スポーツを見ることの面白さについて

竹内　啓

1. スポーツを見る楽しみ

スポーツの楽しみの中には，自分ですることのほかに，見る楽しみも大きいことは，改めていうまでもない．

スポーツを見る楽しみの中には，もちろん技術的にすばらしいプレーやプレイヤーの間の微妙なかけひきというような，数学的定式化をすることが不可能に近い部分も非常に大きい．しかしその中のかなりの部分はやはり勝敗についての関心にあるといってもよいであろう．そうしてこの部分は数学的に扱うことができるように思われる．そこでこの問題をいろいろな角度から考えてみよう．

たけうち　けい　東京大学　経済学部

2. 「興味度」の定義

スポーツの試合を見ているとき，一方がむやみに強くて必ず勝つのは面白くない．横綱と前頭5枚目のとり組みや，東京6大学のT大学とその他の大学との試合があまり興味をひかないのはこのためである．しかしそれでもたまには前頭5枚目の相撲が横綱に勝って「金星」をあげたり，T大学がW大学に連勝したりすると，これは興奮をまきおこす．

「スポーツのOR PART-II」の特集に当って

東京工業大学　鳩山　由起夫

昨年(Vol. 24) 4月号にスポーツのORを特集しましたところ，手前味噌ながら若干の反響をいただき，それではと柳の下に狙いを定めて第2弾を放つことになりました．ところが，日頃ストラテジーに口うるさいOR関係の方々も，時日を費して書くとなるといささか抵抗を示され，結局半数ほどを前回担当してくださった方に執筆していただくことになりましたことをお断わりしておきます．

スポーツをOR的に分析する観点はいろいろあります．内から眺めた戦術論は多くの読者の興味の中心かも知れません．外から眺めて共通の要素に探りを入れることには山師に似た喜びがあるでしょう．また，健康の面からスポーツを捉えることは，次第にスポーツをしなくなってきている現代の人間にとってきわめて重要な問題と言えるでしょう．

さらに本特集では，プロ野球界で管理を実践し，V9という形で完成された元巨人軍の川上氏とのインタビューを掲載します．ドラフト制により均等化されたチームの中で差を生むことに，管理能力の寄与はますます大となっていくでしょう．程度の差はあれ，平均化された社会にいる(と信じている)私どもにとり，自己管理をも含めて管理の重要性，むずかしさが痛感されます．

そうするとスポーツを見ること，その中でも勝敗についての興味は，「意外なこと」「予想外なこと」あるいは「予測できないこと」が起ることを知るという点にあるといってもよいであろう．

そこで次のように考えよう．いまあることが起る確率を p とするとき，そのことが実際に起ったときに感ずる「驚き」「意外さ」等を p の関数として $\phi(p)$ で表わすことにしよう．いうまでもなく ϕ は p の減少関数であると考えられる．また，

$$\phi(1)=0$$

としてもよいであろう．そうするとすべての $0<p<1$ に対して，

$$\phi(p)>0 \qquad \phi'(p)<0$$

となる．

ここで後の便宜上次の2つのことを仮定する．

1° $\phi(p)$ は凸関数である．すなわち $\phi''(p)>0$

2° $p\phi(p)$ は凹関数である．すなわち，

$$2\phi'(p)+p\phi''(p)<0$$

いま2人のプレイヤーあるいは2つのチームがゲームをするとき，一方が勝つ確率を p，他方が勝つ確率を q とすると，試合結果の「驚き」の期待値は，

$$I(p)=p\phi(p)+q\phi(q)$$

で与えられる．これを「興味度」と呼ぶことにしよう．

そうすると条件2°のもとでは，$(p+q)/2=1/2$ だから，

$$\frac{1}{2}\left\{p\phi(p)+q\phi(q)\right\}\leqq\frac{1}{2}\phi\left(\frac{1}{2}\right)$$

か成り立つ．すなわち，

$$I(p)\leqq\phi\left(\frac{1}{2}\right)=I\left(\frac{1}{2}\right)$$

となる．いいかえれば「興味度」は $p=\frac{1}{2}$ すなわち両者の実力が伯仲している時に最大になる．これはすでに述べたことに対応している．

特に $\phi(p)=-\log p$ とおけば，$I(p)$ はいわゆる情報量になる．

3. ゲームの「途中」

次に1つのゲームについて，その「進行」を考えよう．「最初から一方的でつまらなかった」とか「何度も逆転があって手に汗にぎる熱戦であった」とかいうことがしばしばいわれる．このようなことを数学的にどのように表現できるかを「興味度」の点から考えてみよう．

いまあるゲームの特定の「途中時」(たとえばサッカーの前半終了時，野球の5回裏終了時など)の状態を X で表わすことにしよう．X の値は確率的に変動するものであることはいうまでもない．いま X が与えられたときに，一方が勝つ条件確率を $p(X)$ としよう．そうするとこの時点での「興味度」は，

$$I(p(x))=p(x)\phi(p(x))+q(x)\phi(q(x))$$

となる．したがって「途中時」における「興味度」の期待値は $I(p(x))$ の X に関する期待値をとって，

$$E[I(p(x))]$$

と表わすことができる．ここで $I(p)$ が p に関して凹関数であるという仮定を用いると，

$$E[I(p(x))]\leqq I\{E(p(x))\}=I(p)$$

となる．つまり途中時の「興味度」は最初の興味度より一般に小さくなっている．このことは当然であって，ゲームの途中経過がわかってしまえば，それだけゲームの結果は予測しやすくなっているはずだからである．

もし $E[I(p(x))]$ が非常に小さくなってしまえば，もはや試合の興味はなくなってしまっているといってよいであろう．

いま3セットで1ゲームになっている試合を考えよう．いうまでもなく先に2セット取ったほうが勝になる．A，B 2チームが試合をするとし，あるセットをAがとる確率を π，Bが取る確率を，$1-\pi$ としよう．そうするとゲームにAが勝つ確率は，

$$p=\pi^2+2\pi^2(1-\pi)=\pi^2(3-2\pi)$$

Bが勝つ確率は,

$$q=(1-\pi)^2+2\pi(1-\pi)^2=(1+2\pi)(1-\pi)^2$$

となる.

いま1セット終ったところで考えると，Aが第1セットを取っていれば，Aが勝つ確率は,

$$p(\mathrm{A})=\pi+\pi(1-\pi)=\pi(2-\pi)$$

Bが勝つ確率は,

$$q(\mathrm{A})=(1-\pi)^2$$

もしBが第1セットを取っていれば，A，Bが勝つ確率はそれぞれ,

$$p(\mathrm{B})=\pi^2 \qquad q(\mathrm{B})=1-\pi^2$$

となる.

いま $\pi=0.5$ とすると,

$$p(\mathrm{A})=0.75=q(\mathrm{B}) \qquad q(\mathrm{A})=0.25=p(\mathrm{B})$$

であるから，いま $\phi(p)=-\log_2 p$ とすると，無条件の「興味度」すなわち情報量は,

$$I=-\log_2\frac{1}{2}=1.000$$

である．これに対して第1セット終ったときの情報量は(どちらが第1セットを取ったとしても),

$$I(\mathrm{A})=I(\mathrm{B})=-\frac{1}{4}\log_2\frac{1}{4}-\frac{3}{4}\log_2\frac{3}{4}$$
$$=\frac{1}{4}\times 2+\frac{3}{4}\times 0.4150=0.8113$$

となる．ゲームが終ってしまったときの情報量は0であるから，このことは第1セットが終った段階ではまだ試合の興味はほぼ80%残っているといってよいことになる.

また第2セットを終ったところで考えると，ゲームは終ってしまっている確率は1/2．このときは情報量0．また最初の2セットが1—1ならば，残りの1セットでゲームが決まり一方が勝つ確率は1/2だから，そのときの情報量は1．したがってこの場合の平均情報量は,

$$\frac{1}{2}\times 1=0.5$$

となる.

同様に考えると，5セットゲームについては,

第1セット終了時の情報量は,

$$-\frac{11}{16}\log_2\frac{11}{16}-\frac{5}{16}\log_2\frac{5}{16}=0.8960$$

第2セット終了時の情報量は,

一方が2セットをつづけて取った場合,

$$-\frac{1}{8}\log_2\frac{1}{8}-\frac{7}{8}\log_2\frac{7}{8}=0.5436$$

1—1の場合，1であるから，期待値は,

$$0.5436\times\frac{1}{2}+1\times\frac{1}{2}=0.7718$$

第3セット終了時の情報量は3—0の場合0，2—1の場合,

$$-\frac{1}{4}\log_2\frac{1}{4}-\frac{3}{4}\log_2\frac{3}{4}=0.8113$$

であるから，期待値は,

$$0\times\frac{1}{4}+0.8113\times\frac{3}{4}=0.6085$$

となる.

4. ゲームの「半分」の時点での情報量

同様の考え方で，もっと一般的な形で試合の途中における情報量を考えることができる.

いまゲームが次のような構造を持つと仮定しよう．すなわち双方が何らかの形で点数を取っていって，一定の時間，あるいは「回数」たったとき，点数の多いほうが勝とするのである．サッカー，野球，ゴルフ（これは打数の少ないほうが勝であるが）等はこのような構造を持っている．このとき次のことを仮定する.

1) 両者の得点の分布は互いに独立である.

2) 途中のある時点までの得点と，それ以後の得点は互いに独立である.

このとき，特定の「途中」時点における，両者(AとBとしよう)の得点をそれぞれ X_A，X_B，それ以後の両者の得点をそれぞれ Y_A，Y_B とすると，$X_A+Y_A>X_B+Y_B$ ならばAの勝ち，$X_A+Y_A<X_B+Y_B$ ならばBの勝ちとなる．$X_A+Y_A=X_B+Y_B$ のときはそこまででは勝負がつかないことになるから，何らかの形での「延長」(あるいは「プレイオフ」)が必要になる．ここでは話を簡単にするためにこのときは「くじ」で勝負を決めることにしておこう．両者の実力が等しければこれは「延長戦」にすることと同等であろう.

いま$Z=X_{\mathrm{A}}-X_{\mathrm{B}}$, $W=Y_{\mathrm{A}}-Y_{\mathrm{B}}$ とおき，それぞれの確率分布を,

$$P\{Z=z\}=p(z)$$
$$P\{W=z\}=q(z)$$
$$Q(z)=P\{W>z\}=1-P\{W\leqq z\}$$

と表わせば，$Z=z$ が与えられたとき，A の勝つ確率は，$Q(-z)+\frac{1}{2}q(-z)$ となることは明らかである．そこで改めて,

$$Q^*(z)=Q(-z)+\frac{1}{2}q(-z)$$

とおくと $Z=z$ が与えられたときの情報量は,

$$Q^*(z)\log_2 Q^*(z)+(1-Q^*(z))\log_2(1-Q^*(z))$$

となる．したがって平均情報量は,

$$-\sum_z p(z)\{Q^*(z)\log_2 Q^*(z)+(1-Q^*(z))\log_2(1-Q^*(z))\}$$

となる.

いま実力が「伯仲」している場合を考えよう．そのとき Z および W の分布は原点に関して対称となるはずである．そうしてこのときちょうど中央の時点を考えると，Z と W の分布は等しくなる．いま「同点」の面倒をさけるために，Z および W の分布をほぼ連続分布で近似できるものとし，その密度関数を $f(z)$ とすれば,

$$Q^*(z)=\int_{-z}f(z)\,dz=\int^{z}f(z)\,dz$$

となるから，平均情報量は,

$$I_{\frac{1}{2}}=-\int f(z)[Q^*(z)\log_2 Q^*(z)+(1-Q^*(z))\log_2(1-Q^*(z))]\,dz$$

と与えられるが，ここで $u=Q^*(z)$ と変換すれば,

$$I_{\frac{1}{2}}=-\int_0^1(u\log_2 u+(1-u)\log_2(1-u))\,du$$

となる．ところが,

$$\int_0^1 u\log_2 u\,du=\int_0^1(1-u)\log_2(1-u)\,du$$
$$=\frac{1}{\log_e 2}\int_0^1 u\log u\,du$$
$$=\frac{1}{\log_e 2}\left[\frac{1}{2}u^2\log u-\frac{u^2}{4}\right]_0^1=-\frac{1}{4\log_e 2}$$

したがって,

$$I_{\frac{1}{2}}=\frac{1}{2\log_e 2}=\frac{1}{2\times 0.693147}=0.72148$$

となる.

$I(p)$ がちょうどこの値になるような p を求めると，それは $p=0.19975$ となる．あるいはほぼ $p=0.2$ と考えてよい．いいかえれば実力の伯仲したゲームの間の試合が半分経過したとき，残された部分についての勝敗に関する興味度は，ほぼ一方の勝つ確率が80%となる場合の試合の興味に等しいといってよいことになる.

実はこの場合「リード」しているほうが勝つ確率は,

$$\int_{z>0}f(z)Q(z)\,dz+\int_{z<0}f(z)(1-Q(z))\,dz$$
$$=2\int_{z>0}f(z)Q(z)\,dz=2\int_{\frac{1}{2}}^1 u\,du=u^2\Big|_{\frac{1}{2}}^1=\frac{3}{4}$$

となるから，もしどちらがリードしているかだけが知られたとすれば，このときの情報量は,

$$-\frac{3}{4}\log_2\frac{3}{4}-\frac{1}{4}\log_2\frac{1}{4}=0.8113$$

となるから，この値と上記の0.7215との差は，点差自体が知られたことの情報量に対応すると考えることができる.

5. ゲームの進行による情報量の変化

次にまず一般にゲームの進行にともなう情報量の変化を考えよう.

今度は Z, および W が正規分布にしたがうと想定しよう．そうすると全体のゲームの中でちょうど $t(0<t<1)$ だけの時間が経過したとき,

$$E(Z)=E(W)=0$$
$$V(Z)=t\sigma^2 \qquad V(W)=(1-t)\sigma^2$$

と考えられる．すなわち点数の差の時間的経過は1次元のブラウン運動になると考える.

実際にスポーツの点数は離散的であり，かつその値そのものもあまり大きくならないから，それについて正規分布を想定するのは少々非現実的であるが，しかし第1次の近似としてはこれでもよいであろう.

そうすると t 時点での情報量は,

$$I_t = -2\int \frac{1}{\sqrt{t}}\,\phi\left(\frac{z}{\sqrt{t}}\right)\Phi\left(\frac{z}{\sqrt{1-t}}\right)$$

$$\log_2\left[\Phi\left(\frac{z}{\sqrt{1-t}}\right)\right]dz$$

$$= -2\int \sqrt{s}\,\phi(\sqrt{s}\,z)\Phi(z)\log_2\{\Phi(z)\}\,dz$$

と表わされる．ただし，

$$\phi(z) = \frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} \qquad \Phi(z) = \int_{-\infty}^{z}\phi(u)\,du$$

$$s = (1-t)/t$$

である．

数値的に計算すると，

$I_{\frac{1}{4}} = 0.8748 \qquad I_{\frac{1}{3}} = 0.8275$

$I_{\frac{1}{2}} = 0.7215$(これはすでに求めた)

$I_{\frac{2}{3}} = 0.5928$

等となる．

いいかえると強さが等しい2つのチームの間でゲームが2/3だけ進行したときは，ゲームの勝敗に関する「興味」はほぼ60%くらい残っているということになる．

もう少しモデルを現実化して，たとえば野球などの場合，得点の確率分布から，各回の終了時の情報量を計算することも可能である．

6. 多くのチームの中から「優勝者」を決める

今度は全国高校野球大会や，プロ野球のリーグ戦のように，いくつかのチームの中から，1つの優勝者を決める場合を考えよう．

いまkチームあるとして，何らかの方式により優勝者を決めるとして，それぞれが優勝する確率を$p_1, \cdots, p_k$すると，この場合の「興味度」は，

$$I = \sum_{i=1}^{k} p_i \phi(p_i)$$

で与えることができるであろう．

ここでϕが先の条件を満たすとすると，この場合にもIが最大になるのは，

$$p_1 = \cdots = p_k = 1/k$$

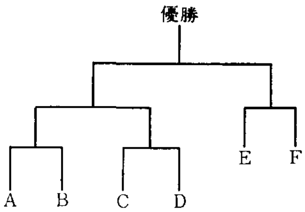

図 1

の場合であることがわかる．

特に$\phi(p) = -\log_2 p$とすれば，

$$I_{\max} = \log_2 k$$

となる．

いますべてのチームが同じ強さであるとすると，事前には情報量は$\log_2 k$であると考えられる．

ところがいま高校野球のような「トーナメント戦」によると図1のように試合を行なうとすると，A，B，C，Dが優勝する確率はそれぞれ1/8，E，Fが優勝する確率は1/4だから，情報量は，

$$-\frac{4}{8}\log_2\frac{1}{8} - \frac{4}{8}\log_2\frac{1}{8} = 2.5$$

となる．これに対して最初の情報量は$\log_2 6 = 2.5850$であった．いいかえれば，

$$2.585 - 2.5 = 0.085$$

だけの情報量あるいは「興味」は，組み合わせの「不平等」のために，組み合わせの決定の段階で失われてしまうことになる．もっと極端にして次の図2のようにすれば情報量は，

$$\frac{1}{2}\log_2 2 + \frac{1}{4}\log_2 4 + \frac{1}{8}\log_2 8 + \frac{1}{16}$$

$$\log_2 16 + \frac{2}{32}\log_2 32 = 1.9375$$

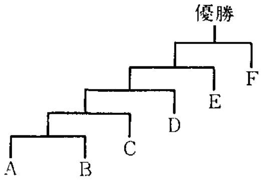

図 2

となって興味の失われたほうの度が大きくなる．

したがってなるべく興味を失わないようにする

には組み合わせはできるだけ公平であるほうがよいということになる．チームの数kが2^mという形で，どのチームも優勝までにk回勝たねばならないようにすれば，情報量の損失は起らない．

またこのとき第j回戦$(1\leqq j\leqq m)$の結果によって減少する情報量は簡単な計算によって$\left(\frac{1}{2}\right)^{j-1}$であることがわかる．したがって$j$が大きいほど試合の興味がそれだけ大きくなるはずである．

7. 強さが一様でない場合──「公正」の原理

これまでは強さが伯仲している場合を主として考えてきた．しかし強さが一様でない場合，上記の議論によれば，なるべく勝つ確率が一様にでるように試合の組み合わせやルールを決めるのが興味を増す方法であると考えられるであろう．いわゆる「ハンディキャップゲーム」はそのための工夫であると考えられる．

しかしながら他方，「強いほうが勝つ」のはスポーツの「公正さ」の基本原則であるとすれば「ハンディをつけるのはあまりよいこととは思われないし，また勝敗がほとんど偶然，すなわち「運」で決まることも「面白くない」と思われるにちがいない．

実際いわゆる「シード制」あるいは極端な形では「チャンピオン制」では，強いものが有利な立場におかれている．チャンピオンに1度なってしまえば次の選手権戦には無条件で出られるのに対して，挑戦者は予選を勝ち抜いてこなければならない．このような制度は「誰が次のチャンピオンになるか」を知る情報量の観点からは，損失の大きい方式であるといわねばならない．しかしながらこのような方式は実力第1のものをチャンピオンにするという観点からは，有効な方式であると思われる．すなわち実力が時間的に急激に変化するものでないかぎり，1度勝ったもの，好成績を収めたものを，次の機会には優遇するということは，実力相応の順位と定めるという点で合理的であるといえる．

しかしこのことはいわば「過去の実績にもとづく予測があたる」ことを望ましいとする点で「意外なことが起る」ことを「面白さ」とする考え方とは正相対のように思われる．

現実のスポーツにおける関心は，このような2つの相反する原則のトレードオフから成り立っているのであろうか，あるいはこれを1つの基準にまとめる方法があり得るであろうか．この点についていろいろと考えてみたが，私にはどうも明確な定式化ができなかった．スポーツと数理の双方に関心ある方々に考えていただきたいと思う．

●ニュース

●昭和55年国際通信研究奨励金

対象：国際電気通信の進歩改善および合理的利用のための独創的な内容であって，その成果が国際電気通信業務の将来の発展に寄与するであろうと認められるその．

奨励金額：総額　1500万円（5件）

申込方法：所定用紙（学会事務にあります）を使用のこと．締切：昭和55年10月31日．

発表：昭和56年3月上旬予定．

●新刊雑誌の紹介

雑誌名：IIASA REPORTS (International Institute for Applied Systems Analysis)

季刊雑誌で，IIASA, National Member Organizations の報告，関連文献等を内容に200～300ページが予定されている．

●日本学術会議 第12期会員選挙候補者の推薦について

今秋行なわれます日本学術会議第12期会員選挙にあたり，当学会理事会は，京都大学教授椹木義一氏を候補者として推薦することを決定いたしました．