2013 年 06 月 19 日

表が出るまで投げ続ける確率の問題について

新潟工科大学 情報電子工学科 竹野茂治

1 はじめに

確率に以下のような有名な問題がある (「サンクトペテルブルグのパラドックス」と呼ばれているらしい)。

> 「10 円玉を、表が出るまで投げ続けてくれ。1 回で表が出れば 10 円あげるが、2 回目に初めて表が出たら倍の 20 円やろう。3 回目に初めて表が出たら、さらにその倍の 40 円やる、という具合に、裏が出続ける度に賞金を倍にする。これは、期待値は無限大になるから、100 万円払っても必ずあなたが得をするはずだが、これを 1 万円払ってやらないかね。」

確かに、この場合の賞金の期待値は、

$$\frac{1}{2}\times 10+\frac{1}{4}\times 20+\frac{1}{8}\times 40+\cdots=5+5+5+\cdots=\infty$$

となるのだが、1 万円でやるかと言われると、たいていの人が断わるのではないか、という話である。

それは、半分の確率で 10 円、3/4 の確率で 20 円以下にしかならないので、1 万円だと参加費が高いと感じるからだと思うが、例えば宝くじもほとんどの人が当たらないのに高額賞金を夢見て多くの人が実際に宝くじを買っていること、およびこの賭けの方が期待値ははるかに高いことを考えると、やや微妙なところである。

それで、これについて、

- 1 万円、および 1000 万円もうかる人はどれくらいの割合でいるのか、それを宝くじと比較するとどうか
- この賭けを連続して何回続けても構わない場合、10 円玉を例えば 1000 回投げ続けた場合はいくら位もらえることになるか

などの観点で少し考えてみたいと思う。

2 もうかる人の割合

この問題は、**幾何分布** と呼ばれる確率分布の問題で、パラメータを少し一般化して考えてみる。今、幾何分布 $G(p)$ $(p>0)$ に従う確率変数を X とする。すなわち、X は 0 以上の整数値を取り、

$$P(X=n)=(X \text{ が } n \text{ に等しい確率})=p(1-p)^{n-1}$$

であるとする。今回の問題では、$p=$ 表が出る確率 $=1/2$ で、X は「表が出るまで投げた回数 -1」に対応し、賞金は $Y=A\cdot B^X$ 円 $(A=10,\ b=2)$ となっている。また、1 回の賭けにかかる参加料を $c=1$ 万円とする。

まず賞金が d 円以上となる割合であるが、これは、

$$Y=A\cdot b^X\geq d,\quad X\geq \log_b\frac{d}{A}$$

より、$N=\lceil\log_b(d/A)\rceil$ ($\lceil x\rceil$ は、x 以上の最小の整数) に対して $P(X\geq N)$ を求めればよい。

$$P(X\geq N)=\sum_{j=N}^{\infty}P(X=j)=\sum_{j=N}^{\infty}p(1-p)^{j-1}$$

は、初項 $p(1-p)^{N-1}$、公比 $(1-p)$ の無限等比級数であるから、

$$P(X\geq N)=\frac{p(1-p)^{N-1}}{1-(1-p)}=\frac{p(1-p)^{N-1}}{p}=(1-p)^{N-1}$$

となる。

およそで考えれば、$N\approx\log_b(d/A)$ であり、今の場合は $1-p=1/2=1/b$ に等しいので、

$$P(X\geq N)\approx(1-p)^{\log_b(d/A)-1}=b^{-\log_b(d/A)+1}=\frac{Ab}{d}$$

となり、よって d 円以上当たる人の割合は、d に反比例することがわかる。

$A = 10, b = 2$ で $d = 1$ 万円の場合は $P(X \geq N) \approx 20/10000$ だからほぼ 500 人に 1 人、$d = 1$ 千万円の場合は $P(X \geq N) \approx 20/10^7$ だからほぼ 50 万人に 1 人、となる。

ちなみに、ジャンボ宝くじだと、元が取れる確率は 1/10 だから、こちらの 1/500 はそれと比べてかなり低いが、ジャンボ宝くじが 3000 万円以上当たる確率は 300 万人に 1 人だそうで、こちらは 3000 万円以上だと 150 万人に 1 人なので、桁が違うというほど離れてはいない。

だから、参加料が宝くじと同様、例えば 1 回 200 円程度ならばやるという人は増えるかもしれないが、1 回 1 万円だと高く感じるのは無理はなさそうである。

3 ある程度の回数やり続けた場合

次に、ある程度の回数やり続けることを考える。

1 回投げて 1 度の賭けが終わる場合もあるし、10 回投げてやっと 1 度の賭けが終わる場合もある。しかし、実際に行える時間は「賭けの回数」というよりも、「投げる回数」の方だろうから、例えば賭けを続けてやることにし、10 円玉を $N = 1000$ 回連続して投げたとして、その結果いくら位もらえることになるかを考えてみる。

この場合は、参加料 $c = 1$ 万円も毎度の賭けで支払うことになるので、それを賭けの回数分引かなければいけないことに注意する。また、丁度 N 回目で賭けが終わらない場合もあるが、その場合は賭けの参加料は払うが賞金は貰えない、と考えることにする。

例えば、$N = 4$ で (ウ、オ、オ、ウ) (「ウ」=裏、「オ」=表) となった場合は、1 回目の賭けでは 20 円、2 回目は 10 円の賞金がもらえ、3 回目は終了していないので賞金はなし、参加料は 3 回分払うので、結局 $30 - 30000$ 円の収入ということになる。

今、例として $N = 3$ の場合を、最後に結果が確定した賭けまでに投げた回数で表 1 に分類してみる。表がでる確率は p、賞金は $Y = Ab^X$ 円、賭けの参加料は 1 回 c 円とする。

N に対する収入の期待値を A_N とし、賞金の期待値を A_N^+、参加料の期待値を A_N^- とする。また、A_N, A_N^+, A_N^- のうち、最後が表で丁度終わっているもの (表 1 でいえば回数が 3 の 4 通り) の期待値部分をそれぞれ B_N, B_N^+, B_N^- とする。

この場合、$A_N = A_N^+ - A_N^-$, $B_N = B_N^+ - B_N^-$ で、$N = 3$ の場合は表 1 より

$$B_3^+ = Ab^2p(1-p)^2 + 2(A + Ab)p^2(1-p) + 3Ap^3,$$

回数	並び	確率	賞金	参加料
0	(ウ、ウ、ウ)	$(1-p)^3$	0	c
1	(オ、ウ、ウ)	$p(1-p)^2$	A	$2c$
2	(オ、オ、ウ)	$p^2(1-p)$	$2A$	$3c$
	(ウ、オ、ウ)	$p(1-p)^2$	Ab	$2c$
3	(ウ、ウ、オ)	$p(1-p)^2$	Ab^2	c
	(オ、ウ、オ)	$p^2(1-p)$	$A+Ab$	$2c$
	(オ、オ、オ)	p^3	$3A$	$3c$
	(ウ、オ、オ)	$p^2(1-p)$	$Ab+A$	$2c$

表 1: $N=3$ の場合の一覧

$$\begin{aligned}
B_3^- &= cp(1-p)^2 + 4cp^2(1-p) + 3cp^3,\\
A_3^+ &= B_3^+ + (A+Ab)p(1-p)^2 + 2Ap^2(1-p),\\
A_3^- &= B_3^- + c(1-p)^3 + 4cp(1-p)^2 + 3cp^2(1-p)
\end{aligned}$$

となる。また、$N=1$ のときは、容易に

$$B_1^+ = Ap,\quad B_1^- = cp,\quad A_1^+ = B_1^+,\quad A_1^- = B_1^- + c(1-p) \tag{1}$$

となることがわかる。

4 参加料の期待値

まず、参加料の期待値である A_N^- を求める。

B_N^- は、A_N^- のうち N 回目が表で終わっているものに対する和で、それをさらに以下のように分割する。

- C_0: 最初の $(N-1)$ 回が裏で、N 回目が表
- C_k: k 回目が表で、$(k+1)$ 回目から $(N-1)$ 回目までが裏で、N 回目が表 $(k=1,2,\dots,N-1)$

例えば表 1 の $N=3$ の例で言えば、B_3^- は回数 3 の 4 通りのものに対応し、それは上から順に C_0, C_1, C_2, C_2 と分類されることになる。なおこれは、最後の表を裏と見れば、回数が 2 以下の分類に等しいことがわかるだろう。

各 C_k の最初の k 回は、表で終わる k 回の任意の並びだから丁度 B_k^- と同じ状況で、その後賭け 1 回分 (確率は $(1-p)^{N-1-k}p$) だけをその中のすべての事象に追加することになる。B_k^- に含まれる事象の確率の総和は p (k 回目が表で $(k-1)$ 回目までは任意) なので、結局 C_k に対応する B_N^- の部分の和は

$$(B_k^- + cp)(1-p)^{N-1-k}p$$

に等しいことがわかる。C_0 は賭け 1 回分だけで確率は $(1-p)^{N-1}p$ だから、

$$c(1-p)^{N-1}p$$

となる。よって、B_N^- は、漸化式

$$B_N^- = c(1-p)^{N-1}p + \sum_{k=1}^{N-1}(B_k^- + cp)(1-p)^{N-1-k}p \quad (N \geq 2) \tag{2}$$

を満たすことになる。ここから B_N^- を求めてみよう。

$N \geq 3$ に対し、$B_N^- - (1-p)B_{N-1}^-$ を考えると、(2) より

$$(1-p)B_{N-1}^- = c(1-p)^{N-1}p + \sum_{k=1}^{N-2}(B_k^- + cp)(1-p)^{N-1-k}p$$

なので、

$$B_N^- - (1-p)B_{N-1}^- = (B_{N-1}^- + cp)p$$

となることがわかり、よって、

$$B_N^- = (1-p)B_{N-1}^- + (B_{N-1}^- + cp)p = B_{N-1}^- + cp^2$$

となるので、B_N^- は $N \geq 2$ では公差 cp^2 の等差数列となる。B_2^- は、(1), (2) より、

$$B_2^- = cp(1-p) + (B_1^- + cp)p = B_1^- p + cp = cp + cp^2 = B_1^- + cp^2$$

なので、ここも同じ公差であり、結局

$$B_N^- = cp + (N-1)cp^2 \quad (N \geq 1) \tag{3}$$

となることがわかる。

A_N^- は、B_N^- に N 回目が裏のものを追加すればよいが、その追加分は C_0, C_k の N 回目を裏にしたものだから、賭けの回数は同じで確率だけが最後の 1 回分変わる。よって、

$$A_N^- = B_N^- + c(1-p)^N + \sum_{k=1}^{N-1}(B_k^- + cp)(1-p)^{N-k}$$

となることがわかる。(2) より、和の部分を B_N^- で表せば、

$$A_N^- = B_N^- + c(1-p)^N + \frac{1-p}{p}(B_N^- - c(1-p)^{N-1}p) = B_N^- + \frac{1-p}{p}B_N^- = \frac{1}{p}B_N^-$$

となるので、結局 (3) より

$$A_N^- = c + (N-1)cp \quad (N \geq 1) \tag{4}$$

と表されることになる。

なお、上では「最後に」1 回追加する形で漸化式を考えたのでだいぶ複雑になったが、「先頭に」1 回追加すると考えればむしろやさしくなる。1 回目に裏が出れば、参加回数はその後の $(N-1)$ 回の参加回数と同じで、1 回目に表が出れば、参加回数は 1 回増えることになるので、A_N^- と A_{N-1}^- の差は cp となり、よって (1) より (4) が得られることになる。

5 賞金の期待値

次は、賞金の期待値 A_N^+ の方を考える。なお、A_N^+ は、4 節の最後に書いたように先頭に 1 回追加すると考えても最後に追加すると考えても、賞金はその続きのものとの関係で変わってしまうので、どちらの形で考察してもさほど難しさに違いはない。

4 節と同じく、C_0, C_k に場合分けして、B_N^+ を求めることから始める。C_0 の場合は、賞金は Ab^{N-1} で、確率は $p(1-p)^{N-1}$ なので、

$$Ab^{N-1}p(1-p)^{N-1} = ApR^{N-1}$$

が C_0 に対応する B_N^+ の部分となる。ここで $R = b(1-p)$ としたが、元の問題では $b = 2, p = 1/2$ だから、その場合は $R = 1$ であることに注意する。

C_k の場合は、賞金は最初の k 回の B_k^+ の分のすべての事象に Ab^{N-k-1} が追加されるが、B_k^+ の確率の和は p であるから、よって C_k に対する B_N^+ の部分は

$$(B_k^+ + Ab^{N-k-1}p)p(1-p)^{N-k-1} = B_k^+ p(1-p)^{N-k-1} + Ap^2R^{N-k-1}$$

となる。よって、B_N^+ の漸化式は

$$B_N^+ = ApR^{N-1} + \sum_{k=1}^{N-1} B_k^+ p(1-p)^{N-1-k} + Ap^2 \sum_{j=0}^{N-2} R^j \quad (N \geq 2) \tag{5}$$

となる。

B_N^- の場合と同様に、$N \geq 3$ に対して $B_N^+ - (1-p)B_{N-1}^+$ を考えれば、

$$\begin{aligned} & B_N^+ - (1-p)B_{N-1}^+ \\ &= ApR^{N-1} - Ap(1-p)R^{N-2} + B_{N-1}^+ p + Ap^2 \sum_{j=0}^{N-2} R^j - Ap^2(1-p) \sum_{j=0}^{N-3} R^j \end{aligned}$$

となり、整理すれば、

$$\begin{aligned} B_N^+ - B_{N-1}^+ &= ApR^{N-2}(R - 1 + p) + Ap^2(1-p)R^{N-2} + Ap^3 \sum_{j=0}^{N-2} R^j \\ &= ApR^{N-2}(R - (1-p)^2) + Ap^3 \sum_{j=0}^{N-2} R^j \end{aligned} \tag{6}$$

となる。

B_2^+ は、(オ、オ) と (ウ、オ) なので、

$$B_2^+ = 2Ap^2 + Abp(1-p) = 2Ap^2 + ApR$$

となり、よって

$$B_2^+ - B_1^+ = 2Ap^2 + ApR - Ap = Ap(R - 1 + 2p)$$

となるが、これは (6) の $N = 2$ の式に等しいので、(6) は $N \geq 2$ で成り立つことになる。

今後この一般の式 (6) のまま計算すると面倒なので、元の条件の $R = 1$ を代入すると、

$$B_N^+ - B_{N-1}^+ = Ap(1 - (1-p)^2) + Ap^3(N-1) = Ap^3N + 2Ap^2(1-p) \tag{7}$$

となり、これにより、B_N^+ は

$$\begin{aligned} B_N^+ &= B_1^+ + \sum_{k=2}^{N}(B_k^+ - B_{k-1}^+) = Ap + \sum_{k=2}^{N}(Ap^3k + 2Ap^2(1-p)) \\ &= Ap + Ap^3\left(\frac{N}{2}(N+1) - 1\right) + 2Ap^2(1-p)(N-1) \\ &= \frac{Ap}{2}\{p^2(N-1)(N-2) + 4p(N-1) + 2\} \end{aligned}$$

と書けることがわかる。

さて次は A_N^+ だが、これは B_N^+ に C_0, C_k の最後のものを裏に変えた場合のものを加えればよいが、その場合は最後の回は賞金がないので、結局

$$A_N^+ = B_N^+ + \sum_{k=1}^{N-1} B_k^+(1-p)^{N-k} \tag{8}$$

となる。この和の部分を B_N^+ を使って表せば、

$$\begin{aligned} A_N^+ &= B_N^+ + \frac{1-p}{p}\sum_{k=1}^{N-1} B_k^+(1-p)^{N-k-1}p \\ &= B_N^+ + \frac{1-p}{p}\left(B_N^+ - ApR^{N-1} - Ap^2\sum_{j=0}^{N-2} R^j\right) \\ &= B_N^+ + \frac{1-p}{p}(B_N^+ - Ap - Ap^2(N-1)) \\ &= \frac{1}{p}B_N^+ - A(1-p)(1+p(N-1)) \\ &= \frac{A}{2}\{p^2(N-1)(N-2) + 4p(N-1) + 2\} - A(1-p)(1+p(N-1)) \end{aligned}$$

となり、結局

$$A_N^+ = \frac{A}{2}Np(p(N-1)+2) \tag{9}$$

となる。

元の問題に戻って考察を行うと、$p = 1/2$ であるから、収入の期待値は (4), (9) より

$$A_N = A_N^+ - A_N^- = \frac{A}{8}N(N+3) - \frac{c}{2}(N+1)$$

となる。$A = 10,\ N = 1000$ だと、賞金は

$$\frac{10 \times 1000 \times 1003}{8} \approx \frac{10^7}{8} = 125 \text{ 万円}$$

で、参加費用は $c = 10000$ 円より

$$\frac{10000 \times 1001}{2} \approx \frac{10^7}{2} = 500 \text{ 万円}$$

となって、380 万円位の損失となる。

N が大きいときは A_N はほぼ

$$A_N \approx \frac{A}{8}N^2 - \frac{c}{2}N = \frac{N}{8}(AN - 4c) \tag{10}$$

で近似されるので、賞金と参加料が釣り合うのは、$N = 4c/A = 4000$ 回位となる。

(10) は N の 2 次式なのでそれほど早くは増大しないが、5 千回だと、$A_{5000} \approx 5\times10^7/8 =$ 625 万円、1 万回だと、$A_{10000} \approx 6 \times 10^8/8 = 7500$ 万円となる。

1 回投げるのに 5 秒かかるとすれば、1 万回で 5 万秒 = 13 時間 53 分だから、約半日で終わる。意外に時間がかからずに一財産は得られることになるわけである。

6 最後に

この問題については、以前知人に「時間制限が問題なのでは」、という意見をもらったことがあり、今回それを思い出して、実際に繰り返し行った場合にはどれくらいで元が取れることになるのかなどを考察してみた。

理論的には無限大の期待値でも、ジャンボ宝くじと似たような程度の賞金と確率であるから、1 万円だと高くて参加しづらいだろうこと、また繰り返しやるのであれば、思ったよりも短い時間で元が取れ、一財産かせげることになることもわかった。

インターネット上には、賭けを受ける人が無限に支払えるはずはないとして金額に上限を設けて損であることを説明するものや、「直感的に損をすることは明らか」のような感覚的な意見、あるいは幾何平均による計算、人間の感覚の対数性などを用いた説明などが見られるが、いずれも「儲からない」という立場の結論が多い。よって、そうでもない、という今回のこのような考察も、それなりに意味があるのではないかと思う。