Journal of the Korean Data & Information Science Society
2014, **25**(4), 697–703

http://dx.doi.org/10.7465/jkdi.2014.25.4.697
한국데이터정보과학회지

딜또는노딜 게임에서 딜금액 결정 모형†

송설희[1] · 안수한[2]

[12]서울시립대학교 통계학과

접수 2014년 6월 10일, 수정 2014년 6월 29일, 게재확정 2014년 7월 4일

요 약

본 논문에서 다루는 딜또는노딜 게임은 미국 NBC 방송국의 인기 오락 프로그램으로 여러 개의 단계로 구성되어 있다. 본 논문에서는 각 단계에서 제시되는 딜금액 결정 모형을 제한이 있는 회귀모형과 이차계획법을 이용하여 규명한다. 최종적으로 딜금액 결정 모형식과 관련하여 각 단계별로 선형식을 유도하고 딜금액이 선형식으로부터 얻어지는 값을 반올림한 정수값과 일치함을 시뮬레이션 자료를 이용하여 밝혔다.

주요용어: 딜또는노딜 게임, 이차계획법, 제한이 있는 회귀모형.

1. 머리말

딜또는노딜 (Deal-or-No-Deal) 게임은 대중적인 오락 TV 프로그램 중 하나로 미국, 영국, 호주 등 많은 나라에서 진행되어지고 있으며 온라인 버전과 텔레비전 버전이 있다. 본 논문에서는 미국 NBC 방송국의 딜또는노딜 게임 TV쇼를 다루고자 한다.

딜또는노딜 게임은 최대 10단계로 이루어져 있고 게임진행자, 은행가와 참가자로 구성되어져 있다. 게임진행에 필요한 26개의 서류가방이 제시되어 있고, 26개의 서류가방 안의 금액은 \$0.1 ∼ \$1,000,000 사이로 서로 다른 금액을 가지고 있으며 서류 가방 안에 제시되어 있는 금액을 아무도 모르는 상태에서 게임은 진행된다. 다음은 서류가방 안에 들어있는 26개의 금액 (단위 \$)이다.

Table 1.1 Prices in 26 suitcases

0.1	1	5	10	25	50	75	100	200
300	400	500	750	1,000	5,000	10,000	25,000	50,000
75,000	100,000	200,000	300,000	400,000	500,000	750,000	1,000,000	

게임이 시작되면 26개의 서류 가방 중에서 참가자는 일정한 개수의 가방을 매 단계마다 선택하여 열어볼 수 있게 된다. 매 단계마다 열어볼 수 있는 가방 수는 차례대로 6개, 5개, 4개, 3개, 2개, 1개, 1개, 1개, 1개, 1개이다. 매 단계가 끝났을 때마다 은행가는 참여자가 열지 않은 서류 가방을 구매하기 위해 일정한 금액을 제시하게 된다. 이 때 참가자는 은행가가 제시한 가격을 듣고 난 뒤, 자신이 가진 서류 가방과 거래하거나 거래하지 않을 수 있으며 거래를 한다면 딜금액 (deal price)을 받고 게임은 종료된다. 만약 참가자가 은행가와 거래하지 않게 된다면 그 다음 단계로 진행되는 방식으로 게임은 진행된다.

† 이 논문은 2013년도 서울시립대학교의 교내학술연구비에 의하여 수행되었음.
[1] (130-743) 서울시 동대문구 전농동 90번지, 서울시립대학교 통계학과, 석사과정.
[2] 교신저자: (130-743) 서울시 동대문구 전농동 90번지, 서울시립대학교 통계학과, 교수.
E-mail: sahn@uos.ac.kr

만약 9 단계가 끝난 후까지 거래하지 않게 된다면 참가자는 자신이 10단계에서 선택하지 않은 서류 가방 속의 금액만큼을 가지게 되며 게임은 종료된다. 딜또는노딜 게임에 대한 자세한 진행방식은 Post 등 (2008)의 Figure 1에서도 확인할 수 있다.

본 논문에서는 매 단계가 끝날 때 은행가가 제시하는 금액에 일정한 공식이 있을 것으로 보고 이에 대한 정확한 공식을 밝히고자 한다. 본 논문은 다음과 같이 구성되어 있다. 먼저 2절에서는 분석을 위한 자료와 은행가가 제시하는 금액에 대한 모형의 가정을 소개한다. 3절에서는 모형에서의 모수 추정을 위하여 이용되는 이차계획법 (quadratic program)에 대하여 소개한다. 4절에서는 모형의 유도 과정을 설명하고 시뮬레이션을 통하여 은행가가 제시하는 금액에 대한 정확한 공식을 유도하고 마지막으로 본 논문의 의미에 대한 간단한 결론을 제공한다.

2. 자료와 모형의 가정

매 단계가 끝났을 때 은행가가 참가자에게 제시하는 금액, 즉 딜금액의 결정모형 분석을 위하여 딜또는노딜 게임쇼 홈페이지 `http://www.nbc.com/Deal\_or\_No\_Deal/`에서 게임 시뮬레이션을 통하여 26개의 자료를 습득하였으며 습득한 자료의 구조는 Table 4.6을 참고하기 바란다. 분석에 앞서 시뮬레이션 자료로부터 다음과 같은 사실을 확인하였다. 첫째, 딜금액은 소수점이 없으므로 금액의 절사가 존재할 것이다. 둘째, 남은 금액의 총액이 큰 경우의 딜금액은 작은 경우의 금액보다 항상 크다. 이러한 사실로부터, 딜금액이 단순히 열지 않은 상자 속의 금액들의 선형결합으로 나타내어진다고 가정한다면 다음의 모형들을 고려하는 것이 타당하겠다.

$$Y_i = \Big[\sum_{j=1}^{26} a_{ij}x_j I(x_j \in L(i))\Big],\ a_{ij} \geq 0 \tag{2.1}$$

여기서 $L(i), i = 1, \cdots, 9$는 i번째 단계 후 남은 서류 가방들 속에 들어 있는 금액들의 집합을 나타내고, 변수 $x_j, j = 1, \cdots, 26$은 26개의 금액들을 작은 금액부터 순서대로 나열했을 때, j번째 금액을 나타내도록 한다. 또한 $[b]$는 소수점 첫째 자리에서 반올림한 값을 나타내고 함수 $I(\cdot)$은 지표함수 (indicator function)을 나타낸다. 참고로 10단계 이후는 한 개의 가방만이 남아 있고 가방 안에 있는 금액이 참가자의 상금으로 제공되므로 모형을 고려할 필요가 없다.

계수 결정과 관련하여 식 (2.1)의 모형들을 고려하여 계수들 a_{ij}를 직접 구하는 것은 어려운 문제일 뿐만 아니라 이러한 문제의 해법을 제시한 연구결과는 저자들이 아는 범위에서는 찾을 수가 없었다. 따라서 절사부분을 오차항으로 놓은 다음의 제한이 있는 회귀모형을 고려하고

$$Y_i = \sum_{j=1}^{26} a_{ij}x_j I(x_j \in L(i)) + \epsilon_i,\ a_{ij} \geq 0 \tag{2.2}$$

계수들 a_{ij}를 오차제곱합을 최소로 하는 값으로 추정하고 이를 이용하여 계수들의 값을 결정하고자 하였다. 여기서 위의 식에서 오차제곱합을 이용한 계수추정 방법은 다음 절에서 소개하는 이차계획법의 특수한 경우이다.

3. 이차계획법

웹사이트 `http://www.nbc.com/Deal\_or\_No\_Deal/`에서 n번의 시뮬레이션을 했다고 가정했을 때, $L(i,j), i = 1, \cdots, 9; j = 1, \cdots, n$을 j번째 시뮬레이션에서 i단계 이후에 남은 금액들의 집합이

라 하고, y_{ij}를 j번째 시뮬레이션에서 i단계 이후 은행가가 참가자에 제시한 딜금액, $z_{ijk} = x_k I(k \in L(i,j)), k = 1, \cdots, 26$으로 정의하면 식 (2.2)로 부터 다음의 식을 얻는다.

$$y_{ij} = \sum_{k=1}^{26} a_{ik} z_{ijk} + \epsilon_{ij},\ a_{ik} \geq 0,\ j = 1, \cdots, n.$$

여기서 n 차원 열벡터 $\vec{y_i} = (y_{ij})$, 26 차원 열벡터 $\vec{a}_i = (a_{ik})$ 그리고 $n \times 26$ 차원 행렬 $Z_i = (z_{ijk})$를 정의하면 오차제곱합을 최소로 하는 문제는 다음의 이차계획법 문제로 귀결된다.

$$\begin{aligned} \text{minimize} \quad & \vec{a}_i^T Z_i^T Z_i \vec{a}_i^T - 2\vec{y}_i^T Z_i \vec{a}_i + \vec{y}_i^T \vec{y}_i \\ \text{subject to} \quad & \vec{a}_i \geq 0. \end{aligned} \tag{3.1}$$

일반적인 이차계획법 문제에 대해서는 다양한 접근방법이 소개되어 있고 또한 많은 알고리즘이 제공되어 있다. 대표적으로 primal method (Beale, 1955), dual method (Lemke, 1962; Goldfarb와 Idnani, 1983), 제약이 있는 최소제곱법 (Stoer, 1971) 등이 있다. 이차계획법은 서포트 벡터 머신 (support vector machine)에도 응용되며 (Hwang과 Shim, 2010, 2012) 좀 더 자세한 내용은 주어진 참고문헌을 참고하기 바란다.

본 논문에의 자료분석은 Goldfarb와 Idnani (1983)의 쌍대방법 (dual method)에 기반을 둔 R 함수 solve.QP를 이용하였다. 함수 solve.QP 에 대한 내용은 `http://cran.r-project.org/web/package/quadprog/quadprog.pdf`를 참조하기 바란다.

4. 딜금액 모형

4.1. 시뮬레이션 자료를 이용한 모형 분석

분석에 앞서 기호를 추가적으로 소개한다. 먼저 $n(i), i = 1, \cdots, 9$는 i번째 단계 (stage) 후 남은 서류 가방의 개수, $d(i)$는 i번째 단계에서 참가자가 선택해야 할 서류 가방의 개수를 나타낸다. 다음 표는 $n(i)$와 $d(i)$를 정리한 것이다.

Table 4.1 Number of suitcases that participants choose and remaining unchosen suitcases at each stage

	stage 1	stage 2	stage 3	stage 4	stage 5	stage 6	stage 7	stage 8	stage 9
$n(i)$	20	15	11	8	6	5	4	3	2
$d(i)$	6	5	4	3	2	1	1	1	1
$\frac{1}{n(i)}$	0.05	0.0667	0.0909	0.125	0.1667	0.2	0.25	0.3333	0.5

다음의 Table 4.2의 값들은 식 (2.2)의 모형에 대하여 3절에서의 QP 알고리즘을 이용하여 오차제곱합을 최소로 하는 계수들을 구한 것이다. Table 4.2에서 볼 수 있듯이 금액이 작은 경우에 해당하는 계수의 추정치는 변동 폭이 심하지만 금액이 커질수록 계수의 변동 폭이 없어짐을 알 수 있다. 이는 금액이 작을수록 절사의 영향을 많이 받는 것으로 추측할 수 있다. Table 4.2에서 발견한 또 다른 중요한 사실들은 다음과 같다. 첫째, 첫 번째에서 19번째까지의 금액에 해당하는 계수는 각 단계에서 동일할 수 있다. 이는 특히 5단계에서의 계수 추정치로부터 추측할 수 있다. 둘째, i단계에서 첫 번째에서 19번째까지의 금액에 해당하는 계수는 Table 4.1에서의 $\frac{1}{n(i+1)}$과 관계가 있다. 또한 계수가 비교적 안정적으로 나타나는 16번째부터 19번째 금액에 해당하는 계수 추정치를 고려하면 소수점 둘째자리까지 유효하게 보이고 이는 $\frac{1}{n(i+1)}$값의 소수점 셋째자리에서 반올림한 값과 비슷하다. 셋째, 9단계 이후는 두 개의 금액만 남게 되는데 이 경우 첫 번째에서 19번째까지의 금액에 해당하는 계수는 0.5임을 쉽게 확인할 수

있다. 넷째, 소수점 넷째자리까지 유효하게 보이는 매 단계에서 20번째부터 26번째 금액에 해당하는 계수는 동일할 수 있으며 이는 첫 번째에서 19번째까지의 금액에 해당하는 계수와는 다른 규칙을 보인다. 이러한 다른 규칙이 사실이라고 했을 때, 이에 대한 이유는 정확히 예측할 수 없으나 단지 액수의 단위가 달라지면서 다른 규칙을 적용했을 거라고 추측할 수는 있다.

Table 4.2 Regression coefficients result

	stage 1	stage 2	stage 3	stage 4	stage 5	stage 6	stage 7
1	0.00000	0.00000	0.00000	33.9130	0.00000	0.00000	372.569
2	0.10376	0.00000	0.06826	0.00000	0.00000	1.49660	6.54758
3	0.24126	0.00000	0.21407	0.00000	0.20000	0.41545	3.85512
4	0.01223	0.14057	0.07265	0.22734	0.20000	0.23043	0.34478
5	0.05590	0.20454	0.05787	0.23398	0.20000	0.25253	0.00000
6	0.09838	0.12870	0.12171	0.08675	0.20000	0.25473	0.00000
7	0.03387	0.06769	0.13967	0.15931	0.20000	0.25567	0.31105
8	0.07146	0.06607	0.13332	0.32287	0.20000	0.24254	0.19089
9	0.07243	0.08538	0.12597	0.33454	0.20000	0.25043	0.43689
10	0.07011	0.09464	0.13172	0.15730	0.20000	0.25122	0.48438
11	0.06995	0.08949	0.12877	0.19843	0.20000	0.24754	0.22691
12	0.06999	0.09983	0.13056	0.15263	0.20000	0.24990	0.36149
13	0.06560	0.10435	0.12932	0.15423	0.20000	0.25039	0.29777
14	0.06806	0.09167	0.13049	0.16782	0.20000	0.24936	0.27328
15	0.07145	0.08936	0.12956	0.17040	0.20000	0.24982	0.32841
16	0.07008	0.08982	0.13010	0.17005	0.20000	0.25000	0.32577
17	0.07001	0.08987	0.13004	0.16963	0.20000	0.25000	0.33200
18	0.07000	0.08996	0.12997	0.16985	0.20000	0.25002	0.33044
19	0.06999	0.09000	0.13003	0.16996	0.20000	0.25000	0.33035
20	0.00770	0.01437	0.02731	0.04426	0.06200	0.10250	0.16812
21	0.00770	0.01441	0.02729	0.04429	0.06200	0.01250	0.16871
22	0.00770	0.01439	0.02730	0.04417	0.06200	0.10250	0.16805
23	0.00770	0.01439	0.02730	0.04419	0.06200	0.10250	0.16847
24	0.00770	0.01441	0.02730	0.04417	0.06200	0.10250	0.16826
25	0.00770	0.01440	0.02730	0.04420	0.06200	0.10250	0.16833
26	0.00770	0.01440	0.02730	0.04419	0.06200	0.10250	0.16826

위의 발견으로부터 모형분석을 위해 먼저 각 단계에서 첫 번째에서 19번째까지의 금액에 해당하는 계수는 $\frac{1}{n(i+1)}$ 값의 소수점 셋째자리에서 반올림한 값이고 20번째부터 26번째 금액에 해당하는 계수는 동일하다라는 가정에서 출발하여 20번째에서 26번째까지의 금액에 해당하는 계수를 결정하고자 하였다. 물론 액수가 작은 금액에 해당하는 계수들의 변동폭이 커서 이러한 가정이 무리할 수도 있으나, 모형에 영향이 적을 수밖에 없는 작은 액수의 금액들을 위해서 다른 규칙을 쓰지는 않았을 것이라고 가정하였다. 각 단계에서 첫 번째에서 19번째까지의 금액에 해당하는 계수를 $w(1,i), i = 1, \cdots, 9$ 그리고 20번째에서 26번째까지의 금액에 해당하는 계수를 $w(2,i), i = 1, \cdots, 9$로 나타내었을 때, 계수 $w(1,i)$의 값들을 표로 정리하면 다음과 같다.

Table 4.3 Model coefficient for prices x_j, $j = 1, \cdots, 19$ at each stage

$w(1,i)$	$w(1,1)$	$w(1,2)$	$w(1,3)$	$w(1,4)$	$w(1,5)$	$w(1,6)$	$w(1,7)$	$w(1,8)$	$w(1,9)$
	0.07	0.09	0.13	0.17	0.2	0.25	0.33	0.5	0.5

나머지 20번째에서 26번째까지의 금액에 해당하는 계수 $w(2,i), i = 1, \cdots, 9$의 결정을 위해서 다양한 시도를 하였으며, 최종적으로, Table 4.2에서의 20번째부터 26번째 금액에 해당하는 계수들을 Table 4.3에서의 $w(1,i)$의 값으로 나눈 후 $w(1,i)$에서와 똑같이 소수점 셋째자리에서 반올림하고 이를 $\beta(i)$로 나타내면 Table 4.4에서와 같은 간단한 결과와 규칙을 얻을 수 있다. 또한 이러한 규칙이 8단계 이후에

도 적용된다고 가정하면 $\beta(8) = 0.61, \beta(9) = 0.71$이고 $w(2,i), i = 1, \cdots, 9$에 대한 값은 Table 4.5에서와 같이 얻을 수 있다.

Table 4.4 Values of $\beta(i)$ for $i = 1, \cdots, 7$

$\beta(i)$	$\beta(1)$	$\beta(2)$	$\beta(3)$	$\beta(4)$	$\beta(5)$	$\beta(6)$	$\beta(7)$
	0.11	0.16	0.21	0.26	0.31	0.41	0.51
		0.05+	0.05+	0.05+	0.05+	0.1+	0.1+

Table 4.5 Model coefficient for prices x_j, $j = 20, \cdots, 26$ at each stage

$w(2,i)$	$w(2,1)$	$w(2,2)$	$w(2,3)$	$w(2,4)$	$w(2,5)$	$w(2,6)$	$w(2,7)$	$w(2,8)$	$w(2,9)$
	0.0077	0.0144	0.273	0.0442	0.062	0.1025	0.1683	0.3050	0.3550

위의 결과로부터 $Y_i, i = 1, \cdots, 9$의 결정 모형으로 다음의 모형을 고려할 수 있다.

$$Y_i = \Big[\sum_{j=1}^{19} w(1,i) x_j I(x_j \in L(i)) + \sum_{j=20}^{26} w(2,i) x_j I(x_j \in L(i)) \Big] \tag{4.1}$$

다음 절에서는 계수 $w(1,i), w(2,i), i = 1, \cdots, 9$들이 Table 4.3과 Table 4.5에서와 같이 주어지는 경우 모형 (4.1)의 타당성 유무를 시뮬레이션을 통하여 확인한다.

Table 4.6 Simulation and deal price at each stage

	price ($)	stage 1	stage 2	stage 3	stage 4	stage 5	stage 6	stage 7	stage 8	stage 9
1	0.1			x	x	x	x	x	x	x
2	1				x	x	x	x	x	x
3	5		x	x	x	x	x	x	x	x
4	10	x	x	x	x	x	x	x	x	x
5	25	x	x	x	x	x	x	x	x	x
6	50		x	x	x	x	x	x	x	x
7	75								x	x
8	100					x	x	x	x	x
9	200			x	x	x	x	x	x	x
10	300									
11	400					x	x	x	x	x
12	500		x	x	x	x	x	x	x	x
13	750				x	x	x	x	x	x
14	1,000		x	x	x	x	x	x	x	x
15	5,000									
16	10,000	x	x	x	x	x	x	x	x	x
17	25,000			x	x	x	x	x	x	x
18	50,000	x	x	x	x	x	x	x	x	x
19	75,000	x	x	x	x	x	x	x	x	x
20	100,000		x	x	x	x	x	x	x	x
21	200,000			x	x	x	x	x	x	x
22	300,000							x	x	x
23	400,000				x	x	x	x	x	x
24	500,000									x
25	750,000						x	x	x	x
26	1,000,000	x	x	x	x	x	x	x	x	x
Deal		19,662	33,824	54,096	69,509	97,175	83,344	85,924	155,150	2,650

4.2. 시뮬레이션을 이용한 모형의 선택

이 절에서는 시뮬레이션을 이용한 딜금액 자료와 모형 (4.1)을 이용한 딜금액을 비교하고자 한다. 참고로 Table 4.6은 시뮬레이션 자료와 각 단계에서의 Deal 가격을 포함하고 있으며, 여기서 x 표시는 각 단계까지 해당 금액이 들어있는 상자가 참가자의 선택에 의해 열렸음을 나타낸다. 이 시뮬레이션과 관련하여 Table 4.7에서의 값들은 Table 4.3 그리고 Table 4.5에서의 계수값과 식 (4.1)에서의 모형을 이용하여 구한 매 단계 이후의 딜금액이다.

Table 4.7 Deal prices obtained using our model in (4.1)

	stage 1	stage 2	stage 3	stage 4	stage 5	stage 6	stage 7	stage 8	stage 9
Deal	19,662	33,824	54,096	69,509	97,175	83,344	85,924	155,150	2,650

위의 Table 4.7과 Table 4.6에서의 딜금액을 비교하면 정확히 일치함을 알 수 있다. 또 다른 25번의 시뮬레이션에서도 똑같은 결과를 확인할 수 있으며 이를 통하여 딜또는노딜 게임에서의 딜금액 결정모형은 식 (4.1)에서 주어지는 모형으로 계수 값은 Table 4.3과 Table 4.5에 주어진 값임을 결론내릴 수 있다.

5. 결론

본 논문에서는 딜또는노딜 게임에서의 금액 결정모형을 제한이 있는 회귀모형과 이차계획법을 이용하여 밝혀내었다. 본 논문의 결과는 사회에서 이용되는 모형과 관련하여 모형에 대한 사전지식과 가정 그리고 관련 자료에 대한 적절한 통계적 기법 적용 등을 이용하면 모형에서 사용되는 공식을 밝힐 수 있음을 보여주는 좋은 예이다.

References

Beale, E. M. L. (1955). On minimaizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society B*, **17**, 173-184.

Goldfarb, D. and Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. *Mahmatical Programming*, **27**, 1-33.

Hwang, C. H. and Shim, J. Y. (2010). Semiparametric support vector machine for accelerated failure time model. *Journal of the Korean Data & Information Science Society*, **21**, 765-775.

Hwang, C. H. and Shim, J. Y. (2012). Smoothing Kaplan-Meier estimate using monotone support vector regression. *Journal of the Korean Data & Information Science Society*, **23**, 1045-1054.

Lemke, C. E. (1979). A method for solving vertain constrained least squares problem. *Mathmatical Programming*, **16**, 141-158.

Stoer, J. (1971). On the numerical solution of constrained least squares problem. *SIAM journal on Numerical Analysis*, **8**, 382-411.

Post, T., Assem, M. J. V. D., Baltussen, G. and Thaler, R. H. (2008). Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review*, **98**, 38-71.

Journal of the Korean Data & Information Science Society
2014, **25**(4), 697–703

http://dx.doi.org/10.7465/jkdi.2014.25.4.697
한국데이터정보과학회지

Deal price model in Deal-or-No-Deal game†

Seolhee Song[1] · Soohan Ahn[2]

[12]Department of Statistics, University of Seoul

Received 10 June 2014, revised 29 June 2014, accepted 4 July 2014

Abstract

Deal-or-No-Deal game is a famous TV show program of NBC, USA, which is composed of 10 stages at most. At each stage from the first and the ninth, a banker suggests a deal price to participants. In this paper, we intend to reveal the banker's deal price model using a constrained linear model and quadratic program. As results, we provide a linear model in relation to the deal price at each stage and then show using simulation data that the deal price is equal to the nearest integer of the value to be obtained by the provided linear model.

Keywords: Constrained linear model, Deal-or-No-Deal game, quadratic program.

---

† This work was supported by the 2013 Research Fund of the University of Seoul.

1 Graduate student, Department of Statistics, University of Seoul, Seoul 130-743, Korea.
2 Corresponding author: Professor, Department of Statistics, University of Seoul, Seoul 130-743, Korea.
E-mail: sahn@uos.ac.kr