学籍番号		論文題目	わたしの アモーレ 君の名は？ —最適停止の数理—
氏名	上坂 健		

1　n 人最良選択問題-無情報問題-

無情報の n 人最良選択問題とは次のような問題である．これからあなたは 1 人ずつお見合いをする．最大 n 人とお見合いできることになっているが,「これだ」という人（相対ランク 1 位）が現れた時点でその人を選び，その後の人とは会わずにお見合いを終える．できれば n 人中で最も好みの人（絶対ランク 1 位）を選びたい．そこで問題はどのような戦略をとれば絶対ランク 1 位の人を選ぶ確率が最大にできるかということである．最良選択問題の場合,「$r-1$ 番目までの人は選択せずに別れて，それ以降での相対ベストの人を選択しろ」という戦略 s_{r-1} をとる．戦略 s_{r-1} に従って選択を停止する方法を**停止規則**と呼び，N_r と表す．停止規則 N_r の下で絶対ベストを選択できる確率を最大化する r を r^* と表し，これを**閾値**と呼ぶ．また停止規則 N_{r^*} を**最適停止規則**と呼ぶ．

定理 1.1　n 人最良選択問題の最適停止規則 N_{r^*} は「r^*-1 人目までは見送り，r^* 人以降での相対ベストを選択せよ」である．また絶対ベストを選択する確率は $P_{r^*}=\dfrac{r^*-1}{n}\displaystyle\sum_{k=r^*}^{n}\frac{1}{k-1}$，ただし $r^*=\min\left\{r\geq 1 \mid \displaystyle\sum_{k=r+1}^{n}\frac{1}{k-1}\leq 1\right\}$ である．

十分大きな n に対して，$1\geq\displaystyle\sum_{k=r+1}^{n}\frac{1}{k-1}\approx\int_r^n\frac{1}{y}dy=\log\frac{n}{r}$ と近似され，この不等式を満たす最小の r を r^* とすると，$r^*=\left\lceil\frac{n}{e}\right\rceil$ となり，十分大きなに対しては相手の $e^{-1}\approx 36.8\%$ までは選択せずにそれ以後の相対ベストを選択するのが最適であり，そのとき絶対ベストを得る確率は，$\log\dfrac{n}{r^*}=1$ を用いて，$P_{r^*}\approx\dfrac{r^*}{n}\log\dfrac{n}{r^*}=e^{-1}$ となる．すなわち N_{r^*} に従った際に絶対ベストを選択する確率は約 37% となる．

2　n 人最良選択問題-完全情報問題-

逐次に出逢う n 人の相手の価値 $X_1,X_2,\ldots,X_n$ は値域を $[a,b]$ で確率分布 F に従う独立同分布確率変数とする．そのとき n 人の中で最も高い価値をもつ相手を選ぶ確率を最大にする最適停止規則 N を求める．$n-j$ 人目の相手の価値が $X_{n-j}=x$ であり，今までに付き合った $n-j$ 人中最も高い，すなわち相対ベストであるとき，この人を選択したときこの人が n 人中の絶対ベストである確率 y_{n-j} は，F の密度関数を $f(x)$ とすると,

$$y_{n-j}=P(X_{n-j}=\max\{X_1,\ldots,X_n\}|X_{n-j}=x)=\prod_{k=1}^{j}P(x>X_{n-j+k})=\prod_{k=1}^{j}\int_a^x f(x)dy=F(x)^j$$

となる．一方で，$n-j$ 人目の相手が価値 $X_{n-j}=x$ の値をもつ相対ベストで，この人と付き合って最適にふるまい絶対ベストを選択する最大の確率を $V_{n-j}(x)$ とする．このとき $X_{n-j}=x$ の値をもつ相対ベストの選択を見送り，これ以降最適にふるまって絶対ベストを選択する最大確率 z_{n-j} は

$$\begin{aligned}z_{n-j}&=\sum_{k=1}^{j}P(n-j\text{ 人目の相対ベスト以降，初めての相対ベストが }n-j+k\text{ 人目に出る}\\&\qquad\qquad|n-j\text{ 人目が相対ベスト})\times EV_{n-j+k}(X_{n-j+k})\\&=\sum_{k=1}^{j}\int_a^x f(y)\cdots\int_a^x f(y)\int_x^b f(y)V_{n-j+k}(y)dy\cdots dy=\sum_{k=1}^{j}F(x)^{k-1}\int_x^b f(y)V_{n-j+k}(y)dy\end{aligned}$$

となる．これより $j=1,\ldots,n-1$ に対して，最適方程式は

$$V_{n-j}(x)=\max\left\{F(x)^j,\sum_{k=1}^{j}F(x)^{k-1}\int_x^b f(y)V_{n-j+k}(y)dy\right\},V_n(x)=1$$

となる．したがって $n-j$ 人目の相手が値 x をもつ相対ベストである状態を $(n-j,x)$ としたとき OLA 停止領域 B は

$$B=\left\{(n-j,x)\;\middle|\;F(x)^j\geq\sum_{k=1}^{j}F(x)^{k-1}\int_x^b f(y)F(y)^{j-k}dy\right\}=\left\{(n-j,x)\;\middle|\;G_j(x)=\sum_{k=1}^{j}\frac{1}{k}\left(\frac{1}{F(x)^k}-1\right)\leq 1\right\}$$

で与えられる．

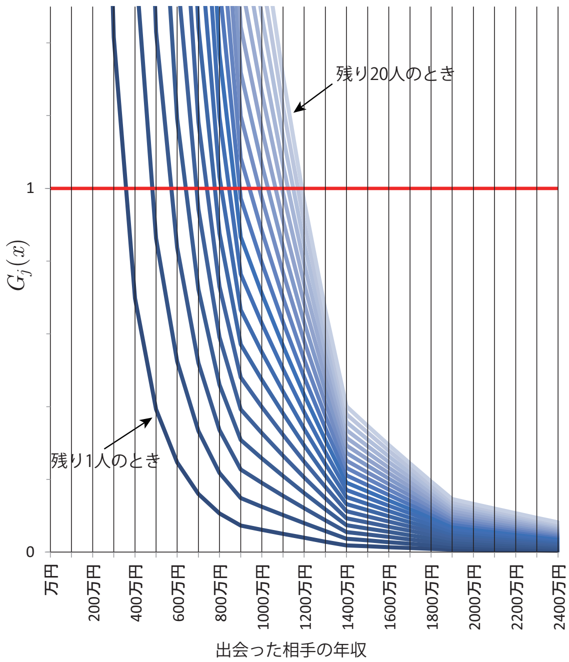
定理 2.1　逐次に出逢う n 人の相手の価値 $X_1, X_2, \ldots, X_n$ は値域が区間 $[a,b]$ で $f(x)$ を密度関数に持つ分布 F に従う独立な確率変数とする．j を残り人数として，n 人の中で最も高い価値をもつ相手を選ぶ確率を最大にする最適停止規則 N は

$$\min\{n-j \geq 1 \mid G_j(x) \leq 1\} = \min\{n-j \geq 1 \mid x \geq t_j \text{かつ } X_{n-j} = \max\{X_1, \ldots, X_{n-j}\} = x\}$$

である．ただし，t_j は x の方程式 $G_j(x) = 1$ の区間 $[a,b]$ における解である．

この結果を現実の例に当てはめ，$G_j(x)$ の値を用いて相手の期待年収を求めてみよう．今回は主人公を女性とし，n として [1] に基づき日本人女性の初婚までの恋人人数の期待値である 3 をとる．さらに恋人の価値が従う分布 F として 2015 年日本人男性の平均年収 [2] を採用する．数値計算においては，年収を 100 万円単位に分割し，年収の上限を 2500 万円としてモデル化する．日本人男性における各階層 $100(k-1) \leq x \leq 100k, k = 1, \ldots, 25$ の割合を $f(k)$ として年収の分布を $F(x) = \sum_{k=1}^{[x/100]} f(k)$ と近似し，実際に実行したものが表 1 である．表 1 によれば $x = 500$ のときに $G_2(x) \leq 1$ が成り立っている．すなわち 1 人目で年収 500 万円以上の人に出会えればそこで停止した方が良いということである．同様に 2 人目では年収 400 万円以上の人に出会えればそこで停止した方が良い．そして今回の OLA 停止領域に従った際に得られる相手の年収の期待値は 679 万円となった．今回のモデルの日本人男性の年収の期待値は 501 万円であるので，それを 178 万円上回る結果が得られた． 同様に n を 10 とした場合を考えてみた結果，得られる相手の年収の期待値は 951 万円となった．

平均年収分布 $F(x)$	
階層 (万円)	割合 (%)
0	3.1
100	7.3
200	12.7
300	18.3
400	17.4
500	12.9
600	8.4
700	6.0
800	4.2
900	2.8
1000-1500	5.0
1500-2000	1.1
2000-2500	0.4



n	(万円)	n	(万円)
1	501	11	974
2	607	12	996
3	679	13	1017
4	738	14	1036
5	787	15	1054
6	827	16	1071
7	864	17	1087
8	895	18	1101
9	924	19	1115
10	951	20	1128

表 2　出逢う最大人数 n と年収の期待値．

表 1　2015 年の日本人男性の年収分布 $F(x)$ と 3 人までと決めた場合の女性の停止規則．

$n = 3$ のときと比較すると得られる相手の年収の期待値が 272 万円上がるので相手の年収にこだわる人にとっては 3 人より 10 人と出逢う方が良い．更に 1 人ずつ n を増やした場合の相手の年収の期待値を算出したものが表 2 である．一般に結婚までに 20 人以上と出逢う人はごくわずかだと考えるので n として 20 までを算出すると，n の増加に伴い期待年収の伸びは悪くなることが分かる．n を増やせば期待年収が上がるものの，その分出会いに要する時間もかかってしまうので，表 2 を参考にして n の値を決めたほうが良いだろう．また n によらず残り 1 人になった際に停止するか否かの指標は年収 400 万円以上なので，一般的に私達大学生は残り 0 人，つまり最後の 1 人として登場しなければ選択してもらえないということである．

参考文献

[1] 小林盾, 大林裕子, 恋愛経験は結婚の前提条件か:2015 年家族形成とキャリア形成についての全国調査による量的分析, 成蹊大学大学院文学研究科 第 24 号, pp. 1-15, 2016.

[2] 国税庁 2015 年民間給与実態統計調査, 国税庁, 2016.

[3] D.V.Lindley, Dynamic programming and decision theory, Appl Statist, pp. 39-51, 1961.