國立交通大學

經營管理研究所

碩士論文

利用 BG/NBD 模型預測顧客未來行為，
以提升流失管理效率--以藥品產業為例

Forecasting Customer Buying Behavior Using BG/NBD Model to Efficient Churn Management-A Case Study of a Pharmaceutical Company

研究生：黃悅慈

指導教授：唐瓔璋 教授

中華民國 九十八年六月

利用 BG/NBD 模型預測顧客未來行為，以提升流失管理效率--以藥品產業為例

Forecasting Customer Buying Behavior Using BG/NBD Model to Efficient Churn Management-A Case Study of a Pharmaceutical Company

研 究 生：黃悅慈　　Student：Yueh-Tzu Huang
指導教授：唐瓔璋 博士　　Advisor：Dr. Edwin Tang

國立交通大學
經營管理研究所
碩士論文

A Thesis
Submitted to Institute of Business and Management
College of Management
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Business Administration

July 2009
Taipei, Taiwan, Republic of China

中華民國 九十八 年 六 月

利用 BG/NBD 模型預測顧客未來行為，以提升流失管理效率--以藥品產業為例

研究生：黃悅慈　　　　　　　　　　　　　　指導教授：唐瓔璋 博士

國立交通大學經營管理研究所

摘 要

本研究利用顧客交易資料，包含顧客重複購買次數、最近一次購買與購買歷史總時間，透過 BG/NBD 模型(Fader et al., 2005)來預測顧客未來交易次數與顧客活躍機率。此研究可利用模型預測結果，對顧客作一對一行銷方案，將高交易金額的顧客與其活躍機率作比對，找出需要重新制定新行銷方案的顧客，以增加顧客權益與提高顧客流失管理效率。

本研究之資料為台灣某人類肝臟用藥的交易資料，交易時間為 97 年 1 月至 98 年 2 月共六十週的交易資料，本研究將其資料分為兩部份，前 30 週作為建模資料，後 30 週資料作為檢驗模型準確度。交易顧客為非契約性的醫院、診所與藥局，總共為 953 個顧客。非契約顧客之研究主因為其非購買行為難以判斷為交易關係流失或者只是在兩次交易時間的暫時停止交易，因此本模型計算出的顧客活躍機率能計算每位顧客於特定時間其交易存活機率。

本研究結果，根據顧客重複購買次數的預測值與實際值，其 MAPE (mean absolute percentage error) = 42.6% < 50%為可靠預測；且透過變異數分析，顧客活躍機率之高低對顧客總金額有顯著影響，而顧客活躍機率之高低對顧客購買的平均金額並無顯著影響。

關鍵字：BG/NBD；顧客活躍機率；顧客流失管理

Forecasting Customer Buying Behavior Using BG/NBD Model to Efficient Churn Management-A Case Study of a Pharmaceutical Company

Student: Yueh-Tzu Huang Advisor: Dr. Edwin Tang

Institution of Business and Management

National Chiao Tung University

Abstract

In the study, we use a transaction database containing information on the frequency and timing of transactions for a list of customers in order to forecast about future purchasing. The forecasting results help us develop one to one marketing plans to increase customer equity and churn management efficiency.

The data for a sample of 935 non-contractual customers who made their purchase of some particular hepatic capsule with the pharmaceutical company is during the first 30 weeks of 2008. We have information on their repeat purchasing behavior up to the lasting 30 weeks. The first 30 weeks purchasing data is used for modeling. The last 30 weeks information is used to measure accuracy. The types of customers are hospitals, clinics and pharmacies in Taiwan.

The result shows the model reliable for the mean absolute percentage error 42.6% < 50%. The level of customer active probability affects customer transaction total monetary significantly by analysis of variance, while the customer active probability has no significant effect on customer average transaction monetary.

Keywords: BG/NBD model, customer active probability, churn

誌 謝

感謝唐瓔璋教授對學生不遺餘力的指導，老師的學識涵養與國際視野讓學生終身獲益，並感激口試委員蕭中強教授、陳美芳教授與曾芳代教授的寶貴意見，使本論文更臻完整，謹此致謝。

學生 黃悅慈謹上

目錄

圖目錄

表目錄

一、 緒論

1.1 研究背景與動機

1995年Woodruff提出顧客價值是公司競爭優勢來源的觀點，顧客關係不僅受到行銷領域的關注，同時也成了策略探討競爭優勢的一個新焦點。顧客關係管理的最終目的就是產生更高的顧客權益(customer equity)，根據Rust、Lemon與Zeithaml學者的觀點，顧客權益是所有顧客總終身價值的折現值，而三項驅動顧客權益的來源為：價值權益、品牌權益與關係權益(Rust, Lemon, & Zeithaml, 2005)。而台灣藥品市場中，在非直接對消費者市場銷售情形之下，與醫院、診所、藥局通路等的關係權益相顯重要。

本篇研究利用台灣藥廠之人類肝臟用藥的顧客交易資料，其銷售對象為台灣之醫院、診所與藥局，我們所研究對象為其中無訂立契約之客戶。對於契約型態的顧客，我們很容易可以透過顧客在時點是否訂立契約來判斷此顧客是否仍持續與本公司的交易關係；但對於非契約型態顧客而言，在某一時間點下，並不容易判斷此時點的非購買行為，表示為正式與公司的交易關係終止，或僅為兩次交易之間的短暫停止購買。因此契約型客戶，其忠誠行為的判定是比無訂立契約的客戶相較容易。因此顧客活躍機率在這些無訂立契約的顧客中相顯重要，顧客活躍機率即在某一時間點下，顧客交易關係仍活躍的機率。

本研究利用顧客歷史交易資料建模，並預測這些無訂立契約之顧客對本藥品的未來購買行為與顧客活躍機率，以利決策者制定正確的行銷策略與計畫，以提升顧客流失(customer churn)管理效率。

1.2 研究範圍與對象

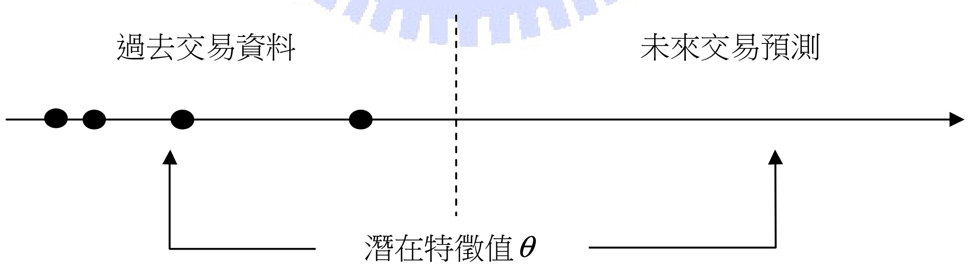
本研究為實證研究，以台灣某藥廠之人類肝臟用藥之銷售交易為資料來源，其中顧客為非訂立契約的醫院、診所與藥局。資料時間為2008年1月至2009年2月，總共60週之顧客交易資料。我們將此資料分為前30週與後30週，前30週資料用以建立模型，後30週資料作以檢驗模型預測結果。前30週資料之顧客數為953筆，總交易次數為2176筆。

1.3 研究內容

傳統的顧客分析預測模型中，多是透過顧客歷史交易資料直接建立預測模型，例如迴歸模型，並對未來之購買行為作預測；本研究採用概似函數模型，其探討模式之不同處為透過顧客歷史之交易資料，求得潛在特徵值θ，再根據此特徵值θ對未來交易行為作預測。

傳統研究

未來交易預測＝f （過去交易資料）



概似函數模型研究

$\hat{\theta} = f$(過去交易資料) $\rightarrow$ 未來交易預測 $= f(\hat{\theta})$

資料來源：Fader, P.S. and B. G. S. Hardie, *Probability Models for Customer-Base Analysis,* 2006, pp.3

圖 1 概似函數模型與傳統模型之比較圖

本研究所採用的模型為，2005年由Fader、Hardie 和Lee三位學者提出之BG/NBD模型，此模型為Pareto/NBD 模型(Schmittlein, Morrison and Colombo，1987)之進化模型。BG/NBD模型主要能根據每位顧客的歷史交易資料，根據顧客的重複購買次數、最後一次的購買時間(t_x)與總購買歷史時間(T)。我們採用(x,t_x,T)來標記，x為在觀察期間$(0,T]$所觀察到的重複購買數量，t_x為此段時間內最後一次交易時間，而$(0<t_x \le T]$，即可對顧客未來之購買次數作預測，亦可計算出個別顧客活躍機率。根據顧客個別活躍機率配合顧客總金額，可以辨認即將流失的關鍵顧客。

1.4 研究目的

本篇研究將這些無訂立契約之客戶，透過BG/NBD模型估計出顧客未來購買次數與顧客的個別活躍機率，以利行銷人員制定一對一的行銷策略與計畫，來提高顧客忠誠度，進而增加顧客權益，以提升流失管理(churn management)效率。

二、文獻回顧

2.1 顧客權益(customer equity)

顧客關係管理的目的是為了產生更高的顧客權益(customer equity)，根據 Rust、Lemon 與 Zeithaml 學者的觀點，顧客權益是所有顧客的總終身價值的折現值，而三項驅動顧客權益的來源為：價值權益(value equity)、品牌權益(brand equity)與關係權益(relationship equity)(Rust, Lemon, & Zeithaml, 2005)。上述三種驅動來源會分別形成以顧客為中心的價值管理、品牌管理與關係管理。以本文之所採用的人類肝臟用藥資料，是以關係權益為主的例子。以顧客權益的觀點，決策人員會透過財務指標，如投資報酬率，來檢驗與重新制定更適宜的策略與行銷方案(Rust, Lemon, & Zeithaml, 2004)。

Hagan、Lemon 與 Rust 也說明顧客權益為公司的策略性資產，是行銷領域的新方向，過去總把行銷費用支出視為短期支出，透過顧客權益觀點後，行銷費用被視為長期投資以創造企業長期利潤，最終回饋股東利益(Hagan, Lemon, & Rust, 2002；Rust, Lemon, & Zeithaml, 2004)。

另一種顧客權益的看法是來自 Blattberg 等學者所提出的驅動模式，透過顧客取得(acquisition)、顧客保留(retention)與附加銷售(add-on selling)三種行銷活動來極大化顧客終身價值(CLV, customer time value)，上述活動皆須考慮支出與回饋(Blattberg, Getz, & Thomas, 2001)。而本研究即採用其顧客保留(retention)之觀點，以適當的行銷活動，如甦醒顧客方案，將有價值的顧客保留以創造公司長期利潤並增加顧客權益。

2.2 顧客流失(churn)管理

如何降低關鍵顧客流失是一個重要的課題，當重要顧客流失時，不僅公司會減少收益，也增加了競爭者的市佔率與顧客所帶來的附加效益，例如口耳相傳(word of mouth)的效果。因此，在顧客關係管理中，有效降低顧客流失是一個重要的行銷活動。

關於顧客流失(churn)與保留(retention)的文獻中，有些目的在建立預測模型，並以預測準確度為目標(Lemmens & Croux, 2006; Neslin, Gupta, Kamakura, Lu, Mason, 2006)。除了探討顧客模型也研究導致顧客流失的因素或之間的影響，如在訂立契約型態的顧客中， Schweidel、Fader 和 Bradlow 提出一個顧客保留(retention)模型說明:1.持續(duration)的相關性 2.宣傳效果 3.用戶異質性 4.群眾互相影響(cross-cohort)5.日期影響(Schweidel, Fader, & Bradlow, 2008)。2003 年 Reinartz 和 Kumar，將顧客的獲利性(profitability)納入顧客關係持續時間(lifetime duration)的計算中(Reinartz & Kumar, 2003)。本研究採用相似的想法，將原始顧客關係持續時間模型上為考慮的獲利性納入檢驗。

2.3 一對一行銷

根據Peppers等學者對一對一行銷的觀點，一對一行銷指的是透過顧客表達或者其資料顯示的訊息，公司根據這些訊息去做適當的行銷活動。這是以顧客導向的觀點。過去市場導向或產品導向的方式，會跟據顧客的背景資料做分群，找出適合產品特性的目標族群，再針對目標顧客設計行銷計畫，希望獲得市場佔有率；當一對一行銷時，透過與顧客接觸的第一線接觸，如通路，接收顧客的個別訊息，並透過這些個別訊息制訂個別計畫，以期望獲得顧客佔有率（Peppers, Rogers, & Dorf, 1999）。

2.4 NBD 模型

Ehrenberg(1959)提出負二項分配(Negative Binomial Distribution)模型來估計與預測顧客的購買次數,此篇文章奠定了未來學者對顧客消費行為探討的文獻基礎。

2.4.1 模型假設

假設顧客隨機購買頻率為平均交易率λ,且λ隨著每位顧客而變動。

令隨機變數$X(t)$為在t時間內之交易量。而顧客個體之交易量$X(t)$假設為Poisson分配,其平均為λt:

$$P(X(t)=x\mid\lambda)=\frac{(\lambda t)^{x}e^{-\lambda t}}{x!}$$

而母體間的交易率(λ)則隨著Gamma分配:

$$g(\lambda\mid r,\alpha)=\frac{\alpha^{r}\lambda^{r-1}e^{-\alpha\lambda}}{\Gamma(r)}$$

而隨機選擇之顧客個體之交易量分配為:

$$P(X(t)=x\mid r,\alpha)=\int_{0}^{\infty}P(x(t)=x\mid\lambda)g(\lambda)d\lambda=\frac{\Gamma(r+x)}{\Gamma(r)x!}(\frac{\alpha}{\alpha+t})^{r}(\frac{t}{\alpha+t})^{x}$$

此即為負二項分配(NBD; Negative binomial distribution)。

2.4.2 期望交易量

我們希望能根據過去顧客購買歷史資料,來求得在T時附近$(T,T+t]$之期望交易量$E(Y(t)\mid data)$。對於NBD而言,最直接的應用即為Bayes理論:

$$E[Y(t)\mid r,\alpha,x,T]=(\frac{r+x}{\alpha+T})t$$

在交易模型建立中,我們可分成兩區的過程,交易過程(Transaction Process)與退出過程(Dropout Process):

交易過程:當顧客仍存活時,顧客之交易會隨機地在其平均交易率附近出現。且交易率(Transaction rates)會隨著每位顧客間不同。

退出過程：每位顧客皆有一段無法觀察之存活時間。且退出率(Dropout rates)會隨著每位顧客間不同。

2.5 Pareto/NBD 模型

Schmittlein, Morrison and Colombo，1987年提出Pareto/NBD 模型，在交易模型建立中，依然分成兩區的過程，交易過程(Transaction Process)與退出過程(Dropout Process)。

2.5.1 模型假設

1. 當顧客存活時，其顧客交易量會隨著交易率 λ 而Poisson分配。
2. 顧客之間交易率的異質性為Gamma分配，隨著形狀參數(shape parameter) r，與測量單位參數(scale parameter) α，$Gamma(r,\alpha)$
3. 每位顧客存在一段無法觀察之存活時間 τ，這期間內的每個時間點顧客顯的不活躍，因此其退出率 μ，且退出率 μ 為指數分配。
4. 顧客間之退出率的異質性為Gamma分配，隨著形狀參數(shape parameter) s，與測量單位參數(scale parameter) β，$Gamma(s,\beta)$。
5. 交易率 λ 與退出率 μ 在顧客間獨立變化。

2.5.2 模型概似函數之推導

假設每位顧客的 x 次交易發生在 $(0,T]$ 期間內，其標記為 $t_1, t_2 \ldots,\ t_x$。

有兩種可能會導致交易次數提升：

1. 觀察期結束時，此顧客仍然存活，也就是 $\tau > T$。
2. 在 τ 時點時，顧客即失去活性，而 $\tau \in (t_x, T]$。

當條件落於λ，其模型概似函數為：

$$L(\lambda \mid t_1,\ldots,t_x,T,\tau>T)=\lambda e^{-\lambda t_1}\lambda e^{-\lambda(t_2-t_1)}\ldots\lambda e^{-\lambda(t_x-t_{x-1})}e^{-\lambda(T-t_x)}=\lambda^x e^{-\lambda T}$$

$$L(\lambda \mid t_1,\ldots,t_x,T,\tau\in(t_x,T])=\lambda e^{-\lambda t_1}\lambda e^{-\lambda(t_2-t_1)}\ldots\lambda e^{-\lambda(t_x-t_{x-1})}e^{-\lambda(\tau-t_x)}=\lambda^x e^{-\lambda \tau}$$

當移除τ之條件，其模型概似函數為：

$$L(\lambda,\mu \mid x,t_x,T)=L(\lambda \mid x,T,\tau>T)P(\tau>T \mid \mu)+\int_{t_x}^{T}L(\lambda \mid x,T,\tau\in(t_x,T])f(\tau \mid \mu)d\tau$$
$$=\frac{\lambda^x\mu}{\lambda+\mu}e^{-(\lambda+\mu)t_x}+\frac{\lambda^{x+1}}{\lambda+\mu}e^{-(\lambda+\mu)T}$$

我們並不需要得知每個交易發生的時點，我們只需要得知每個顧客的最近一次的購買時間(recency)與購買頻率(frequency)。我們採用(x,t_x,T)來標記，x為在觀察期間$(0,T]$所觀察到的交易數量，t_x為此段時間內最後一次交易時間，$(0<t_x\le T]$。

高斯超幾何函數(Gaussian Hypergeometric Function)

$$_2F_1(a,b;c,;z)=\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\sum_{j=0}^{\infty}\frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)}\frac{z^j}{j!}$$

利用尤拉積分展開為：

$$_2F_1(a,b;c,;z)=\frac{1}{B(b,c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a}dt$$

其中$c>b>0$且B為Beta函數。

2.5.3 Pareto/NBD 概似函數

若$\alpha \geq \beta$，

$$L(r,\alpha,s,\beta \mid x,t_x,T)$$

$$=\frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)}\left\{\left(\frac{r+x}{r+s+x}\right)\frac{{}_2F_1(r+s+x,s+1;r+s+x+1;\frac{\alpha-\beta}{\alpha+t_x})}{(\alpha+t_x)^{r+s+x}}+\left(\frac{r+x}{r+s+x}\right)\frac{{}_2F_1(r+s+x,s;r+s+x+1;\frac{\alpha-\beta}{\alpha+T})}{(\alpha+t_x)^{r+s+x}}\right\}$$

若$\alpha \leq \beta$，$L(r,\alpha,s,\beta \mid x,t_x,T)$

$$=\frac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)}\left\{\left(\frac{r+x}{r+s+x}\right)\frac{{}_2F_1(r+s+x,r+x;r+s+x+1;\frac{\beta-\alpha}{\beta+t_x})}{(\beta+t_x)^{r+s+x}}+\left(\frac{r+x}{r+s+x}\right)\frac{{}_2F_1(r+s+x,r+x+1;r+s+x+1;\frac{\beta-\alpha}{\beta+T})}{(\beta+T)^{r+s+x}}\right\}$$

2.5.4 Pareto/NBD 小結

Pareto/NBD 能預測：

1. $E[X(t)]$，為在$(o,t]$之交易量期望值：

$$E[X(t)\mid r,\alpha,s,\beta]=\int_0^\infty\int_0^\infty E[X(t)\mid\lambda,\mu]g(\lambda\mid r,\alpha)g(\mu\mid s,\beta)d\lambda d\mu=\frac{r\beta}{\alpha(s-1)}\left[1-\left(\frac{\beta}{\beta+t}\right)^{s-1}\right]$$

2. $P(alive)$ 為每個顧客其伴隨著過去購買資料在T時仍存活的機率；也就是 $P(\tau > T)$，τ為顧客失去活性的時間點；即為顧客活躍機率。

$$P(alive\mid\lambda,\mu,x,t_x,T)=\frac{\dfrac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)(\alpha+T)^{r+x}(\beta+T)^s}}{L(r,\alpha,s,\beta\mid x,t_x,T)}$$

3. $E(Y(t)\mid x,t_x,T)$，在$(T,T+t]$期間，帶著(x,t_x,T)歷史交易資料之個體條件期望交易量。

令$Y(t)$為$(T,T+t]$期間之購買次數，則.

$$E(Y(t)\mid\lambda,\mu,\tau>T)=\lambda t P(\tau>T+t\mid\mu,\tau>T)+\int_T^{T+t}\lambda\tau f(\tau\mid\mu,\tau>T)d\tau=\frac{\lambda}{\mu}-\frac{\lambda}{\mu}e^{-\mu t}$$

而 $E(Y(t)\mid\lambda,\mu,x,t_x,T)=E(Y(t)\mid\lambda,\mu,\tau>T)\times P(\tau>T\mid\lambda,\mu,x,t_x,T)$

將 $P(alive\mid\lambda,\mu,x,t_x,T)=\dfrac{\dfrac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)(\alpha+T)^{r+x}(\beta+T)^s}}{L(r,\alpha,s,\beta\mid x,t_x,T)}$ 和 λ 與 μ 之共同後驗機率分配(Joint posterior distribution)代入，

$$E(Y(t)\mid\lambda,\mu,x,t_x,T)=\frac{\dfrac{\Gamma(r+x)\alpha^r\beta^s}{\Gamma(r)(\alpha+T)^{r+x}(\beta+T)^s}}{L(r,\alpha,s,\beta\mid x,t_x,T)}\times\frac{(r+x)(\beta+T)}{(\alpha+T)(s-1)}\left[1-\left(\frac{\beta+T}{\beta+T+t}\right)^{s-1}\right]$$

2.6 BG/NBD 模型

2005年Fader，Hardie 和Lee三位學者提出BG/NBD(Beta-Geometric/NBD)模型，此模型有效解決原本Pareto/NBD模型計算不易的問題，僅以Excel軟體即可計算出顧客未來之購買次數。

2.6.1 模型假設

1. 當顧客具備活性時，其顧客交易量會隨著交易率 λ 而Poisson分配。也就是說在兩交易時間中交易率是隨著指數分配：

   $$f(t_j \mid t_{j-1};\lambda) = \lambda e^{-\lambda(t_j - t_{j-1})}, t_j > t_{j-1} \geq 0$$

2. 顧客之間交易率的異質性為Gamma分配：

   $$f(\lambda \mid r,\alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda\alpha}}{\Gamma(r)}, \lambda > 0$$

3. 在每次交易後，顧客會變得不具備活躍性的機率為 p ，因此顧客退出時間點是隨著幾何分配，其機率質量函數(pmf)為：

   P(在第j次交易後，立即不活躍的退出機率)= $p(1-p)^{j-1}$, j=1, 2, 3…

4. 不活躍機率 p 的異質性服從Beta分配，其機率密度函數：

   $$f(p \mid a,b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, 0 \leq p \leq 1，B(a,b)=\frac{\Gamma(\mathrm{a})\Gamma(\mathrm{b})}{\Gamma(\mathrm{a+b})}。$$

5. 顧客交易率 λ 與退出率 p 在顧客間獨立變化。

2.6.2 模型概似函數之推導

假設每位顧客的 x 次交易發生在 $(0,T]$ 期間內，其標記為 $t_1, t_2 \ldots,\ t_x$ 。

有兩種可能會導致交易次數提升：

1. 觀察期結束時，此顧客仍然存活，也就是 $\tau > T$ 。

2. 在第x次交易後，顧客立即失去活性，也就是$\tau=t_x$。

當條件落於λ，其模型概似函數為：

$$L(\lambda \mid t_1,...,t_x,T,\tau>T)=\lambda e^{-\lambda t_1}\lambda e^{-\lambda(t_2-t_1)}...\lambda e^{-\lambda(t_x-t_{x-1})}e^{-\lambda(T-t_x)}=\lambda^x e^{-\lambda T}$$

$$L(\lambda \mid t_1,...,t_x,T,\tau=T)=\lambda e^{-\lambda t_1}\lambda e^{-\lambda(t_2-t_1)}...\lambda e^{-\lambda(t_x-t_{x-1})}=\lambda^x e^{-\lambda t_x}$$

當移除τ之條件時，其模型概似函數為：

$$L(\lambda,p \mid x,t_x,T)=L(\lambda \mid x,T,\tau>T)P(\tau>T \mid p)+L(\lambda \mid x,T,\tau=t_x)P(\tau=t_x \mid p)$$

$=(1-p)^x\lambda^x e^{-\lambda T}+\delta_{x>0}p(1-p)^{x-1}\lambda^x e^{-\lambda t_x}$，若$x>0$，則$\delta_{x>0}=1$，否則$\delta_{x>0}=0$。

當移除λ和p的條件時：

$$\begin{aligned}
&L(r,\alpha,a,b \mid x,t_x,T)\\
&=\int_0^1\int_0^\infty L(\lambda,p \mid x,t_x,T)f(\lambda \mid r,\alpha)f(p \mid a,b)d\lambda dp\\
&=\int_0^1\int_0^\infty (1-p)^x\lambda^x e^{-\lambda T}\frac{\alpha^r\lambda^{r-1}e^{-\lambda\alpha}}{\Gamma(r)}\frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}d\lambda dp\\
&+\delta_{x>0}\int_0^1\int_0^\infty\left\{p(1-p)^{x-1}\lambda^x e^{-\lambda t_x}\frac{\alpha^r\lambda^{r-1}e^{-\lambda\alpha}}{\Gamma(r)}\frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}\right\}d\lambda dp
\end{aligned}$$

2.6.3 BG/NBD模型概似函數

因此，我們可以表示BG/NBD模型之概似函數為

$$L(r,\alpha,a,b \mid x,t_x,T)=A_1A_2(A_3+\delta_{x>0}A_4)$$

其中

$$A_1=\frac{\Gamma(r+x)\alpha^r}{\Gamma(r)},A_2=\frac{\Gamma(a+b)\Gamma(b+x)}{\Gamma(b)\Gamma(a+b+x)},A_3=\left(\frac{1}{\alpha+T}\right)^{r+x},A_4=\frac{a}{b+x-1}\left(\frac{1}{\alpha+t_x}\right)^{r+x}$$

2.6.4 顧客活躍機率

顧客活躍機率其定義為此顧客在 t_x 不活躍之機率加上此段時間間隔內仍然活躍卻未買的機率：

$$P(alive \mid x, t_x, T, r, \alpha, a, b) = \left. \frac{B(a,b+x)}{B(a,b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)(\alpha+T)^{r+x}} \middle/ L(r,\alpha,a,b \mid x,t_x,T) \right.$$

$$= 1 \Big/ \left\{ 1 + \delta_{x>0} \frac{a}{b+x-1} \left(\frac{\alpha+T}{\alpha+t_x} \right)^{r+x} \right\}$$

但根據此方法，當顧客在 $(0,T]$ 期間之重複購買次數為0的時候，其存活機率=1。這事依循此模型之假設，當顧客剛完成購買行為時，視此顧客為死亡狀態，而在觀察階段的初始期，所有顧客為存活狀態(alive)。為了解決此問題，有以下兩個變化型的方式來解決。(Fader, Hardie, & Lee, 2008)

i. 1969年Morrison提出的方法。他鬆綁了全部顧客再觀察的初期階為存活狀態的假設。其假設於觀察初期於全部顧客有 $\pi \times 100\%$ 為死亡狀態，剩下 $(1-\pi) \times 100\%$ 在BG/NBD中為存活。

則此時之顧客活躍機率為：

$$P(alive \mid x, t_x, T, r, \alpha, a, b, \pi) = \begin{cases} \dfrac{1-\pi}{\pi(\frac{\alpha+T}{\alpha})^r + 1 - \pi}, x = 0 \\ \dfrac{1}{1 + \frac{a}{b+x-1}(\frac{\alpha+T}{\alpha+t_x})^{r+x}}, x > 0 \end{cases}$$

ii. 另一個解決方式，將在0時，此顧客之死亡機率為丟擲骰子決定，而此時之BG/NBD模型之概似函數為：

$$L(r,\alpha,a,b \mid x,t_x,T) = \frac{B(a,b+x+1)}{B(a,b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)(\alpha+T)^{r+x}} + \frac{B(a+1,b+x)}{B(a,b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)(\alpha+t_x)^{r+x}}$$

則其顧客活躍機率為：

$$P(alive \mid x, t_x, T, r, \alpha, a, b, \pi) = \frac{1}{1 + \frac{a}{b+x}\left(\frac{\alpha+T}{\alpha+t_x}\right)^{r+x}}$$

三、研究方法

本研究之資料來源為台灣某藥廠之人類肝臟用藥之銷售資料，其交易時間為2008年1月到2009年2月之間共60週的交易資料。我們將此交易資料分成前後兩部分，前半段30週之顧客交易資料用以建立BG/NBD模型，後30週之資料作為檢驗模型預測結果之準確度分析。用以建立模型的前30週資料共有953個顧客，交易筆數為2176筆交易資料。

本研究之顧客鎖定為非契約型顧客，本文之非契約型顧客包含：無訂立契約之醫院、診所、藥局，而各類顧客之比例，醫院佔整體14%，診所47%，藥局39%。

此研究採用之BG/NBD模型，除了可利用顧客歷史交易資料來預測未來此顧客之購買次數外，利用整體顧客之概似函數極大值之方法求出的最適參數，利用此參數計算出個別顧客活躍機率。BG/NBD能有效解決Pareto/NBD之最大突破之處為其模型建構方法與工具之簡便，我們能利用Excel即能建立概似函數模型與求出最適參數，並作預測分析。再利用MAPE(mean absolute percentage error)來檢驗此模型所預測之結果與實際值之誤差比例。

3.1 資料整理

先將每位顧客資料整理成 $X=(x,t_x,T)$ 之模式，x 為在觀察期間 $(0,T]$ 所觀察到的交易數量，t_x 為此段時間內最後一次交易時間，$(0<t_x\leq T]$。T 為此顧客的總歷史交易時間。本次研究之時間單位為週次計算。如圖2，顧客(ID=1)其 $X=(x,t_x,T)=(4,\ 25.43,\ 28.4286)$，$T$＝28.4286表示其第一次購買時間在30－28.4286＝$1\frac{4}{7}$ 週時間，最後一次購買時間為25.43週的時間，此顧客之重複交易次數為4次，表示總購買次數為5次。我們可以發現ID={3, 5, 6, 7}之顧客，其重複交易次數為0，表示這些顧客並不再觀察期間內再次購買，故 $t_x=0$。整體資料中，無重複購買行為的顧客有493位，佔整體資料51.7%。

	A	B	C	D
1	ID	x	t_x	T
2	1	4	25.43	28.4286
3	2	1	12.29	27.7143
4	3	0	0.00	12.5714
5	4	1	16.43	26.4286
6	5	0	0.00	29.5714
7	6	0	0.00	29.4286
8	7	0	0.00	27.7143
9	8	2	26.14	25.4286
10	9	3	25.71	23.4286
11	10	5	26.29	27.2857
12	11	3	14.57	27.5714
13	12	0	0.00	21.2857
14	13	1	21.71	14.7143
15	14	2	27.43	8.85714
16	15	0	0.00	5.85714
17	16	0	0.00	14.8571
18	17	1	29.71	24.8571
19	18	2	29.14	25.7143
20	19	2	25.14	15.8571
21	20	0	0.00	21.5714
22	21	0	0.00	4.42857
951	950	0	0.00	18.8571
952	951	1	21.29	23.7143
953	952	1	11.43	26.2857
954	953	1	23.29	21.2857

圖 2 前 30 週顧客購買資料整理圖

3.2 BG/NBD 模型建立

3.2.1 建立 BG/NBD 對數概似函數

BG/NBD概似函數：

$$L(r,\alpha,a,b \mid x,t_x,T)=A_1A_2(A_3+\delta_{x>0}A_4) \qquad (1)$$

其中

$$A_1=\frac{\Gamma(r+x)\alpha^r}{\Gamma(r)},A_2=\frac{\Gamma(a+b)\Gamma(b+x)}{\Gamma(b)\Gamma(a+b+x)},A_3=\left(\frac{1}{\alpha+T}\right)^{r+x},A_4=\frac{a}{b+x-1}\left(\frac{1}{\alpha+t_x}\right)^{r+x}$$

我們為了概似函數之操作便利，我們將公式(1)等號兩邊取對數，則

$$\ln L(r,\alpha,a,b \mid x,t_x,T)=\ln\left[A_1A_2(A_3+\delta_{x>0}A_4)\right]=\ln A_1+\ln A_2+\ln(e^{\ln A_3}+\delta_{x>0}e^{\ln A_4}) \qquad (2)$$

其中，

$$\ln A_1=\Gamma\ln(r+x)-\Gamma\ln(r)+r\ln(\alpha)$$

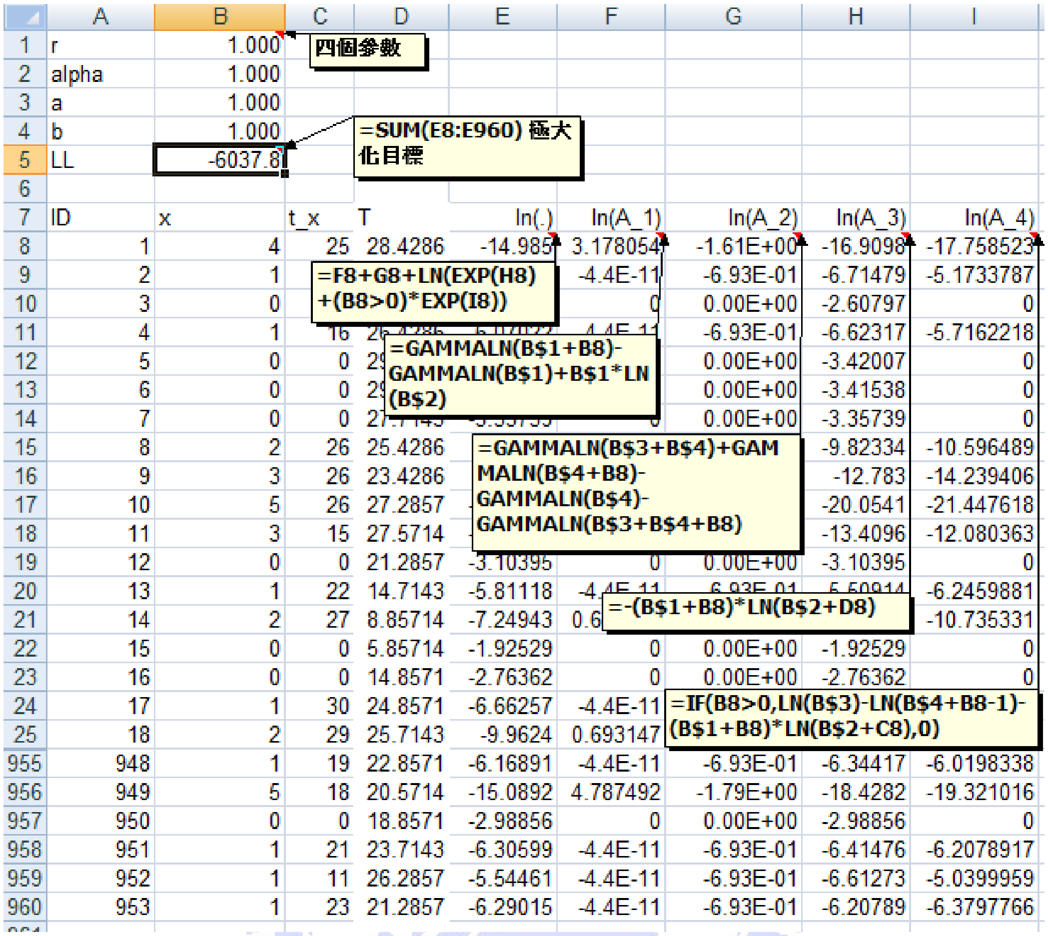
$$\ln A_2=\Gamma\ln(a+b)+\Gamma\ln(b+x)-\Gamma\ln(b)-\Gamma\ln(a+b+x)$$

$$\ln A_3=-(r+x)\ln(\alpha+T)$$

$$\ln A_4=\ln a-\ln(b+x-1)-(r+x)\ln(\alpha+t_x)$$

我們可以透過我們在圖2中的前30週顧客交易資料，利用Excel中表單的公式功能與規劃求解功能，來建立公式(2)。我們先設定四個參數之起始值=1。如圖3所示。而B5即為我們公式(2)之值，也是我們要透過四個最適參數達成極大化之目標。

因此每個顧客皆可求得一個對數概似函數ln(.)，再將所有的顧客概似函數作總合，即可求得此953顧客之整體對數概似函數 $LL(r,\alpha,a,b \mid x,t_x,T)$ ，對 $LL(r,\alpha,a,b \mid x,t_x,T)$ 求極大值，即可 r,α,a,b 獲得最佳估計值。

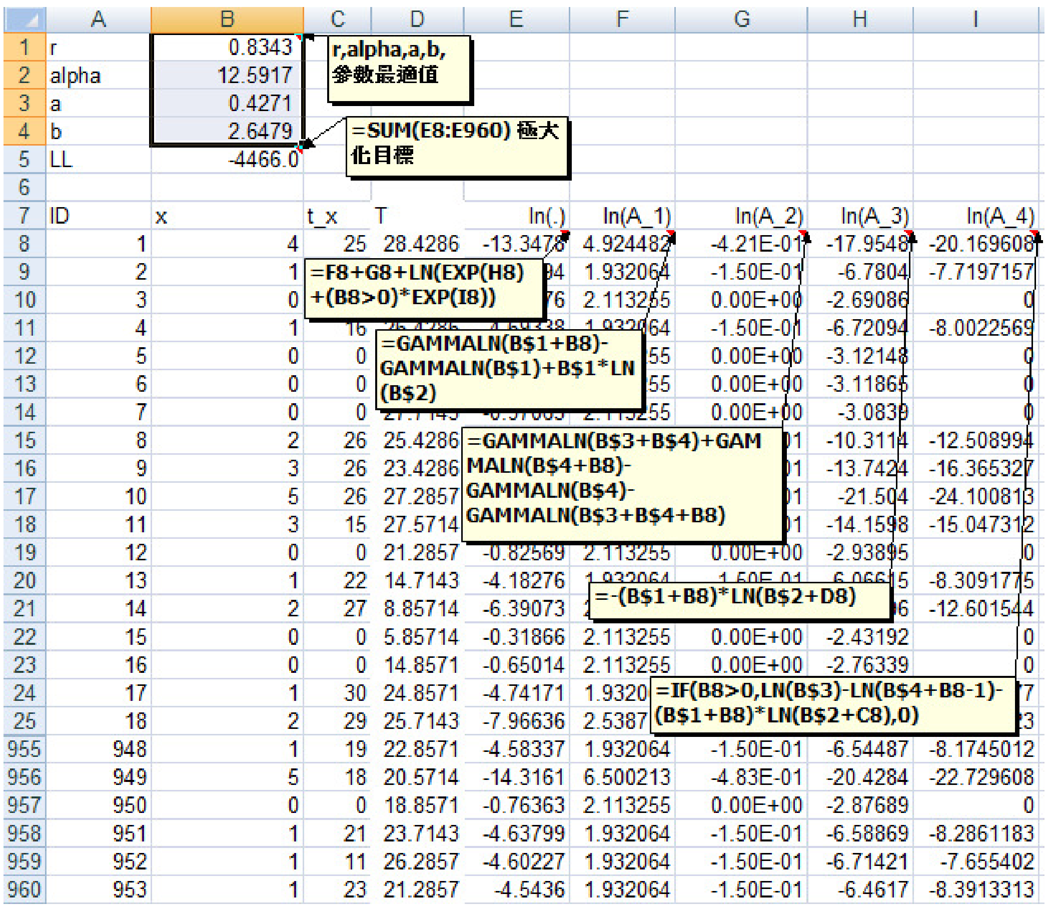


	A	B	C	D	E	F	G	H	I
1	r	1.000							
2	alpha	1.000							
3	a	1.000							
4	b	1.000							
5	LL	-6037.8							
6									
7	ID	x	t_x	T	ln(.)	ln(A_1)	ln(A_2)	ln(A_3)	ln(A_4)
8	1	4	25	28.4286	-14.985	3.178054	-1.61E+00	-16.9098	-17.758523
9	2	1	[illegible]	[illegible]	[illegible]	-4.4E-11	-6.93E-01	-6.71479	-5.1733787
10	3	0	[illegible]	[illegible]	[illegible]	0	0.00E+00	-2.60797	0
11	4	1	16	[illegible]	[illegible]	[illegible]	-6.93E-01	-6.62317	-5.7162218
12	5	0	0	[illegible]	[illegible]	[illegible]	0.00E+00	-3.42007	0
13	6	0	0	[illegible]	[illegible]	[illegible]	0.00E+00	-3.41538	0
14	7	0	0	[illegible]	[illegible]	0	0.00E+00	-3.35739	0
15	8	2	26	25.4286	[illegible]	[illegible]	[illegible]	-9.82334	-10.596489
16	9	3	26	23.4286	[illegible]	[illegible]	[illegible]	-12.783	-14.239406
17	10	5	26	27.2857	[illegible]	[illegible]	[illegible]	-20.0541	-21.447618
18	11	3	15	27.5714	[illegible]	[illegible]	[illegible]	-13.4096	-12.080363
19	12	0	0	21.2857	-3.10395	0	0.00E+00	-3.10395	0
20	13	1	22	14.7143	-5.81118	[illegible]	[illegible]	[illegible]	-6.2459881
21	14	2	27	8.85714	-7.24943	[illegible]	[illegible]	[illegible]	-10.735331
22	15	0	0	5.85714	-1.92529	0	0.00E+00	-1.92529	0
23	16	0	0	14.8571	-2.76362	0	0.00E+00	-2.76362	0
24	17	1	30	24.8571	-6.66257	-4.4E-11	[illegible]	[illegible]	[illegible]
25	18	2	29	25.7143	-9.9624	0.693147	[illegible]	[illegible]	[illegible]
955	948	1	19	22.8571	-6.16891	-4.4E-11	-6.93E-01	-6.34417	-6.0198338
956	949	5	18	20.5714	-15.0892	4.787492	-1.79E+00	-18.4282	-19.321016
957	950	0	0	18.8571	-2.98856	0	0.00E+00	-2.98856	0
958	951	1	21	23.7143	-6.30599	-4.4E-11	-6.93E-01	-6.41476	-6.2078917
959	952	1	11	26.2857	-5.54461	-4.4E-11	-6.93E-01	-6.61273	-5.0399959
960	953	1	23	21.2857	-6.29015	-4.4E-11	-6.93E-01	-6.20789	-6.3797766

圖 3 利用 Excel 表單建立概似函數公式

3.2.2 求參數之最適值

我們可以使用Excel之規劃求解的功能，求對數概似函數的極大值。對Excel表格B5位置之LL值設定為目標儲存格，變數儲存格為B1:B4，並保持變數大於0，即可得出LL最大值，與參數的最適值。可得 $r=0.8343, \alpha=12.5917, a=0.4271, b=2.6479$ 而 $LL(r,\alpha,a,b \mid x,t_x,T)=-4466.0$。如圖4所示。



	A	B	C	D	E	F	G	H	I
1	r	0.8343							
2	alpha	12.5917							
3	a	0.4271							
4	b	2.6479							
5	LL	-4466.0							
6									
7	ID	x	t_x	T	ln(.)	ln(A_1)	ln(A_2)	ln(A_3)	ln(A_4)
8	1	4	25	28.4286	-13.3478	4.924482	-4.21E-01	-17.9548	-20.169608
9	2	1	[illegible]	[illegible]	[illegible]	1.932064	-1.50E-01	-6.7804	-7.7197157
10	3	0	[illegible]	[illegible]	[illegible]	2.113255	0.00E+00	-2.69086	0
11	4	1	16	[illegible]	[illegible]	[illegible]	-1.50E-01	-6.72094	-8.0022569
12	5	0	0	[illegible]	[illegible]	[illegible]	0.00E+00	-3.12148	0
13	6	0	0	[illegible]	[illegible]	[illegible]	0.00E+00	-3.11865	0
14	7	0	0	[illegible]	[illegible]	[illegible]	0.00E+00	-3.0839	0
15	8	2	26	25.4286	[illegible]	[illegible]	[illegible]	-10.3114	-12.508994
16	9	3	26	23.4286	[illegible]	[illegible]	[illegible]	-13.7424	-16.365327
17	10	5	26	27.2857	[illegible]	[illegible]	[illegible]	-21.504	-24.100813
18	11	3	15	27.5714	[illegible]	[illegible]	[illegible]	-14.1598	-15.047312
19	12	0	0	21.2857	-0.82569	2.113255	0.00E+00	-2.93895	0
20	13	1	22	14.7143	-4.18276	[illegible]	[illegible]	[illegible]	-8.3091775
21	14	2	27	8.85714	-6.39073	[illegible]	[illegible]	[illegible]	-12.601544
22	15	0	0	5.85714	-0.31866	2.113255	0.00E+00	-2.43192	0
23	16	0	0	14.8571	-0.65014	2.113255	0.00E+00	-2.76339	0
24	17	1	30	24.8571	-4.74171	[illegible]	[illegible]	[illegible]	[illegible]
25	18	2	29	25.7143	-7.96636	[illegible]	[illegible]	[illegible]	[illegible]
955	948	1	19	22.8571	-4.58337	1.932064	-1.50E-01	-6.54487	-8.1745012
956	949	5	18	20.5714	-14.3161	6.500213	-4.83E-01	-20.4284	-22.729608
957	950	0	0	18.8571	-0.76363	2.113255	0.00E+00	-2.87689	0
958	951	1	21	23.7143	-4.63799	1.932064	-1.50E-01	-6.58869	-8.2861183
959	952	1	11	26.2857	-4.60227	1.932064	-1.50E-01	-6.71421	-7.655402
960	953	1	23	21.2857	-4.5436	1.932064	-1.50E-01	-6.4617	-8.3913313

圖 4 利用 Excel 求參數之最適值

3.2.3 個別顧客之未來期望購買次數預測

對於每個顧客，我們都能透過之前所獲得之四個最適參數值 $r=0.8343, \alpha=12.5917, a=0.4271, b=2.6479$ 與每位顧客之購買資料 (x, t_x, T) 來預測本顧客之未來購買次數 $E(Y(t)\,|\,x, t_x, T)$

$$E(Y(t)\,|\,r,\alpha,a,b,x,t_x,T)=\frac{\dfrac{a+b+x-1}{a-1}\left[1-\left(\dfrac{\alpha+T}{\alpha+T+t}\right)^{r+x}{}_2F_1(r+x,b+x;a+b+x-1;\dfrac{t}{\alpha+T+t})\right]}{1+\delta_{x>0}\dfrac{a}{b+x-1}\left(\dfrac{\alpha+T}{\alpha+t_x}\right)^{r+x}} \tag{3}$$

首先先建立 ${}_2F_1(r+x,b+x;a+b+x-1;\dfrac{t}{\alpha+T+t})$ 高斯超幾何函數(Gaussian hyper-geometric function)。

此函數為冪級數，其型式為：

$${}_2F_1(a,b;c;z)=\sum_0^\infty \frac{(a)_j(b)_j}{(c)_j}\frac{z^j}{j!}, c\neq 0,-1,-2,...,$$

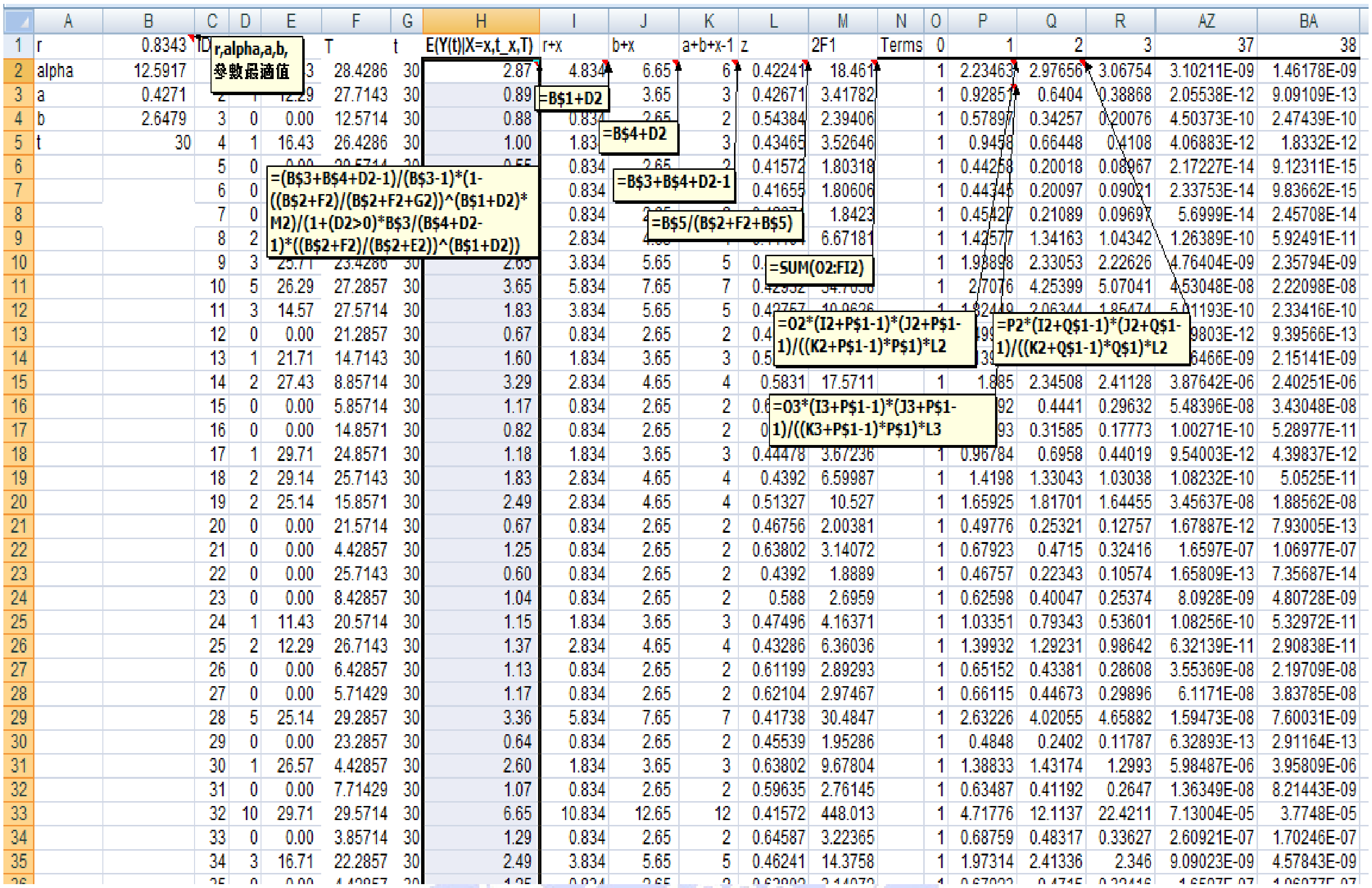
其 $(a)_j$ 為降乘冪符號，$(a)_j=a(a+1)(a+2)...(a+j-1)=\dfrac{\Gamma(a+j)}{\Gamma(a)}$。當|z|<1，則級數收斂；當|z|>1，則此級數發散；當|z|=1時，此級數收斂於c-a-b>0。

$${}_2F_1(a,b;c;z)=\sum_0^\infty u_j \text{，其中 } u_j=\frac{(a)_j(b)_j}{(c)_j}\frac{z^j}{j!} \text{。}$$

我們可用遞迴關係式展開：

$$\frac{u_j}{u_{j-1}}=\frac{(a+j-1)(b+j-1)}{c+j-1}z, j=1,2,3,...,u_0=1$$

在本次研究中，在 ${}_2F_1$ 我們計算前38項(j=38)，因38項之數列對整體之預測只有小數點後四位之差異，而預測次數為整數值，故此數列僅計算至第38項。如圖5所示，對於每位顧客都能根據公式(3)來建立其未來預測購買次數期望值，也就是圖5中的H欄。



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	AZ	BA
1	r	0.8343	ID	[illegible]	[illegible]	T	t	E(Y(t)|X=x,t_x,T)	r+x	b+x	a+b+x-1	z	2F1	Terms	0	1	2	3	37	38
2	alpha	12.5917	[illegible]	[illegible]	[illegible]	28.4286	30	2.87	4.834	6.65	6	0.42241	18.461		1	2.23463	2.97656	3.06754	3.10211E-09	1.46178E-09
3	a	0.4271	2	1	12.29	27.7143	30	0.89	[illegible]	3.65	3	0.42671	3.41782		1	0.92851	0.6404	0.38868	2.05538E-12	9.09109E-13
4	b	2.6479	3	0	0.00	12.5714	30	0.88	0.834	2.65	2	0.54384	2.39406		1	0.57897	0.34257	0.20076	4.50373E-10	2.47439E-10
5	t	30	4	1	16.43	26.4286	30	1.00	1.834	[illegible]	3	0.43465	3.52646		1	0.9458	0.66448	0.4108	4.06883E-12	1.8332E-12
6			5	0	[illegible]	[illegible]	[illegible]	[illegible]	0.834	2.65	2	0.41572	1.80318		1	0.44258	0.20018	0.08967	2.17227E-14	9.12311E-15
7			6	0	[illegible]	[illegible]	[illegible]	[illegible]	0.834	[illegible]	[illegible]	0.41655	1.80606		1	0.44345	0.20097	0.09021	2.33753E-14	9.83662E-15
8			7	0	[illegible]	[illegible]	[illegible]	[illegible]	0.834	[illegible]	[illegible]	[illegible]	1.8423		1	0.45427	0.21089	0.09697	5.6999E-14	2.45708E-14
9			8	2	[illegible]	[illegible]	[illegible]	[illegible]	2.834	[illegible]	[illegible]	[illegible]	6.67181		1	1.42577	1.34163	1.04342	1.26389E-10	5.92491E-11
10			9	3	[illegible]	[illegible]	[illegible]	[illegible]	3.834	5.65	5	[illegible]	[illegible]		1	1.93898	2.33053	2.22626	4.76404E-09	2.35794E-09
11			10	5	26.29	27.2857	30	3.65	5.834	7.65	7	[illegible]	[illegible]		1	2.7076	4.25399	5.07041	4.53048E-08	2.22098E-08
12			11	3	14.57	27.5714	30	1.83	3.834	5.65	5	0.42757	10.9626		1	1.82449	2.06344	1.85474	5.01193E-10	2.33416E-10
13			12	0	0.00	21.2857	30	0.67	0.834	2.65	2	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	9.39566E-13
14			13	1	21.71	14.7143	30	1.60	1.834	3.65	3	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.15141E-09
15			14	2	27.43	8.85714	30	3.29	2.834	4.65	4	0.5831	17.5711		1	1.885	2.34508	2.41128	3.87642E-06	2.40251E-06
16			15	0	0.00	5.85714	30	1.17	0.834	2.65	2	[illegible]	[illegible]		[illegible]	[illegible]	0.4441	0.29632	5.48396E-08	3.43048E-08
17			16	0	0.00	14.8571	30	0.82	0.834	2.65	2	[illegible]	[illegible]		[illegible]	[illegible]	0.31585	0.17773	1.00271E-10	5.28977E-11
18			17	1	29.71	24.8571	30	1.18	1.834	3.65	3	0.44478	3.67236		1	0.96784	0.6958	0.44019	9.54003E-12	4.39837E-12
19			18	2	29.14	25.7143	30	1.83	2.834	4.65	4	0.4392	6.59987		1	1.4198	1.33043	1.03038	1.08232E-10	5.0525E-11
20			19	2	25.14	15.8571	30	2.49	2.834	4.65	4	0.51327	10.527		1	1.65925	1.81701	1.64455	3.45637E-08	1.88562E-08
21			20	0	0.00	21.5714	30	0.67	0.834	2.65	2	0.46756	2.00381		1	0.49776	0.25321	0.12757	1.67887E-12	7.93005E-13
22			21	0	0.00	4.42857	30	1.25	0.834	2.65	2	0.63802	3.14072		1	0.67923	0.4715	0.32416	1.6597E-07	1.06977E-07
23			22	0	0.00	25.7143	30	0.60	0.834	2.65	2	0.4392	1.8889		1	0.46757	0.22343	0.10574	1.65809E-13	7.35687E-14
24			23	0	0.00	8.42857	30	1.04	0.834	2.65	2	0.588	2.6959		1	0.62598	0.40047	0.25374	8.0928E-09	4.80728E-09
25			24	1	11.43	20.5714	30	1.15	1.834	3.65	3	0.47496	4.16371		1	1.03351	0.79343	0.53601	1.08256E-10	5.32972E-11
26			25	2	12.29	26.7143	30	1.37	2.834	4.65	4	0.43286	6.36036		1	1.39932	1.29231	0.98642	6.32139E-11	2.90838E-11
27			26	0	0.00	6.42857	30	1.13	0.834	2.65	2	0.61199	2.89293		1	0.65152	0.43381	0.28608	3.55369E-08	2.19709E-08
28			27	0	0.00	5.71429	30	1.17	0.834	2.65	2	0.62104	2.97467		1	0.66115	0.44673	0.29896	6.1171E-08	3.83785E-08
29			28	5	25.14	29.2857	30	3.36	5.834	7.65	7	0.41738	30.4847		1	2.63226	4.02055	4.65882	1.59473E-08	7.60031E-09
30			29	0	0.00	23.2857	30	0.64	0.834	2.65	2	0.45539	1.95286		1	0.4848	0.2402	0.11787	6.32893E-13	2.91164E-13
31			30	1	26.57	4.42857	30	2.60	1.834	3.65	3	0.63802	9.67804		1	1.38833	1.43174	1.2993	5.98487E-06	3.95809E-06
32			31	0	0.00	7.71429	30	1.07	0.834	2.65	2	0.59635	2.76145		1	0.63487	0.41192	0.2647	1.36349E-08	8.21443E-09
33			32	10	29.71	29.5714	30	6.65	10.834	12.65	12	0.41572	448.013		1	4.71776	12.1137	22.4211	7.13004E-05	3.7748E-05
34			33	0	0.00	3.85714	30	1.29	0.834	2.65	2	0.64587	3.22365		1	0.68759	0.48317	0.33627	2.60921E-07	1.70246E-07
35			34	3	16.71	22.2857	30	2.49	3.834	5.65	5	0.46241	14.3758		1	1.97314	2.41336	2.346	9.09023E-09	4.57843E-09

圖 5 利用 Excel 作個別顧客之未來購買次數之預測

3.2.4 對顧客購買預測值與實際值作誤差計算

利用平均絕對百分誤MAPE來計算此預測是否為良好預測。

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

因後30週無實質購買之顧客為377筆，若採用MAPE之方式，其 $A_t = 0$。故此部份之顧客我們採用*RMSE*方法來檢驗這部分之預測誤差。

$$RMSE = \left(\frac{1}{n}\sum_{t=0}^{n}\left(A_t - F_t\right)^2\right)^{\frac{1}{2}}$$

3.3 顧客活躍機率

顧客活躍機率其定義為此顧客在 t_x 不活躍之機率加上此段時間間隔內仍然活躍卻未發生購買行為的機率：

$$
\begin{aligned}
P(alive \mid x, t_x, T, r, \alpha, a, b) &= \left. \frac{B(a, b+x)}{B(a,b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)(\alpha+T)^{r+x}} \middle/ L(r, \alpha, a, b \mid x, t_x, T) \right. \\
&= 1 \middle/ \left\{ 1 + \delta_{x>0} \frac{a}{b+x-1} \left(\frac{\alpha+T}{\alpha+t_x} \right)^{r+x} \right\}
\end{aligned} \tag{4}
$$

但根據此方法，當顧客在 $(0,T]$ 期間之重複購買次數為0的時候，其存活機率=1。如圖6中之K欄。

為了避免當顧客在 $(0,T]$ 期間之重複購買次數為0的時候，其存活機率=1，我們採用2008年12月Fadar所提出的新方法，此方法鬆綁了顧客在觀察初期全部存活狀態的假設，且將顧客活躍機率作調整。如圖6中之J欄。

則此時之顧客活躍機率為：

$$
P(alive \mid x, t_x, T, r, \alpha, a, b, \pi) = \frac{1}{1 + \frac{a}{b+x} \left(\frac{\alpha+T}{\alpha+t_x} \right)^{r+x}} \tag{5}
$$

=1/(1+(B$3/(B$4+B8))*(((B$2+D8)/(B$2+C8))^(B$1+B8)))

=exp(G8+H9+I9-F9)

=F8+G8+H8-E8

	A	B	C	D	E	F	G	H	I	J	K	L
1	r	0.8343										
2	alpha	12.5917										
3	a	0.4271										
4	b	2.6479										
5	LL	-4465.97										
6												
7	ID	x	t_x	T	ln	ln(A_	ln(A_	ln(A_	ln(A_	P(ali	Palive(o	)+ln(A3)-ln(A
8	1	4	25	28.42857	-13.3478	4.924482	-4.21E-01	-17.9548	-20.1696	0.9151	0.9016	-1.04E-01
9	2	1	12	27.71429	-4.66794	1.932064	-1.50E-01	-6.7804	-7.71972	0.7790	0.7190	-3.30E-01
10	3	0	0	12.57143	-0.5776	2.113255	0.00E+00	-2.69086	0	0.7767	1.0000	0.00E+00
11	4	1	16	26.42857	-4.69338	1.932064	-1.50E-01	-6.72094	-8.00226	0.8323	0.7827	-2.45E-01
12	5	0	0	29.57143	-1.00823	2.113255	0.00E+00	-3.12148	0	0.6934	1.0000	0.00E+00
13	6	0	0	29.42857	-1.0054	2.113255	0.00E+00	-3.11865	0	0.6940	1.0000	0.00E+00
14	7	0	0	27.71429	-0.97065	2.113255	0.00E+00	-3.0839	0	0.7014	1.0000	0.00E+00
15	8	2	26	25.42857	-7.92767	2.538714	-2.60E-01	-10.3114	-12.509	0.9198	0.9000	-1.05E-01
16	9	3	26	23.42857	-10.44	3.580501	-3.48E-01	-13.7424	-16.3653	0.9436	0.9323	-7.01E-02
17	10	5	26	27.28571	-15.4153	6.500213	-4.83E-01	-21.504	-24.1008	0.9392	0.9307	-7.19E-02
18	11	3	15	27.57143	-10.5827	3.580501	-3.48E-01	-14.1598	-15.0473	0.7469	0.7084	-3.45E-01
19	12	0	0	21.28571	-0.82569	2.113255	0.00E+00	-2.93895	0	0.7308	1.0000	0.00E+00
20	13	1	22	14.71429	-4.18276	1.932064	-1.50E-01	-6.06615	-8.30918	0.9285	0.9040	-1.01E-01
21	14	2	27	8.857143	-6.39073	2.538714	-2.60E-01	-8.68896	-12.6015	0.9846	0.9804	-1.98E-02
22	15	0	0	5.857143	-0.31866	2.113255	0.00E+00	-2.43192	0	0.8184	1.0000	0.00E+00
23	16	0	0	14.85714	-0.65014	2.113255	0.00E+00	-2.76339	0	0.7639	1.0000	0.00E+00
24	17	1	30	24.85714	-4.74171	1.932064	-1.50E-01	-6.64554	-8.69366	0.9144	0.8858	-1.21E-01
25	18	2	29	25.71429	-7.96636	2.538714	-2.60E-01	-10.3327	-12.7204	0.9328	0.9159	-8.79E-02
26	19	2	25	15.85714	-7.15978	2.538714	-2.60E-01	-9.48946	-12.4349	0.9604	0.9500	-5.12E-02
27	20	0	0	21.57143	-0.8327	2.113255	0.00E+00	-2.94595	0	0.7294	1.0000	0.00E+00
28	21	0	0	4.428571	-0.25142	2.113255	0.00E+00	-2.36468	0	0.8282	1.0000	0.00E+00
29	22	0	0	25.71429	-0.92819	2.113255	0.00E+00	-3.04144	0	0.7102	1.0000	0.00E+00
30	23	0	0	8.428571	-0.42752	2.113255	0.00E+00	-2.54078	0	0.8017	1.0000	0.00E+00
31	24	1	11	20.57143	-4.38431	1.932064	-1.50E-01	-6.42261	-7.6554	0.8254	0.7743	-2.56E-01

圖 6 顧客活躍機率求算方式之 Excel 工作單

3.4 顧客活躍機率與購買金額的變異數分析

Fadar 所提出的 BG/NBD 模型是由顧客交易資料中的最後一次交易時間、重複購買次數、交易歷史時間，三個變數所建構的模型。相較於顧客關係管理中的 RFM 理論，RFM 模型的三個變數，為最後一次交易時間(Recency)、購買頻率(Frequency)、購買金額(Monetary)，BG/NBD 缺少了購買金額的部分。對於藥商而言，個別求算出顧客存活機率固然重要，但顧客對公司購買金額的貢獻也是同等重要。

根據 20/80 法則，公司 80%的利潤來源，是來自於 20%的重要顧客，我們會利用這份交易資料，檢驗是否符合 20/80 法則，再將每位顧客的顧客活躍機率對購買金額作變異數分析，檢驗活躍機率之高低是否會對顧客購買金額造成影響。我們會將顧客活躍機率作分群，再對購買金額作變異數分析，此時之購買金額我們採用了總金額與平均金額兩個變數。

四、實證分析

4.1 建模資料整理

我們將每位顧客資料整理成 $X=(x,t_x,T)$ 之模式，x 為在觀察期間 $(0,T]$ 所觀察到的重複購買數量，t_x 為此段時間內最後一次交易時間，$(0<t_x\le T]$。T 為此顧客的歷史交易時間。本次研究之時間單位為週次計算。關於此三變數的敘述統計量可參閱表 6 各變數的敘述統計量總表，而變數之間的相關係數可參閱表 7 個變數的敘述統計量相關係數總表。

4.1.1 重複購買次數(x)

我們可透過資料得知，前 30 週中 953 位顧客之重複購買次數，其中有 51.7% 顧客之重複購買次數為 0，也就是這些顧客在這 30 週當中只有過 1 次購買行為，即停止交易；而最高之重複購買次數為 14 次，有三位顧客有 14 次之重複購買行為。

表 1 根據顧客重複購買次數分類之顧客數目表

	重複購買次數(次)															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	總計
顧客數目	493	182	117	45	32	34	11	10	8	5	6	3	3	1	3	953
比例(%)	51.7%	19.1%	12.3%	4.7%	3.4%	3.6%	1.2%	1.1%	0.8%	0.5%	0.6%	0.3%	0.3%	0.1%	0.3%	100%

4.1.2 觀察期中之歷史交易時間(T)

關於顧客之歷史交易時間(T)與重複購買次數(x)之間的關係，由圖 7 中，我

們可以得知當重複購買次數較高的顧客，其交易歷史時間亦較長，換句話說，這些高重複購買行為之顧客在我們觀察 30 週的前幾週即開始有購買行為。而 0 次重複購買行為之顧客，其第一次購買行為，則是在我們所觀察的 30 週當中平均分布，而交易歷史時間長但重複購買行為低的顧客，有可能已經喪失交易關係，也就是此段顧客關係已經死亡；但對於歷史交易時間短而重複購買行為低的顧客，也許此顧客只是將其重複購買行為延遲到我們觀察期之後才會發生，也就是說這段顧客關係尚未死亡。對於此部份之研究，我們會在後續之顧客活躍機率作深入探討。

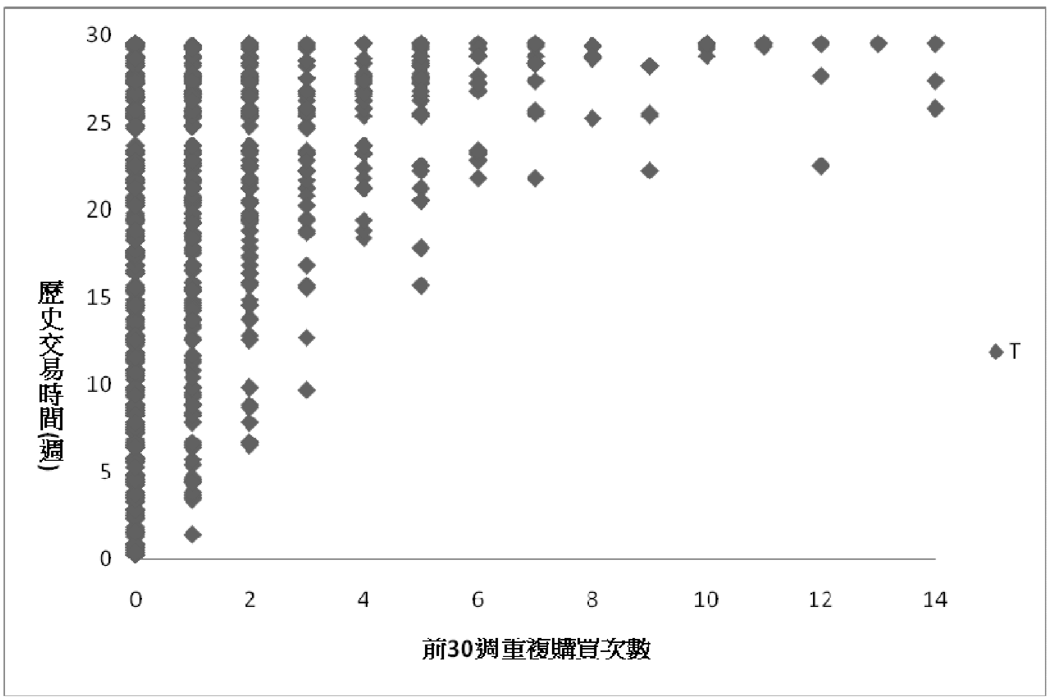


圖 7 顧客之重複購買次數與歷史交易時間關係之分布圖

4.1.3 最後一次交易時間(t_x)

關於顧客之最後一次交易時間(t_x)與重複購買次數(x)之間的關係，我們可以得知當重複購買次數較高的顧客，最後一次交易時間(t_x)也較晚，這些顧客不僅早開始購買並持續交易行為到觀察期的最後數週，我們可以說明這些是具有忠

誠度之顧客；而 0 次重複購買行為之顧客，因其並未具有重複購買行為，所以其最後一次交易時間亦為 0。

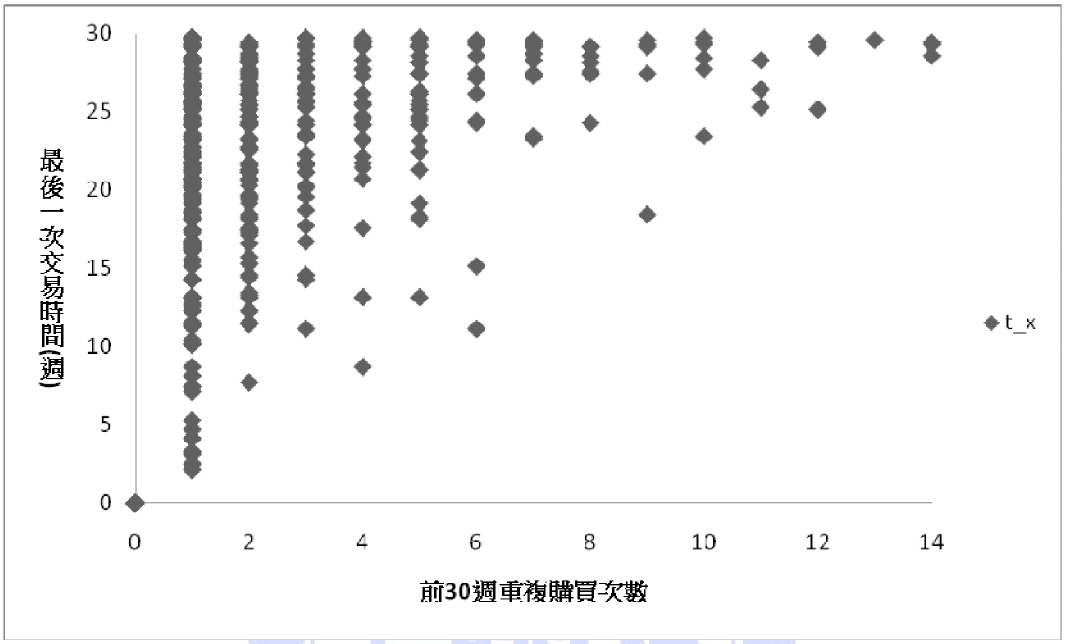


圖 8 顧客之重複購買次數與最後一次交易時間關係之分布圖

4.2 個別顧客之未來期望購買次數

我們能根據

$$E(Y(t)\mid r,\alpha,a,b,x,t_x,T)=\frac{\dfrac{a+b+x-1}{a-1}\left[1-\left(\dfrac{\alpha+T}{\alpha+T+t}\right)^{r+x}{}_2F_1(r+x,b+x;a+b+x-1;\dfrac{t}{\alpha+T+t})\right]}{1+\delta_{x>0}\dfrac{a}{b+x-1}\left(\dfrac{\alpha+T}{\alpha+t_x}\right)^{r+x}}$$

的公式，求得個別顧客之未來期望購買次數E(Y(t))，t=30。

根據前30週顧客購買次數來作分類，我們可獲得預測之平均值。並與後30週之實際購買次數做比較。我們可以得知，當購買次數低時，其未來30週之預測值與實際值之預測較準確，原因為樣本筆數較多筆，我們依據前30週之購買筆數分類列出每組之未來期望購買次數與實際購買次數，如表2與圖9所示。

表 2 根據前 30 週重複購買次數之分類，後 30 週之重複購買次數實際值與預測值

	前30週重複購買次數														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
顧客數目	493	182	117	45	32	34	11	10	8	5	6	3	3	1	3
後30週預測重複購買次數	0.99	1.82	2.53	3.25	3.86	4.58	5.29	5.88	6.45	7.67	7.74	8.42	9.78	9.81	10.91
後30周實際重複購買次數	0.79	1.42	2.04	2.20	3.19	3.69	5.20	6.90	7.14	19.20	11.83	9.33	14.00	13.00	26.67
誤差	0.2	0.4	0.49	1.05	0.67	0.89	0.09	-1.02	-0.69	-11.53	-4.09	-0.91	-4.22	-3.19	-15.76

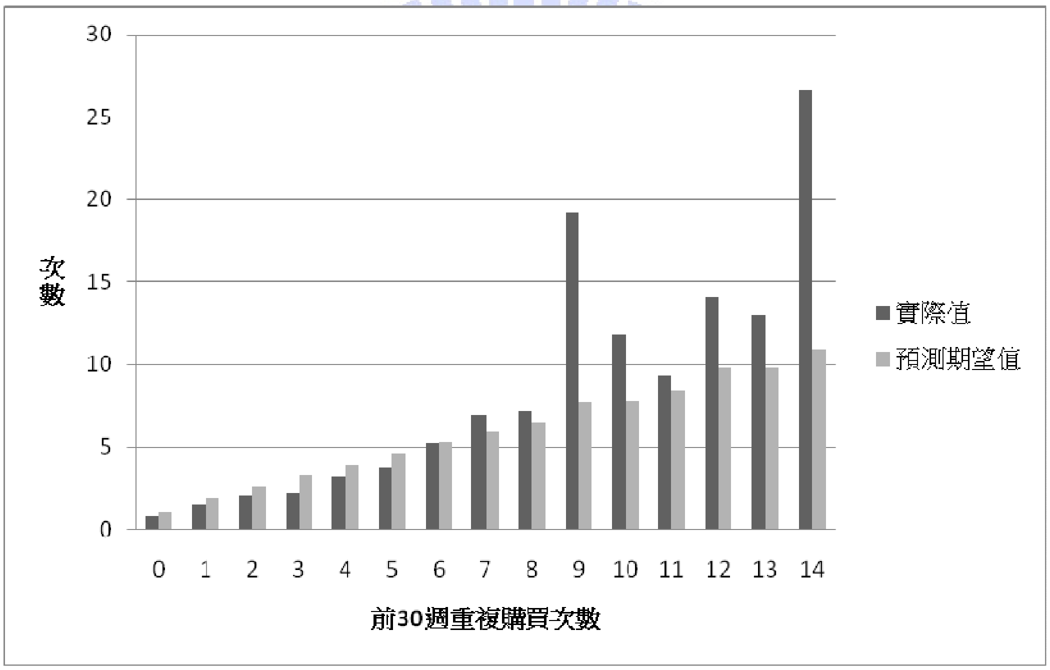


圖 9 後 30 週之重複購買次數實際值與預測值

4.3 個別顧客未來重複購買次數預測之誤差

4.3.1 *MAPE* (mean absolute percentage error)

用*MAPE*來檢驗此模型預測結果。因某些顧客之後30週之重複購買實際值為 $A_t = 0$，這會使得 $MAPE \to \infty$。

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

因此本部份之*MAPE*計算並不將未來無實質重複購買之顧客計算在內。

我們針對後30週實質重複購買次數來作分類，將計算其*MAPE*值作檢驗。檢驗結果為下表中可見。整體之*MAPE*為42.67%，根據MAPE<50%即為可靠預測，故此研究之整體預測為可靠預測。在整個研究當中，當未來實際重複購買次數較小時，如 $A_t = 1$，因其誤差所佔之百分比會提高，故其MAPE會提高；而當未來實際重複購買次數大時，如 $A_t = 18, 19, 20$，因其樣本數少，故其MAPE亦會較大。但整體之MAPE仍顯示此預測為可靠預測。如表3與圖10所示。

表 3 根據後 30 週實際重複購買次數之分類之 MAPE 值

後30週實際購買次數	1	2	3	4	5	6	7	8	9	10
MAPE	75.56%	50.70%	39.93%	38.61%	35.86%	41.89%	32.34%	51.69%	53.68%	42.71%
後30週實際購買次數	11	12	13	14	15	16	17	18	19	20
MAPE	26.75%	48.72%	36.61%	23.58%	49.98%	51.03%	38.30%	59.93%	59.46%	68.14%

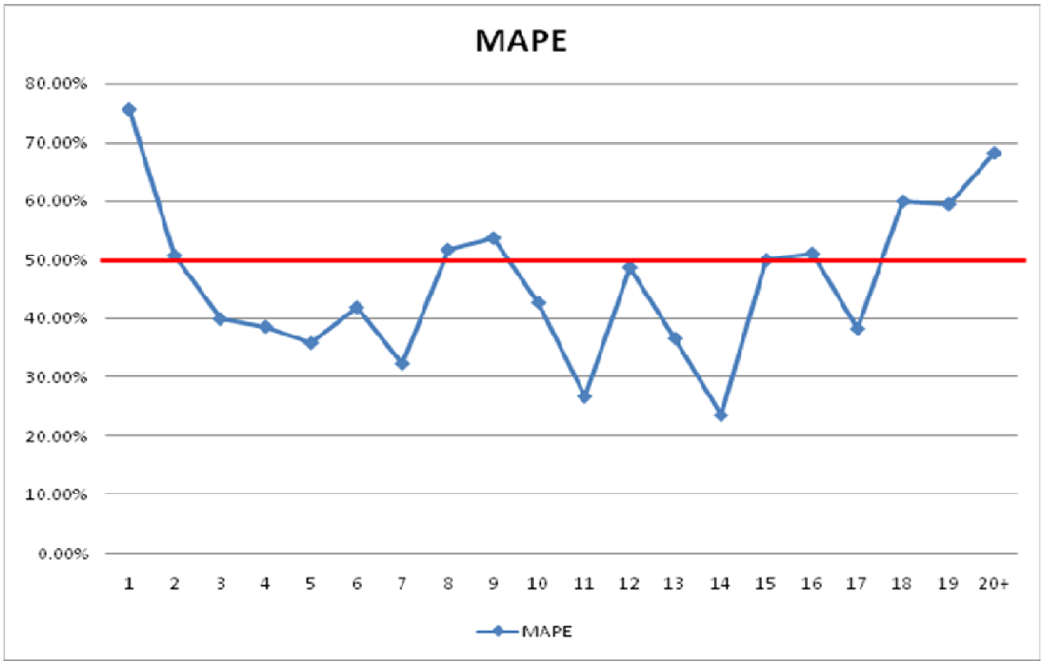


圖 10 根據後 30 週實際重複購買次數之分類之 MAPE 值

4.3.2 *RMSE* (Root mean squared error)

因後30週無實質購買之顧客為377筆，此部份之顧客需用*RMSE*方法來檢驗這部分之預測誤差。

$$RMSE = \left(\frac{1}{n} \sum_{t=0}^{n} (A_t - F_t)^2 \right)^{\frac{1}{2}}$$

而此 377 筆之 *RMSE* 為 1.57。我們也將其他後 30 週購買次數有 575 位顧客之預測值與實際值利用 RMSE 方式做檢驗。整體之 RMSE 為 1.36。

表 4 根據後 30 週實際重複購買次數之分類之 RMSE 值

後30週實際購買次數	0	1	2	3	4	5	6	7	8	9	10
RMSE	1.57	0.87	1.01	1.09	1.24	1.34	1.59	1.50	2.03	2.20	2.07
後30週實際購買次數	11	12	13	14	15	16	17	18	19	20	
RMSE	1.72	2.42	2.18	1.82	2.74	2.86	2.55	3.28	3.36	4.68	

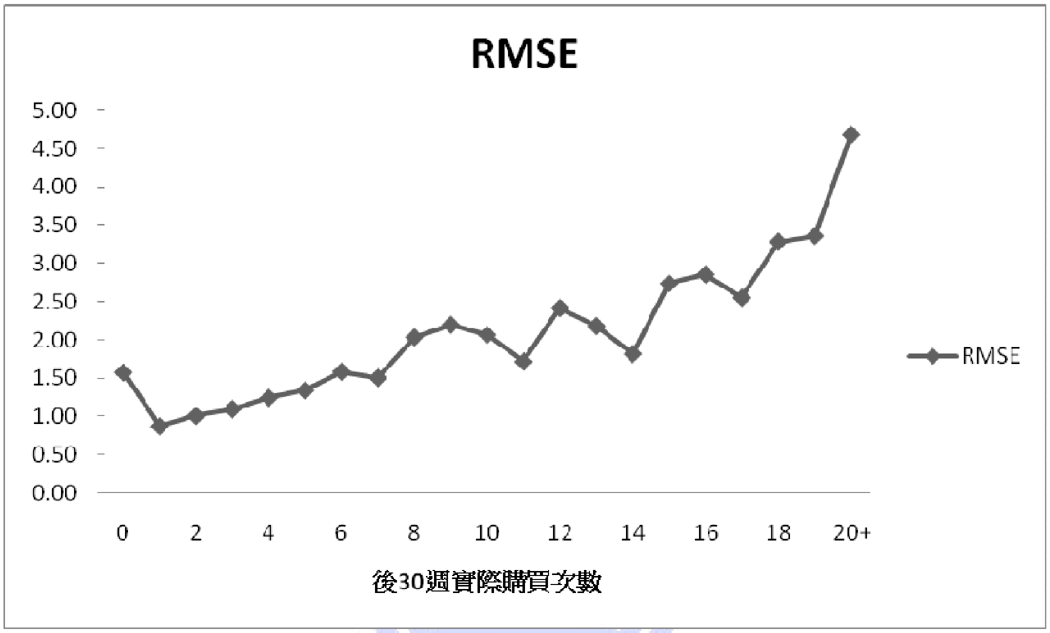


圖 11 根據後 30 週實際重複購買次數之分類之 RMSE 值

4.4 顧客活躍機率

透過本研究之顧客活躍機率結果，根據原始BG/NBD顧客活躍機率公式(4)與修改過後之公式(5)，做出結果如圖12之比較，無重複購買次數(x=0)之顧客，採用修改過後之顧客活躍機率，能夠避免原始機率等於1的情形，也更能有效在無重複購買次數之顧客中，作活躍機率之區別。因此，在本文研究即採用修改過後之活躍機率的結果作後續應用。

ID	x	t_x	T	P(alive)	Palive(old)
1	4	25	28.42857	0.915119	0.901569
2	1	12	27.71429	0.778973	0.718961
3	0	0	12.57143	0.776749	1
4	1	16	26.42857	0.832256	0.782674
5	0	0	29.57143	0.693428	1
6	0	0	29.42857	0.694029	1
7	0	0	27.71429	0.701358	1
8	2	26	25.42857	0.919814	0.90003
9	3	26	23.42857	0.943632	0.932325
10	5	26	27.28571	0.939169	0.930654
11	3	15	27.57143	0.746946	0.708378
12	0	0	21.28571	0.730809	1
13	1	22	14.71429	0.928469	0.904048
14	2	27	8.857143	0.984554	0.980403
15	0	0	5.857143	0.818432	1
16	0	0	14.85714	0.763918	1
17	1	30	24.85714	0.914393	0.885757
18	2	29	25.71429	0.932769	0.91589
19	2	25	15.85714	0.960367	0.950045
20	0	0	21.57143	0.729428	1
21	0	0	4.428571	0.828211	1
22	0	0	25.71429	0.710175	1
23	0	0	8.428571	0.801691	1
24	1	11	20.57143	0.825372	0.774307
25	2	12	26.71429	0.748498	0.700223

圖 12 兩種顧客活躍機率之結果比較

顧客活躍機率結果，如圖12所示，最高之活躍機率為0.9942(ID=806)，最低之顧客活躍機率為0.3095(ID=625)。我們試著將顧客重複購買次數(x)、最後一次交易時間(tx)、總歷史交易時間(T)，分別與顧客活躍機率作相關性分析。如圖13、圖14、圖15為其與各變數之間的分布圖。圖16為整體之顧客活躍機率。

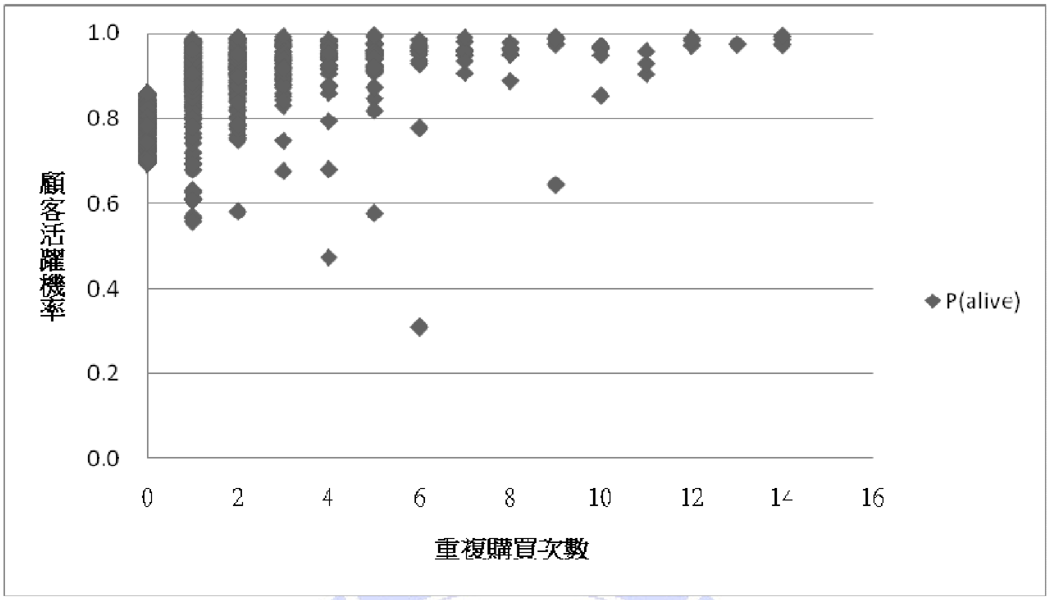


圖 13 顧客活躍機率與重複購買次數分布圖

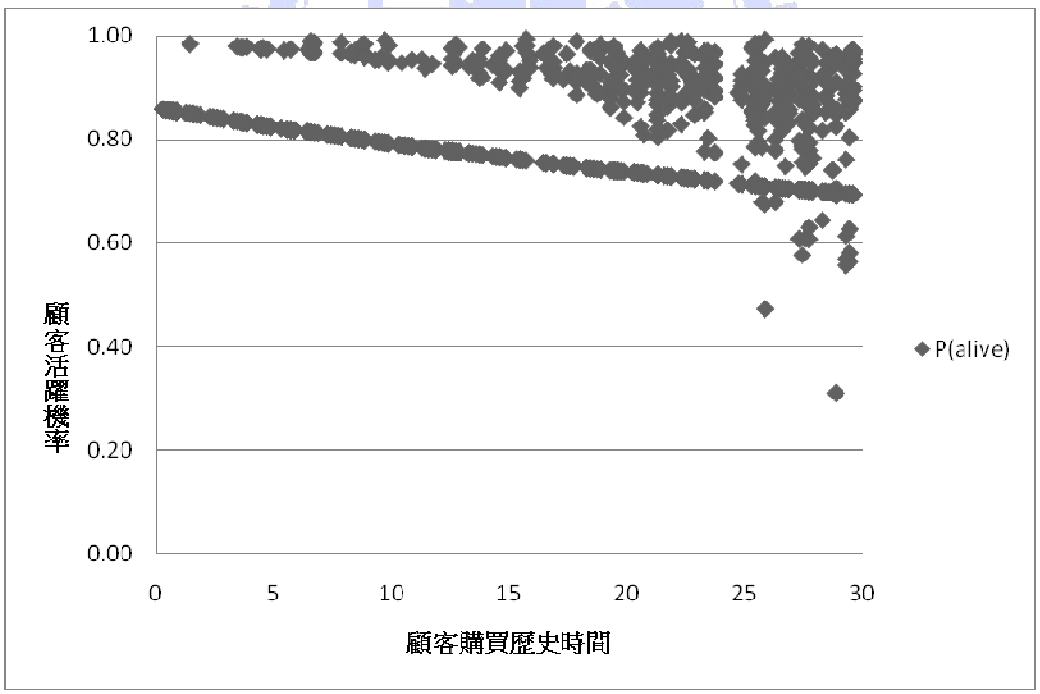


圖 14 顧客活躍機率與購買歷史時間分布圖

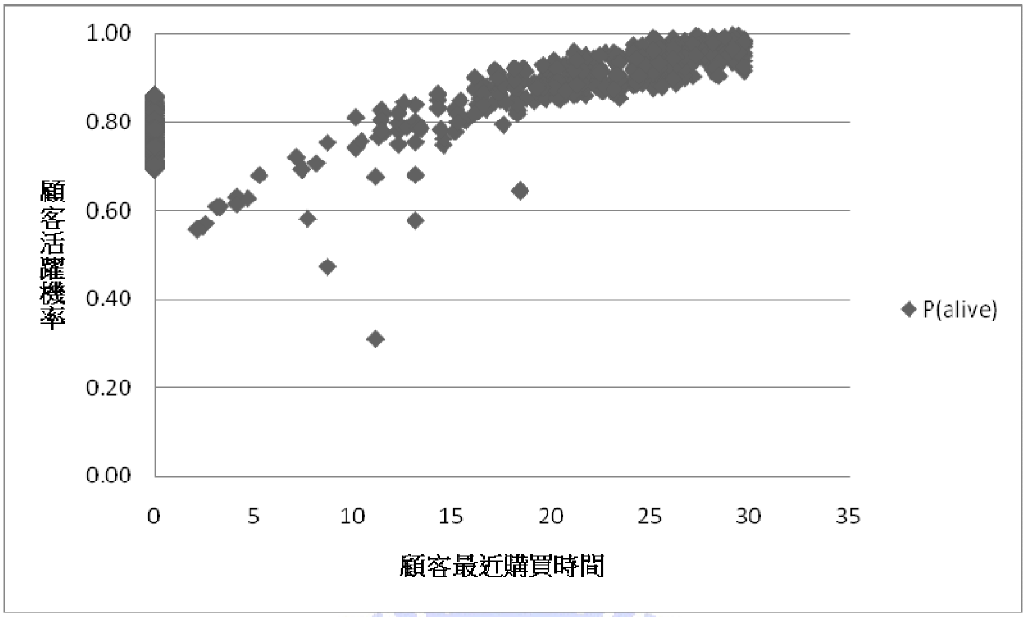


圖 15 顧客活躍機率與最近購買時間分布圖

	A	B	C	D	E
1	ID	x	t_x	T	P(alive)
2	806	5	29	15.71429	0.994235571
3	778	14	29	25.85714	0.993176646
4	626	3	27	9.714286	0.991913749
5	673	9	27	22.28571	0.990607799
6	453	2	29	6.571429	0.9899799542
7	766	5	28	17.85714	0.9898794674
8	930	7	29	21.85714	0.989434655
9	223	12	25	22.57143	0.988350321
10	741	2	29	7.857143	0.9879794503
11	623	2	26	6.714286	0.987390525
12	652	2	25	6.714286	0.9864333907
13	792	9	29	25.42857	0.9855487913
14	322	9	29	25.57143	0.9855020474
15	335	1	30	1.428571	0.9847921755
16	14	2	27	8.857143	0.984554088
17	470	4	29	18.85714	0.984417018
18	500	3	25	12.71429	0.9841459119
19	252	2	26	8.714286	0.983393817
20	53	14	29	27.42857	0.983383623
21	48	6	27	21.85714	0.9825787648
949	639	5	13	27.42857	0.5766274538
950	363	1	3	29.28571	0.5698892943
951	88	1	2	29.42857	0.564092845
952	753	1	2	29.28571	0.5569539158
953	573	4	9	25.85714	0.4727965290
954	625	6	11	28.85714	0.30952575
955					

圖 16 顧客活躍機率

4.5 顧客活躍機率對顧客購買金額的影響

BG/NBD 模型缺少了購買金額的部分，對於藥商而言，個別求算出顧客存活機率固然重要，但顧客對公司購買金額的貢獻也是同等重要。

將顧客活躍機率作四分位差，分成「HH」、「H」、「L」、「LL」，四個群組，「HH」範圍為 0.920~0.994，「H」範圍為 0.826~0.919，「L」範圍為 0.750~0.825，「LL」範圍為 0.310~0.749，再對顧客購買總金額、平均金額作變異數分析，檢驗活躍機率之高低是否會對顧客購買總金額、平均金額造成影響。

利用 SAS 軟體之 Proc Anova 指令，分析結果對總平均金額(TM)有顯著影響，其 F 值=19.48，其 P 值<0.0001，表示顧客活躍機率的高低對顧客購買總金額有顯著影響。對平均金額(AM)無顯著影響，其 F 值=1.43，P 值=0.2321，表示顧客活躍機率之高低對顧客購買平均金額無顯著影響。

關於對總金額影響部分，將顧客活躍機率各組間的差異，利用最小顯著差異法(LSD; least significant difference)作事後檢驗，結果在顧客活躍率最高(HH)與其他三組(H、L、LL)的總購買金額差都處於顯著，剩餘之組間差距皆不顯著。如表五所示：

表 5 活躍機率組間對總金額(LSD)事後檢驗

	平均差 (Difference between Means)
HH – H	24421**
HH – LL	30265**
HH – L	33156**
H – LL	5845
H – L	8736
L– LL	2891

**P<0.05

根據我們上述顧客活躍機率與購買金額的關係，在總金額方面，我們可以透過顧客活躍機率之高低，來說明顧客對企業的價值。但因其 R-Square=5.79%，說明利用顧客活躍機率的高低來說明顧客總購買金額只有 5.79%的解釋能力，因此我們仍須參考顧客的總購買來判斷顧客的重要程度。關於 20/80 法則，公司 80%的利潤來源，是來自於 20%的重要顧客，根據此份交易資料，總交易金額為新台幣 16,456,118 元，而購買總金額前 20%的顧客所交易的金額為新台幣 12,807,526 元，佔總交易金額之 77.83%，非常接近 20/80 法則，因此我們可以得知總交易金額為重要指標，我們可透過顧客總交易金額配合顧客活躍機率，共同檢驗顧客目前的交易狀態，並對此顧客制定新的行銷計畫。

表 6 各變數的敘述統計量

	個數	平均數	標準差	最小值	最大值
顧客活躍機率P(alive)	953	0.83	0.0960	0.30953	0.99424
重複購買次數(x)	953	1.34	2.2120	0	14
最後一次交易時間(t_x)	953	11.03	12.1413	0	29.71
總歷史交易時間(T)	953	18.86	8.3939	0.28571	29.57
顧客購買總金額(TM)	953	17304	53466	1276	793335
顧客購買平均金額(AM)	953	5795	11214	1276	169436

表 7 各變數之間的相關係數

	Palive	x	t_x	T	TM
顧客活躍機率P(alive)					
重複購買次數(x)	0.5518***				
最後一次交易時間(t_x)	0.8426***	0.6822***			
總歷史交易時間(T)	-0.1046**	0.4180***	0.3518***		
顧客購買總金額(TM)	0.20056***	0.5452***	0.2709***	0.1348***	
顧客購買平均金額(AM)	0.05658	0.1499***	-0.0934**	0.0669**	0.6560***

***P<0.001；**P<0.05

4.6 管理意涵

2005 年 Kumar 提出七個顧客階層的行銷戰略以提高顧客權益：1.篩選對的顧客 2.與這些顧客保持聯繫 3.在正確的時間傳遞正確的訊息 4.管理多通路之購物模式 5.管理高成本的顧客 6.找尋與維持對的客戶 7.同時管理顧客忠誠度與收益力(profitability) (Kumer & Petersen, 2005)。當我們將顧客活躍機率，配合顧客購買總金額後，我們可以得知哪些關鍵客戶，目前顧客活躍機率有較低之情形，如圖所示，顧客(ID=479)為最高的總交易金額的客戶，其 30 週內的總交易金額為 793,335 元，其重複購買次數有 9 次，總購買歷史時間為 28.29 週，而其最後一次購買時間為在第 18.43 週的時間點，其活躍機率卻只剩下 0.64，因此行銷人員此時可獲得警訊，並針對顧客(ID=479)做客製化的行銷計畫。顧客(ID=479)的購買習慣為多次(x=9)且大量購買，其平均交易金額為 79,333.5 元，根據其購買頻率與最後一次購買時間，此顧客以有一段時間無購買行為，此時行銷人員可針對顧客(ID=479)作訪談，了解此關鍵顧客的需求是否有變動或其它降低購買意願之因素，並重新制定新的客製化計畫，以求維持長遠交易關係。

1	ID	x	t_x	T	Palive	TM	AM
2	479	9	18.428571	28.285714	0.643853876	793335	79333.5
3	778	14	29.428571	25.857143	0.993176646	620411	41360.7
4	793	12	29.428571	29.571429	0.970441484	525713	40439.5
5	792	9	29.142857	25.428571	0.985548791	519803	51980.3
6	234	2	17.428571	25.428571	0.847803424	400000	133333
7	347	10	29.428571	29.428571	0.967331593	400000	36363.6
8	463	8	27.428571	28.857143	0.94814977	363619	40402.1
9	87	14	29.285714	29.571429	0.972402906	359616	22476
10	433	2	20.714286	26.714286	0.871870035	299005	99668.3
11	322	9	29.285714	25.571429	0.985502047	284001	28400.1
12	53	14	28.571429	27.428571	0.983383623	244296	16286.4
13	889	7	29.428571	29.571429	0.956511404	231527	28940.9
14	659	7	29.571429	29.428571	0.958670555	204286	25535.8
15	252	2	26.142857	8.7142857	0.983393817	203810	67936.7
16	170	0	0	21.857143	0.728054943	169436	169436
17	332	4	29.714286	21.285714	0.978522235	160000	32000

圖 17 顧客活躍機率配合顧客購買總金額

五、結論

5.1 研究結論

我們利用 BG/NBD 模型的方法，所預測之顧客未來購買次數與顧客活躍機率，根據顧客購買次數的預測值與實際值誤差檢驗，整體相對誤差 MAPE=42.6%小於 50%，雖不到 25%以下的良好預測，但仍為可靠預測。因此我們將原本 Fader 等學者用以預測網路消費的 BG/NBD 模型，應用於藥品產業的非契約顧客的未來購買行為預測上，仍為一個可靠的預測模型。

在關係行銷著重的藥品產業顧客關係上，BG/NBD 模型所預測出來的顧客活躍機率，是行銷人員檢驗顧客關係的良好指標。根據本研究所採用的人類肝臟用藥資料，顧客之購買總金額會受到 BG/NBD 模型所預測之顧客活躍率影響，當顧客活躍率高的顧客(P(alive)>0.92)，與其他較低活躍機率的顧客相比，顧客活躍機率高的顧客，其購買總金額顯著較高。

本研究資料滿足 20/80 法則，本藥品前 30 週的非契約性顧客交易金額為新台幣 16,456,118 元，而購買總金額前 20%的顧客所交易的金額為新台幣 12,807,526 元，佔總交易金額之 77.83%。合乎 80%的公司利益來自 20%的關鍵顧客。透過總購買金額的排序，配合顧客活躍機率值，對於行銷人員可以偵察到關鍵顧客的活躍情形，並針對較低活躍率之顧客探討與制訂適當之顧客甦醒方案。

5.2 研究限制與未來發展

本研究由於時間、空間等因素，具有幾項研究限制與未來發展，詳述如下：

1. 因為資料庫資料的限制，有些顧客背景資料我們無法獲得，如本研究只能針對忠誠程度(level)作辨別，但無法對忠誠類型(type)作辨認。
2. 本研究假設顧客所面臨的購買條件相同。
3. 因本次研究的建模資料為 2008 年 1 月~2008 年 7 月中 30 週的顧客交易資料，而驗證資料為 2008 年 7 月底~2009 年 2 月初的 30 週資料，因日期與季節性影響，會對顧客購買有所影響，後續研究應採用年份為單位作建模資料。
4. 本研究並無將顧客維持成本納入考慮，顧客維持成本與顧客終身價值決定關鍵顧客的選取，可納入未來研究考量。
5. 顧客之間具有口碑效應(word of mouth)，口碑價值於本研究並無納入考慮。
6. 顧客具有成長性，本研究並無考慮顧客的成長性，其對企業的潛在貢獻度無衡量。
7. 未來研究可針對顧客類型(type)與顧客維持(retention)計畫的有效性作相關研究。

參考文獻

Blattberg, R.C., G Getz, JS Thomas (2001),"Customer Equity: Building and Managing Relationships and Valuable Assets," *Harvard Business School Press*

Colombo, R. and W. Jiang (1999), "A Stochastic RFM Model," *Journal of Interactive Marketing*, 13(3), 2-12

Fader, P. S. and B. G.S. Hardie (2001), "Forecasting Repeat Sales at CDNOW: A Case Study," *Interfaces*, 31(2), 94-107

Fader, P. S., B. G. S. Hardie, and K. L. Lee (2005), "Counting Your Customers: the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science,* 24(2), 275-284.

Fader, P. S. and B. G. S. Hardie (2007)," How to project customer retention," *Journal of Interactive Marketing*, 21, 76-90

Fader, P. S., B. G. S. Hardie, and K. L. Lee (2008)"Computing P(alive) Using the BG/NBD Model,"

Fader, P.S. and B. G. S. Hardie (2006)," Probability Models for Customer-Base Analysis"

Hogan, J.E., K. N. Lemon, and R. T. Rust (2002), "Customer equity management: Charting new directions for the future of marketing," *Journal of Service Research* ,5(1), 4-12

Kumar, V and J. Andrew Petersen (2005), "Using a Customer-Level Marketing Strategy to Enhance Firm Performance: A Review of Theoretical and Empirical Evidence," *Journal of the Academy of Marketing Science,* 33(4), 504-519

Lemmens, A. and C. Croux (2006), "Bagging and Boosting Classification Trees to Predict Churn," *Journal of Marketing Research,* 43(2), 276-286

Malthouse, E and F. Mulhern (2008),"Understanding and Using Customer Loyalty and Customer Value," *Journal of Relationship Marketing*, 6(3), 59-86

Neslin, S. A., S. Gupta, W. Kamakura, J. Lu, and C.H. Mason (2006), **"**Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models,**"** *Journal of Marketing Research,* 43(2), 204-221

Peppers, D., M. Rogers, & B. Dorf (1999), "Is Your Company Ready for One-to-one Marketing?" *Harvard Business Review,* 77(1), 151-156

Reinartz, W., J. S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing,* 69, 63–79

Rust, R. T., K.N. Lemon, and D. Narayandas (2005), "Customer Equity Management," Pearson Prentice Hall

Rust, R. T., K.N. Lemon, and V. A. Zeithaml (2004)," Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68, 109–127

Schweidel, D. A., P. S. Fader, & E. T. Bradlow (2008),"Understanding Service Retention Within and Across Cohorts Using Limited Information," *Journal of Marketing*, 72, 82-94

Woodruff, R. B.(1997),"Customer value: The next source for competitive advantage," *Academy of Marketing Science. Journal,* 25(2), 139-153

Werner J. Reinartz & V. Kumar (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing,* 67, 77–99