(t).$$

Notice that $\frac{dP}{dt}$has a very simple form$\alpha t + constant$ in its non-Brownian part. As a consequence we only need to multiply σ (estimated standard deviation of the supply prices) with our defined base hazard ratio to fulfill the inherent Brownian process (Shreve, 2004).

$p = \bar{p} + \sigma Ln\left[\frac{\hat{H}_0}{H_0}\right] + \frac{1}{2}\alpha(t-\tau)t$In which σ is the estimated standard deviation of the supply prices and $\bar{p}$ is the weighted initial average full retail price. $\boldsymbol{\sigma}$ Measures the inherent volatility of the supply prices just as it does for the prices in financial applications.

$\alpha(t-\tau)$Applies if t > τ otherwise $\dot{p} = 0$ and only the stochastic component $\bar{p} + Ln\left[\frac{\hat{H}_0}{H_0}\right]$ rests.

In Figure 5 we see what it looks like for the price development in calendar time.

It is important to notice that we cannot ignore the fluctuations in the supply price of the items; the fluctuations could be substantial and outweigh the possible optimal markdowns, especially in the time interval around the 'knot' moment of the markdown.

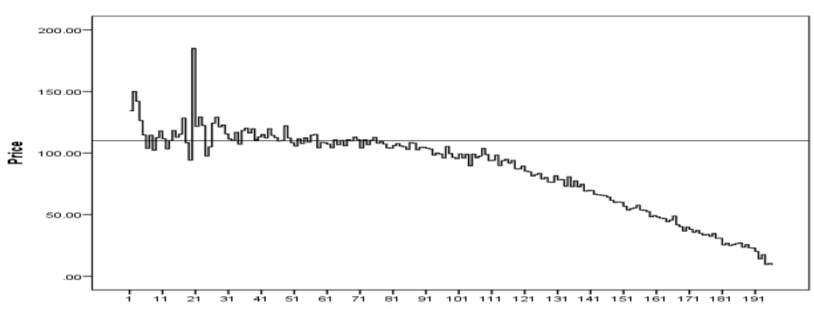

Figure 5

5. Revenue optimization

Now all the ingredients to optimize the revenue function are formulated. As was already given in the introduction, our task is now relatively simple. For the univariate case that we formulated, there is no need to use the more advanced mathematical tools than those found in the Bellman-Jacobi-Hamilton formulation (Liberzon, 2012). We can develop a simple algorithm that searches for the optimum. It is only the markdown moment in time that varies and influences the sell through and realized retail price. Given that the revenue function is concave and the restriction of the selling period in 200 days, we find an optimum that is global in the interior or an optimum that is restricted by the end of the selling period.

We give in Figure 6 an example of the development of the optimization algorithm for the color white and define the algorithm that gives the optimum. Notice that we could also define our problem as $min[P - R + cp]$.

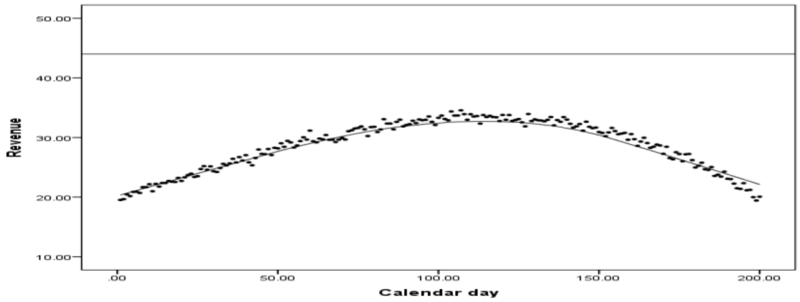

Figure 6

For color 11 (white T-shirts) the optimal revenue is when the markdowns are applied after 104 selling days. Note that Figure 6 gives the path of the algorithm; it does not give the cumulative path of the revenue for the specific color. We find the cumulative path in Figure 7

We can now translate the revenue integral into an algorithm that gives us numerical results.

$R_t = R(S_t, p, \tau)$.

The revenue at time t is a function fromS_t,p and τ (the markdown moment)

$$1 - S_t = \begin{cases} 1 - \hat{S}_0^{\exp(\alpha(t-\tau))} & ,t \geq \tau \\ 1 - S_0^{\exp(\beta)} & ,t < \tau \end{cases}$$ the sell through at time t.

$$p_t = \begin{cases} \sigma Ln\left[\frac{H_0}{H_0}\right] + \bar{p} + \frac{1}{2}\alpha(t-\tau)t, \ t \geq \tau \\ \sigma Ln\left[\frac{\hat{H}_0}{H_0}\right] + \bar{p}, \ t < \tau \end{cases}.$$

The revenue becomes the weighted price at t by the sell through minus cost price.

$$\dot{p}_t = \begin{cases} \alpha(t-\tau) + Ln\left[\frac{H_0}{H_0}\right], \ t \geq \tau \\ 0, \ t < \tau \end{cases}.$$

For color 11 the estimated parameters are:

α = -0.004

β = -0.177

σ = *N(0, 0.*116)

$\bar{p}$ = 71.50

$Ln\left[\frac{H_0}{H_0}\right]$ = 0.0597

cp = = 27.41

For the exhaustive search we define 200 vectors with 200 entries for every i = (1,…,200) τ^i has i as the first markdown. For example τ^5 = (0,0,0,0,5,…,200)

The revenue vectors are

$R_t^i = \left[1 - S_t^i\right] * p_t^i - cp_t^i$.

Cumulate the vectors over t to find$R^i = \sum_t R_t^i$.

The results for color 11 we find in Figure 6. The maximum is reached with an applied markdown starting at day 104.
$ArgMax\ R^i$ = 104.

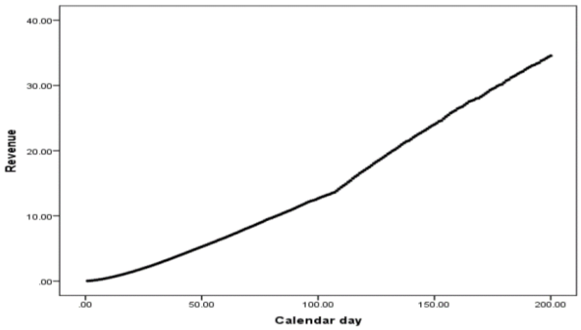

Figure 7

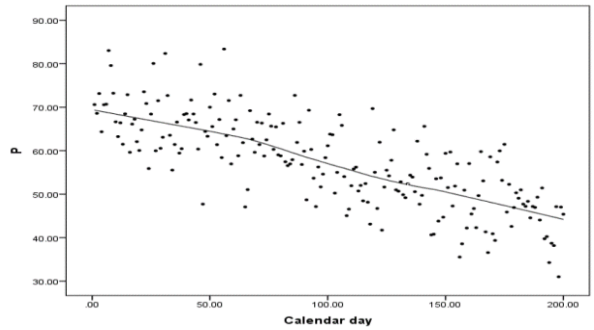

Figure 8

For R^{104} we find in Figure 7 the course of the revenue per calendar day and in Figure 8 the development of the price. As expected the price development is decreasing while the price elasticity α is negative and the law of demand holds.

6. Conclusions

Using the results from Survival analysis and financial calculus we can derive in the univariate case a revenue curve, appropriate for an optimal markdown policy in retail. We showed that the law of demand holds and during the selling period the expected price development is decreasing as is common in retail and find as such the optimal markdown moment.

6.1. Further research

The line of arguments that we followed guarantees expansion to the application of the multivariate cases and also a further research in optimal pricing that arises in the question of the optimal initial price at the start of the selling period.

References

Aalen, O. O. et al., 2009. History of applications of martingales in survival analysis. *Journal Électronique d'Histoire des Probabiltés et de la Statistique,* Volume 5, pp. 1-28.

Flemming, T. R. & Harrington, D. P., 2005. *Counting Processes and Survival Analysis.* sl:Wiley.

Gill, R. D., 1984. Understanding Cox's regression model: A martingale approach. *Journal of the American Statistical Association,* Volume 79, pp. 441-447.

Gjessing, H. K., Aalen, O. O. & Borgan, O., 2008. *Survival and Event History Analysis: A Process Point of View.* sl:Springer Verlag.

Gosavi, A., 2003. *Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning.* sl:Kluwer Academic Publishers.

Kalbfleisch, J. D. & Prentice, R. L., 2002. *The Statistical Analysis of Failure Time Data.* sl:Wiley Interscience.

Liberzon, D., 2012. *Calculus of Variations and Optimal Control Theory: A Concise Introductioin.* sl:Princeton University Press.

Shreve, S. E., 2004. *Stochastic Calculus for Finance II, Continuous-Time Models.* sl:Springer.

SPSS Inc. , 2010. *IBM SPSS Statistics 19 Algorithms.* sl:sn

Therneau, T. M. & Grambsch, P. M., 2000. *Modeling survival data: extending the Cox model.* sl:Springer verlag.
Tijms, H. C., 2003. *A first Course in Stochastic Models.* sl:John Wiley & Sons, Ltd..
Zhang, D., 2005. *Analysis of Survival Data, lecture notes.* [Online]
Available at: http://www4.stat.nscu.edu/%7Edzhang2/st754/chap~.pdf