行政院國家科學委員會專題研究計畫 成果報告

吃角子老虎問題之最佳貝氏策略
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 99-2118-M-004-004-
執 行 期 間 ：99 年 08 月 01 日至 100 年 07 月 31 日
執 行 單 位 ：國立政治大學統計學系

計畫主持人：洪英超

計畫參與人員：碩士班研究生-兼任助理人員：謝至芬
碩士班研究生-兼任助理人員：劉世璿

處 理 方 式 ：本計畫可公開查詢

中 華 民 國 100 年 10 月 20 日

行政院國家科學委員會補助專題研究計畫

■成果報告　□期中進度報告

（計畫名稱）
吃角子老虎問題之最佳貝氏策略

計畫類別：■個別型計畫　□整合型計畫
計畫編號：NSC 99－2118－M－004－004－
執行期間：　99年 8 月 1 日至 100年 7月 31日

執行機構及系所：政治大學統計系

計畫主持人： 洪英超
共同主持人：
計畫參與人員：謝至芬　劉世瑢

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本計畫除繳交成果報告外，另須繳交以下出國心得報告：
□赴國外出差或研習心得報告
□赴大陸地區出差或研習心得報告
□出席國際學術會議心得報告
□國際合作研究計畫國外研究報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢
　　□涉及專利或其他智慧財產權，■一年□二年後可公開查詢

中　華　民　國 100 年 10 月 20 日

中、英文摘要及關鍵詞

多拉桿吃角子老虎問題(multi-armed bandit problem)可以應用在許多領域如臨床試驗,線上工業實驗(on-line industrial experimentations),可調性網路路由(adaptive network routing)等. 本計畫將以貝氏的角度探討“無窮多拉桿之吃角子老虎問題“.我們假設未知的白努利參數為相互獨立且來自同一個機率分配F,而我們的目的是找出一如何選擇拉桿的策略使得長時間操作下的失敗率為最低. 在本計畫的第一部份,我們假設F為一 任意但已知的機率分配.接著介紹1996年由Berry等人提出的三種策略,並証明當試驗次數趨近無窮大時,此三種策略皆可以使長時間操作下的失敗率為最低.此外,我們也利用電腦模擬來比較此三種策略的實際表現. 在本計畫的第二部份,我們假設F為一未知的機率分配.在此假設下,我們提出一個新的策略叫做”empirical non-recalling m-run策略”,並証明此策略亦為一近似最佳策略. 此外,我們也將利用電腦模擬與Herschkorn等人於1995年提出的二個策略進行比較.

關鍵字: 多拉桿吃角子老虎問題; 貝氏策略.

Multi-armed bandit problems have a wide area of applications such as clinical trials, on-line industrial experimentations, adaptive network routing, etc. In this study, we examine the bandit problem with infinitely many arms from a Bayesian perspective. We assume the unknown Bernoulli parameters are independent observations from a common distribution F, and the objective is to provide strategies for selecting arms at each decision epoch so that the expected long run failure rate is minimized. In the first part of this study, we assume the common distribution F is arbitrary but known. We introduce three strategies proposed by Berry et al. (1996) and show that they asymptotically minimize the expected long run failure rate. Numerical results from computer simulations are also provided to evaluate the performance of the three strategies. In the second part of this study, we assume the common distribution F is unknown. For this setting, we propose a strategy called the “empirical non-recalling m-run strategy” and prove that this strategy is asymptotically optimal. Numerical results from computer simulations will also be provided to evaluate the proposed strategy and two other strategies by Herschkorn et al. (1995).

Key words: Multi-armed bandit problem; Bayesian strategy.

報告內容

（一）前言

We consider the following settings for the bandit problem: Suppose there are an infinite number of Bernoulli arms, of which the unknown parameters p_i (success probabilities) are i.i.d. random variables with a non-degenerate common distribution F defined on [0,1]. At each decision stage, the decision maker chooses an arm for observation. Our goal here is to select arms at each decision stage so that the expected long-run failure rate (failure proportion) over n choices (with $n \to \infty$) is minimized.

A strategy of selecting arms that minimizes the long-run failure rate is called the optimal strategy. It is obvious that under the current settings the optimal strategies depends greatly on the prior distribution F and the number of trials n. Intuitively, when n is large, the decision maker is more inclined to sacrifice immediate gain and select new arms in hoping to find one that has a larger Bernoulli parameter (thus producing a substantial low failure rate). However, when n is small, new information has less value and it might be prudent to use an arm that has a small failure rate. Therefore, during the course of an experiment and as the horizon nears, arms with a low failure rate are more appealing even though they have less potential for providing information. The same reasoning reveals that, when F has a perceivable weight near 1, the decision maker will be more optimistic and aggressive in searching for a new arm with a very low failure rate. On the other hand when F has a negligible weight near 1, one becomes more conservative and reluctant in searching for a new arm, and has the propensity to use an arm which has the best performance so far.

（二）研究目的

Bandit problems have a wide area of applications in clinical trials, on-line industrial experimentation, machine learning, inter-temporal allocation in an economics environment, etc (for real examples, see Banks and Sundaram, 1992; Berry and Fristedt, 1985; Gittins, 1989; Lai and Robbins, 1984; Wang et al., 2005; and the references there in). For the problem under current settings, we introduce its one important application to the control of data routing networks. Consider a typical computer network (such as Internet) or telecommunications network that is determined by connections (or links) between nodes (or stations). The network routes digital data in small pieces (called packets) and each of which is transmitted independently through the links between nodes to the correct destination. At each intermediate node, there could be a large number of downlinks from which the router can select one to transmit a particular packet (called unicast transmission). However, the transmission of packets between nodes may not succeed due to factors such as channel congestion, corrupted packets rejected in-transit, faulty networking hardware, faulty network drivers, etc. These “unreliable” links may result in packet loss, thus affecting the quality of service (QoS) performance of network. Therefore, a good routing strategy must be able to select, at any point in time, the best downlink to transmit the packet so as to possibly minimize the packet loss rate. Note that such a routing scheme can be simply modeled as the bandit problem, where each possible downlink of a particular node can be viewed as an "arm", and the corresponding probability of successfully transmitting a packet is denoted by p_i (or the probability of losing a packet is $1 - p_i$). Figure 1 illustrates a basic component of the described network topology with one router and n parallel downlinks. However, when information about the reliability of downlinks is not available (this is often the real situation), constructing an optimal routing strategy becomes a challenging task.

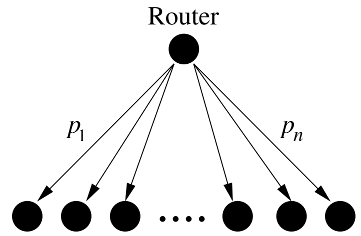

Figure 1: An illustration of data network with one router and n parallel downlinks.

（三）文獻探討

The primary study of bandit problems can be traced back to the work by Mallows and Robbins (1964), Robbins (1952), and Thompson (1933). Afterwards, there exists a fairly rich literature discussing one-armed bandit problems with various settings (Clayton, 1989; Sarkar, 1991; Woodroofe, 1979; Zoubeidi, 1994). Berry and Fristedt (1985) described the general setting of multi-armed bandit problems and provided extensive treatment to them. Narendra and Thathachar (1989) treated bandit problems from the engineering perspective, providing a good discussion of the various theoretical traditions that have focused on them, Pandey et al. (2007) provided a framework to exploit dependencies among arms. Other pertinent references include the work by Auer and Cesa-Bianchi (2002), Gittins (1979, 1989), Guha et al. (2010), Lai (1987), Lai and Robbins (1984), Powell (2007), and Whittle (1982, 1983), just to name a few. A remarkable study of bandit problems from the Bayesian perspective is the work by Berry et al. (1997), wherein three appealing strategies (called the m-run strategy, the non-recalling m-run strategy, and the N-learning strategy) were shown to, as the number of trials goes to infinity, achieve the best lower bound of the expected failure rate when F is the $uniform(0, 1)$ distribution (i.e., they are asymptotically optimal). Later on, Lin and Shiau (2000) conducted a simulation study to evaluate the numerical performance of these three strategies when F is the beta distribution. On the other hand, the study of bandit problems regardless of the common distribution F is rather limited. Herschkorn et al. (1995) proposed a strategy that pulls the ith arm until i failures in a row are observed (called the "$i - i$ strategy" in later analysis) and showed that this strategy minimizes the expected long-run failure rate without requiring knowledge of F. However, numerical evidence shows that this particular strategy does not perform well due to its slow convergence rate (see Berry et al. (1997) and Section 4 for examples).

（四）研究方法

Optimal Strategies for Arbitrarily Known Priors

We start with introducing two sets of strategies that are closely related to the optimal strategies described later. The first strategy is called a "k-failure strategy", which calls for using the same arm until that arm produces k failures, and when this happens, it calls for switching to a new arm. Note that this strategy never recalls arms that have produced failures. With the possible exception of the arm being used when the total number of trials n is reached, every arm used yields a total of k failures. The second strategy is called a "β-rate strategy", which stays on the same arm until that arm has produced a failure rate greater than β, $\beta \in [0,1]$, and when this happens, the arm is discarded and a new arm is used. Analogously, the discarded arms are never recalled.

One might suspect that the β-rate strategy can do better than the 1-failure strategy for some particular choices of β. However, as shown by Berry et al. (1997), if $F(\alpha) < 1$ for all $\alpha < 1$, the expected failure rate of the β-rate strategy is always greater than a positive constant as $n \to \infty$ for any $\beta > 0$. On the other hand, we see that if $\sum_{j=0}^{\infty} \int_0^1 \alpha^j dF(\alpha) = \infty$, the expected failure rate of the 1-failure strategy goes to 0 as $n \to \infty$. This means that if n is large enough, the 1-failure strategy outperforms the β-rate strategy for any choices of β when F has the property that $\sum_{j=0}^{\infty} \int_0^1 \alpha^j dF(\alpha) = \infty$.

It is noted that the 1-failure strategy can perform poorly when $\sum_{j=0}^{\infty} \int_0^1 \alpha^j dF(\alpha) < \infty$. A simple example is when F is a $beta(a, b)$ distribution with the shape parameter $b > 1$. Therefore, we seek for other strategies that can possibly achieve the lower bound given in Eq. (4) when n is large. We next introduce three appealing strategies proposed by Berry et al. (1997).

- The m-run strategy: This strategy uses the 1-failure strategy until either the current arm has produced a success run of length m or m arms are used. If the former obtained, then the current arm will be used for the all remaining trials. If the latter obtained, then the arm with lowest failure proportion among the m arms used so far will be used for the all remaining trials. So an m-run strategy uses at most m arms. If it does use m arms, then the best performing arm is recalled and will be used for all remaining trials.
- The non-recalling m-run strategy: This strategy uses the 1-failure strategy until an arm has produced a success run of length m at which this arm will be used for the all remaining trials. If no arm has produced a success run of length m, then the 1-failure strategy is used for all n trials.
- The N-learning strategy ($N \leq n$): This strategy follows the 1-failure strategy for the first N trials (the arm used at the N-th trial will be used until such time that it yields a failure), and then it calls for using the arm that performed best during the learning period for all remaining trials.

An Optimal Strategy for Unknown Priors

Motivated by the optimality of the "non-recalling m-run strategy" from Section 2 and its nice numeric performance given by Lin and Shiau (2000), we next introduce a new strategy called the "empirical non-recalling m-run strategy" when the prior distribution F is unknown.

• The empirical non-recalling m-run strategy: For each positive integer n, let k be the smallest integer such that $k \geq \sqrt{2\sqrt{n}}$. Choose k independent arms and perform k trials for each arm (thus k^2 trials in total). Record the observed proportion of success $\hat{p}_i$ for each arm i and construct an empirical distribution F_k for $\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_k$. Implement the non-recalling m-run strategy for the remaining $n - k^2$ trials based on the empirical distribution F_k.

Note that this new strategy is comprised of two stages – learning and implementation. In the learning stage, the k^2 trials are used to construct the empirical distribution of the observed Bernoulli parameters $\hat{p}_i$ so as to obtain a good estimate for the underlying unknown distribution F. In the implementation stage, the non-recalling m-run strategy is used for the remaining $n - k^2$ trials. The following theorem shows the asymptotic optimality of the empirical non-recalling m-run strategy.

（五）結果與討論

When F is known, we prove that:

Theorem 1. *Suppose that $0 < x^* = inf\{t : F(t) = 1\} \leq 1$, then the expected failure rate of the non-recalling m-run strategy can be arbitrarily close to $1 - x^*$.*

Theorem 2. *The non-recalling u_n-run strategy is asymptotically optimal.*

Theorem 3. *The k_n-run strategy is asymptotically optimal.*

Theorem 4. *The N-learning strategy is asymptotically optimal.*

When F is unknown, we prove that:

Theorem 5. *Suppose that $0 < x^* = inf\{t : F(t) = 1\} \leq 1$, then the expected failure rate of the empirical non-recalling m-run strategy can be arbitrarily close to $1 - x^*$.*

Theorem 6. *The empirical non-recalling u_n-run strategy is asymptotically optimal.*

Now we proceed to evaluate the numerical performance of the proposed empirical non-recalling m-run strategy via computer simulation. Since now we assume the underlying distribution F is unknown, for comparison purpose we also evaluate the performance of (i) the $i - i$ strategy proposed by Herschkorn et al. (1995), which pulls the ith arm until i failures in a row are observed; and (ii) the 1-failure strategy, which is a popular strategy and easy to implement. The "simulated" long-run failure rates for these strategies with different numbers of trials are given in Table 1, for which F is generated from five different beta distributions: $beta(0.5, 0.5)$, $beta(1, 1)$, $beta(2, 2)$, $beta(1, 2)$, and $beta(2, 1)$. It should be noted that, the choices of F in the simulation study represent a fairly wide range of shapes including possibly symmetrical, right-skewed, and left-skewed distributions over $(0, 1)$. In addition, each failure rate is estimated by the average of 2,000 Monte Carlo simulation trials that were executed on 2.53GHz Intel Core i5 processor with 4GB of cache under the operating system of Mac OS 10.6. The computer programs were written in R, where the function "rbeta" was used to generate beta random numbers.

As can be seen from Table 1, our proposed strategy significantly outperforms the other two strategies for almost all simulation scenarios. On the other hand, the $i - i$ strategy, although proven optimal, performs poorly due to its slow convergence rate. It is worth noting that the 1-failure strategy reveals to be fairly competitive when F is $beta(0.5, 0.5)$ or $beta(2, 1)$. The numerical evidence supports the fact that the 1-failure strategy, although proven not optimal, performs well when F has a lot of its probability mass near one.

Table 1: The estimated failure rates of the three strategies for different choices of n and F. Note that $\overline{m}$ represents the average of the best chosen values of m for the empirical non-recalling m-run strategy.

F	n	Empirical non−recalling m−run strategy $\overline{m}$		The i−i strategy	The 1−failure strategy
beta (0.5,0.5)	100	7.5	0.117	0.176	0.104
	300	10.3	0.065	0.137	0.063
	500	13.9	0.055	0.125	0.049
	1000	18.1	0.032	0.110	0.035
	1500	25.7	0.030	0.105	0.029
	2500	30.6	0.024	0.092	0.024
beta (1,1)	100	6.2	0.176	0.262	0.219
	300	14.8	0.109	0.227	0.180
	500	16.0	0.090	0.209	0.167
	1000	21.0	0.066	0.192	0.149
	1500	27.5	0.054	0.182	0.141
	2500	33.0	0.046	0.173	0.132
beta (2,2)	100	9.2	0.246	0.348	0.355
	300	17.1	0.204	0.312	0.344
	500	22.3	0.197	0.298	0.340
	1000	28.9	0.154	0.283	0.336
	1500	33.8	0.135	0.273	0.336
	2500	40.3	0.126	0.261	0.334
beta (1,2)	100	9.1	0.347	0.472	0.520
	300	16.2	0.335	0.421	0.508
	500	23.5	0.306	0.393	0.504
	1000	29.1	0.294	0.373	0.502
	1500	31.3	0.288	0.358	0.501
	2500	35.6	0.258	0.343	0.501
beta (2,1)	100	11.0	0.147	0.182	0.136
	300	15.8	0.090	0.156	0.108
	500	19.4	0.075	0.147	0.098
	1000	22.6	0.053	0.134	0.088
	1500	26.8	0.040	0.129	0.083
	2500	31.5	0.035	0.123	0.078

The Control of Data Routing Networks - Revisited

Let's recall the network routing problem introduced in Section 1. A well-known routing strategy called "round-robin", which selects downlinks to transmit the packets in circular order and in equal proportion, is extensively used and has been shown to be optimal in many situations. However, it is straightforward to see that under our current settings, such a routing strategy will result in a successful transmission rate close to the "mean" of the prior distribution F as the number of downlinks becomes large (this is simply the result of the law of large numbers). Therefore, the expected "failure" (or packet loss) rates under the round-robin strategy are 0.5, 0.5, 0.5, 0.67, and 0.33 when F is chosen to be $beta(0.5, 0.5)$, $beta(1, 1)$, $beta(2, 2)$, $beta(1, 2)$, and $beta(2, 1)$, respectively. This means that the round-robin strategy is far from "optimal" when the number of "arms" is large. In this case, the performance of our proposed strategy has been shown promising.

參考文獻

[1] Auer, P., Cesa-Bianchi, N., Fischer, P., 2002. Finite time analysis of the multi-armed bandit problem. Mach. Learn. 47, 235–256.

[2] Banks, J.S., Sundaram, R.K., 1992. Denumerable-armed bandits. Econometrica 60, 1071–1096.

[3] Berry, D.A., Chen, R.W., Heath, D.E., Shepp, L.A., Zame, A., 1997. Bandit Problems with Infinitely Many Arms. Ann. Statist. 25, 2103–2116.

[4] Berry, D.A., Fristedt, B., 1985. Bandit Problems: Sequential Allocations of Experiments. Chapman–Hall, London.

[5] Clayton, M.K., 1989. Covariate models for Bernoulli bandits. Sequential Anal. 8, 405–426.

[6] Gittins, J.C., 1989. Multi-armed Bandit Allocation Indices. John Wiley & Sons, Chichester.

[7] Gittins, J.C., 1979. Bandit processes and dynamic allocation indices. J. Roy. Statist. Soc. B 41, 148–177.

[8] Gittins, J.C., 1979. A dynamic allocation index for the discounted multi-armed bandit problem. Biometrika 66, 561–565.

[9] Guha, S., Munagala, K., Shi, P., 2010. Approximation algorithms for restless bandit problems. J. ACM 58, Article No. 3.

[10] Herschkorn, S.J., Peköz, E., Ross, S.M., 1995. Policies without memory for the infinite-armed Bernoulli bandit under the average-reward criterion. Probab. Eng. Inf. Sci. 10, 21–28.

[11] Lai, T.L., 1987. Adaptive treatment allocation and multi-armed bandit problem. Ann. Statist. 15, 1091–1114.

[12] Lai, T.L., Robbins, H., 1984. Asymptotically optimal allocation of treatments in sequential experiments, in Santner, T.J., Tamhane, A.C. (Eds.), Design of Experiments: Ranking and Selection, New York: Dekker.

[13] Lai, T.L., Yakowitz, S., 1995. Machine learning and nonparametric bandit theory. IEEE Trans. Autom. Control 40, 1199–1209.

[14] Lin, C.T., Shiau, C.J., 2000. Some Optimal Strategies for Bandit Problems with Beta Prior Distributions. Ann. Inst. Statist. Math. 52, 397–405.

[15] Mallows, C.L., Robbins, H., 1964. Some problems of optimal sampling strategy. J. Math. Anal. Appl. 8, 90–103.

[16] Narendra, K.S., Thathachar, M.A.L., 1989. Learning Automata: An Introduction. Prentice-Hall, Inc., Upper Saddle River, New Jersey.

[17] Pandey, S., Chakrabarti, D., Agarwal, D., 2007. Multi-armed Bandit Problems with Dependent Arms. Proc. Int. Conf. Mach. Learn.

[18] Powell, W.B., 2007. Approximate Dynamic Programming: Solving the Curses of Dimensionality (Chapter 10), John Wiley and Sons, New York.

[19] Robbins, H., 1952. Some aspects of the sequential design of experiments. Bull. Amer. Math. Soc. 58, 527–536.

[20] Sarkar, J., 1991. One-armed bandit problems with covariates. Ann. Statist. 19, 1978–2002.

[21] Thompson, W.R., 1933. On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. Biometrika 25, 285–294.

[22] Wang, C.C., Kulkarni, S.R., Poor, H.V., 2005. Bandit Problems with Side Observations. IEEE Trans. Autom. Control 50, 1–16.

[23] Whittle, P., 1982. Optimization over Time. Vol. I, John Wiley & Sons, Chichester.

[24] Whittle, P., 1983. Optimization over Time. Vol. II, John Wiley & Sons, Chichester.

[25] Woodroofe, M., 1979. A one-armed bandit problem with a concomitant variable. J. Amer. Statist. Assoc. 74, 799–806.

[26] Zoubeidi, T., 1994. Optimal allocations in sequential tests involving two populations with covariates. Comm. Statist. Theor. Meth. 23, 1215–1225.

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

---

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■ 達成目標

□ 未達成目標（請說明，以100字為限）

□ 實驗失敗

□ 因故實驗中斷

□ 其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 □無

技轉：□已技轉 □洽談中 □無

其他：（以100字為限）

Journal of Statistical Planning and Inference 142 (2012) 86–94.

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以500字為限）

本計畫借由統計之貝式假設與機率方法解決吃角子老虎之最佳化問題，其成果兼具學術理論與實際應用之價值--可以應用在許多領域如臨床試驗,線上工業實驗(on-line industrial experimentations),可調性網路路由（adaptive network routing)等。未來研究可針對模型假設適用性進行評估或與文獻中現有之方法進行比較。

國科會補助計畫衍生研發成果推廣資料表

日期:2011/10/16

<table>
<tr><td rowspan="3">國科會補助計畫</td><td>計畫名稱: 吃角子老虎問題之最佳貝氏策略</td></tr>
<tr><td>計畫主持人: 洪英超</td></tr>
<tr><td>計畫編號: 99-2118-M-004-004- 學門領域: 機率</td></tr>
<tr><td colspan="2">無研發成果推廣資料</td></tr>
</table>

99 年度專題研究計畫研究成果彙整表

計畫主持人：洪英超			計畫編號：99-2118-M-004-004-				
計畫名稱：吃角子老虎問題之最佳貝氏策略							
成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	2	2	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	1	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無。

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

Journal of Statistical Planning and Inference 142 (2012) 86－94.

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

本計畫借由統計之貝式假設與機率方法解決吃角子老虎之最佳化問題，其成果兼具學術理論與實際應用之價值--可以應用在許多領域如臨床試驗，線上工業實驗(on-line industrial experimentations)，可調性網路路由 (adaptive network routing)等。未來研究可針對模型假設適用性進行評估或與文獻中現有之方法進行比較。